

Trans Chromosomal Methylation in *Arabidopsis* hybrids

Ian K. Greaves^{a,b,1}, Michael Groszmann^{a,c,1}, Hua Ying^a, Jennifer M. Taylor^a, W. James Peacock^{a,2}, and Elizabeth S. Dennis^a

^aCommonwealth Scientific and Industrial Research Organization Plant Industry, Canberra, ACT 2600, Australia; ^bDepartment of Genome Biology, John Curtin School of Medical Research, Australia National University, Canberra, ACT 0200, Australia; and ^cNSW Agricultural Genomics Centre, Wagga Wagga, NSW 2650, Australia

Contributed by W. James Peacock, January 23, 2012 (sent for review December 13, 2011)

The heterotic hybrid offspring of *Arabidopsis* accessions C24 and Landsberg erecta have altered methylomes. Changes occur most frequently at loci where parental methylation levels are different. There are context-specific biases in the nonadditive methylation patterns with ^mCG generally increased and ^mCHH decreased relative to the parents. These changes are a result of two main mechanisms, Trans Chromosomal Methylation and Trans Chromosomal deMethylation, where the methylation level of one parental allele alters to resemble that of the other parent. Regions of altered methylation are enriched around genic regions and are often correlated with changes in siRNA levels. We identified examples of genes with altered expression likely to be due to methylation changes and suggest that in crosses between the C24 and Ler accessions, epigenetic controls can be important in the generation of altered transcription levels that may contribute to the increased biomass of the hybrids.

DNA methylation | heterosis | RNA-directed DNA methylation | epigenetics | hybrid vigor

In the formation of a hybrid, the genome and epigenome of each of the parents are brought together within the one nucleus. The interactions of these two sets of genetic instructions result in the unique characteristics of the hybrid, including superior performance. Both the level and pattern of expression of many genes are altered in hybrids (1–3). Altered transcription levels have mostly been explained by the interaction between the alleles of a gene delivered by the parents involving a range of interactions such as dominance, overdominance, and epistatic interactions between loci (1). Despite these genetic analyses, there is a lack of understanding of the molecular mechanisms underpinning heterosis.

It has been suggested that the magnitude of hybrid vigor is positively correlated with the genetic distance or amount of sequence variation between the parental genomes (4, 5). However, crosses between genetically similar parents such as *Arabidopsis* accessions or subspecies of rice can produce hybrids displaying significant heterosis, apparently breaking down the relationship between genetic distance and extent of hybrid vigor (6, 7). It has been reported that the epigenome evolves at a significantly faster rate than the genetic sequence (8–10), consistent with genetically similar parents having markedly different epigenomes (11–20). These epigenomic systems, such as DNA methylation and small RNAs, play a vital role in genomic stability, development, and the regulation of genes within a plant. The epigenome may contribute the allelic variability needed to generate heterosis in crosses between genetically similar parents. We previously reported that the *Arabidopsis* C24 and Ler accessions have different siRNA distributions and that the reciprocal heterotic hybrids show a 27% reduction in the levels of 24-nt siRNAs (18). The major reduction in these 24-nt siRNA sequences corresponded to those segments of the genome, primarily the gene bodies and their flanking sequences, where the two parents had unequal levels of 24-nt siRNAs. Given the importance of 24-nt siRNAs in directing DNA methylation through targeted sequence homology, it is likely that these same hybrids would have methylation patterns distinct to those found in the parents. The altered methylome could have consequences for gene expression and play a role in the development of heterosis.

In this study, we showed that the methylation patterns of the parental lines are largely inherited faithfully in the hybrids (additive methylation), with exceptions applying mostly to those regions of the genome where the parents have unequal levels of DNA methylation (nonadditive methylation). The differences in the nonadditive methylation frequencies apply to all three cytosine sequence contexts (^mCG, ^mCHG, and ^mCHH) with ^mCG predominantly showing an increase and ^mCHH a general decrease. In some cases, we were able to track the parental chromosomes by virtue of SNPs in the DNA sequence. We identified processes causing the nonadditive methylation levels in the hybrids that we termed Trans Chromosomal Methylation (TCM) and Trans Chromosomal deMethylation (TCdM). These processes are frequently associated with the action of 24-nt siRNAs resulting in changes of methylation according to the contributions of siRNAs from both parental sources. We describe several examples showing the effects the altered methylome of the hybrids can have on patterns of gene expression.

Results

We compared the methylomes of C24, Ler, and their reciprocal F1 hybrids by generating ≈ 90 million reads on bisulphite-treated DNA from each genome. The data provided at least two-read coverage of $\approx 68\%$ of cytosines (Cs) in the genome (Dataset S1, Table S1). Both parental genomes have similar frequencies of highly methylated C residues (^mCs) in each of the three ^mC contexts— $\approx 26\%$ ^mCG, $\approx 6\%$ ^mCHG, and $\approx 2\%$ ^mCHH (Dataset S1, Table S1)—that are similar to levels found in Columbia (21). These frequencies increased to $\approx 30\%$ ^mCG, $\approx 15\%$ ^mCHG, and $\approx 8\%$ ^mCHH (Dataset S1, Table S2) when accounting for all ^mCs with a level of methylation above the false-positive methylation rate, which was determined by correcting against the unmethylated chloroplast genome sequences (SI Materials and Methods). Both hybrids showed slight increases in the frequency of ^mC (above parental levels) to $\approx 37\%$ ^mCG, $\approx 19\%$ ^mCHG, and $\approx 10\%$ CHH (Dataset S1, Tables S1 and S2), consistent with the hybrids inheriting ^mCs common to both parents and sites unique to one of the parents.

In both parents and hybrids, the majority of ^mCG sites are highly methylated ($>80\%$), whereas ^mCHG and ^mCHH sites display a broader distribution of methylation (Fig. S14). Chromosomal distributions of ^mC mirrored the patterns of 24-nt siRNA density with the highest levels associated with the pericentric regions of the chromosomes (Fig. 14 and Fig. S2). In genic regions, the frequency of ^mCG sites is high in the gene body and lower in the flanking regions, whereas ^mCHG and ^mCHH sites are highest in flanking regions and lower throughout the

Author contributions: I.K.G., M.G., W.J.P., and E.S.D. designed research; I.K.G. and M.G. performed research; I.K.G., M.G., H.Y., and J.M.T. analyzed data; and I.K.G., M.G., W.J.P., and E.S.D. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE35542).

¹I.K.G. and M.G. contributed equally to this work.

²To whom correspondence should be addressed. E-mail: jim.peacock@csiro.au.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1201043109/-DCSupplemental.

gene body (Fig. S1B). The bulk of the chromosomal distribution and levels of methylation are similar in parents and hybrids, paralleling the data for siRNA in this hybrid system (18). Localized differences were observed where siRNA expression in the hybrid deviated from the expected midparent value (MPV; average of the two parents), most frequently at loci with large differences in siRNA levels between parents. To examine localized changes in mC s, we used a cluster approach identifying regions containing several mC s within localized segments (*SI Materials and Methods*). Two mC cluster datasets were produced for each of the parents and F1 hybrids; the first used a more stringent cluster criterion (*SI Materials and Methods*), defining a cluster to have a minimum of six mC s with the distance between two adjacent mC s no greater than 20 nt. The second dataset involved more of the mC -seq data by defining clusters containing at least three mC s with no greater than 20 nt between neighboring mC s. Both approaches yielded similar results (see *Dataset S2* for a comparison) and, as such, the 3 mC cluster dataset was used for further analyses so as to use the greatest mC -seq coverage. Within this dataset, the parental and hybrid samples contained an average of $\approx 152,000$ clusters (*Dataset S1, Table S3*) located in transposons (TEs), gene bodies (transcribed regions of genes), flanking regions of gene bodies (± 1 kb), or in intergenic regions (Fig. S3A and *SI Materials and Methods*). TEs have the most extensive stretches of methylation (Fig. S3B), whereas the flanking regions show elevated mC densities for both intergenic and TE clusters (Fig. S3B). Intragenic regions generally contain comparatively short and sparsely methylated clusters (Fig. S3B).

C24 and Ler Accessions Have Localized Differences in Cytosine Methylation. To compare parental methylomes we identified $\approx 114,000$ mC clusters with coverage in both parents (*Dataset S1, Table S4*). Each cluster was allocated a methylation score (0 = Low to ≥ 5 = High) dependent on its average methylation for all mC s and for each of the methylation contexts (*Dataset S1, Table S5*; see *SI Materials and Methods*). Clusters were considered differentially methylated if their scores differed by two or more units (≥ 2).

Between the two parents, $\approx 23\%$ of mC clusters were differentially methylated based on total mC levels (*Dataset S1, Table S4*). When individual contexts were analyzed independently, $mCHH$ was most frequently differentially methylated (35% of clusters) with mCG and $mCHG$ showing 21% and 18%, respectively (*Dataset S1, Table S4*). The preponderance of CG methylation is sufficient in many instances to mask the changes in $mCHH$ and $mCHG$ in clusters. Clusters differentially methylated between parents were enriched in gene bodies and their

flanking regions and were underrepresented in TEs (Fig. 1B and Fig. S3C). These findings are consistent with an earlier tiling array analysis, which showed that methylation differed between the Col and Ler accessions in gene bodies but not in TEs (13).

F1 Hybrids Show Locus Specific Nonadditive Methylation. We identified 76,496 C24 mC clusters that had coverage in Ler and at least one hybrid and 68,893 Ler mC clusters that had coverage in C24 and at least one hybrid (*Dataset S1, Table S6*; see *SI Materials and Methods*). Hybrid clusters were assigned a methylation score for both the expected MPV and observed mC level with nonadditive methylation identified if scores between the MPV and observed value differed by ≥ 2 (*Dataset S1, Table S5* and *SI Materials and Methods*). The majority of clusters identified as either additive or nonadditive showed a similar state in both reciprocal hybrids (*Dataset S1, Table S7*). Seven percent of the clusters showed nonadditive methylation in total mC levels, and 20% showed nonadditive methylation in at least one mC context (*Dataset S1, Table S6*). $mCHH$ levels were nonadditive in $\approx 22\%$ of clusters, $mCHG$ in $\approx 8\%$, and mCG in $\approx 5\%$ (*Dataset S1, Table S6*). Nonadditive mC clusters were enriched in gene bodies and flanking regions and underrepresented in TEs (Fig. 1C and Fig. S3C). Differential methylation between parents was also enriched in gene bodies and flanking regions (Fig. 1B), suggesting a correlation between differences in parental methylation and nonadditive methylation in hybrids. To determine whether this correlation was the case, hybrid clusters were divided into 10 groups based on differences in methylation between parents (Fig. 2 and *Dataset S2*). The frequency of nonadditive methylation in the hybrids increased when parental methylation levels were different. Clusters showing nonadditive $mCHG$ had approximately equal frequencies above and below the MPV, whereas nonadditive mCG clusters were predominantly above MPV and nonadditive $mCHH$ clusters predominantly below MPV.

Alterations to the Hybrid Methylome Occur Through Trans Chromosomal Interactions. Nonadditive methylation in the hybrids indicates alterations to the mC levels of one or both parental epialleles. Nonadditive methylation in the hybrid may be achieved in a number of ways, for instance, a nonadditive increase in methylation can arise either through an increase in methylation of the low parent, the high parent, or both parental epialleles. To determine what patterns of change in allelic methylation most frequently lead to nonadditive methylation in the hybrid, we tracked 5,479 mC clusters where the methylation pattern for each parental allele could be followed by using SNPs within the sequenced reads of the C24 $\times$ Ler hybrid (*Dataset S1, Table S8*;

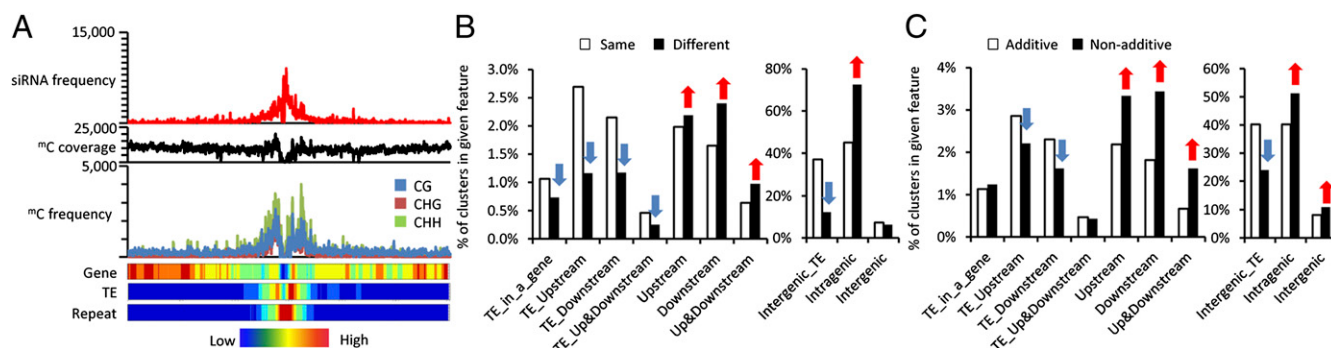


Fig. 1. Distribution of siRNA and DNA methylation frequencies across chromosomes and features of the genome. (A) Depicted here is chromosome 1 from C24 showing a representative pattern consistent across all chromosomes in both parents and hybrids (refer to Fig. S2). Cytosine DNA methylation and 24-nt siRNA densities were plotted by using a 50-kb window stepped in 25-kb increments. The heat maps show gene, TE, and repeat densities across the chromosome. Densely methylated regions correlate with higher levels of siRNAs, mainly around the pericentric region and associated with high TE densities. (B) A comparison of the C24-identified mC clusters between C24 and Ler parents. For Ler-identified clusters, refer to Fig. S3. (C) Additively and nonadditively methylated C24-identified mC clusters in the C24 $\times$ Ler hybrid. Arrows indicate statistical enrichment (red) or underrepresentation (blue) by using the Fisher exact test ($P < 0.05$). For Ler-identified cluster and Ler $\times$ C24 hybrid data refer to Fig. S3.

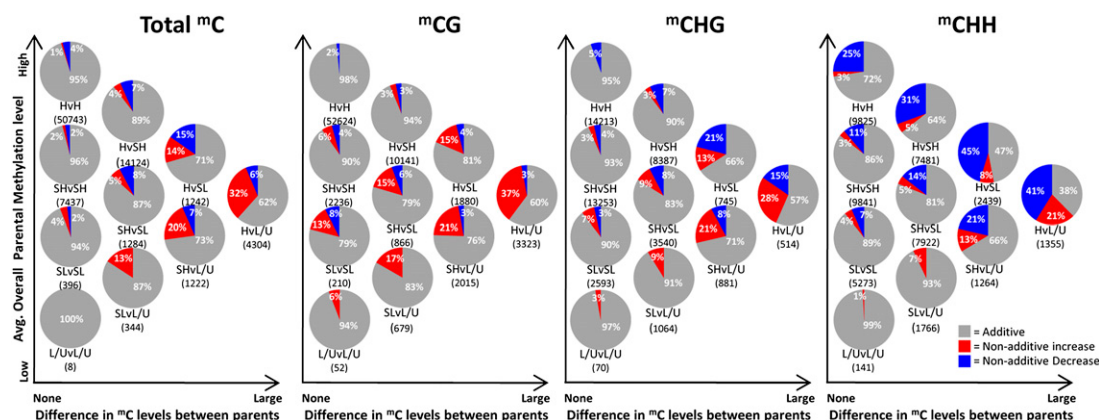


Fig. 2. The frequency of additive and nonadditive mC levels in hybrids. Nonadditive methylation in the hybrids occurs more frequently when parental methylation levels are different. Represented here are the C24-identified clusters in the C24 x Ler hybrid. The C24 x Ler hybrid clusters were divided into 10 groups based on the similarity of methylation levels between parents (H, highly methylated; SH, semi-high; SL, semi-low—refer to Dataset S1, Table S5 for respective index scores). The methylation category of each parent (C24 v. Ler) and the number of clusters in each group are shown below each pie chart. The y axis separates groups on the overall parental methylation levels. The x axis separates groups based on the degree of difference in methylation levels between the parents. The Ler x C24 hybrid and Ler-identified clusters produced similar patterns (Dataset S2).

see *SI Materials and Methods*). Consistent with the general cluster analysis, the majority of SNP clusters showing non-additive methylation in the hybrid occurred most frequently when parents had different levels of methylation (Fig. 3). Additive methylation in the hybrids was almost exclusively due to parental epialleles retaining their methylation pattern in the hybrid (Fig. 3 and Fig. S44). Of the possible changes to allelic methylation that would lead to nonadditive methylation, two predominant patterns emerged (Fig. 3 and Fig. S44). The first involved the high parent mC epiallele retaining its methylation pattern, whereas the low parent epiallele increased in methylation, leading to an overall gain in methylation levels. We called this Trans Chromosomal Methylation (TCM; Fig. 3 and Fig. S4). The second involves a decrease in methylation of the high parent epiallele, causing a reduction in overall methylation, which we called Trans Chromosomal deMethylation (TCdM; Fig. 3 and Fig. S4). These two trans chromosomal events accounted for 681 of the 789 (86%) nonadditive mC allelic inheritance events and represent the dominant processes leading to nonadditive methylation in the hybrids (Fig. 3).

TCM events often lead to increases in methylation levels above MPV that can only be obtained through methylation of the previously unmethylated or low methylated parental allele. TCM events can be indirectly observed in the absence of SNPs by

studying nonadditive mC clusters where one parent is highly methylated, the other parent has no methylation, and the hybrid a higher than MPV mC level (Fig. S4B). Conversely, TCdM can also be observed without SNPs in loci where the high parent methylation level drops significantly in the hybrids (Fig. S4B). Often a corresponding TCM can be scored in the allele from the low methylation parent. Among $\approx 5,800$ mC clusters where one parent was methylated and the other unmethylated, $\approx 2,700$ gained the observed mC levels through TCM or TCdM (Fig. S4B). All clusters showing nonadditive methylation are included within this subset of clusters. Extrapolation from our dataset provides an estimate of $\approx 8,500$ TCM/TCdM events genome wide (Dataset S1, Table S9; see *SI Materials and Methods*).

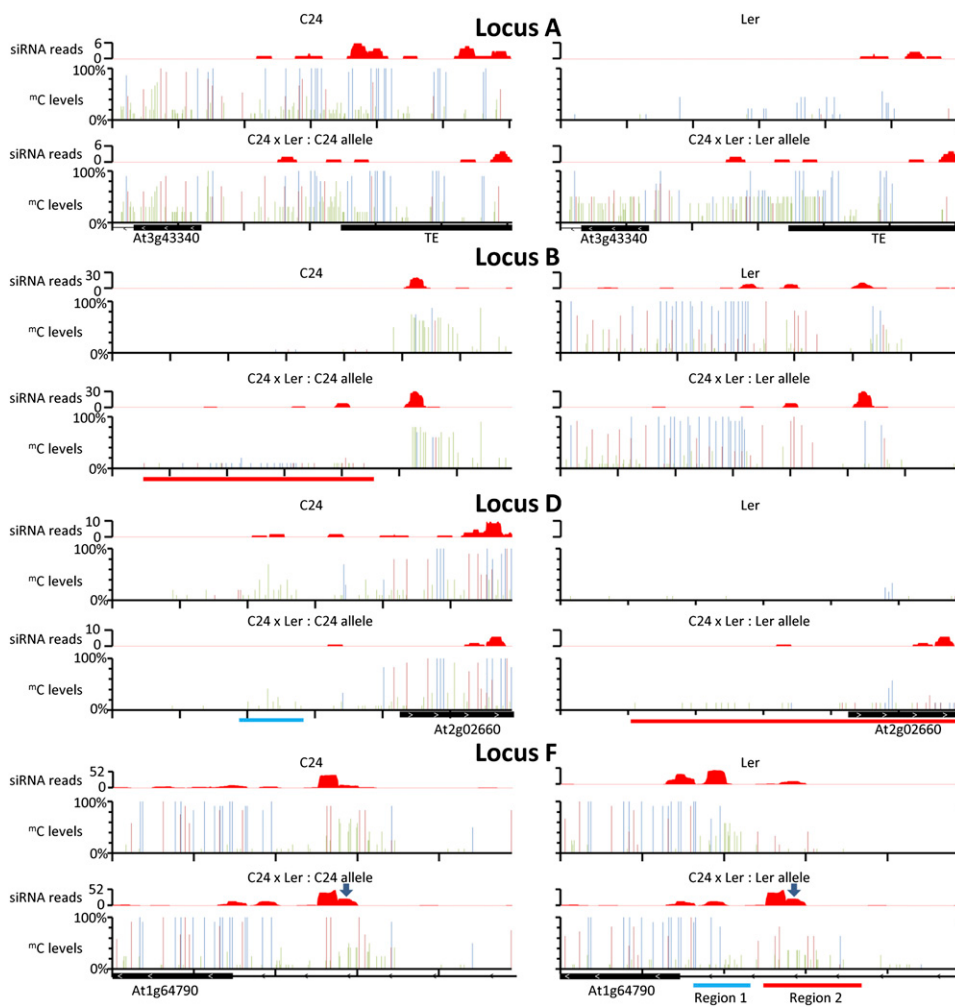
Bisulphite PCR followed by Sanger sequencing showed examples of TCM and TCdM over segments longer than could be tracked through the shorter deep sequencing reads around SNPs (Fig. 4 and Fig. S5). Locus A was methylated in all C residues in C24 and had only low CG methylation in Ler (Fig. 4 and Fig. S5A). In both reciprocal hybrids, the C24 allele retained the level and distribution of methylation of the C24 parent. The Ler allele showed a dramatic change in its methylation state, the mC pattern resembling the C24 epiallele with an increase to existing mC residues by de novo methylation of many CG, CHG, and CHH sites throughout the 700-nt region (Fig. 4 and Fig. S5A). Not all TCM events lead to large changes and some loci show TCM at low levels predominantly associated with CG and CHG sites (Loci B and C; Fig. 4 and Fig. S5B). TCdM is obvious at Locus D where a cluster of CHH and CHG sites in the C24 parent shows decreased mC levels in the C24 allele of both reciprocal hybrids (Fig. 4 and Fig. S5A). A TCM event also occurs at the Ler allele at the same locus (Fig. 4 and Fig. S5A).

These trans chromosomal processes operate in all chromosomal regions—including gene bodies (Loci A, C, and D) and their flanking sequences (Loci D and F), in transposable elements (Locus A) and in intergenic sequences (Locus B). Both TCM and TCdM processes can occur within the one chromosomal segment and even within a relatively limited neighborhood of a segment sequence (Locus F; Fig. 4 and Fig. S5).

In the above examples, siRNAs are present in at least one parent and these siRNAs are expected to associate with homologous sequences of both parental alleles in the hybrids. In TCM events, de novo methylation leads to increases in existing mC s and methylation of previously unmethylated Cs. In the low methylated parental allele the increases coincide with the distribution of siRNAs inherited from the more methylated parent (Loci A–C; Fig. 4 and Fig. S5). In TCdM events (Locus D), the

		Parental alleles with similar levels of methylation (n = 4,626)			Parental alleles with different levels of methylation (n = 853)		
		High mC Parent allele			High mC Parent allele		
Low mC parent allele	Increased mC level in hybrid	12 (0.3%)	207 (4.5%)	12 (0.3%)	3 (0.4%)	132 (15.5%)	23 (2.7%)
	Unchanged mC level in hybrid	-	4,144 (89.5%)	234 (5.0%)	20 (2.3%)	546 (64.0%)	108 (12.7%)
	Decreased mC level in hybrid	-	-	17 (0.4%)	1 (0.1%)	12 (1.4%)	8 (0.9%)

Fig. 3. Changes to parental allelic methylation levels in the C24 x Ler hybrid. Permutation tables of all possible changes to parental allelic methylation levels in the hybrid. The majority of allelic methylation is stably inherited (gray), whereas the majority of nonadditive methylation occurs through TCM (red) and TCdM (blue). In the “parents similar” table, the parents were separated into high parent and low parent based on the minor differences of methylation levels.



serves to highlight that nonadditive methylation in the hybrids apparently leads to alterations in gene expression.

Nine examples are shown in Fig. 5 and Fig. S6C, which were matched against Columbia transcriptome and methylome data from wild-type (WT) and mutants with altered DNA methylation patterns (24, 26, 27). At2g32160, At3g43340, and At5g47590 show TCM extending over the promoter region and into the gene body and having an extensive siRNA distribution correlated with a decrease in gene expression (Fig. 5). At5g38720, At4g19690, and At4g22310 show TCM in either the downstream region or in the last exon of the gene correlated with a decrease in expression (Fig. S6C). Consistent with a role for DNA methylation in controlling gene expression in At2g32160, the Columbia WT epiallele resembles the unmethylated Ler epiallele. In the *rdd* mutant, which lacks demethylases, this allele becomes methylated in a pattern similar to C24 and the hybrids, and gene activity is suppressed (Dataset S1, Table S11). At3g43340 is transcriptionally active in a *met1* Col mutant, and shows a loss of methylation in the upstream region and first exon (Dataset S1, Table S10). In the hybrids, methylation is gained on the Ler allele in this region of At3g43340 and is consistent with a decreased transcript level of the Ler allele and of the overall expression level (Figs. 4 and 5). At5g38720 shows a similar up-regulation in Col *met1* and *ddc* mutants with the affected methylated region corresponding to the TCM region in the hybrid and which presumably causes

the decrease in gene expression (Fig. S6C and Dataset S1, Table S11).

At4g15920, At4g09490, and At3g26612 all showed increased expression levels in the hybrid correlating with localized TCdM events and are associated with reductions in levels and distribution of siRNAs (Fig. 5). A loss of methylation in the region upstream of At4g09490 in the Col *met1* mutant coincides with increased expression of this gene (Dataset S1, Table S11), paralleling the TCdM event identified in the hybrid over this same region (Fig. 5).

Discussion

In our analysis of the epigenome in the progeny of crosses between the *Arabidopsis* accessions C24 and Ler, we explored the relationship between two epigenetic systems, DNA methylation and the production of siRNAs. We examined the mC distribution in the genetically similar C24 and Ler accessions and in their hybrid progeny and asked whether alterations in the methylome influenced the hybrid transcriptome, which ultimately must be involved in the characteristic increase in biomass of the F1 hybrid. The parents showed significant differences between their methylomes with one-quarter of mC regions differentially methylated. These regions were enriched for genes and their flanking regions and underrepresented in transposable elements consistent with comparisons between other *Arabidopsis* accessions (13, 14). The hybrids had significant changes in methylation levels at localized regions of the genome, frequently corresponding to those chromosomal segments where methylation levels were markedly different in the C24 and Ler parents. Seven percent to 20% of mC clusters had nonadditive changes either above or below the expected MPV in either Total mC or at least one mC context. We identified two major processes associated with nonadditive methylation, TCM, which led to methylation levels greater than the MPV, and TCdM, which produced methylation levels lower than the MPV. Both of these processes involved a change in methylation of one parental allele to resemble that of the other parental allele with the de novo methylation or demethylation events superimposed on the efficiently and accurately maintained background inherited from the parental genotypes. These processes affected all mC contexts, with TCM-generated increases in methylation more frequently associated with mCG and TCdM-generated decreases in methylation most obvious in mCHH .

Transallelic interactions have also been reported in paralogous loci (28, 29), certain cases of self-incompatibility (30), silencing of Flowering WAGENINGEN A (FWA; refs. 31 and 32), *PAI* locus transmethylation (33), and remethylation and demethylation of loci in epiRILs (34–36). In most cases, siRNAs appear to play a central role. Many alterations to the hybrid methylome through TCM and TCdM are associated with siRNAs. The previously reported reduction in 24-nt siRNA levels in the C24/Ler hybrid system correlates with the decrease in mCHH levels (18), which depends on siRNAs for continued maintenance. The reduction of siRNA levels in the hybrid has a lesser effect on mCG and mCHG , which have siRNA-independent maintenance of methylation via METHYLTRANSFERASE 1 (MET1; ref. 37) and CHROMOMETHYLASE 3 (CMT3; ref. 38). MET1 and CMT3 activity may also explain the more frequent non-additive increases in CG and CHG methylation compared with CHH, because even low levels of de novo RNA-directed DNA methylation in the hybrid would be efficiently maintained by these methyltransferases.

Many siRNA-associated loci show additive inheritance of allelic methylation, suggesting that other factors may determine the ability of a locus to undergo TCM or TCdM; the chromatin state of a region may impose the level of siRNAs needed to establish or maintain DNA methylation. TCM events occur when the threshold is exceeded and TCdM events occur below threshold levels, possibly resulting from a dilution of siRNAs over the two alleles in the hybrid (39). We also found a large subset of mC clusters exhibiting TCM/TCdM in the absence of

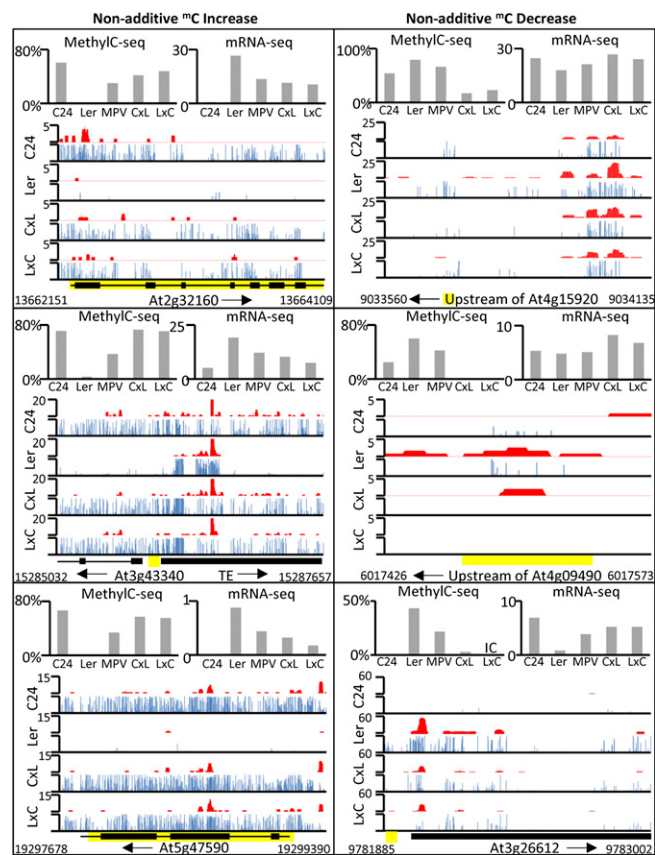


Fig. 5. Altered methylation levels in the hybrids associated with changes in gene expression. mC clusters located in or within ± 1 kb of the listed genes. Gray histograms show average methylation level of the mC cluster (highlighted in yellow) and levels of gene expression obtained from a mRNA-seq dataset. The next set of graphs depict the distribution and abundance of siRNAs (red) and mC (blue) in both parents and each reciprocal hybrid. The six examples are of inverse correlations between mC and gene expression. Gene annotations are located at the bottom of each image with the arrows indicating orientation. Numbers at each extreme of the regions show genomic position. IC, insufficient coverage.

siRNAs, these being primarily associated with gene body methylation. Mechanisms independent of siRNAs also may function to alter the hybrid methylome.

In large part, the chromosomal segments with altered methylation levels and patterns in the hybrid were concentrated in genes and in their flanking regions. We were able to identify a number of genes where altered levels and distribution of siRNAs and resulting DNA methylation corresponded with altered gene transcription. Most of these examples fit the general rule that methylation is suppressive of transcription. Although we described only a small number of examples, these paralleled results in mutants affecting methylation levels. It is not clear from our limited data what proportion of the altered hybrid transcriptome is directly caused by changes to siRNA and methylation levels. Limitations of our study are that we sampled only one developmental time point and that the data were from a mix of tissues.

Other reports indicate that alterations to epigenetic systems occur in hybrids (7, 17, 18). If these systems contribute to heterosis, then a loss of maximum vigor in generations beyond the F1 would be due not only to the segregation of alleles and epialleles but also to possible alterations to the epialleles in the F2 and beyond. For example, a proportion of loci we examined had low levels of TCM, which may be either lost or increased in subsequent generations reminiscent of transgenerational changes in epiRILs (17, 34–36).

Our analysis indicates that changes in gene expression in a hybrid are influenced not only by differences in DNA sequence between parents, but also by variation in gene-associated siRNA and DNA methylation marks. It is likely that TCM and TCdM occur in all hybrid systems. These epigenetically controlled changes may be among the initial events in the cascades of altered gene expression that contribute to hybrid heterosis. It will be necessary in different hybrid combinations to analyze specific tissues at a number of developmental stages to gain a better understanding of how the epigenome contributes to hybrid vigor.

Materials and Methods

Experimental designs and plant material used are described in *SI Materials and Methods*. MethylC-seq libraries were prepared from immature floral buds with sequencing performed by Geneworks on the Genome Analyzer IIx (Illumina). mRNA-seq was carried out by the Biomolecular Resource Facility at The Australian National University on the Genome Analyzer IIx (Illumina). Processing of sRNA, methylation, transcriptome sequences, and methylated regions are described in *SI Materials and Methods*. Raw and processed mapped MethylC-seq sequences are deposited in GEO with accession number GSE35542.

ACKNOWLEDGMENTS. We thank Aihua Wang and Limin Wu for technical assistance in the laboratory. This work was supported by National Collaborative Research Infrastructure Strategy and Science and Industry Endowment Fund funded by the Australian Government and by the Agricultural Genomics Centre, New South Wales Government BioFirst Initiative.

- Lippman ZB, Zamir D (2007) Heterosis: Revisiting the magic. *Trends Genet* 23:60–66.
- Hochholdinger F, Hoecker N (2007) Towards the molecular basis of heterosis. *Trends Plant Sci* 12:427–432.
- Birchler JA, Yao H, Chudalayandi S, Vaiman D, Veitia RA (2010) Heterosis. *Plant Cell* 22:2105–2112.
- Melchinger AE (1999) Genetic diversity and heterosis. *Genetics and Exploitation of Heterosis in Crops* (CSSA, Madison, WI), pp 99–118.
- Chen ZJ (2010) Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci* 15:57–71.
- Meyer RC, Törjék O, Becher M, Altmann T (2004) Heterosis of biomass production in Arabidopsis. Establishment during early development. *Plant Physiol* 134:1813–1823.
- He GM, et al. (2010) Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell* 22:17–33.
- Ossowski S, et al. (2010) The rate and molecular spectrum of spontaneous mutations in Arabidopsis thaliana. *Science* 327:92–94.
- Becker C, et al. (2011) Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. *Nature* 480:245–249.
- Schmitz RJ, et al. (2011) Transgenerational epigenetic instability is a source of novel methylation variants. *Science* 334:369–373.
- Herrera CM, Bazaga P (2010) Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. *New Phytol* 187:867–876.
- Li YD, et al. (2008) Utility of the methylation-sensitive amplified polymorphism (MSAP) marker for detection of DNA methylation polymorphism and epigenetic population structure in a wild barley species (*Hordeum brevisubulatum*). *Ecol Res* 23: 927–930.
- Vaughn MW, et al. (2007) Epigenetic natural variation in Arabidopsis thaliana. *PLoS Biol* 5:e174.
- Zhang X, Shiu SH, Cal A, Borevitz JO (2008) Global analysis of genetic, epigenetic and transcriptional polymorphisms in Arabidopsis thaliana using whole genome tiling arrays. *PLoS Genet* 4:e1000032.
- Zhai JX, et al. (2008) Small RNA-directed epigenetic natural variation in Arabidopsis thaliana. *PLoS Genet* 4:e1000056.
- Jin H, et al. (2008) Alterations in cytosine methylation and species-specific transcription induced by interspecific hybridization between *Oryza sativa* and *O. officinalis*. *Theor Appl Genet* 117:1271–1279.
- Ha M, et al. (2009) Small RNAs serve as a genetic buffer against genomic shock in Arabidopsis interspecific hybrids and allopolyploids. *Proc Natl Acad Sci USA* 106: 17835–17840.
- Groszmann M, et al. (2011) Changes in 24-nt siRNA levels in Arabidopsis hybrids suggest an epigenetic contribution to hybrid vigor. *Proc Natl Acad Sci USA* 108: 2617–2622.
- Makarevitch I, et al. (2007) Natural variation for alleles under epigenetic control by the maize chromomethylase zmet2. *Genetics* 177:749–760.
- Lira-Medeiros CF, et al. (2010) Epigenetic variation in mangrove plants occurring in contrasting natural environment. *PLoS ONE* 5:e10326.
- Cokus SJ, et al. (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452:215–219.
- Matzke M, Kanno T, Huettel B, Daxinger L, Matzke AJM (2007) Targets of RNA-directed DNA methylation. *Curr Opin Plant Biol* 10:512–519.
- Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* 11:204–220.
- Zhang XY, et al. (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. *Cell* 126:1189–1201.
- Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S (2007) Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* 39:61–69.
- Kurihara Y, et al. (2008) Identification of the candidate genes regulated by RNA-directed DNA methylation in Arabidopsis. *Biochem Biophys Res Commun* 376:553–557.
- Lister R, et al. (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* 133:523–536.
- Arteaga-Vazquez M, et al. (2010) RNA-mediated trans-communication can establish paramutation at the b1 locus in maize. *Proc Natl Acad Sci USA* 107:12986–12991.
- Erhard KF, Jr., Hollick JB (2011) Paramutation: A process for acquiring trans-generational regulatory states. *Curr Opin Plant Biol* 14:210–216.
- Tarutani Y, et al. (2010) Trans-acting small RNA determines dominance relationships in Brassica self-incompatibility. *Nature* 466:983–986.
- Cao XF, Jacobsen SE (2002) Role of the Arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr Biol* 12:1138–1144.
- Chan SWL, Zhang X, Bernatavichute YV, Jacobsen SE (2006) Two-step recruitment of RNA-directed DNA methylation to tandem repeats. *PLoS Biol* 4:e363.
- Luff B, Pawlowski L, Bender J (1999) An inverted repeat triggers cytosine methylation of identical sequences in Arabidopsis. *Mol Cell* 3:505–511.
- Johannes F, et al. (2009) Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genet* 5:e1000530.
- Reinders J, et al. (2009) Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. *Genes Dev* 23:939–950.
- Teixeira FK, et al. (2009) A role for RNAi in the selective correction of DNA methylation defects. *Science* 323:1600–1604.
- Finnegan EJ, Dennis ES (1993) Isolation and identification by sequence homology of a putative cytosine methyltransferase from Arabidopsis thaliana. *Nucleic Acids Res* 21: 2383–2388.
- Lindroth AM, et al. (2001) Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science* 292:2077–2080.
- Groszmann M, et al. (2011) Epigenetics in plants-vernalisation and hybrid vigour. *Biochim Biophys Acta* 1809:427–437.