High-density genetic map of *Miscanthus sinensis* reveals inheritance of zebra stripe

**SIYAO LIU**1, **LINDSAY V. CLARK**1, **KANKSHITA SWAMINATHAN**2, **JUSTIN M. GIFFORD**1, **JOHN A. JUVIK**1 and **ERIK J. SACKS**3

1Department of Crop Sciences, University of Illinois at Urbana-Champaign, 1201 W Gregory Dr., Urbana, IL 61801, USA, 2Energy Biosciences Institute, Institute for Genomic Biology, University of Illinois at Urbana-Champaign, 1206 W Gregory Dr., Urbana, IL 61801, USA

**Abstract**

*Miscanthus* is a perennial C₄ grass that has recently become an important bioenergy crop. The efficiency of breeding improved *Miscanthus* biomass cultivars could be greatly increased by marker-assisted selection. Thus, a high-density genetic map is critical to *Miscanthus* improvement. In this study, a mapping population of 261 F₁ progeny was developed from a cross between two diploid *M. sinensis* cultivars, ‘Strictus’ and ‘Kaskade’. High-density genetic maps for the two parents were produced with 3044 newly developed single nucleotide polymorphisms (SNPs) obtained from restriction site-associated DNA sequencing, and 138 previously mapped GoldenGate SNPs. The female parent (‘Strictus’) map spanned 1599 cM, with 1989 SNPs on 19 linkage groups, and an average intermarker spacing of 0.8 cM. The length of the male parent (‘Kaskade’) map was 1612 cM, with 1821 SNPs, and an average intermarker spacing of 0.9 cM. The utility of the map was confirmed by locating quantitative trait loci (QTL) for the zebra-striped trait, which was segregating in this population. Three QTL for zebra-striped presence/absence (zb₁, zb₂ on LG 7, and zb₃ on LG 10) and three for zebra-striped intensity (zb₁₁, zb₁₂, zb₁₃ on LGs 7, 10, 3) were identified. Each allele that caused striping was recessive. Incomplete penetrance was observed for each zb QTL, but penetrance was greatest when two or more zb QTL were homozygous for the causative alleles. Similarly, the intensity of striping was greatest when two or more zbi QTL were homozygous for alleles that conferred the trait. Comparative mapping indicated putative correspondence between zbi and/or zb₂ on LG 10 to previously sequenced genes conferring zebra stripe in maize and rice. These results demonstrate that the new map is useful for identifying marker–trait associations. The mapped markers will become a valuable community resource, facilitating comparisons among studies and the breeding of *Miscanthus*.

**Keywords:** bioenergy, chlorophyll, chloroplast, gene, linkage QTL, mapping, Poaceae, restriction site-associated DNA sequencing, single nucleotide polymorphisms, synteny

**Received 10 February 2015; accepted 26 March 2015**

**Introduction**

*Miscanthus* has more than a 100-year history as an ornamental grass in the USA (Anonymous, 1876; Bailey & Miller, 1901), with many cultivars currently sold by the horticulture trade, but it has only recently become the subject of modern breeding efforts as a bioenergy feedstock crop (Scurlock, 1999; Clifton-Brown *et al.*, 2004; Somerville *et al.*, 2010; Sacks *et al.*, 2013). *Miscanthus* is a perennial, self-incompatible, C₄ grass that is closely related to sugarcane but can produce high biomass yield in temperate climates, making it a good choice for much of the USA (Lewandowski *et al.*, 2000; Clifton-Brown *et al.*, 2008). So far, the most promising *Miscanthus* biofuel candidate for the emerging US bioenergy industry is a single, high-yielding, sterile genotype of *M. × giganteus* (2n = 3x = 57), a nothospecies derived from a cross between diploid *M. sinensis* (2n = 2x = 38) and tetraploid *M. sacchariflorus* (2n = 4x = 76) (Linde-Laursen, 1993; Hodkinson *et al.*, 2002; Heaton *et al.*, 2008; Glowacka *et al.*, 2014a,b). Dependence on a single clonally propagated genotype for biomass production poses risks because a newly emerged disease or pest strain could result in severe damage to plantings. Moreover, there is a need to breed new cultivars with improved cold tolerance and high yield potential for cold-temperate environments such as the Midwest USA.

Substantial genetic variation has been found among the two parental species of *M. × giganteus*, *M. sinensis* and *M. sacchariflorus* (Jorgensen & Muhs, 2001; Clark *et al.*, 2014, 2015; Glowacka *et al.*, 2014a,b). By choosing parents with desired traits, we expect that superior *Miscanthus* cultivars may be bred. Thus, an understanding of the
genetics behind key traits in Miscanthus is critical to enabling selection of parents. For many crops, genetic maps have provided a useful tool for identifying associations between traits of interest and molecular markers, and such associations have enabled marker-assisted selection (MAS; Collard & Mackill, 2008). MAS would be especially valuable for increasing the efficiency of breeding Miscanthus because phenotypic selection for most of this perennial crop's traits must typically be done in the second and third years of field trials, whereas MAS can be completed on 1- to 2-month-old seedlings before they are transplanted to the field.

Genetic mapping in Miscanthus is challenging due to its obligate outcrossing nature that results in high levels of heterozygosity. However, the two-way pseudo-test-cross mapping strategy (Grattapaglia & Sederoff, 1994) has provided a useful approach for outcrossing species and has been applied to construct genetic maps in grass species such as sugarcane (Saccharum spp.; Al-Janabi et al., 1993; Grivet et al., 1996; Ming et al., 1998), ryegrass (Lolium spp.; Inoue et al., 2004), and switchgrass (Panicum virgatum L.; Missaoui et al., 2005; Okada et al., 2010). To date, several genetic maps for Miscanthus have been published using this approach (Atienza et al., 2002; Kim et al., 2012; Ma et al., 2012; Swaminathan et al., 2012). The first Miscanthus genetic map, published by Atienza et al. (2002), used 257 randomized amplified polymorphic DNA (RAPD) markers spanning over 28 linkage groups (LGs); however, these maps were incomplete because Miscanthus has a basic chromosome number of 19. Moreover, the low reproducibility of RAPD markers would be disadvantageous to applying MAS in breeding, although quantitative trait loci (QTL) mapping was performed with this map (Atienza et al., 2003a,b,c,d). Kim et al. (2012) developed simple sequence repeat (SSR)-based genetic maps of M. sinensis and M. sacchariflorus, which mitigated the low repeatability problem in the initial map, and allowed them to identify a whole-genome duplication in Miscanthus relative to sorghum. However, the Kim et al. (2012) maps remained incomplete (23 LGs detected for M. sinensis and 40 LGs for M. sacchariflorus).

A framework genetic map for M. sinensis was recently developed at the University of Illinois, using 658 single nucleotide polymorphism (SNP) markers from GoldenGate assays and 210 SSR markers derived from sugarcane expressed sequence tags (Swaminathan et al., 2012). This map (Swaminathan et al., 2012) successfully identified all 19 LGs, confirmed the recent genome duplication in Miscanthus, revealed substantial synteny between Miscanthus and sorghum, and identified the chromosome fusion that resulted in \( n = 19 \) for Miscanthus from the \( n = 10 \) that is typical of the Andropogoneae; however, modest marker density of GoldenGate SNPs was a limitation to QTL mapping and MAS. Moreover, the ascertainment bias of GoldenGate SNPs in Swaminathan et al.'s study (2012) limited the usefulness of the markers for other populations. The most marker-dense map for Miscanthus published previously, included 3745 SNP markers on a composite genetic map of M. sinensis, obtained from genotyping-by-sequencing (Ma et al., 2012), which also supported the conclusions of genome duplication and synteny with sorghum. However, the methods and markers in Ma et al.'s study were proprietary (although their raw sequence data were made available through NCBI's Sequence Read Archive). Thus, a high-density genetic map with numerous fully accessible SNP markers is needed.

Zebra stripe is a leaf trait characterized by transverse yellow-green or white bands (perpendicular to the leaf axis). Zebra stripe has been observed in many grass species, including maize (Zea mays L.), rice (Oryza sativa L.), and sorghum [Sorghum bicolor (L.) Moench] (Demerec, 1921; Hayes, 1932; Quinby & Karper, 1942; Kinoshita & Takeamura, 1984; Kinoshita & Takahashi, 1991). In the ornamental horticulture trade, the alternating pattern of yellow and green is considered desirable, with M. sinensis 'Zebrinus' among the first Miscanthus genotypes to be imported into the USA in the 1870s (Anonymous, 1876). In Miscanthus, zebra-striped ornamental cultivars differ in the intensity of striping, with the most highly striped cultivars, such as 'Super Stripe' and 'Gold Bar' especially valued. This nonuniform pigment-deficiency has also been observed to be affected by environmental factors such as light, temperature, and the diurnal light/dark cycle, as well as developmental stage (He et al., 1999; Kusumi et al., 2000; Huang et al., 2009; Lu et al., 2012). In maize, three types of zebra striping have been described: zebra striping only in mature plants, zebra striping only in seedlings, and zebra necrosis (Demerec, 1921; Stroman, 1924; Hayes, 1932; Giesbrecht, 1965); all three produce yellow-green transverse stripes on the leaves as a result of the chlorophyll deficiency, but they differ in timing of appearance. In rice, zebra phenotypes can appear in a variety of colors, from pale green, yellow-green, white-green to albino (Kusumi et al., 2000; Zhao et al., 2014). Natural zebra-striped mutants have, to the best of our knowledge, not been reported in sorghum but have been induced by X-rays (Quinby & Karper, 1942), with the zebra-striped sorghum plants less vigorous than wild type green plants, perhaps due to lower photosynthetic capacity. The incomplete, often temporal, chlorosis in zebra-striped expression provides a valuable opportunity to study the genetics of chlorophyll and plastid development because complete knockout mutations would typically disable photosynthesis and thus be lethal.

Where studied, a single recessive gene model has been proposed for the inheritance of zebra striping, in
maize, sorghum, and rice (Demerec, 1921; Quinby & Karper, 1942; Wang et al., 2009). Only four genes for zebra stripe, two from maize and two from rice, have been cloned and their function and mechanism understood. A recent map-based cloning study has shown that zebra crossbands7 (zb7) gene in maize encodes 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate reductase (IspH, also called HDR or LytB), a key enzyme essential for early stage chloroplast development (Lu et al., 2012). Huang et al. (2009) characterized maize camouflage 1 (cfl) mutant, which displays a zebra banding pattern of nonclonal yellow-green sectors on the leaves and the yellow sectors can progress to bundle sheath cell-specific death. Huang et al. (2009) also cloned the cfl gene and determined that it encodes porphobilinogen deaminase, an enzyme in the tetrapyrrole biosynthesis pathway, which is responsible for producing heme and chlorophyll. In rice, zebra2 (z2), the first cloned gene responsible for yellow- and green-striped leaves, has been shown to encode carotenoid isomerase (CRTISO), which serves a key role in the carotenoid biosynthesis pathway (Fang et al., 2008; Chai et al., 2011; Zhao et al., 2014). Previous studies have reported seven mutations in the CRTISO gene in rice, and all had unique mutation sites, but each affected chloroplast development and photoprotection (Fang et al., 2008; Wei et al., 2010; Chai et al., 2011; Han et al., 2012; Liu et al., 2013; Zhao et al., 2014). Map-based cloning of zebra necrosis (zn) gene in rice revealed that zn encodes a thylakoid-bound protein that protects developing chloroplasts from photodamage during early stages of leaf development (Li et al., 2010).

In M. sinensis, Touchell et al. (2007) studied phenotypic segregation in F1, F2, and backcross populations of ‘Strictus’ (yellow striping perpendicular to the leaf axis) × ‘Variegatus’ (white striping parallel with the leaf axis) and concluded that zebra-striped presence/absence from the ‘Strictus’ parent was likely conferred by homozygous recessive alleles at a single locus, although variation in expression of the trait was likely due to other genetic and environmental factors. Until this study, zebra stripe has not been mapped in Miscanthus, although selection by horticulturalists for Miscanthus cultivars that differ in zebra-striped intensity may thus have unintentionally provided grass geneticists with a useful trove of mutants.

In this study, we report on using restriction site-associated DNA sequencing (RAD-seq) to develop high-density genetic maps for M. sinensis. High-throughput sequencing of RAD tags is a genotyping method that integrates SNP discovery and genotype calling into one step (Baird et al., 2008). RAD-seq has been used to construct high-density genetic maps for many crops, such as barley and wheat (Poland et al., 2012), rice (Spindel et al., 2013) and perennial ryegrass (Pfender et al., 2011). The objectives of this study were (1) to identify thousands of SNP markers for genotyping M. sinensis, (2) to construct high-density genetic maps of M. sinensis that integrate new SNPs from RAD-seq with the previously mapped but less numerous SNPs from GoldenGate assays, and (3) to confirm the utility of this map by locating QTL for zebra stripe in Miscanthus. This high-density genetic map for Miscanthus provides a community resource, enabling others to reproduce the mapped markers for their own research purposes. Moreover, this map is useful for identifying robust marker–trait associations, which will facilitate genetic improvement of Miscanthus.

Materials and methods

Mapping population and growing conditions

A cross between two diploid (2n = 2x = 38) M. sinensis cultivars, ‘Strictus’ (sourced from Emerald Coast Growers, Pensacola, FL) and ‘Kaskade’ (sourced from Digging Dog Nursery, Albion, CA, USA) produced 293 individuals that were used in this study. The cross was made in an isolated greenhouse bay at the University of Illinois in 2010 with ‘Strictus’ as the female parent. To initially confirm the parentage of the progeny, 30 random individuals along with the two parents were screened with 12 SSRs derived from sugarcane expressed sequence tags and intergenic sequences (Swaminathan et al., 2012), and the test result confirmed that ‘Strictus’ and ‘Kaskade’ were the true parents of this population. Because Miscanthus is self-incompatible, the female parent was not emasculated; however, infrequent progeny from selfing are possible, as the self-incompatibility system has been shown to be leaky (Hirayoshi et al., 1955). Analysis of GoldenGate and RAD-seq SNPs confirmed that 261 of the seedlings were F1 progeny, but 32 seedlings were identified as the product of self-pollination. Typically, Miscanthus individuals are highly heterozygous due to self-incompatibility and thus F1 progeny are expected to segregate for many traits. ‘Strictus’ was characterized by its yellow zebra-striped leaves, whereas the leaves of ‘Kaskade’ were entirely green.

Seeds were germinated in trays containing a peat-based potting mix (Metro-Mix 900, Sun Gro Horticulture, Agawam, MA, USA; 15–25% Canadian sphagnum peat moss, composted bark, perlite, vermiculite, dolomite lime, and ureaform fertilizer) in a greenhouse. When the seedlings reached sufficient size, each individual was divided into ramets and grown in 36-cell plug trays (PL-36-STAR®, T.O. Plastics, Clearwater, MN, USA) for subsequent planting in a replicated field trial. The field trial was a randomized complete block design with three replications. Plots consisted of single plants on 1.5 m centers. The trial was planted at the University of Illinois’ Energy Farm in Urbana, IL (40°3′57″ N, 88°11′43″ W; USDA hardiness zone 6) in Flanagan silt loam (fine, smectic, mesic, Aquic, Argudollis, 4–5% organic matter) on June 9, 2011. Irrigation was provided on an as-needed basis during the first (establishment) year to prevent drought stress. No supplemental irrigation was provided in subsequent years. In the spring of each year,
100.8 kg N ha$^{-1}$ fertilizer was applied. Pre-emergence and post-emergence herbicides were applied at planting and in the spring of each year (Atrazine, 2.8 kg ha$^{-1}$; S-metolachlor, 1.5 L ha$^{-1}$; and 2, 4-D, 1.8 L ha$^{-1}$), and additional hand-weeding was carried out as needed.

**Marker development and genetic map construction**

Genomic DNA was extracted from freeze-dried leaves via a CTAB (cetyltrimethylammonium bromide) protocol adapted from Kabelka et al. (2002). Extracted DNA was quantified with a Quant-iT dsDNA Picogreen Kit (Life Technologies Inc., Carlsbad, CA, USA), and DNA concentration was normalized to 100 ng μl$^{-1}$ for GoldenGate and RAD-seq analysis. RAD libraries for Illumina sequencing were prepared based on a protocol for sorghum, barley, and wheat (Poland et al., 2012), using 96 barcoded adapter sequences from Thurber et al. (2013). In brief, each DNA sample was digested with the rare-cutting enzyme PstI-HF (High-Fidelity; New England Biolabs Inc., Ipswich, MA, USA) and the common-cutting enzyme MspI (New England Biolabs Inc., Ipswich, MA, USA), then ligated to a unique barcoded adapter and a common adapter. Samples were pooled, and the 200-500 bp base pair (bp) size fraction was extracted from a 2% agarose gel after electrophoresis and purified using a Qiagen Gel Extraction Kit (Qiagen, Hilden, Germany). The purified DNA was PCR amplified using Kapa-HF Master Mix (New England Biolabs Inc., Ipswich, MA, USA), and the PCR product was extracted as above to eliminate primer-dimers. All RAD libraries were sequenced on a HiSeq 2000 (Illumina, San Diego, CA, USA) with 100 bp single-end reads at the DNA Sequencing Unit of the Roy J. Carver Biotechnology Center at the University of Illinois. Most F1 individuals from this mapping population were duplicated in a second round of libraries to obtain at least 500 000 read counts per sample, except for 36 individuals with low concentrations (< 50 ng μl$^{-1}$). Both parents were included in four libraries to obtain high read depth. All sequence data from this study have been deposited in the NCBI Sequence Read Archive (BioProject SRP053003).

To discover SNPs and call genotypes from RAD-seq, we used the Universal Network Enabled Analysis Kit (UNEAK) pipeline in TASSEL (version 3.0) standalone, due to its ability to distinguish heterozygous SNPs from paralogous loci in organisms lacking a reference genome (Lu et al., 2013), especially given the recent genome duplication in Miscanthus (Ma et al., 2012; Swaminathan et al., 2012). Additionally, three M. sinensis doubled haploid lines (Głowacka et al., 2012; Swaminathan et al., 2012) were used to further distinguish paralogous loci from heterozygous loci. To ensure that genotypes were called accurately, a minimum of five reads was required to call a homoygote for a given SNP if only a single allele was detected for a given individual. If fewer than five reads were present, a missing genotype was assigned to that individual to avoid the possible error of calling a genotype a homozygote when in fact it was a heterozygote.

RAD-seq yielded 1192.6 million sequences reads, with each RAD library of 96 DNA samples yielding close to 200 million reads. With no more than 10% missing genotype calls allowed per SNP, 10 398 SNPs were discovered via the UNEAK pipeline. The average read depth per F1 individual was 52 read counts per SNP, and each parent had 129 read counts per SNP. A total of 672 previously identified GoldenGate SNP markers, including 658 previously mapped SNPs (Swaminathan et al., 2012), were designed in an oligonucleotide pool assay (OPA) on the Illumina website. The 658 SNPs were used to link prior Miscanthus genetic maps with the new RAD-seq SNPs maps. The two parents and 76 F1 individuals were genotyped with this OPA at the Functional Genomic Unit of the Roy J. Carver Biotechnology Center at the University of Illinois. Genotypes were called manually based on signal intensities using Illumina Genome Studio software. Only markers showing clear polymorphism were used in the linkage analysis.

Because heterozygote under-calling was anticipated within parental genotypes, we confirmed RAD-seq SNP genotype of the parents based on the allele frequency distribution in the F1 population. Specifically, expected allele frequencies (0.25, 0.5, or 0.75) corresponded to segregation ratios that resulted from one of three possible combinations of parental genotypes: (1) the crosses of AA x Aa with 1 : 1 segregation of AA and Aa progeny, (2) Aa x Aa with 1 : 2 : 1 segregation of AA, Aa, and aa progeny, and (3) Aa x aa with 1 : 1 segregation of Aa and aa progeny. The coding scheme for both RAD-seq SNP and GoldenGate SNP markers followed the cross pollinator (CP) population type in JoinMap 4.1 (Van Ooijen, 2011).

Among the 10 398 RAD-seq SNPs identified, 2229 were eliminated from the data set because they appeared heterozygous in at least one of the three doubled haploid lines evaluated. In addition, 263 SNPs were removed due to missing genotypes in either of the two parents. To maintain the missing rate for each SNP of < 10% (via fewer than five reads for a given SNP if only a single allele detected), 1892 SNPs were eliminated, with 6014 SNPs retained for further filtering. Segregation of each SNP was tested for goodness of fit using x$^2$ values from the Locus Genotype Frequency calculation in JoinMap4.1 to identify departure from expected ratios of 1 : 1 or 1 : 2 : 1 depending on the marker class. SNPs that showed segregation distortion with P < 0.05 were eliminated from map construction. An additional 2634 SNPs were eliminated due to segregation distortion (P < 0.05), with 641 markers from the marker class AA x Aa, 1265 markers from the marker class Aa x Aa, and 737 markers from the marker class Aa x aa. In all, 3371 non-distorted SNPs were retained, of which the marker class was compared with the two parental genotypes. Among these 3371 SNPs, 86% matched their corresponding parental genotypes, and among the mismatched SNPs, 32% can be explained by heterozygote under-calling in the parental genotypes. Thus, a final set of 3056 RAD-seq SNPs were retained for genetic map construction. Of the 672 SNPs in the GoldenGate assay, 241 were informative (segregated in one or both parents) in our population. After 32 segregation distorted SNPs (P < 0.05), and 38 SNPs that had similarity > 0.99 were eliminated, 171 GoldenGate SNPs were retained for map construction.

Map construction was based on the 261 F1 individuals. The set of 3056 RAD-seq SNPs and 171 GoldenGate SNPs (3227 total) were separated into three segregation types: (1) 1399 SNPs heterozygous in ‘Strictus’, (2) 1219 SNPs heterozygous in ‘Kaskade’, and (3) 609 SNPs heterozygous in both parents.
Separate maps were produced for the female (‘Strictus’) and male (‘Kaskade’) parents. Genetic maps were constructed using JoinMap4.1 with CP population type (Van Ooijen, 2011). A minimum of independence logarithm of odds (LOD) score of 11 was used to define linkage groups. The use of GoldenGate SNPs from a previously published Miscanthus genetic map (Swaminathan et al., 2012) allowed for direct identification of linkage groups on the newly developed RAD-seq SNPs map. To ensure markers were assigned to the correct linkage groups, markers in suspect linkages (pairwise recombination frequency estimates larger than 0.6 calculated in JoinMap4.1) were not used for marker ordering, and thus 10 RAD-seq SNP markers and 13 GoldenGate SNPs were eliminated. Maternal and paternal maps were first constructed using the regression mapping algorithm in JoinMap4.1 (Van Ooijen, 2011). Marker orders in both parental maps were calculated using the Kosambi mapping function. A composite map was subsequently generated using the multipoint maximum likelihood mapping algorithm implemented in JoinMap4.1 (Van Ooijen, 2011), and map distance was estimated using the Haldane mapping function.

**Phenotyping of zebra stripe and QTL analysis**

Two sets of phenotypic data were taken on June 2, 2013: (1) zebra-striped presence/absence and (2) zebra-striped intensity. Zebra-striped intensity was visually scored with the following scale: 0, 0.05, 0.1, 0.3, 0.5, 0.7, 0.9 (Fig. 1); the two parents were used as controls with the green-leafed parent ‘Kaskade’ scored as 0, and the highly striped parent ‘Strictus’ scored as 0.9. The zebra-striped intensity score estimated the relative leaf area that was striped rather than green; it does not refer to any differences in color saturation, or hue of the stripe, which was typically yellow, although late in the season sometimes included brown areas (presumably damage from exposure to full sun over the course of the growing season). Analyses of variance (ANOVA) were conducted, and least squares means (LS-means) were calculated using SAS Procedure GLM (SAS® 9.3, SAS Institute Inc., Cary, NC, USA). Mean separation among genotypic classes was performed using Tukey–Kramer honest significant difference (HSD) test in JMP Genomics Version 7.0 (SAS® 9.3, SAS Institute Inc.).

QTL analysis was performed on the LS-means for both zebra-striped presence/absence and zebra-striped intensity data. Single marker analysis (SMA) was initially conducted using a customized R script, then interval mapping (IM) using Haley–Knott regression algorithm (Lander & Botstein, 1989; Haley & Knott, 1992) was conducted with a step size of 1 cM using R/qtl (Broman et al., 2003); both methods are based on a single-QTL model, which assumes a single QTL segregating in each genome scan. Additionally, the IM approach assumes the residual variation in the phenotype follows a normal distribution, although the regression-based interval mapping method is robust against non-normal trait distribution (Rebai, 1997) and application of interval mapping will still provide reasonable results for non-normal traits if the significance threshold was determined via a permutation test (Churchill & Doerge, 1994; Doerge, 2002). Thus, nonparametric interval mapping that uses the Wilcoxon rank-sum test was performed in R/qtl for both nonparametric interval mapping that uses the Wilcoxon rank-sum test was performed in R/qtl for both chromosome-wide significance threshold (P < 0.05) to call QTL was determined based on permutation tests with 1000 permutations (Churchill & Doerge, 1994). The interaction among QTL was explored using the ‘scantwo’ function in R/qtl, which is based on a two-QTL model, assuming two QTL segregating when searching for QTL (Fig. S2, S3). Significant interactions among QTL were tested in R/qtl after fitting all QTL terms in the model. The position and effect of significant QTL were refined using Haley–Knott regression method, and percent variation explained was calculated in R/qtl by fitting a model containing all QTL identified with their interactions if existed. Confidence intervals were calculated using the 1.5-LOD support interval method (Lander & Botstein, 1989; Dupuis & Siegmund, 1999).

**Results**

**Genetic maps**

The female and male maps each had the expected 19 LGs, but to resolve LG 15 on the male map, it was
necessary to include 14 segregation distorted markers \((P < 0.05)\). The female map spanned 1599 cM, with 1887 RAD-seq SNPs and an average intermarker spacing of 0.8 cM (Fig. 2a). There were only two gaps > 10 cM on LGs 3 and 9 of the female map, with between marker distances of 16.6 cM and 10.6 cM, respectively. Length of the male parent map was 1612 cM, with 1821 markers (90 GoldenGate SNPs and 1731 RAD-seq SNPs) and an average intermarker spacing of 0.9 cM (Fig. 2b). Two gaps > 10 cM were found on LG 2 and LG 4 of the male map, with between marker distances of 10.9 cM and 14.4 cM, respectively. A composite genetic map was constructed from 3044 RAD-seq SNPs and 138 GoldenGate SNPs (3182 total) with all 19 LGs successfully resolved at a minimum LOD score of 11 (Fig. S4).

**Phenotypic analysis**

The broad-sense heritability of zebra-striped presence/absence was 1.0, because the phenotype was consistent across all three replications. zebra-striped presence/absence segregated approximately 1 : 1 in the F1 progeny (138 striped plants: 123 nonstriped, \(x^2 < x^2_{0.05} = 3.84; P > 0.05, \text{Fig. S2a})\), which indicated that inheritance was recessive and suggested a single locus. The 32 individuals that resulted from self-fertilization of the striped female parent, ‘Strictus’, were all striped, further confirming recessive gene action, as a hypothetically heterozygous dominant striped parent would have been expected to produce some nonstriped progeny from selfing, but this was not observed. Thus, we determined that the striped female parent, ‘Strictus’, was homozygous recessive for the striping trait whereas the nonstriped male parent, ‘Kaskade’, carried a recessive allele for striping on at least one locus. QTL mapping for zebra stripe was performed entirely on the male parent (‘Kaskade’) map. Among the progeny that had zebra stripes, the intensity of striping varied, and most of this variation was due to genetics. The broad-sense heritability of zebra-striped intensity was calculated on a plot-mean basis as follows:

\[
H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2} = \frac{0.149}{0.149 + \frac{0.006}{3}} = 0.99
\]

where \(\sigma_g^2\) is the genetic variance, \(\sigma_e^2\) is the environmental variance, and \(r\) is the number of replications.

**QTL mapping for zebra stripe**

An initial SMA for zebra-striped presence/absence identified the most significant SNP on LG 7 (Fig. 3a). After accounting for this significant SNP on LG 7 in the model, a significant SNP on LG 10 was detected. When both SNPs on LG 7 and 10 were included in the model, a second significant SNP on LG 7 was identified. Based on a permutation determined LOD threshold, IM using
Haley–Knott regression algorithm also identified two peaks above threshold on LG 7 and one on 10 for zebra-striped presence/absence (Fig. 3b). However, the IM method makes the assumption that only one QTL is segregating when searching for QTL and does not account for QTL elsewhere in the genome. CIM was subsequently applied based on the IM results. After controlling for the major peak on LG 7, a slight increase of the LOD score was observed in the peak of the QTL on LG 10, suggesting that the QTL on LG 10 was real.
Then, by taking into account both QTL on LG 7 and 10, a small second peak on LG 7 and a peak on LG 14 were observed just above the significance threshold (Fig. 3b), suggesting that additional QTL needed to be accounted for. To explore whether or not a second QTL on LG 7 existed, a two-dimensional two-QTL genome scan was conducted on LG 7 while controlling for the QTL on LG 10, and the second QTL on LG 7 was confirmed to be significant (Fig. S2a). Including the second QTL on LG 7 increased the LOD score by 6 if epistasis was allowed and by 4 if no epistasis was allowed (Fig. S2a). After controlling for the two QTL on LG 7 and one on LG 10, the peak on LG 14 disappeared and no additional QTL appeared to be significant (Fig. 3c), which suggested the three-QTL model for zebra-striped presence/absence was appropriate. Independent from the CIM method, the automated selection procedure of MQM analysis also identified three significant QTL (zb1, zb2, and zb3) for zebra-striped presence/absence, with two of them (zb1 and zb2) detected on LG 7 at position 51 cM and 36 cM, respectively, and one on LG 10 (zb3) at position 4 cM (Fig. 3d,e). A significant interaction between zb1 and zb3 was also detected by MQM (Fig. 3d, Fig. S2b). The 1.5-LOD support interval for each QTL extended to a 2 cM, 8 cM, and 3 cM region around the point estimates of zb1, zb2, and zb3, respectively (Fig. 2b). An additional interaction term between zb1 and zb2 was detected after the model was fitted and therefore was included in the final QTL model. The final three-QTL model for zebra-striped presence/absence explained 63% of the phenotypic variation (Table 1).

For zebra-striped presence/absence, each QTL could independently confer the zebra-striped phenotype, but penetrance was not 100%, and it was often low for

| Trait                          | QTL* | LG† | QTL peak (cM) | Closest marker to QTL peak | 1.5-LOD support interval (cM) | Flanking markers | LOD | PVE‡ |
|-------------------------------|------|-----|---------------|-----------------------------|------------------------------|-----------------|-----|------|
| Zebra-striped presence/absence | zb1  | 7   | 51.0          | TP107979                    | 50.0                         | X83             | 19.5| 19.8 |
|                               | zb2  | 7   | 36.0          | TP78983                     | 34.0                         | TP47894         | 4.9 | 6.9  |
|                               | zb3  | 10  | 4.0           | TP74459                     | 2.0                          | TP79286         | 22.3| 18.6 |
|                               | zb1:zb2 |     |               |                             |                              | TP79286         | 4.8 | 3.3  |
|                               | zb1:zb3 |     |               |                             |                              | TP65474         | 12.6| 9.2  |
|                               | zb_Model |     |               |                             |                              |                 | 63.0|      |
| Zebra-striped intensity       | zb1  | 7   | 47.0          | X83                         | 46.0                         | TP28957         | 26.7| 45.5 |
|                               | zb2  | 10  | 6.0           | TP65474                     | 3.0                          | TP47894         | 27.4| 47.3 |
|                               | zb3  | 3   | 24.4          | TP75257                     | 7.0                          | TP43225         | 7.0 | 8.3  |
|                               | zb1:zb2 |     |               |                             |                              | TP79286         | 10.8| 13.8 |
|                               | zb_Model |     |               |                             |                              |                 | 68.3|      |

*The numbering of QTL was based on the linkage group number first and then LOD score from largest to smallest.
†Linkage group.
‡Percent variation explained calculated from the drop-one-QTL-at-a-time ANOVA analyses (at α = 0.05) in R/qtl. Estimates from the QTL models are subject to some degree of inflation due to the Beavis effect (Beavis, 1998).

Table 2 Number of striped and nonstriped F1 progeny of Miscanthus sinensis 'Strictus' × 'Kaskade', and penetrance (proportion) within each genotypic class† among three loci‡ for zebra-striped presence/absence (zb1, zb2, and zb3)

| QTL | 11 | 21 | 12 | 22 |
|-----|----|----|----|----|
| zb1 | 1  | 2  | 3  | 39 |
| zb2 | 58 | 9  | 9  | 2  |
| zb3 | 0.18 cd | 0.25 bcd | 0.95 a* |

| QTL | 21 | 12 | 22 |
|-----|----|----|----|
| zb1 | 28 | 10 | 7  | 46 |
| zb2 | 30 | 1  | 4  | 2  |
| zb3 | 0.48 bc | 0.91 a*** | 0.64 ab | 0.96 a |

†Genotypic classes: 1 = heterozygous; 2 = homozygous recessive.
‡Counts based on nearest markers to the QTL.
§Penetrance estimated as number of striped F1 progeny divided by total number of F1 progeny within each genotypic class. Means separation by Tukey–Kramer HSD; means with the same letter were not significantly different. Interaction between QTL indicated, where *, *** = significant at P = 0.05, 0.001, respectively.
single loci (Table 2). For example, in the absence of other homozygous recessive \( zb \) loci, \( zb1 \) (on LG 7) was only 18\% penetrant and \( zb2 \) (on LG 7) was only 25\% penetrant. However, there were significant interactions among some of the QTL and different combinations gave greater penetrance than others (Table 2). For example, the combinations, \( zb1 \) and \( zb2 \), or \( zb1 \) and \( zb3 \), or \( zb1 \), \( zb2 \), and \( zb3 \) had > 90\% penetrance. In contrast, the

**Fig. 4** QTL results for zebra-striped intensity in a *Miscanthus sinensis* ‘Strictus’ × *M. sinensis* ‘Kaskade’ population: (a) Manhattan plot of single marker association analysis. Each of the 1821 SNPs is plotted. The x-axis represents the chromosomal location on the male parent map, and the y-axis represents the -log10 \( P \) values. The horizontal dotted line represents the Bonferroni-corrected \( P \) value significance threshold via 1000 permutations. The arrows indicate significant SNPs associated with zebra-striped intensity, which were detected in three rounds of single marker analyses after controlling for the significant SNPs from the previous rounds. (b, c) LOD curves from interval mapping (IM) and composite interval mapping (CIM) methods, controlling for the first two QTL in (b), and all three QTL in (c). The horizontal dotted line represents the significance threshold via 1000 permutations. (d) The sequence of models tested by the stepwise procedure in multiple-QTL model (MQM), with circles corresponding to QTL (chromosome number indicated within circles). The best model was identified at step 3 with the largest penalized LOD score (pLOD). (e) Profile LOD score curves for a three-QTL model for zebra-striped intensity. The numbering of QTL was based on the linkage group number first and then LOD score from largest to smallest.
combination of zbi2 and zbi3 had 64% penetrance, which although not as high as other combinations, exceeded the penetrance of any single locus.

For zebra-striped intensity, a series of single marker analyses also identified three significant SNPs on the male parent map, with two on LG 7 and one on LG 10 (Fig. 4a). IM identified two peaks above the significance threshold on LG 7 and 10, respectively (Fig. 4b). CIM was subsequently applied based on the IM results and showed an increase of 8 for the LOD score in the peak of the QTL on LG 10, after controlling for the major QTL on LG 7, which strongly suggested that the QTL on LG 10 was real (Fig. 4b). However, after taking into account the QTL on LG 7 and 10, an additional QTL on LG 3 appeared to be significant, and a weak peak on LG 7 exceeded the significance threshold (Fig. 4b). Because the QTL on LG 3 was more significant than the one on LG 7, the QTL on LG 3 was taken into account along with the QTL on LG 7 and 10, and no more additional significant QTL were found, suggesting the weak peak on LG 7 was probably an artifact and the three-QTL model was appropriate for zebra-striped intensity (Fig. 4c). In addition, a significant interaction between QTL on LG 7 and 10 was detected in the two-dimensional two-QTL genome scan (Fig. S3). Consistent with CIM results, MQM analysis identified three significant QTL (zbi1, zbi2, and zbi3) on LG 7 at position 47 cM, LG 10 at position 6 cM, and LG 3 at position 24 cM, respectively (Fig. 4d,e). However, no significant interaction was detected in the MQM automated selection procedure (Fig. 4d). The 1.5-LOD support interval for each QTL extended to a 4 cM, 6 cM, and 21 cM region around the point estimates of zbi1, zbi2, and zbi3, respectively (Fig. 2b). The final three-locus model for zebra-striped intensity explained 68% of the phenotypic variation (Table 1). ANOVA results for zebra-striped intensity indicated that only zbi1 and zbi2 (on LG 7 and 10) interacted to give a greater intensity than would be expected from additive effects alone, but zbi3 (on LG 3) was a modifier with only additive effect (Table 3, Fig. 5). The best combination was all three QTL, which gave more than 50% of the leaf area striping (Table 3, Fig. 5).

Discussion

High-density genetic map

High-density genetic maps for two M. sinensis cultivars, ‘Strictus’ and ‘Kaskade’, were produced with more than 3000 newly developed RAD-seq SNP markers. We have made these mapped markers available for use by others (Dataset S1), and we expect that they will become a valuable community resource, enabling comparisons of QTL across populations, environments and studies. Moreover, we linked our RAD-seq SNP maps with the previously published maps of M. sinensis ‘Grosse Fontaine’ × M. sinensis ‘Undine’ (Swaminathan et al.,...
2012), and all four maps had similar lengths. The maps developed from the current RAD-seq study will enable detection of markers tightly linked to QTL, which is a first step toward applying MAS and thereby improving breeding efficiency. Given the especially high marker density in some regions of the current maps, and the high degree of synteny between Miscanthus and sorghum, it will also be possible to hypothesize causative (candidate) genes underlying traits of interest. In addition, the sequenced and mapped RAD tags in this study will facilitate ongoing efforts to obtain a whole-genome assembly of Miscanthus by resolving uncertainties in assembling large scaffolds caused by high heterozygosity and the recent genome duplication.

**Mapping zebra-striped QTL**

To confirm the utility of the new high-density genetic map, we mapped QTL for zebra stripe in Miscanthus. Consistent with the study of Touchell et al. (2007), we confirmed the recessive inheritance of zebra stripe in Miscanthus. However, we identified three loci for both zebra-striped presence/absence (two on LG 7 and one on LG 10) and intensity (one each on LGs 3, 7, and 10), which differed from the Touchell et al. (2007) conclusion that a single locus conferred zebra-striped presence/absence, based on the segregation ratio in a ‘Strictus’ (striping perpendicular to the leaf axis) × ‘Variegatus’ (striping parallel with the leaf axis) backcross population. The different QTL mapping methods we employed showed consistent results, indicating strong evidence for the identified QTL (Fig. 3). To further test whether the second QTL on LG 7 (zb2) was real rather than a false association caused by the heterozygote under-calling of genotypes within zb2 QTL region, graphical genotypes were evaluated to identify suspect heterozygote under-calls, and modified QTL analyses were conducted using two approaches: (1) eliminate the individuals which had suspected heterozygote under-called genotypes and (2) correct the suspected heterozygote under-called genotypes. The second QTL on LG 7 was still detected as significant using both approaches, which suggested that there were more than one QTL on LG 7 associated with zebra-striped presence/absence.

The apparent discrepancy between the conclusion of Touchell et al. (2007) and our QTL analyses can be reconciled by the observations of incomplete penetrance for each of the three zebra-striped presence/absence QTL (Table 2). Without the marker data and the subsequent discovery of incomplete penetrance, phenotypic data alone could also have led us to the erroneous conclusion that only a single locus conferred the trait. Segregation of the F2 in the Touchell et al. (2007) study was also inconsistent with a single recessive locus, as only about half of the expected number of horizontally zebra-striped plants were observed. Penetrance for zebra-striped presence/absence was greatest when two or more zb QTL were homozygous for the causative alleles (Table 2). Similarly, the intensity of zebra striping was greatest when two or more zbi QTL were homozygous for alleles that conferred the trait (Table 3). Notably, two pairs of QTL for presence/absence and intensity on LG 7 (zb1 and zbi1) and on LG 10 (zb3 and zbi2) had overlapping 1.5-LOD support intervals (Fig. 2b), so it was not possible to determine whether the QTL within each pair were in fact the same locus or tightly linked loci. Although we might expect that individuals with multiple mutations for zebra striping would be rare and at a selective disadvantage in nature, horticulturalists appear to have preferentially selected Miscanthus individuals with multiple zebra-striped mutations because their phenotypes were more consistent and more visually striking than single mutants (Figs. 1 and 5). Lastly, the three-locus models explained 63% and 68% of the total phenotypic variation for zebra-striped presence/absence and intensity, respectively (Table 1), which suggested that other genetic or environmental factors affected zebra striping. Moreover, estimates from the QTL models are subject to some degree of inflation due to the Beavis effect (Beavis, 1998; Xu, 2003). Chai et al. (2011) and Lu et al. (2012) have shown that many of the zebra mutants in maize and rice are temperature and/or light dependent, which would be consistent with our observations of incomplete penetrance of zb QTL in Miscanthus. Although the high estimates of broad-sense heritabilities indicated a consistent phenotype among ramets of a given genotype within the tested environment, we do not know whether penetrance would be considerably higher or lower in different environments (e.g., temperature and day length). Additionally, the average zebra-striped intensity score among the selfed progeny of ‘Strictus’ was not significantly different from the average score among F1 progeny with all three homozygous zebra-striped intensity loci, indicating that there was no evidence that the female parent contributed dominant alleles from one or more unmapped zebra-striped loci. Thus, the unexplained phenotypic variation in the current study may have been due to the environment and/or yet to be identified genes.

**Comparative mapping**

Genes controlling zebra stripe in maize and rice are critical to chlorophyll and/or carotenoid biosynthesis and chloroplast development, which are of fundamental importance for plant growth and crop productivity (Chai et al., 2011; Lu et al., 2012; Zhao et al., 2014). In
prior studies, at least 12 genes controlling zebra stripe have been found in maize and 16 have been found in rice (Table 4, S1; Szalma et al., 1999; Hayes, 1938; Singh, 1934; Hayes & Chang, 1939; Lu et al., 2012; Nelson, 1991; Horovitz, 1948; Giesbrecht, 1965; Neuiffer & England, 1995; Huang et al., 2009; Zhao et al., 2014; Li et al., 2010). Genomic comparisons among species within the Poaceae have demonstrated that synteny and collinearity are a common feature of this family (Devos & Gale, 1997; Gaut, 2002), and recent comparisons of the Miscanthus and sorghum genomes were consistent with this broader finding (Kim et al., 2012; Ma et al., 2012; Swaminathan et al., 2012). Thus, we would expect to find correspondence among loci underlying zebra stripe across Miscanthus, maize, and rice genomes.

In maize, zb7 is a recessive mutant of the IspH gene on LG 1 that confers transverse green-/yellow-striped leaves in young plants (Lu et al., 2012). Intriguingly, the IspH gene on sorghum LG 1 was found to be 0.2 Mb from a RAD tag that mapped to Miscanthus LG 10 within 15 cM (or 0.9% of the 1612 cM genome) of zb3 and zbi2 detected in the present study (Table 4). This finding suggests that Miscanthus zb3 and/or zbi2 may be a mutant of IspH or part of a cluster of genes involved in chloroplast development that includes IspH. However, it should be noted that sorghum LG 1 does not typically correspond with Miscanthus LG 10 but rather LGs 1 and 2; the closest RAD tags to IspH were on Miscanthus LGs 1, 13, and 10 (0.1, 0.1, and 0.2 Mb, respectively), whereas the closest tag on LG 2 was 3.0 Mb distant (Table 4). It is possible that a translocation of this genomic region has occurred in Miscanthus with respect to sorghum, or alternatively that the alignment position of this marker was incorrect. Lu et al. (2012) found that the maize zb7 gene has a single nucleotide mutation in a highly conserved region, and gene silencing of IspH results in a complete albino phenotype, indicating that zb7 is important for chlorophyll development. Lu et al. (2012) also noted that the expression of zb7 is under 16-h light/8-h dark cycle regulation, and low temperature inhibits the mutant expression. This is consistent with our observation in Miscanthus that zebra-striped plants usually do not develop transverse yellow bands on the leaves during early spring in central Illinois, and the number and area of the horizontal yellow bands increase later in the season as the temperature becomes warmer and days get longer (data not shown).

Table 4  Cloned zebra-striped genes in maize and rice, and their corresponding orthologous regions in Miscanthus

| Gene name             | Gene symbol | Species | Linkage group for the species in which gene was identified | LG | Range (Mb) | Orthologous region(s) in sorghum based on sorghum physical map | Expected linkage groups in Miscanthus based on Swaminathan et al. (2012) | The three markers that were nearest to the zebra-striped gene in sorghum and mapped in Miscanthus (Miscanthus LG* shown) | Distance to gene† (Mb) | Distance to QTL‡ (cM) |
|-----------------------|-------------|---------|----------------------------------------------------------|----|------------|-------------------------------------------------------------|-----------------------------------------------------------------|-----------------------------------------------------------------|----------------------|----------------------|
| Zebra crossbands7     | zb7         | Zea     | 1_long arm                                              | 1  | 7.894–7.897| 1                                                           | 13 TP42765 0.100                                                  | 13.2 (zb3) 11.2 (zbi2)                                           |                     |                      |
|                       |             |         |                                                         | 2  |            |                                                             | 1 TP94641 0.132                                                  |                     |                      |
|                       |             |         |                                                         | 10 | 0.236      |                                                             | 10 TP54034 0.236                                                 |                     |                      |
| Camouflage1           | cf1         | Zea     | 5_short arm                                             | 4  | 4.443–4.447| 7                                                           | 8 TP105102 0.064                                                 |                     |                      |
|                       |             |         |                                                         | 8  |            |                                                             | 1 TP122214 0.067                                                |                     |                      |
|                       |             |         |                                                         | 8  |            |                                                             | 8 TP40456 0.084                                                  |                     |                      |
| Zebra2                | z2          | Oryza   | 11                                                      | 5  | 53.890–53.895| 9                                                           | 10 TP27302 0.001                                                 | 57.9 (zb3) 55.9 (zbi2)                                          |                     |                      |
|                       |             |         |                                                         | 10 |            |                                                             | 9 TP20195 0.496                                                  |                     |                      |
|                       |             |         |                                                         | 19 | 0.496      |                                                             | 9 TP112083 0.684                                                 |                     |                      |
|                       |             |         |                                                         | 19 |            |                                                             | 19 TP742 0.013                                                   |                     |                      |
|                       |             |         |                                                         | 19 |            |                                                             | 19 TP86555 0.050                                                 |                     |                      |
|                       |             |         |                                                         | 19 |            |                                                             | 19 TP90066 0.263                                                 |                     |                      |

*Linkage group.
†Physical distance of marker to the zebra-striped gene in sorghum was calculated based on the marker position in sorghum to the midpoint of the corresponding genomic region of the gene in sorghum.
‡Genetic distance of marker to the Miscanthus QTL if the marker was found on the same linkage group where QTL was located.

© 2015 The Authors. Global Change Biology Bioenergy Published by John Wiley & Sons Ltd., 8, 616–630
In addition, maize z6b gene on LG 4 mapped to sorghum LG 4 between 51 Mb and 57 Mb, which corresponds to Miscanthus LG 7 (as expected) within the genomic region that includes the 1.5 support intervals for zb1, zb2, and zbi1 (Table S1). The finding suggests evidence for colocalization between Miscanthus zb1, zb2, zbi1 and maize z6b gene; however, this correspondence merits further investigation as maize z6b gene has not been fine mapped or cloned.

In rice, z2 is a zebra-striped gene that was mapped to rice LG 11 (Chai et al., 2011), and the subsequently cloned gene was found to be a mutant of CRTISO, a carotenoid biosynthesis gene. Based on syntenic relationships across rice, sorghum and Miscanthus, z2 was found to be on Miscanthus LG 10 (as expected) within 60 cm (or 3.7% of the 1612 cm genome) of Miscanthus zbi3 and zbi2 (Table 4). No other genes mapped in maize or rice were found near any of the identified loci in Miscanthus. However, different Miscanthus populations may have additional zebra-striped QTL that correspond to some of the other genes mapped in maize and rice. Thus, multiple lines of evidence, including QTL mapping in Miscanthus and cloned genes for zebra stripe in maize and rice, indicate that LG 10 in Miscanthus likely has several genes that are important for chlorophyll development, and some alleles of these genes can result in the zebra-striped phenotype.

Leveraging the extensive gene annotations available for sorghum, maize, rice, and Arabidopsis (http://www.maizegdb.org; http://www.gramene.org/; The Arabidopsis Information Resource (TAIR); http://www.phytozome.net/), we also identified candidate genes in genomic regions of these grasses that corresponded to RAPD tags in the 1.5-LOD support interval of each Miscanthus QTL for zebra-striped presence/absence and intensity (Table S2). In total, 44 genes related to chloroplast development, or chlorophyll biosynthesis were found from sorghum, maize, rice, and Arabidopsis, that corresponded to the QTL regions identified in Miscanthus. Of these 44 genes, 4 genes were identified in the zbi3 & zbi2 region on LG 10, 6 in the zb2 region on LG 7, 1 in the zb1 & zbi1 region on LG 7, and 33 in the zbi3 region on LG 3. Notably, the zb1 & zbi1 region harbored a gene in Arabidopsis that encodes 2-C-methyl-D-erythritol 2,4-cyclophosphosphate (IsPF) synthase, which is an especially promising candidate for zebra stripe because Arabidopsis plants defective in this gene display an albino lethal phenotype (Hsieh & Goodman, 2006). Moreover, homologs to the IsPF gene in sorghum, maize and rice have a conserved function (Table S2). A similar candidate gene approach using synteny has been employed in fava bean (Vicia faba L.) by Khaazaei et al. (2014) to identify a ribose-phosphate pyrophosphokinase gene from the sequenced model legume, Medicago truncatula, as a candidate for QTL underlying canopy temperature under water-deficit.

Some of the QTL for zebra stripe that we found in Miscanthus corresponded to genomic regions in maize and rice that have not previously been associated with genes for this trait but which intriguingly contained genes for chlorophyll, carotenoid, and plastid development. At least 32 maize genes associated with photosynthesis were found on maize LG 2 within corresponding Miscanthus QTL regions, of which, 5 genes were found in the zb1 region, 2 genes in the zb3 & zbi2 region, and 25 genes in the zbi3 region (Table S2). In rice, 15 genes associated with photosynthesis were identified on rice LG 9 in the zbi3 region and 2 were found on rice LG 12 in the zb3 & zbi2 region (Table S2). To our best knowledge, there have not been any zebra-striped genes mapped to maize LG 2 or rice LG 9 & 12. Thus, past selection for zebra stripe in Miscanthus by horticulturists has likely provided geneticists and physiologists with novel mutants for studying photosynthesis in the Poaceae. zebra-striped genes in Miscanthus, maize, rice, and sorghum offer a valuable research opportunity to elucidate the complex mechanisms of photosynthesis and productivity in humanity’s most important crops.

Acknowledgements

This work was supported with the funding from the Energy Biosciences Institute, HATCH project IUU-802-311, and the DOE Office of Science, Office of Biological and Environmental Research (BER), Grant No. DE-SC0006634. We thank Katarzyna Dubiel, Megan Swanson, and Melina Salgado for their assistance with leaf tissue collection and DNA extraction. We thank the staff at the Energy Farm for maintaining the field trial and Pat Brown for providing PstI adapters in RAD-seq library preparation. We also thank Brian Diers and Pat Brown for their valuable internal reviews.

References

Al-Janabi SM, Honeycutt RJ, McClelland M, et al. (1993) A genetic linkage map of Saccharum spontaneum L. ’SES 208’. Genetics, 134, 1249–1260.

Anonymous (1876) The zebra-striped Eulalia. The American Agriculturist 3, 460.

Atienza SG, Satovic Z, Petersen KK, Dolstra O, Martin A (2002) Preliminary genetic linkage map of Miscanthus sinensis with RAPD markers. Theoretical and Applied Genetics, 105, 946–952.

Atienza SG, Satovic Z, Petersen KK, Dolstra O, Martin A (2003a) Identification of QTLs associated with yield and its components in Miscanthus sinensis Anderss. Euphytica, 132, 353–361.

Atienza SG, Satovic Z, Petersen KK, Dolstra O, Martin A (2003b) Identification of QTLs influencing agronomic traits in Miscanthus sinensis Anderss. I. Total height, flag-leaf height and stem diameter. Theoretical and Applied Genetics, 107, 123–129.

Atienza SG, Satovic Z, Petersen KK, Dolstra O, Martin A (2003c) Identification of QTLs influencing combustion quality in Miscanthus sinensis Anderss. II. Chlorine and potassium content. Theoretical and Applied Genetics, 107, 857–863.

Atienza SG, Satovic Z, Petersen KK, Dolstra O, Martin A (2013d) Influencing combustion quality in Miscanthus sinensis Anderss. Identification of QTLs for calcium, phosphorus and sulphur content. Plant Breeding, 132, 141–145.

Bailey LH, Miller W (1901) Cyclopaedia of American Horticulture. The MacMillan Co, New York.

© 2015 The Authors. Global Change Biology Bioenergy Published by John Wiley & Sons Ltd., 8, 616–630
Grattapaglia D, Sederoff R (1994) Genetic mapping of QTLs controlling vegetative growth in Populus. *Science* 264: 1361-1363.

Beavis WD (1998) QTL analyses: power, precision, and accuracy. In: Molecular Dissection of Complex Traits (ed. Paterson AH), pp. 145–162. CRC Press, Boca Raton, FL.

Bremner K, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, 19, 889–890.

Chai CL, Fang J, Liu Y et al. (2011) ZEER2, encoding a carotenoid isomerase, is involved in photoprotection in rice. *Plant Molecular Biology*, 75, 211–221.

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics*, 138, 963–971.

Clark LV, Brunner JE, Glowacka K et al. (2014) A footprint of past climate change on the diversity and population structure of Miscanthus sinensis. *Annals of Botany*, 114, 97–107.

Clark LV, Stewart JR, Nishiwaki A et al. (2015) Shattering clones in Miscanthus. *Biology of Insecta*, 65, 279–289. Springer, New York.

Clifton-Brown JC, Stampfl PF, Jones MB (2004) The changes of chlorophyll content and several enzyme activities in zebra-leaf rice. *Journal of Wuhan University (Natural Science Edition)*, 64, 761–765.

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

Hirayoshi I, Nishikawa K, Kato R (1959) Cytogenetical studies on forage plants. (IV) Self-incompatibility in Miscanthus. *Japanese Journal of Breeding*, 5, 167–170.

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

Horovitz S (1948) Zebra-necrosis. *Mazits Genetics Cooperative Newsletter*, 22, 42–43.

Hsieh MH, Goodman HM (2006) Functional evidence for the involvement of Arabidopsis Ishf homolog in the nonmevalonate pathway of plastid isoprenoid biosynthesis. *Plants*, 223, 779–784.

Huang M, Skewinski TL, Baker RE, Janick-Buckner D, Buckner B, Jolab GS, Braun DM (2009) Camouflage patterning in maize leaves results from a defect in porobilinogen deaminase. *Molecular Plant*, 2, 773–789.

Inoue M, Gao ZS, Hirata M, Fujimoto M, Cai HW (2004) Construction of a high-density linkage map of Italian ryegrass (*Lolium multiflorum Lam.*) using restriction fragment length polymorphism, amplified fragment length polymorphism, and telemetric repeat associated sequence markers. *Genome*, 47, 57–65.

Jansen RC (1993) Interval mapping of multiple quantitative trait loci. *Genetics*, 135, 205–212.

Jansen RC, Stam P (1994) High-resolution of quantitative traits into multiple loci via interval mapping. *Genetics*, 136, 1447–1455.

Jorgensen U, Muhs HJ (2001) Miscanthus breeding and improvement. In: *Miscanthus for Energy and Fibre* (eds Jones MB, Walsh M), pp. 68–85. James and James (Science Publishers), London.

Kalbelka E, Franchino B, Francis DM (2002) Two loci from *Lycopersicon hirsutum* encoding an IPP and IspF homolog in the nonmevalonate pathway of plastid isoprenoid biosynthesis. *Plant Molecular Biology*, 50, 504–510.

Khzai H, O’sullivan DM, Sillanpaa MJ, Stoddard FL (2014) Use of syntenry to identify candidate genes underlying QTL controlling stomatal traits in faba bean (*Vicia faba* L.). *Theoretical and Applied Genetics*, 127, 2371–2385.

Kim C, Zhang Z, Auckland SA et al. (2012) SSR-based genetic maps of Miscanthus sinensis and *S. sacchariflorus*, and their comparison to sorghum. *Theoretical and Applied Genetics*, 124, 1325–1338.

Knitola T, Takahashi M-E (1991) The one hundredth report of genetic studies on rice plant: linkage studies and future prospects. *The Journal of the Faculty of Agriculture*, Hokkaido University, 65, 1–61.

Knitola T, Takehara I (1984) Inheritance and linkage relationship on zebra chlorosis and zebra necrosis in rice. *Genetical Studies on Rice Plant*. LXXXVIII. *The Journal of the Faculty of Agriculture*, Hokkaido University, 64, 445–455.

Kusumi K, Komori H, Sato H, Iba K (2010) Characterization of a zebra mutant of rice with increased susceptibility to light stress. *Plant Cell Physiology*, 41, 158–164.

Lander ES, Botstein D (1988) Mapping mendelian factors underlying quantitative traits using RFLP Linkage Maps. *Genetics*, 121, 185–199.

Leweandowski I, Clifton-Brown JC, Scurlock JM, Hussey RW (2008) Miscanthus: European experience with a novel energy crop. *Biomass and Bioenergy*, 32, 209–227.

Li J, Pandey A, Nath K et al. (2010) ZEBRA-NECROSIS, a thylakoid-bound protein, is critical for the photoprotection of developing chloroplasts during early leaf development. *The Plant Journal*, 62, 713–725.

Linde-Laursen I (1993) Cytogenetic analysis of Miscanthus *Giganteus*, an interspecific hybrid. *Hereditas*, 119, 297–300.

Liu S, Wei X, Shao G, Tang Z, Hu P (2013) Map based cloning of a *Zebra*’s leaf mutant gene z2l in rice. *Chinese Journal of Rice Science* (Chinese Version), 27, 231–239.

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

Lu XM, Hu XJ, Zhao YZ et al. (2012) Map-based cloning of *z2l* encoding an IPP and *DMAPP* synthase in the MEP pathway of maize. *Plant Biology*, 15, 1110–1112.

Ma XF, 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.

Meng R, Liu SC, Lin YR et al. (1998) Detailed alignment of *Saccharum* and *Sorghum* chromosomes: comparative organization of closely related diploid and polyploid genomes. *Genetics*, 150, 1663–1682.

Misawa AM, Paterson AH, Beaton JH (2005) Investigation of genomic organization in switchgrass (*Panicum virgatum* L.) using DNA markers. *Theoretical and Applied Genetics*, 110, 1372–1383.

Nelson OE Jr (1991) Locations of new mutations on *S. Mazits Genetics Cooperative Newsletter*, 65, 86.

Neuffer MG, England DJ (1995) Induced mutations with confirmed locations. *Mazits Genetics Cooperative Newsletter*, 69, 43–46.
Okada M, Lanzatella C, Saha MC, Bouton J, Wu RL, Tobias CM (2010) Complete switchgrass genetic maps reveal subgenome collinearity, preferential pairing and multilocus interactions. *Genetics*, 185, 745–760.

Pfender WF, Saha MC, Johnson EA, Slabaugh MB (2011) Mapping with RAD (restriction site-associated DNA) markers to rapidly identify QTL for stem rust resistance in *Lolium perenne*. *Theoretical and Applied Genetics*, 122, 1467–1480.

Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One*, 7, e32253.

Quinby JR, Karper RE (1942) Inheritance of mature plant characters in sorghum induced by radiation. *The Journal of Heredity*, 33, 3–327.

Rebai A (1997) Comparison of methods for regression interval mapping in QTL analysis with non-normal traits. *Genetical Research*, 69, 69–74.

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

Scurlock J (1999) *Miscanthus*: A Review of European Experience with a Novel Energy Crop. pp. 26. Oak Ridge National Laboratory, Oak Ridge, TN.

Singh S (1934) Zebra5 (*S. l. zib5*), which shows in seedlings as a virescent and in mature plants as a zebra stripe (transverse bands of green and yellow tissue) shows no crossing over with *S. l*. *Maize Genetics Cooperation Newsletter*, 8, 1.

Somerville C, Youngs H, Taylor C, Davis SC, Long SP (2010) Feedstocks for lignocellulosic biofuels. *Science*, 329, 791–792.

Spindel J, Wright M, Chen C *et al.* (2013) Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. *Theoretical and Applied Genetics*, 126, 2699–2716.

Stroman GN (1924) The inheritance of certain chlorophyll characters in maize. *Genetics*, 9, 493–511.

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.

Szalma S, Stinar PS, Cone KC (1999) Allelic relationships among the zebra mutants in the Coop’s *zhib*, *zhib2*, and *zhib3* stocks. *Maize Genetics Cooperation Newsletter*, 73, 23.

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

Touchell D, Smith J, Ranney T (2007) Inheritance of leaf variegations in *Sorghum*. *Southern Nursery Association Research Conference*, 52, 349–351.

Van Ooijen J (2011) Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. *Genetics Research*, 93, 343–357.

Wang Q, Sang X, Ling Y, Zhao F, Yang Z, Li Y, He G (2009) Genetic analysis and molecular mapping of a novel gene for zebra mutation in rice (*Oryza sativa L.*). *Journal of Genetics and Genomics*, 36, 679–684.

Wei J, Xu M, Zhang D, Mi H (2010) The role of carotenoid isomerase in maintenance of photosynthetic oxygen evolution in rice plant. *Acta Biochimica et Biophysica Sinica* (Shanghai), 42, 457–463.

Xu SZ (2003) Theoretical basis of the Beavis effect. *Genetics*, 165, 2259–2268.

Zeng ZB (1993) Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proceedings of the National Academy of Sciences of the United States of America*, 90, 10972–10976.

Zeng ZB (1994) Precision mapping of quantitative trait loci. *Genetics*, 136, 1457–1468.

Zeng ZB, Kao CH, Basten CJ (1999) Estimating the genetic architecture of quantitative traits. *Genetical Research*, 74, 279–289.

Zhao J, Fang Y, Kang S *et al.* (2014) Identification and characterization of a new allele for *ZEBRA LEAF 2*, a gene encoding carotenoid isomerase in rice. *South African Journal of Botany*, 95, 102–111.

© 2015 The Authors. *Global Change Biology Bioenergy* Published by John Wiley & Sons Ltd., 8, 616–630.

---

**Supporting Information**

Additional Supporting Information may be found in the online version of this article:

**Table S1.** Mapped but not cloned zebra-striped genes in maize and rice, and their corresponding orthologous regions in *Miscanthus*.

**Table S2.** Candidate genes for *Miscanthus* zebra-striped presence/absence and intensity identified from sorghum, maize, rice and *Arabidopsis* due to their function in chloroplast development and chlorophyll biosynthesis.

**Figure S1.** Frequency distributions of the zebra-striped trait in *F1* progeny of a cross between *Miscanthus sinensis ‘Strictus’* and *Miscanthus sinensis ‘Kaskade’*.

**Figure S2.** Two-dimensional, two-QTL genome scan with the zebra-striped presence/absence data.

**Figure S3.** Two-dimensional, two-QTL genome scan on *Miscanthus* linkage groups 3, 7 & 10 with the zebra-striped intensity data.

**Figure S4.** Composite genetic map for *Miscanthus sinensis* with 3044 RAD-seq SNPs and 138 GoldenGate SNPs (3182 total).

**Figure S5.** Scatterplots of positions of mapped *Miscanthus* markers relative to their positions on *Sorghum bicolor* v1.4 physical map.

**Figure S6.** Principal component analysis plot demonstrating 19 individuals (blue circles) appeared to be the product of self-fertilization of the female parent *Miscanthus sinensis ‘Strictus’* (red circle on the right).

**Figure S7.** Plot of average heterozygosity against the proportion of missing data by individual across 6014 RAD-seq SNPs.

**Figure S8.** Boxplots of zebra-striped intensity distributions for all genotypic combinations (*zb1*, *zb2*, *zb3* on LG 7, 10 & 3, respectively).

**Dataset S1.** A Microsoft Excel file containing information on markers used in the study. All RAD-seq and GoldenGate markers are listed, including sequence information, marker type, linkage groups, marker positions, aligned to sorghum chromosome numbers and sorghum positions.

**Appendix S1.** Materials and methods.