

Aberystwyth University

*Contrasting geographic patterns of genetic variation for molecular markers vs. phenotypic traits in the energy grass *Miscanthus sinensis**

Slavov, Gancho; Robson, Paul; Jensen, Elaine; Hodgson, Edward; Farrar, Kerrie; Allison, Gordon; Hawkins, Sarah; Thomas-Jones, Sian ; Ma, Xue-Feng; Huang, Lin; Swaller, Timothy; Flavell, Richard; Clifton-Brown, John; Donnison, Iain

Published in:
GCB Bioenergy

DOI:
[10.1111/gcbb.12025](https://doi.org/10.1111/gcbb.12025)

Publication date:
2013

Citation for published version (APA):

Slavov, G., Robson, P., Jensen, E., Hodgson, E., Farrar, K., Allison, G., Hawkins, S., Thomas-Jones, S., Ma, X.-F., Huang, L., Swaller, T., Flavell, R., Clifton-Brown, J., & Donnison, I. (2013). Contrasting geographic patterns of genetic variation for molecular markers vs. phenotypic traits in the energy grass *Miscanthus sinensis*. *GCB Bioenergy*, 5(5), 562-571. <https://doi.org/10.1111/gcbb.12025>

Document License CC BY

General rights

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400
email: is@aber.ac.uk

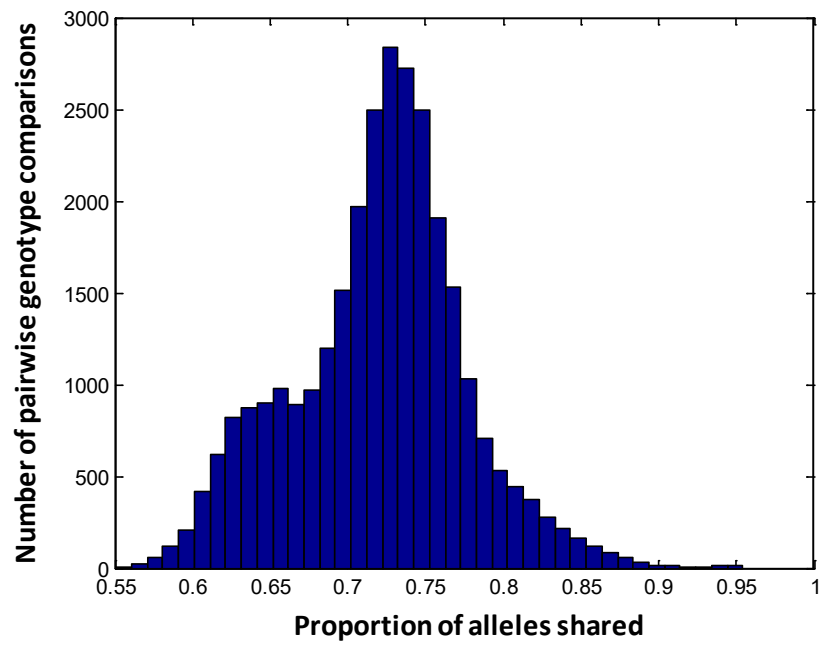


Fig. S1. Proportions of 691 simple sequence repeat (SSR) and 240 single-nucleotide polymorphism (SNP) alleles (i.e., 931 alleles total) shared between pairs of 241 *Miscanthus* genotypes.

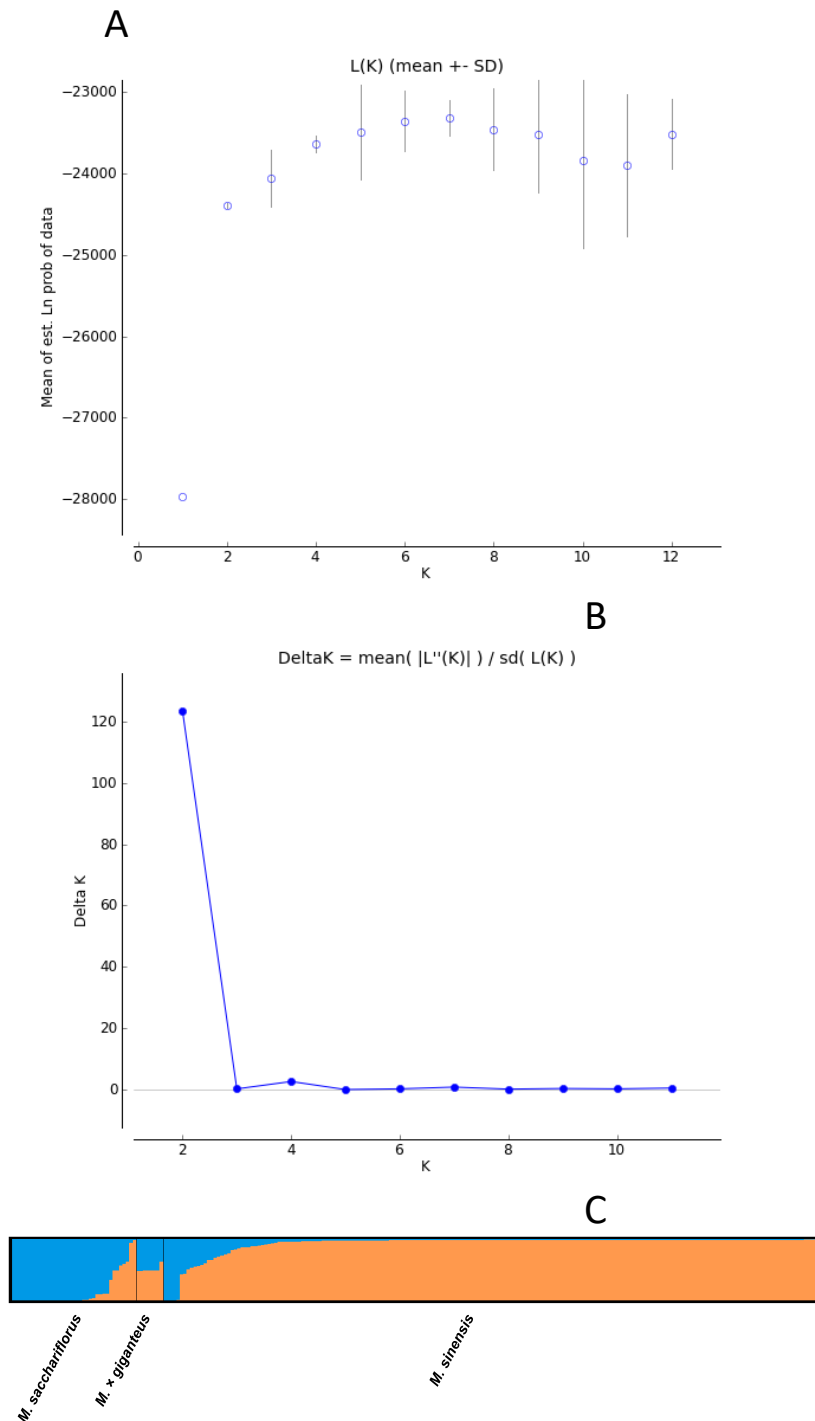


Fig. S2. Model-based clustering of SNP data for 241 *Miscanthus* genotypes using v. 2.3.3 of the STRUCTURE program (Pritchard *et al.* 2000; Falush *et al.* 2003, 2007). (A) Mean log-likelihoods and their standard deviations from runs assuming different numbers of subpopulations (K). (B) Values of the *ad hoc* statistic ΔK , which tends to peak at the value of K that corresponds to the highest hierarchical level of substructure (Evanno *et al.* 2005). (C) Individual proportional memberships of 241 *Miscanthus* genotypes assuming $K = 2$. Clustering results from 10 independent runs of STRUCTURE were aligned using the CLUMPP program (Jakobsson and Rosenberg, 2007) and illustrated using the DISTRUCT program (Rosenberg, 2004). Genotypes were grouped based on *a priori* morphological classification.

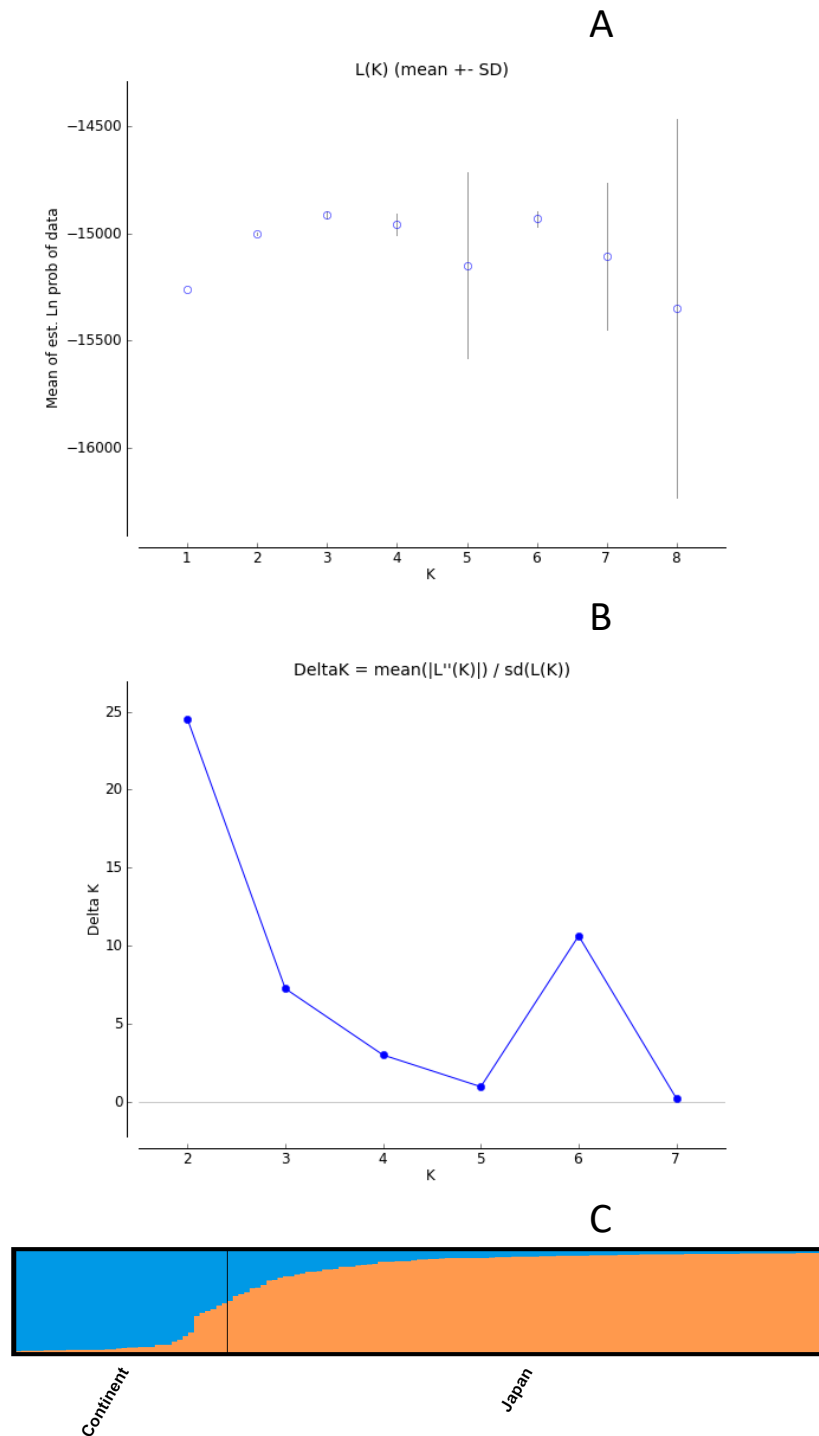


Fig. S3. Model-based clustering of SNP data for 145 *Miscanthus sinensis* genotypes using v. 2.3.3 of the STRUCTURE program (Pritchard *et al.* 2000; Falush *et al.* 2003, 2007). (A) Mean log-likelihoods and their standard deviations from runs assuming different numbers of subpopulations (K). (B) Values of the *ad hoc* statistic ΔK , which tends to peak at the value of K that corresponds to the highest hierarchical level of substructure (Evanno *et al.* 2005). (C) Individual proportional memberships assuming K = 2. Clustering results from 10 independent runs of STRUCTURE were aligned using the CLUMPP program (Jakobsson and Rosenberg, 2007) and illustrated using the DISTRUCT program (Rosenberg, 2004). Genotypes were grouped based on *a posteriori* proportional memberships.

Table S1. Trait loadings from PCA of phenotypic data for 17 traits measured in up to three years in 145 *M. sinensis* genotypes. Only principal components that accounted for more than 5% of the total variation are shown.

Trait	PC1	PC2	PC3	PC4	PC5
<i>DOYFS1.7</i>	-0.21	-0.22	-0.04	0.09	-0.03
<i>DOYFS1.8</i>	-0.21	-0.21	-0.04	0.07	-0.01
<i>DOYFS1.9</i>	-0.22	-0.21	-0.05	0.05	-0.01
<i>AvgSen.7</i>	0.25	0.05	-0.12	0.11	0.00
<i>AvgSen.8</i>	0.25	0.00	-0.08	0.12	0.11
<i>AvgSen.9</i>	0.24	-0.01	-0.04	0.11	0.09
<i>BaseDiameter.7</i>	-0.12	0.21	-0.23	-0.13	0.12
<i>BaseDiameter.8</i>	-0.12	0.19	-0.15	-0.15	0.19
<i>BaseDiameter.9</i>	-0.17	0.20	-0.13	-0.04	0.22
<i>DryMatter.7</i>	-0.27	0.01	-0.16	-0.03	-0.02
<i>DryMatter.8</i>	-0.27	0.04	-0.17	-0.04	0.08
<i>DryMatter.9</i>	-0.25	0.04	-0.17	-0.08	0.08
<i>LeafLength.7</i>	-0.16	-0.23	0.01	0.07	-0.02
<i>LeafWidth.7</i>	-0.15	0.08	0.24	-0.05	0.09
<i>MaxCanopyHeight.7</i>	-0.22	0.17	-0.07	0.16	-0.10
<i>MaxCanopyHeight.8</i>	-0.25	0.14	-0.06	0.15	-0.07
<i>MaxCanopyHeight.9</i>	-0.24	0.16	-0.01	0.14	-0.14
<i>Moisture.7</i>	-0.13	-0.29	0.04	-0.02	0.08
<i>Moisture.8</i>	-0.16	-0.27	-0.07	-0.08	0.00
<i>Moisture.9</i>	-0.09	-0.28	-0.05	-0.08	0.16
<i>StatureCategory.7</i>	0.04	0.01	0.15	-0.29	0.48
<i>StatureLeafAngle.7</i>	0.01	0.06	-0.01	-0.06	0.24
<i>StatureStemAngle.7</i>	0.04	0.00	0.15	-0.29	0.45
<i>StemDiameter.7</i>	-0.20	0.12	0.15	-0.04	0.12
<i>StemDiameter.8</i>	-0.19	0.08	0.17	0.03	0.05
<i>TallestStem.8</i>	-0.10	0.29	0.10	0.11	-0.01
<i>TallestStem.9</i>	-0.12	0.27	0.16	0.02	-0.06
<i>TransectCount.7</i>	0.00	-0.02	-0.44	-0.15	-0.06
<i>TransectCount.8</i>	0.07	0.01	-0.42	-0.14	0.04
<i>TransectCount.9</i>	0.13	-0.01	-0.41	-0.11	-0.03
<i>Cellulose.7</i>	0.07	0.31	-0.11	0.06	0.02
<i>Cellulose.8</i>	0.04	0.26	-0.08	0.14	0.08
<i>Hemicellulose.7</i>	-0.01	0.12	0.11	-0.37	-0.27
<i>Hemicellulose.8</i>	0.07	0.10	0.11	-0.31	-0.28
<i>Lignin.7</i>	0.08	0.00	-0.06	0.37	0.17
<i>Lignin.8</i>	0.02	0.00	-0.01	0.42	0.30
Eigenvalue	10.6	7.3	3.9	3.2	2.3
% Explained	29.5	20.2	10.8	9.0	6.4

Table S2. Trait loadings from PCA of phenotypic data for two phenological traits measured in three years in 145 *M. sinensis* genotypes. Only principal components that accounted for more than 5% of the total variation are shown.

Trait	PC1	PC2
<i>DOYFS1.7</i>	-0.42	-0.39
<i>DOYFS1.8</i>	-0.41	-0.43
<i>DOYFS1.9</i>	-0.43	-0.38
<i>AvgeSen.7</i>	0.42	-0.37
<i>AvgeSen.8</i>	0.39	-0.45
<i>AvgeSen.9</i>	0.39	-0.43
Eigenvalue	4.2	1.4
% Explained	70.0	23.9

Table S3. Trait loadings from PCA of phenotypic data for 12 morphological/biomass-related traits measured in up to three years in 145 *M. sinensis* genotypes. Only principal components that accounted for more than 5% of the total variation are shown.

Trait	PC1	PC2	PC3	PC4
<i>BaseDiameter.7</i>	-0.22	0.06	-0.28	-0.17
<i>BaseDiameter.8</i>	-0.22	0.05	-0.19	-0.25
<i>BaseDiameter.9</i>	-0.26	0.05	-0.15	-0.19
<i>DryMatter.7</i>	-0.28	-0.22	-0.09	0.02
<i>DryMatter.8</i>	-0.29	-0.20	-0.11	-0.05
<i>DryMatter.9</i>	-0.28	-0.19	-0.12	-0.08
<i>LeafLength.7</i>	-0.06	-0.36	0.17	0.09
<i>LeafWidth.7</i>	-0.19	0.04	0.26	-0.15
<i>MaxCanopyHeight.7</i>	-0.31	0.04	-0.04	0.18
<i>MaxCanopyHeight.8</i>	-0.32	-0.01	-0.02	0.16
<i>MaxCanopyHeight.9</i>	-0.32	0.02	0.02	0.20
<i>Moisture.7</i>	-0.01	-0.40	0.18	-0.04
<i>Moisture.8</i>	-0.05	-0.44	0.08	-0.01
<i>Moisture.9</i>	0.02	-0.41	0.09	-0.13
<i>StatureCategory.7</i>	0.04	0.05	0.09	-0.58
<i>StatureLeafAngle.7</i>	0.00	0.08	-0.05	-0.21
<i>StatureStemAngle.7</i>	0.04	0.04	0.10	-0.55
<i>StemDiameter.7</i>	-0.26	0.03	0.16	-0.15
<i>StemDiameter.8</i>	-0.23	0.00	0.20	-0.05
<i>TallestStem.8</i>	-0.23	0.30	0.01	0.06
<i>TallestStem.9</i>	-0.24	0.28	0.06	0.03
<i>TransectCount.7</i>	0.01	-0.16	-0.44	0.00
<i>TransectCount.8</i>	0.06	-0.08	-0.45	-0.08
<i>TransectCount.9</i>	0.13	-0.06	-0.44	-0.01
Eigenvalue	8.1	4.2	3.7	2.5
% Explained	33.7	17.5	15.3	10.6

Table S4. Trait loadings from PCA of phenotypic data for three cell wall composition traits measured in two years in 145 *M. sinensis* genotypes.

Trait	PC1	PC2	PC3
<i>Cellulose.7</i>	0.06	-0.70	0.02
<i>Cellulose.8</i>	0.17	-0.66	0.25
<i>Hemicellulose.7</i>	-0.52	-0.22	-0.26
<i>Hemicellulose.8</i>	-0.46	-0.14	-0.59
<i>Lignin.7</i>	0.45	-0.05	-0.65
<i>Lignin.8</i>	0.53	-0.01	-0.30
Eigenvalue	2.7	1.8	1.0
% Explained	44.5	30.7	15.9

References

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.

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.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574-578.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801-1806.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137-138.