

Extending the Genotype in *Brachypodium* by Including DNA Methylation Reveals a Joint Contribution with Genetics on Adaptive Traits

Steven R. Eichten, Akanksha Srivastava, Adam J. Reddiex, Diep R. Ganguly, Alison Heussler, Jared C. Streich, Pip B. Wilson, and Justin O. Borevitz¹

Australian Research Council Centre of Excellence in Plant Energy Biology, Research School of Biology, Australian National University, Canberra, Acton, Australian Capital Territory 2601, Australia

ORCID IDs: 0000-0003-2268-395X (S.R.E.); 0000-0002-4553-9232 (A.S.); 0000-0002-6376-9592 (A.J.R.); 0000-0001-6746-0181 (D.R.G.); 0000-0003-4861-1188 (J.C.S.); 0000-0001-9548-691X (P.B.W.); 0000-0001-8408-3699 (J.O.B.)

ABSTRACT Epigenomic changes have been considered a potential missing link underlying phenotypic variation in quantitative traits but is potentially confounded with the underlying DNA sequence variation. Although the concept of epigenetic inheritance has been discussed in depth, there have been few studies attempting to directly dissect the amount of epigenomic variation within inbred natural populations while also accounting for genetic diversity. By using known genetic relationships between *Brachypodium* lines, multiple sets of nearly identical accession families were selected for phenotypic studies and DNA methylome profiling to investigate the dual role of (epi)genetics under simulated natural seasonal climate conditions. Despite reduced genetic diversity, appreciable phenotypic variation was still observable in the measured traits (height, leaf width and length, tiller count, flowering time, ear count) between as well as within the inbred accessions. However, with reduced genetic diversity there was diminished variation in DNA methylation within families. Mixed-effects linear modeling revealed large genetic differences between families and a minor contribution of DNA methylation variation on phenotypic variation in select traits. Taken together, this analysis suggests a limited but significant contribution of DNA methylation toward heritable phenotypic variation relative to genetic differences.

KEYWORDS

Brachypodium
epigenomics
missing
heritability

Heritable natural variation has largely been attributed to genetic variation between individuals within and across populations. Novel combinations of alleles and regulatory sequences can influence gene expression and lead to complex changes in downstream phenotypes. Tools such as genome-wide association studies have been highly successful for identifying regions of the genome that contribute to complex trait variation. The aggregate effect of these regions, however, may only explain a small fraction of the expected heritable

variation, a phenomenon referred to as “missing heritability” (Manolio *et al.* 2009). The potential sources of the missing heritability are varied and include; rare variants, structural variation, epistasis, and the focus of this article, epigenomics (Manolio *et al.* 2009; Zuk *et al.* 2012; Trerotola *et al.* 2015).

Recently, there has been great excitement investigating the role of epigenomic variation, in the form of chromatin modifications, which may act alongside, or independently, of traditional genetic variation. A range of population-level studies have reported substantial diversity between different genetic backgrounds (Vaughn *et al.* 2007; Eichten *et al.* 2013, 2016; Schmitz *et al.* 2013b, 2013a; Pignatta *et al.* 2014). One potential source of epigenomic variation arises from DNA methylation, which can vary between populations and geographical locations (Schmitz *et al.* 2013b; Shen *et al.* 2014; Dubin *et al.* 2015; Kawakatsu *et al.* 2016). Results often highlight the covariation of genetic variation with chromatin state, transposable element (TE) methylation, and differential cytosine methylation among accessions. Nonetheless, a small portion of genetically-independent methylation variation, often in the CG context and in promoter regions,

Copyright © 2020 Eichten *et al.*

doi: <https://doi.org/10.1534/g3.120.401189>

Manuscript received December 19, 2019; accepted for publication March 2, 2020; published Early Online March 4, 2020.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material available at figshare: <https://doi.org/10.25387/g3.11918430>.

¹Corresponding author: The Australian National University, Building 134, Linnaeus Way, Canberra, 2601, Australia. E-mail: justin.borevitz@anu.edu.au

may contribute to phenotypic variance (Schmitz *et al.* 2011, 2013b; Hofmeister *et al.* 2017; Schmid *et al.* 2018).

To date, there are few cases showing a direct relationship of variable DNA methylation tied to downstream phenotypes (Ong-Abdullah *et al.* 2015; He *et al.* 2018). In addition, changes in DNA methylation are often confounded with genetic variation present between samples. That is, variable DNA methylation is clearly apparent in cases where genetic variation exists and it has been difficult to disentangle these when associated with a phenotype of interest (Eichten *et al.* 2014). Many studies address this by examining DNA methylation variation through inbreeding generations, so called epi-Recombinant Inbred Lines (epiRILs) or mutation accumulation lines, which display phenotypic variation (Shaw *et al.* 2000; Becker *et al.* 2011; Schmitz *et al.* 2011; Cortijo *et al.* 2014). However, such estimates of the contribution of DNA methylation on phenotypic differences are made in isolation. Therefore, a more holistic approach in which DNA methylation variation is assessed along with genetic polymorphisms and environmental variation to reveal the relative importance of these heritable factors. This can be viewed as a part of the 'Extended Genotype' of an organism, that is, sources of heritable variation that are largely overlooked and/or misinterpreted in relation to more traditional genotype assessments such as SNPs (Eichten *et al.* 2014).

The model grass *Brachypodium distachyon* (*B. distachyon*) provides an ideal system to examine the impact of the extended genotype on natural populations. Advantages include having a relatively small and fully sequenced genome alongside a growing array of genetic resources, displaying a wide climatic distribution resulting in phenotypic diversity, rapid generation times leading to increased rounds for (epi)genetic selection, and a small stature that facilitates systematic study. Pertinently, multiple genetically similar accessions of *Brachypodium* have been identified in different environments across Turkey (Vogel *et al.* 2009; Wilson *et al.* 2019). This set of 'BdTR' accessions provide a unique natural set of germplasm to investigate the impact of 'extended genotype' signals, such as DNA methylation.

MATERIALS AND METHODS

Plant germplasm, growth, and phenotyping

Germplasm for experiments was selected based on previously established SNP-based genotypic relationships as described in (Wilson *et al.* 2019). Briefly, the genetic distance matrix, as defined for all samples, was used to identify clades of individuals with minimal genetic variation (similar to that observed between technical replicates) and for which seed was available. The 95 selected *B. distachyon* accessions were grown in biological triplicate alongside 17 biological replicates of the reference genotype Bd21 for direct comparison to an inbred background (Initiative and The International Brachypodium Initiative 2010). Two seeds per pot were planted ~2.5 cm below surface in 5x5x8 cm pots in a steam-pasteurized 75:25 martin's soil:washed river sand mix. Trays of 14-16 pots were watered with tap water, covered in plastic film, and placed at 4° in the dark for seven days for seed stratification. Trays were subsequently moved into modified Conviron growth chambers which have been fitted with 7-band LED light panels and control light intensity, quality, chamber temperature, and humidity every 5 min (Brown *et al.* 2014). These were planted under simulated conditions for regions near Istanbul, Turkey season-shifted for planting in both northern-hemisphere spring (April) and fall (August) to investigate how plants respond to either a rapid-cycling or overwintering environment. A pair of identical specialized growth chambers were used to simulate these Turkish climates (41.146, 29.026, 72m elevation) modeled using

SolarCalc (version E-2014) (Spokas and Forcella 2006). The chambers updated temperature and 7-band LED light quality information every 5 min (Sup Figure 1) (Brown *et al.* 2014). Spring started with average high temps of 17° and reached a peak summer daily high at ~29°. Night time lows ranged from a season start of ~10° to ~21° by the peak of summer (simulated July). In contrast, fall planting started with average high temps of 19° and reached a minimum winter daily high at ~7°. Night lows ranged from a season start of 14° to 5° (limited by chamber specifications) by mid-winter (simulated February). The two chambers performed the same modeled conditions offset by six months. This allowed for a 'spring' and 'fall' planting in which plants entered the simulated conditions on April 26th and October 16th. Plants were regularly watered using tap water over the course of the experiment when standing water was not observed in trays. All plant measurements were taken by hand, including plant height, third leaf length and width, tiller count, ear count, and flowering time (Sup Table 1) (Rebetzke and Richards 1999; Wilson *et al.* 2015, 2019).

Tissue harvest and DNA extraction

Spring planting flag leaf tissue was harvested 10 weeks after emergence. In contrast, fall planting leaf 3 and leaf 7 tissue was harvested 14 weeks after emergence due to limited overwinter growth (Sup Table 2). Tissue was harvested directly into 96 well plates on liquid nitrogen for DNA extraction. Leaf tissue was ground using the TissueLyser II (Qiagen) and DNA was extracted using the Invisorb DNA Plant HTS 96 Kit (cat 7037300400; Stratec Biomedical) using the manufacturer's protocol. DNA was quantified using the Quant-iT High Sensitivity dsDNA Assay (cat. Q33120; ThermoFisher).

Whole genome bisulfite sequencing and analysis

Whole genome bisulfite sequencing libraries were created using the Accel-NGS Methyl-Seq DNA Library Kit (cat. 30096; Swift Biosciences; Ann Arbor, MI). The standard protocol was modified for third-reaction volumes throughout using 27 ng bisulfite converted gDNA. Initial shearing step of gDNA was omitted. Bisulfite conversion was conducted using the EZ-96 DNA Methylation-Gold MagPrep kit (cat. D5043; Zymo Research). 80 ng of gDNA in 45 µl H₂O was used for low concentration gDNA. Dual indexing of samples was performed with the Methyl-Seq Dual Indexing Kit (cat. 38096; Swift Bioscience) using 11 PCR cycles. The library underwent a final dual-side SPRI cleanup upon completion of the library preparation (0.6x right-side SPRI followed with 0.85x left-side SPRI) to compensate for the lack of physical shearing of the initial DNA. Libraries were quantified using the Caliper LabChip GXII (PerkinElmer) and equal-molar pooling of 96 libraries was performed. Pools of 96 libraries were sequenced using the HiSeq 2500 (Illumina). The subset of samples selected for high-coverage methylome analysis were run as a pool of 96 samples across a flow cell (8 lanes) on the HiSeq 2000. All sequencing was conducted at the ANU Biomolecular Resource Facility.

Raw sequencing reads had 5' trimming of 6 bp to eliminate library bias in methylation state, along with base quality and adapter trimming using *Trim Galore!*. Trimmed reads were mapped to the Bd21 reference (v2.1) using Bismark (v0.13.0) under bowtie1 alignment mode (Krueger and Andrews 2011). Alignments were subsequently deduplicated and all context (CG/CHG/CHH) methylation data were extracted via *bismark_methylation_extractor* using default parameters.

Differentially methylated regions (DMRs) were called using HOME (Srivastava *et al.* 2019). HOME-pairwise module was implemented with the following cutoffs: methylation difference ≥ 20%, number of CGs ≥ 4 and DMR length ≥ 50 bp (parameters: -delta 0.2, -minc 4, and -len 50). Briefly, HOME first computes the P-values using

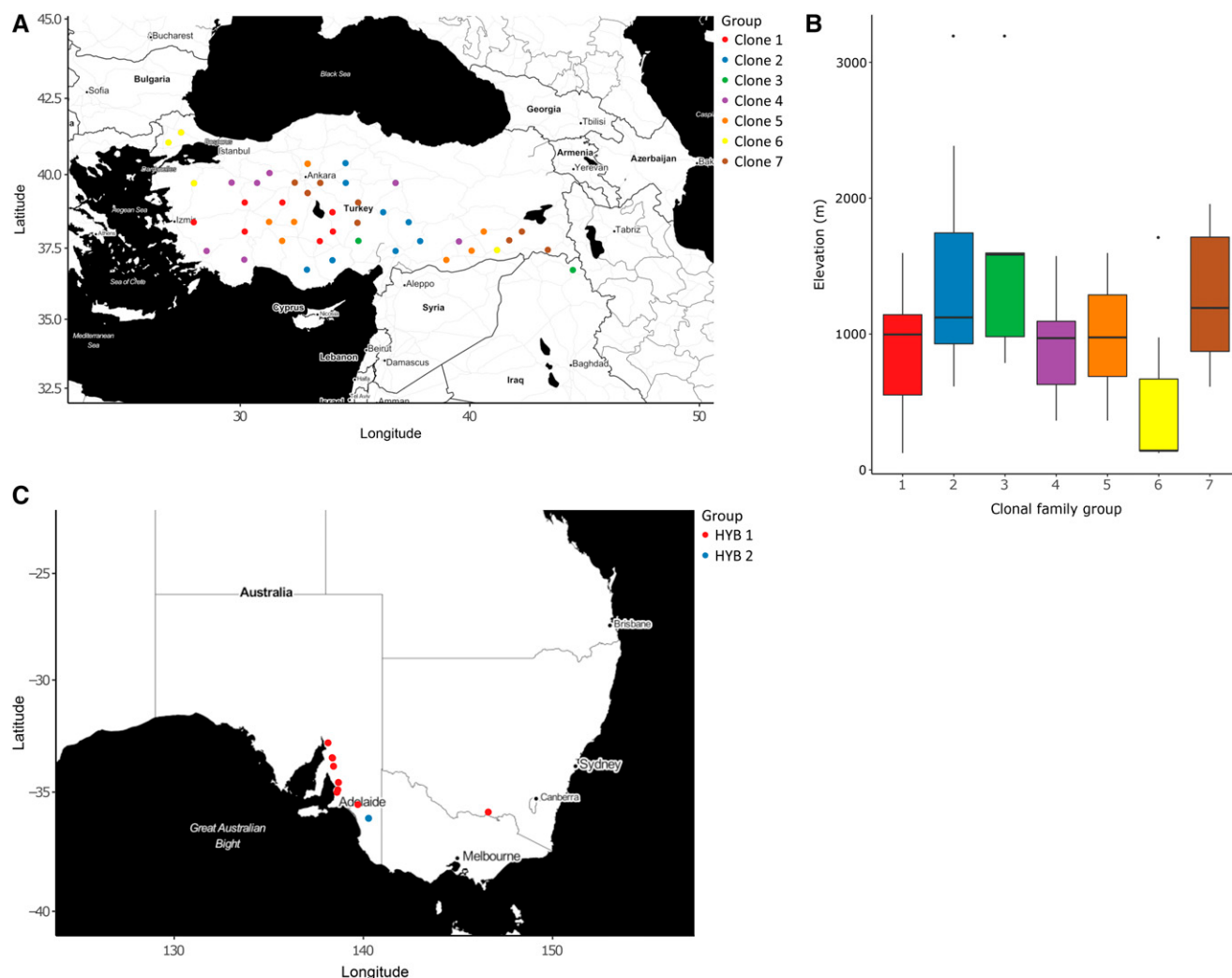


Figure 1 Selection of *Brachypodium* accessions for analysis (A) Map of diploid Turkish accessions colored by genetic relationship. (B) Boxplot of genetic family elevation variation from initial collection sites. (C) Map of *Brachypodium hybridum* accession collection locations. Colors indicate genetic grouping into ‘H’ or ‘H2’ subsets.

weighted logistic regression to model methylation levels and variance between accessions and replicates. Here, the weighted logistic regression uses coverage as a weight. Thereafter, the difference in methylation level is weighted by the p-values to compute the DMRs. Moreover, the spatial correlation present among neighboring cytosine sites is captured by moving average smoothing and the use of weighted voting for histogram based features used by HOME.

Dissection of phenotypic variance

We used linear mixed modeling to estimate the amount of phenotypic variance that can be explained by the differences in CG methylation between accessions while accounting for the genetic effects due to differences between the clone groups. Analyses were performed for each trait and environment separately and was fitted using the R package “Asreml-R”. Initially, we construct a reduced model with one random effect:

$$y = X\mu + Z_g v + \varepsilon$$

Where y is a vector of phenotypes, μ is the population mean, v is the vector of breeding values treated as a random effect with

a $\sim N(0, \sigma^2 G)$ distribution, and ε is the vector of residual effects. Z_g is the design matrix allocating clone groups to individual plants and has been defined as the inverse of the genomic relatedness matrix G . The genomic relatedness matrix was estimated following (VanRaden 2008) that has been adjusted for almost completely homozygous organisms. First, from whole-genome sequencing data from one accession of each clone group, we build the matrix X with the dimensions of number of clone groups ($n = 7$) and number of SNPs ($m = 510,230$). SNPs with at least one copy of the minor allele and no more than one missing value were included in X . X was rescaled to account for allele frequencies to produce W where $W_{ij} = M_{ij} - p_j$, where p_j is the allele frequency for SNP j . Finally, G is calculated as $G = WW' / \sum p_j(1-p_j)$.

We then construct a more complicated model with an added random effect to capture the variance due to differences between the accessions in CG methylation states:

$$y = X\mu + Z_g v + Z_m \rho + \varepsilon$$

Here, ρ is a vector of fitted values ($\sim N(0, \sigma^2 M)$) that we interpret as the methylation equivalent of breeding values for which the

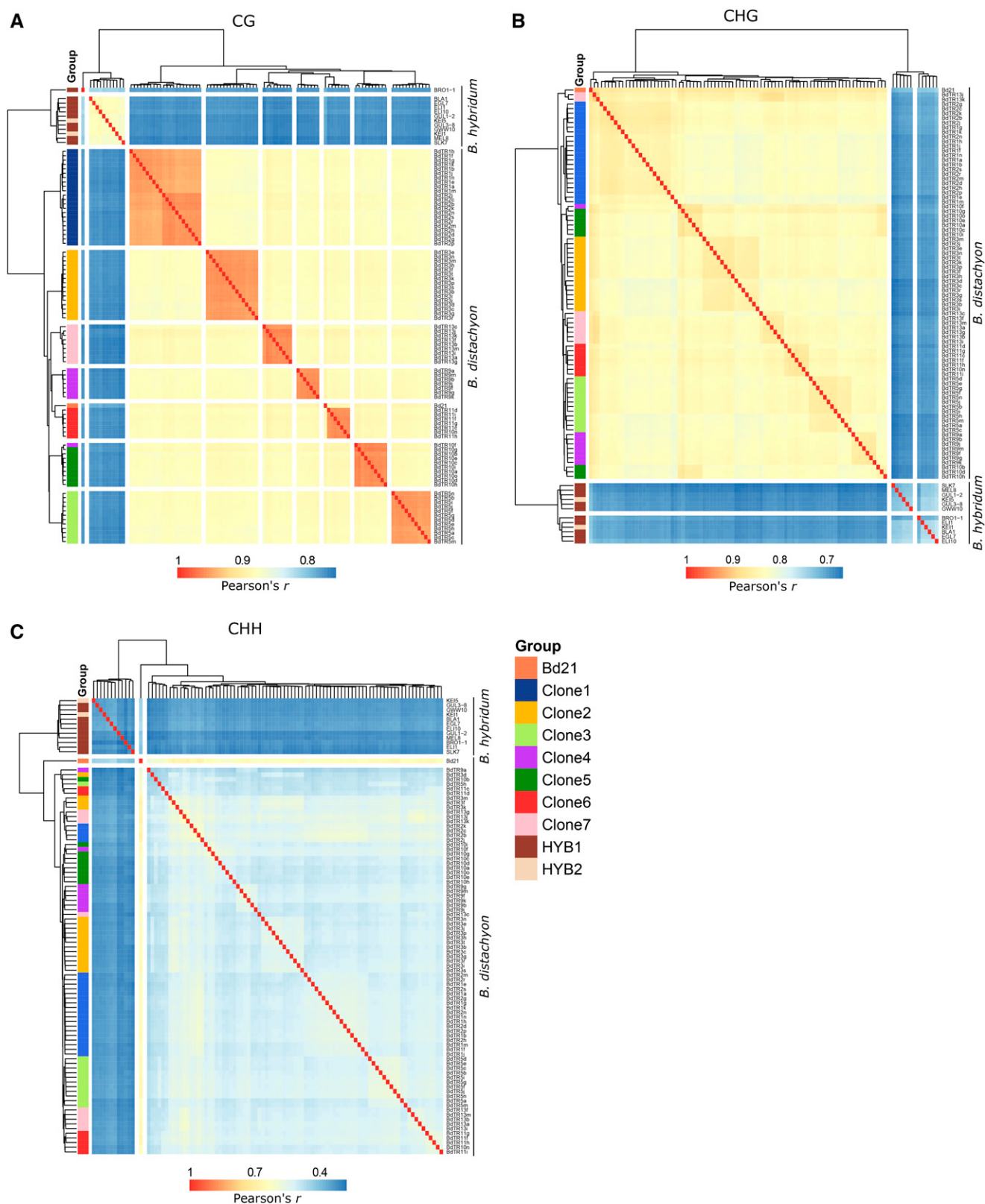


Figure 2 Broad methylation patterns resemble genetic relationships Heat maps representing two-dimensional hierarchical clustering of *Brachypodium* accessions based on correlations (Pearson's r) in genome-wide DNA methylation levels, binned into 100 bp genomic tiles and averaged across all replicates per accession, for (A) CG, (B) CHG, and (C) CHH contexts. Family groups are denoted based on previously established genetic relationships (Wilson et al. 2019).

■ Table 1 CG-DMR analysis using *HOME*

Dataset	Comparison	Mean count	Mean delta	Mean length (bp)	Mean CGs
Low-coverage	Between replicates	505	0.6	199	6
	Between accessions	3,625	0.6	207	8
	Between family groups	15,005	0.7	405	15
High-coverage dataset (family groups 1 and 6)	Between replicates	5,479	0.4	152	9
	Between accession	10,024	0.4	160	11
	Between families/clones	31,980	0.6	390	23

“methylation heritability” can be estimated from the variance in these values. Z_m is the design matrix allocating accessions to plants defined by the inverse of a methylation similarity matrix described below.

We pooled the reads from bisulfite DNA sequencing of the replicates of each accession separately for each environment. For each accession, we scored the proportion of methylated reads compared to the total number of reads at all CG sites. We further pooled CG sites into 200bp windows. The proportion of methylated reads for a specific window was treated as missing data if an accession had less than 10 reads total. Windows with more than 50% of accessions containing missing values were excluded from further analysis, additionally, windows that did not contain a single methylated read were also excluded. In total, 838,231 windows were retained in the spring conditions and 778,183 windows in the fall conditions.

After filtering, we retained a matrix of proportions of methylated reads, Q , with the dimensions of accessions ($n = 83$) by the number of 200bp windows. Columns of Q were scaled to have a mean of zero and unity of variance. The methylation similarity matrix was then calculated as $M = QQ'/N$, where N is a matrix of pairwise number of windows with non-missing data between two accessions.

We used a log likelihood ratio test to determine if the methylation state of the accessions explained a statistically significant amount of phenotypic variance. Here the statistic of 2 times the difference in log likelihood between the models was tested against a chi-squared distribution with one degree of freedom. All variance components were mean-standardized following (Houle 1992).

Data availability

All bisulfite sequencing data are available at the NCBI under BioProject PRJNA349755. All phenotypic and experimental records are accessible at https://github.com/steichten/clonal_brachypodium. Supplemental material available at figshare: <https://doi.org/10.25387/g3.11918430>.

RESULTS

Selection of *Brachypodium* accessions and genetic profiling

Previous genomic analyses of thousands of *Brachypodium* accessions was conducted using genotype-by-sequencing (GBS) culminating in a selection of diverse *B. distachyon* germplasm as a *HapMap* set (Wilson *et al.* 2019). In that study it was observed that some Turkish accessions were highly similar based on SNP assessment, but were located in different geographical regions (Figure 1A). These 83 Turkish accessions (BdTR) were grouped into seven nearly genetically identical ‘families’ based on prior analysis (Sup Table 1). This was generally consistent with the accession’s initial naming based on phenotypic similarity (Vogel *et al.* 2009), though BdTR1 and BdTR2 belonged to the same genotype family. The BdTR accessions within each genetic family were widely dispersed geographically and found

at different elevations (Figure 1B) consistent with wide migration and little recombination of the highly selfing species.

As a genetically distinct global outgroup, we included the allopolyploid relative *Brachypodium hybridum* (*B. hybridum*). Previous studies highlighted two genetically-similar families consisting of 12 accessions collected from across southern Australia. HYB1 consisted of accessions initially collected in South Australia and one location in New South Wales. HYB2 consisted of a single geographic location with multiple individuals (Figure 1C). The selection of germplasm for this study provided a unique, natural system in which to study new heritable variation across a series of genetic families containing minimal *intra*-family variation within, but with substantial *inter*-family genetic and phenotypic variation.

Methylome variance reflects genetic distance

To examine the DNA methylation state of accessions here, low-coverage whole genome bisulfite sequencing was conducted for all 604 experimental plants (Sup Table 2). Samples were correlated over 100 bp genomic tiles to obtain average genome-wide methylation state for all three sequence contexts. CG methylation almost completely reproduced the known genetic relationships - separating *B. hybridum* and grouping the *B. distachyon* accessions based on previously determined family clone groups (Figure 2A). Individual biological replicates were often most similar, however, variation observed near terminal edges may be due to the limitations of low coverage sequencing data. This is apparent in the non-CG methylation contexts (CHG and CHH) in which sample relationships showed less complete recapitulation of genetic relatedness (Figure 2B-C). This was likely due to low coverage preventing accurate measurements for these methylation contexts that occur at lower levels. Therefore we focus the remainder of the analysis on CG sites.

The nested design of this experiment allowed for further quantification of DNA methylation variation at a variety of levels. We identified differentially methylated regions in the CG context (CG-DMRs) using *HOME* for single-sample DMR calling of pairwise comparisons between: (I) replicates of the same accession (basal level of stochastic differences), (II) accessions within families (*intra*-family), and (III) between families (*inter*-family, Table 1, Figure 3). This was performed for the complete dataset (low coverage) and repeated for a subset of samples (family group 1 and 6) sequenced to greater depth (Sup Table 2). Indeed, greater sequencing depth improved the power of *HOME* to detect CG-DMRs. As comparisons were made between samples with increasing genetic distance, a greater mean number of DMRs were identified (of greater magnitude, length, and CG count). On average, fourfold more CG-DMRs were called between, compared to within, family groups, highlighting previous observations that genetically variable samples contain many more DNA methylation variants. Nonetheless, a substantial number of CG-DMRs were still observed within accessions and family groups, which could contribute to heritable and even adaptive phenotypic variation. This warranted further systematic analysis to quantify the phenotypic variance that can be attributed to SNPs and DMRs.

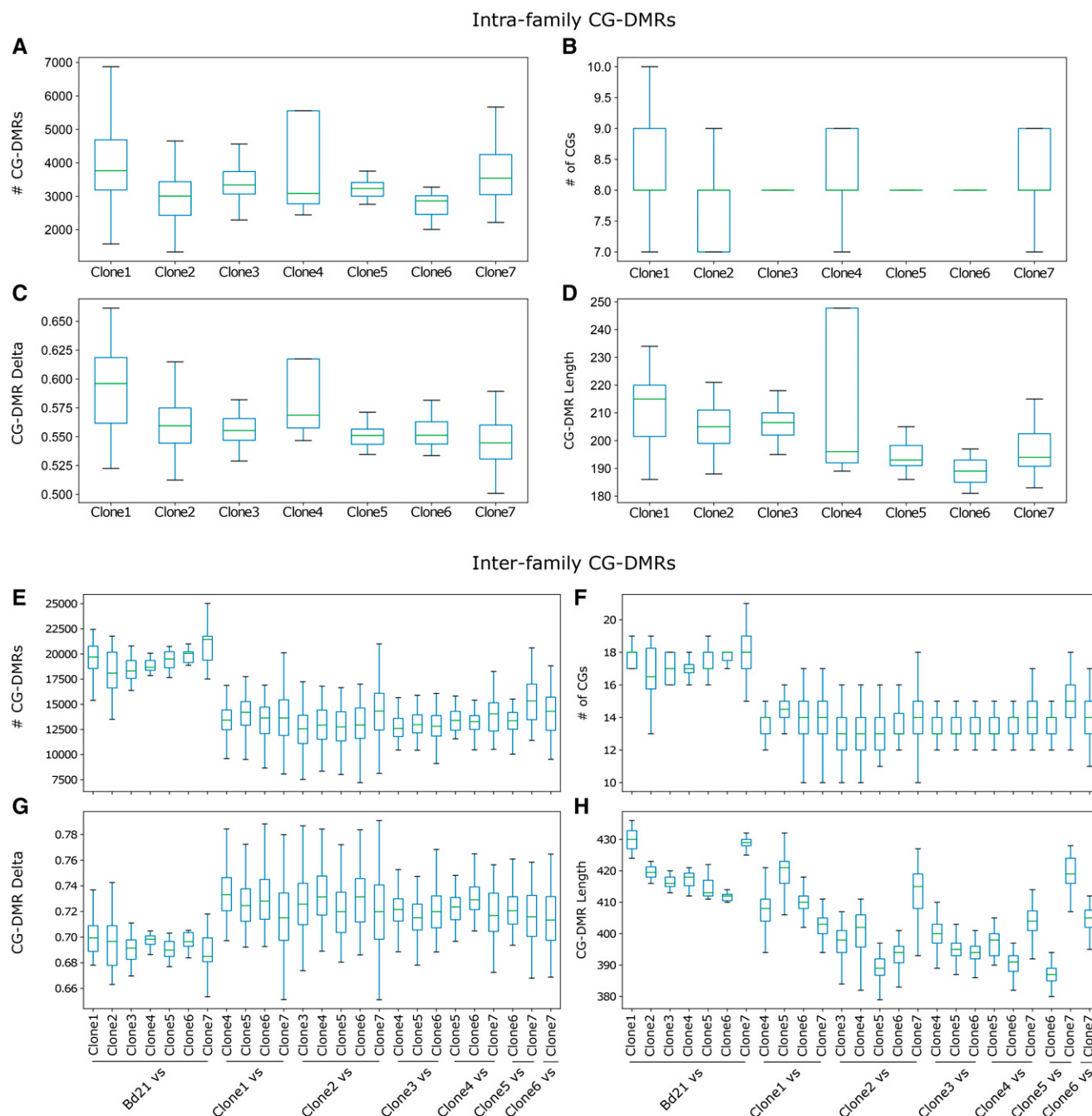


Figure 3 Pairwise intra- and inter-family CG-DMRs Box plots presenting the results of pairwise CG-DMR calling with HOME for intra- (A-D) and inter-family (E-H) comparisons. Plots display the distribution of the number of CG-DMRs (A, E), number of CGs per CG-DMR (B, F), and delta (C, G) and length (D, H) of CG-DMRs.

Dissection of phenotypic variance into genetic and epigenomic components

We used a mixed linear model framework to dissect the phenotypic variance of fitness related traits across multiple levels of biological organization. Our experimental design allowed us to investigate the differences between the species *B. distachyon* and *B. hybridum*, and within our *B. distachyon* samples. We estimated the relative contribution to phenotypic variance of additive genetic effects (as characterized by whole-genome sequence distance among seven clone groups), CG methylation state, which was obtained through

low-coverage bisulfite sequencing of 83 accessions, and the residual variance being estimated via the biological replicates of the accessions. In the analyses that considered all hierarchical levels of variation, the species contrast of *B. distachyon* and *B. hybridum* explained most of the phenotypic variance in flowering time in spring conditions (84%) where *B. distachyon* flowered considerably later than *B. hybridum*. A similar pattern was observed in the fall overwintering conditions; however, this relationship explained less variation overall (20%) as vernalization overwhelms the genetic signal. We also found differences in leaf morphology between the two species, where *B. hybridum*

■ **Table 2** Sources of phenotypic variation Breakdown of the phenotypic variance from the independent effect of SNPs (*i.e.*, broad-sense heritability), the equivalent effect of CG methylation state, and environment variance. Variance components were estimated using linear mixed models and have been mean-standardized and multiplied by 100 for readability. Bold components were found to be statistically significant at $P < 0.05$ using log-likelihood ratio tests between full and reduced models

Trait	Spring			Fall		
	$I_A (H^2_{SNP})$	$I_M (H^2_{CG})$	I_{RES}	$I_A (H^2_{SNP})$	$I_M (H^2_{CG})$	I_{RES}
Flowering time	0.332 (0.485)	0.062 (0.096)	0.269	0.245 (0.894)	0.004 (0.013)	0.026
Height	1.384 (0.387)	0.124 (0.035)	2.073	0.327 (0.373)	0.140 (0.161)	0.407
Tiller count	3.984 (0.633)	0.311 (0.056)	1.992	4.240 (0.597)	0.365 (0.051)	2.494
Leaf length	0.572 (0.483)	0.048 (0.041)	0.563	0.538 (0.523)	0.055 (0.053)	0.437
Leaf width	1.736 (0.627)	<0.001 (<0.001)	1.035	1.015 (0.241)	0.379 (0.090)	2.818
Ear count	0.806 (0.176)	0.308 (0.067)	3.474	0.290 (0.100)	0.232 (0.086)	2.273

had both wider and longer leaves compared to *B. distachyon*. This difference in leaf width and length between the species accounted for 68% and 94% of variance in spring conditions and 76% and 98% in fall conditions for these traits respectively.

Within *B. distachyon*, we observed substantial genetic variance across all traits in both conditions, where heritabilities ranged from 18 to 63% in spring conditions and from 10 to 89% in fall conditions (Table 2). The addition of the random effect of CG methylation state to the models rarely increased the amount of phenotypic variance explained compared to the simpler model where only additive genetic effects were considered. In this study, only two traits had statistical support that the phenotypic variance explained by methylation patterns was greater than zero and, in both cases, this result was dependent on the conditions in which the traits were measured (Table 2). We estimated that an additional 10% of variation in flowering time is due to methylation in spring conditions ($\chi^2 = 7.29$, d.f. = 1, $P = 0.007$), we also found methylation patterns explained an additional 16% of the variation in plant height measured in fall conditions ($\chi^2 = 13.14$, d.f. = 1, $P < 0.001$). We would like to stress the importance of the joint estimation of genetic and methylation contributions to phenotypic variance as these two explanatory variables were highly correlated, *i.e.*, closely related individuals shared similar methylation patterns. In this study, ignoring the underlying DNA sequence variation in our sample population would have led us to grossly overstate the importance of methylation state on phenotypic variance. For example, we found that methylation state only explained 5% of variation in the tiller count in fall conditions ($\chi^2 = 3.39$, d.f. = 1, $P = 0.065$), while additive genetic effects explained 60% of the variation (Table 2). Re-analyzing the data with the removal of the random effect of additive genetics from the model, the amount of variation explained by the methylation state dramatically increased to 51% ($\chi^2 = 133.33$, d.f. = 1, $P < 0.001$). Overall, we show that in some cases significant phenotypic variation can arise due to changes in DNA methylation that can potentially be inherited by subsequent generations, however, we stress that variation in methylation state is largely dependent on the underlying genetic variation and needs to be analyzed together.

DISCUSSION

Epigenomic diversity continues to be considered as a new source of variation in heritable traits that could be harnessed for plant breeding (Springer and Schmitz 2017). However, genetic and epigenomic polymorphisms are often considered independently making it difficult to determine their relative contribution toward heritable phenotypic variation. Here, utilizing a diverse range of clonal *Brachypodium* accessions, grown in two distinct controlled environments (simulating local climates of Istanbul spring and fall),

we systematically quantified the proportion of phenotypic variance, across numerous adaptive plant traits, attributable to genetic polymorphisms (SNPs), variations in DNA methylation, and the environment. Whereas the majority of phenotypic differences across all traits could be attributed to genetic polymorphisms, CG methylation demonstrated an additive effect in particular environments, such as plant height under longer overwintering (fall) conditions.

Although the known population structure of the selected accessions in this study was recapitulated with methylation data, the level of methylation variation was linked to the overall genetic distance between any given sample(s). The methylation variance captured likely reflects an under-estimation of the true difference, given the limitations of low coverage sequencing. Our DMR calling scans for regions of differential methylation larger than single loci, and are therefore more robust to lower coverage. The key of our experimental design is to have higher biological replicates so any observed differences are reproducible. Overall the results are likely conservative given low coverage per sample, while alternative study designs might identify false positives of biological differences among few samples of higher coverage. Nonetheless, these results are similar to reports across natural *Arabidopsis thaliana* populations (Schmitz *et al.* 2013b; Hagmann *et al.* 2015; Kawakatsu *et al.* 2016) and a landrace Wheat collection (Gardiner *et al.* 2018). Methods as described here, can dissect this confounding between genetics and epigenomic components, in particular DNA methylation as studied here, to capture an additional piece of the missing heritability. This highlights the importance of dissecting chromatin variation along with correlated genetic variation to explain phenotypic variation. In the majority of cases where samples are genetically similar, phenotypic variation of quantitative traits is also limited. Despite this, our analyses also revealed that epigenomic distance (CG methylation) could explain additional phenotypic variation of select traits in particular environments.

Going forward, in experimental systems in which genetic sources of variation are not known, it would be advantageous to separate epigenomic (chromatin modifications) from genomic changes (SNPs, SVs) to be able to jointly test effects on a phenotype. Fortunately, advances in genomics, including long reads that also type base modifications, are upon us making this a practical solution (Simpson *et al.* 2017; Kelleher *et al.* 2018; Ni *et al.* 2019). In essence, a joint understanding of genetic and epigenomic relationships could be a new standard for the examination of quantitative trait variation.

ACKNOWLEDGMENTS

The authors would like to thank the Australian Plant Phenomics Facility for providing plant growth facilities and the National

Computational Infrastructure for provision of computing resources, both of which are supported under the National Collaborative Research Infrastructure Strategy of the Australian Government. Illumina sequencing was performed at the Biomolecular Research Facility at the ANU. This project was supported by the Australian Research Council Centre of Excellence in Plant Energy Biology (CE140100008). SRE was funded by an Australian Research Council Discovery Early Career Research Award (DE150101206).

LITERATURE CITED

- Becker, C., J. Hagmann, J. Müller, D. Koenig, O. Stegle *et al.*, 2011 Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. *Nature* 480: 245–249. <https://doi.org/10.1038/nature10555>
- Brown, T. B., R. Cheng, X. R. R. Sirault, T. Rungrat, K. D. Murray *et al.*, 2014 TraitCapture: genomic and environment modelling of plant phenomic data. *Curr. Opin. Plant Biol.* 18: 73–79. <https://doi.org/10.1016/j.pbi.2014.02.002>
- Cortijo, S., R. Wardenaar, M. Colomé-Tatché, A. Gilly, M. Etcheverry *et al.*, 2014 Mapping the epigenetic basis of complex traits. *Science* 343: 1145–1148. <https://doi.org/10.1126/science.1248127>
- Dubin, M. J., P. Zhang, D. Meng, M.-S. Remigereau, E. J. Osborne *et al.*, 2015 DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. *eLife* 4: e05255. <https://doi.org/10.7554/eLife.05255>
- Eichten, S. R., R. Briskine, J. Song, Q. Li, R. Swanson-Wagner *et al.*, 2013 Epigenetic and Genetic Influences on DNA Methylation Variation in Maize Populations. *Plant Cell* 25: 2783–2797. <https://doi.org/10.1105/tpc.113.114793>
- Eichten, S. R., R. J. Schmitz, and N. M. Springer, 2014 Epigenetics: Beyond Chromatin Modifications and Complex Genetic Regulation. *Plant Physiol.* 165: 933–947. <https://doi.org/10.1104/pp.113.234211>
- Eichten, S. R., T. Stuart, A. Srivastava, R. Lister, and J. O. Borevitz, 2016 DNA methylation profiles of diverse Brachypodium distachyon align with underlying genetic diversity. *Genome Res.* 26: 1520–1531. <https://doi.org/10.1101/gr.205468.116>
- Gardiner, L.-J., R. Joynton, J. Omony, R. Rusholme-Pilcher, L. Olohan *et al.*, 2018 Hidden variation in polyploid wheat drives local adaptation. *Genome Res.* 28: 1319–1332. <https://doi.org/10.1101/gr.233551.117>
- Hagmann, J., C. Becker, J. Müller, O. Stegle, R. C. Meyer *et al.*, 2015 Century-scale methylome stability in a recently diverged Arabidopsis thaliana lineage. *PLoS Genet.* 11: e1004920. <https://doi.org/10.1371/journal.pgen.1004920>
- He, L., W. Wu, G. Zinta, L. Yang, D. Wang *et al.*, 2018 A naturally occurring epiallele associates with leaf senescence and local climate adaptation in Arabidopsis accessions. *Nat. Commun.* 9: 460. <https://doi.org/10.1038/s41467-018-02839-3>
- Hofmeister, B. T., K. Lee, N. A. Rohr, D. W. Hall, and R. J. Schmitz, 2017 Stable inheritance of DNA methylation allows creation of epigenotype maps and the study of epiallele inheritance patterns in the absence of genetic variation. *Genome Biol.* 18: 155. <https://doi.org/10.1186/s13059-017-1288-x>
- Houle, D., 1992 Comparing evolvability and variability of quantitative traits. *Genetics* 130: 195–204.
- International Brachypodium Initiative, 2010 Genome sequencing and analysis of the model grass Brachypodium distachyon. *Nature* 463: 763–768. <https://doi.org/10.1038/nature08747>
- Kawakatsu, T., S.-S. C. Huang, F. Jue, E. Sasaki, R. J. Schmitz *et al.*, 2016 Epigenomic Diversity in a Global Collection of Arabidopsis thaliana Accessions. *Cell* 166: 492–505. <https://doi.org/10.1016/j.cell.2016.06.044>
- Kelleher, P., J. Murphy, J. Mahony, and D. van Sinderen, 2018 Identification of DNA Base Modifications by Means of Pacific Biosciences RS Sequencing Technology. *Methods Mol. Biol.* 1681: 127–137. https://doi.org/10.1007/978-1-4939-7343-9_10
- Krueger, F., and S. R. Andrews, 2011 Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27: 1571–1572. <https://doi.org/10.1093/bioinformatics/btr167>
- Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorf *et al.*, 2009 Finding the missing heritability of complex diseases. *Nature* 461: 747–753. <https://doi.org/10.1038/nature08494>
- Ni, P., N. Huang, Z. Zhang, D.-P. Wang, F. Liang *et al.*, 2019 DeepSignal: detecting DNA methylation state from Nanopore sequencing reads using deep-learning. *Bioinformatics* 35: 4586–4595. <https://doi.org/10.1093/bioinformatics/btz276>
- Ong-Abdullah, M., J. M. Ordway, N. Jiang, S.-E. Ooi, S.-Y. Kok *et al.*, 2015 Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature* 525: 533–537. <https://doi.org/10.1038/nature15365>
- Pignatta, D., R. M. Erdmann, E. Scheer, C. L. Picard, G. W. Bell *et al.*, 2014 Natural epigenetic polymorphisms lead to intraspecific variation in Arabidopsis gene imprinting. *eLife* 3: e03198. <https://doi.org/10.7554/eLife.03198>
- Rebetzke, G. J., and R. A. Richards, 1999 Genetic improvement of early vigour in wheat. *Aust. J. Agric. Res.* 50: 291. <https://doi.org/10.1071/A98125>
- Schmid, M. W., C. Heichinger, D. Coman Schmid, D. Guthörl, V. Gagliardini *et al.*, 2018 Contribution of epigenetic variation to adaptation in Arabidopsis. *Nat. Commun.* 9: 4446. <https://doi.org/10.1038/s41467-018-06932-5>
- Schmitz, R. J., Y. He, O. Valdés-López, S. M. Khan, T. Joshi *et al.*, 2013a Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. *Genome Res.* 23: 1663–1674. <https://doi.org/10.1101/gr.152538.112>
- Schmitz, R. J., M. D. Schultz, M. G. Lewsey, R. C. O'Malley, M. A. Urich *et al.*, 2011 Transgenerational epigenetic instability is a source of novel methylation variants. *Science* 334: 369–373. <https://doi.org/10.1126/science.1212959>
- Schmitz, R. J., M. D. Schultz, M. A. Urich, J. R. Nery, M. Pelizzola *et al.*, 2013b Patterns of population epigenomic diversity. *Nature* 495: 193–198. <https://doi.org/10.1038/nature11968>
- Shaw, R. G., D. L. Byers, and E. Darmono, 2000 Spontaneous mutational effects on reproductive traits of Arabidopsis thaliana. *Genetics* 155: 369–378.
- Shen, X., J. De Jonge, S. K. G. Forsberg, M. E. Pettersson, Z. Sheng *et al.*, 2014 Natural CMT2 variation is associated with genome-wide methylation changes and temperature seasonality. *PLoS Genet.* 10: e1004842. <https://doi.org/10.1371/journal.pgen.1004842>
- Simpson, J. T., R. E. Workman, P. C. Zuzarte, M. David, L. J. Dursi *et al.*, 2017 Detecting DNA cytosine methylation using nanopore sequencing. *Nat. Methods* 14: 407–410. <https://doi.org/10.1038/nmeth.4184>
- Spokas, K., and F. Forcella, 2006 Estimating hourly incoming solar radiation from limited meteorological data. *Weed Sci.* 54: 182–189. <https://doi.org/10.1614/WS-05-098R.1>
- Springer, N. M., and R. J. Schmitz, 2017 Exploiting induced and natural epigenetic variation for crop improvement. *Nat. Rev. Genet.* 18: 563–575. <https://doi.org/10.1038/nrg.2017.45>
- Srivastava, A., Y. V. Karpievitch, S. R. Eichten, J. O. Borevitz, and R. Lister, 2019 HOME: a histogram based machine learning approach for effective identification of differentially methylated regions. *BMC Bioinformatics* 20: 253. <https://doi.org/10.1186/s12859-019-2845-y>
- Trerotola, M., V. Relli, P. Simeone, and S. Alberti, 2015 Epigenetic inheritance and the missing heritability. *Hum. Genomics* 9: 17. <https://doi.org/10.1186/s40246-015-0041-3>
- VanRaden, P. M., 2008 Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91: 4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- Vaughn, M. W., M. Tanurdzić, Z. Lippman, H. Jiang, R. Carrasquillo *et al.*, 2007 Epigenetic natural variation in Arabidopsis thaliana. *PLoS Biol.* 5: e174. <https://doi.org/10.1371/journal.pbio.0050174>
- Vogel, J. P., M. Tuna, H. Budak, N. Huo, Y. Q. Gu *et al.*, 2009 Development of SSR markers and analysis of diversity in Turkish populations of

- Brachypodium distachyon. BMC Plant Biol. 9: 88. <https://doi.org/10.1186/1471-2229-9-88>
- Wilson, P. B., G. J. Rebetzke, and A. G. Condon, 2015 Of growing importance: combining greater early vigour and transpiration efficiency for wheat in variable rainfed environments. Funct. Plant Biol. 42: 1107. <https://doi.org/10.1071/FP15228>
- Wilson, P. B., J. C. Streich, K. D. Murray, S. R. Eichten, R. Cheng *et al.*, 2019 Global Diversity of the Brachypodium Species Complex as a Resource for Genome-Wide Association Studies Demonstrated for Agronomic Traits in Response to Climate. Genetics 211: 317–331. <https://doi.org/10.1534/genetics.118.301589>
- Zuk, O., E. Hechter, S. R. Sunyaev, and E. S. Lander, 2012 The mystery of missing heritability: Genetic interactions create phantom heritability. Proc. Natl. Acad. Sci. USA 109: 1193–1198. <https://doi.org/10.1073/pnas.1119675109>

Communicating editor: A. Rokas