

# Inheritance of Trans Chromosomal Methylation patterns from *Arabidopsis* F1 hybrids

Ian K. Greaves<sup>a</sup>, Michael Groszmann<sup>a</sup>, Aihua Wang<sup>a</sup>, W. James Peacock<sup>a,b,1</sup>, and Elizabeth S. Dennis<sup>a,b</sup>

<sup>a</sup>Plant Industry Commonwealth Scientific and Industrial Research Organization, Canberra, ACT 2600, Australia; and <sup>b</sup>University of Technology, Sydney, NSW 2007, Australia

Contributed by W. James Peacock, December 19, 2013 (sent for review September 11, 2013)

Hybridization in plants leads to transinteractions between the parental genomes and epigenomes that can result in changes to both 24 nt siRNA and cytosine methylation (<sup>m</sup>C) levels in the hybrid. In *Arabidopsis* the principle processes altering the hybrid methylome are Trans Chromosomal Methylation (TCM) and Trans Chromosomal deMethylation (TCdM) in which the <sup>m</sup>C pattern of a genomic segment attains the same <sup>m</sup>C pattern of the corresponding segment on the other parental chromosome. We examined two loci that undergo TCM/TCdM in the *Arabidopsis* C24/Landsberg *erecta* (Ler) F1 hybrids, which show patterns of inheritance dependent on the properties of the particular donor and recipient chromosomal segments. At At1g64790 the TCM- and TCdM-derived <sup>m</sup>C patterns are maintained in the F2 generation but are transmitted in outcrosses or backcrosses only by the C24 genomic segment. At a region between and adjacent to At3g43340 and At3g43350, the originally unmethylated Ler genomic segment receives the C24 <sup>m</sup>C pattern in the F1, which is then maintained in backcross plants independent of the presence of the parental C24 segment. In backcrosses to an unmethylated Ler allele, the newly methylated F1 Ler segment may act as a TCM source in a process comparable to paramutation in maize. TCM-derived <sup>m</sup>C patterns are associated with reduced expression of both At3g43340 and At3g43350 in F1 and F2 plants, providing support for such events influencing the transcriptome. The inheritance of the F1 <sup>m</sup>C patterns and the segregation of other genetic and epigenetic determinants may contribute to the reduced hybrid vigor in the F2 and subsequent generations.

heterosis | chromatin | histones | RdDM

Hybrid plants are used extensively in agriculture due to their superior productivity and biomass (1). The gains are assumed to be a consequence of changes to the transcriptome in the hybrid that alter the phenotype of the plant. In 1948, Shull proposed that heterosis had non-Mendelian features (2), Hollick and Chandler (3) have suggested an epigenetic basis to a case of single locus heterosis, and recent studies on the molecular biology of hybrids have suggested that epigenetic systems contribute to the mechanisms generating heterosis. Hybrids have altered small RNA (sRNA) frequencies and cytosine methylation (<sup>m</sup>C) patterns that can affect mRNA abundance (4, 5). In the nuclei of the hybrid plants, the two parental genomes and epigenomes interact, altering the patterns of <sup>m</sup>C across both genomes (6–8). These changes are localized to particular chromosomal segments and occur where the parents have differentially methylated regions (DMRs). Of the nonadditively methylated clusters in the hybrid, 86% are due to Trans Chromosomal Methylation (TCM) and Trans Chromosomal deMethylation (TCdM) (6). TCM involves a less methylated parental segment attaining the same <sup>m</sup>C pattern as the homologous highly methylated parental segment, whereas in TCdM the more highly methylated parental genomic segment shows a reduction in <sup>m</sup>C levels so as to resemble the pattern in the less methylated parental segment.

An <sup>m</sup>C is an epigenetic modification of DNA that has a role in silencing transposable elements and repetitive sequences and in regulating gene expression (reviewed in refs. 9, 10). Despite its

repressive action, <sup>m</sup>C in genic regions does not always result in the silencing of gene activity. In mutants defective in <sup>m</sup>C, only a low number of genes with changed <sup>m</sup>C levels have altered transcriptional activity, implying a lack of a tight correlation of <sup>m</sup>C and gene expression (11–13). This means that although TCM and TCdM are the major epigenetic processes driving changes in the hybrid methylome, their roles in altering transcriptional activity, which may be important for generating the heterotic phenotype, remain to be determined.

The 24 nt short interfering RNAs (siRNAs) operating through the RNA-directed DNA methylation system are involved in de novo <sup>m</sup>C (reviewed in ref. 14). The presence of 24 nt siRNAs in regions undergoing TCM/TCdM suggests that they mediate transchromosomal interactions (6, 15). Not all loci that have different levels of <sup>m</sup>C in segments in the two parents and have associated 24 nt siRNAs undergo TCM/TCdM (6), indicating that other factors such as histone modifications that influence the state of chromatin may also be important for TCM/TCdM processes. Some histone modifications such as histone 3 lysine 9 acetylation/histone 4 acetylation (H3K9ac/H4ac) and di/trimethylation of histone 3 lysine 4 (H3K4me2/3) promote gene activity, whereas other modifications such as histone 3 lysine 9 dimethylation (H3K9me2) and histone 3 lysine 27 trimethylation (H3K27me3) repress gene expression (reviewed in refs. 16, 17). Histone modifications can change during plant development and influence the level of <sup>m</sup>C (18, 19).

Hybrid vigor is restricted to the F1 generation, with the F2 and subsequent generations showing reduced vigor and increased heterogeneity of heterotic traits such as yield and biomass. The traditional view has been that loss of vigor occurs through segregation of the genetic determinants of the vigor phenotype, but it may also be linked to segregation or further alterations to the

## Significance

Hybridization of two epigenetically distinct *Arabidopsis* accessions is associated with an altered methylome in the F1 hybrid brought about by the processes of Trans Chromosomal Methylation and deMethylation. In the loci studied, the altered F1 cytosine methylation (<sup>m</sup>C) patterns are inherited to the F2. These observations of trans-<sup>m</sup>C events are reminiscent of paramutation described in maize and may be common in hybridization events. In one locus, the altered <sup>m</sup>C states are associated with changed mRNA levels in the F1 hybrid and F2 plants. These events may contribute to F1 hybrid vigor and to the phenotypic heterogeneity and loss of vigor in the F2 through segregation of genetic and epigenetic determinants.

Author contributions: I.K.G., M.G., W.J.P., and E.S.D. designed research; I.K.G., M.G., and A.W. performed research; I.K.G., M.G., and A.W. 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.

<sup>1</sup>To 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.1323656111/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1323656111/-DCSupplemental).

hybrid epigenome. In relation to this possibility, we followed the inheritance of  $m^C$  patterns produced by TCM/TCdM events at three genes in the C24/Landsberg *erecta* (Ler) F1 hybrid. We found that in the hybrids, these TCM-derived  $m^C$  patterns can be accompanied by changes to histone acetylation levels and altered gene expression.

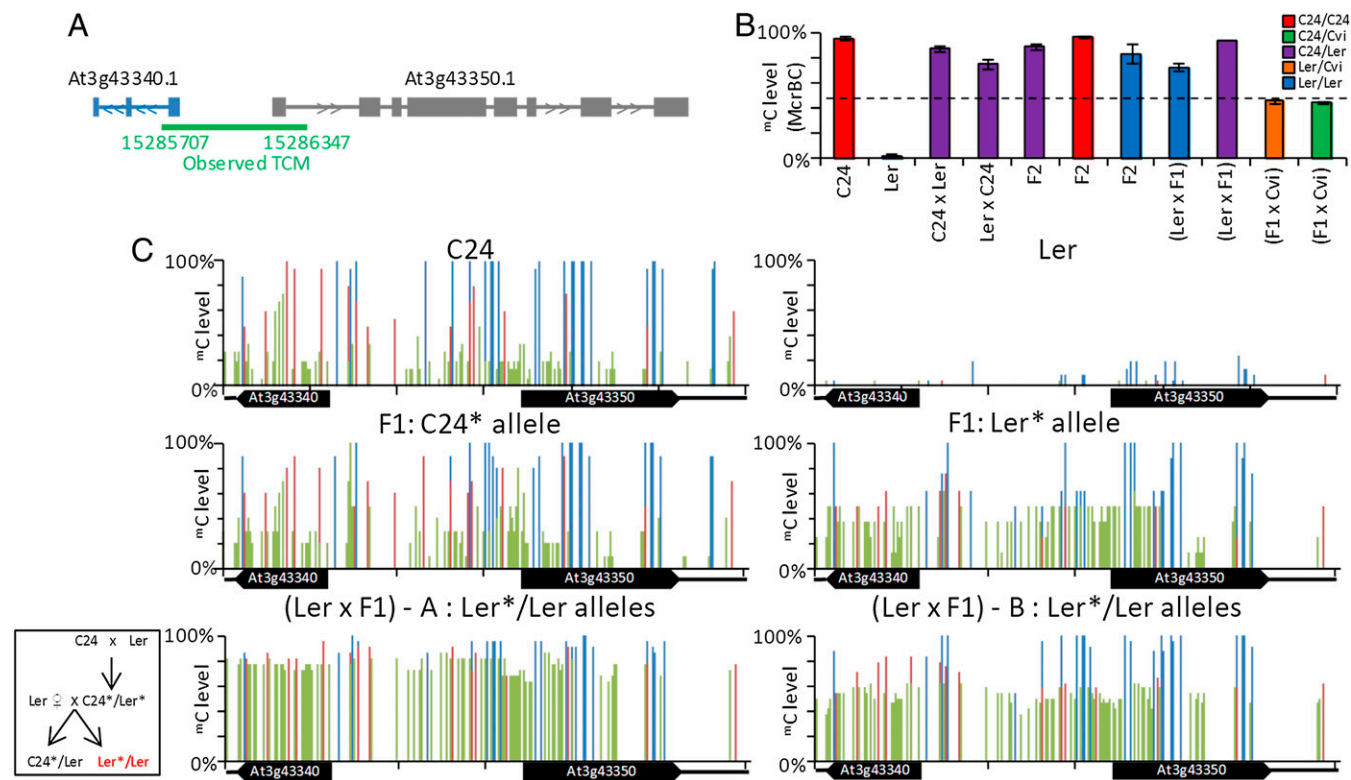
## Results

In hybrids between *Arabidopsis* C24 and Ler accessions, DMRs in the first intron of At1g64790 and a genomic segment spanning At3g43340 and At3g43350 (referred to as At3g43340/50) act as recipient segments in Ler, acquiring the  $m^C$  patterns present on C24 segments. The Ler segments gain methylation at CG, CHG, and CHH sites not methylated in the Ler parent (TCM), changing  $m^C$  levels at these three genes in the hybrid. A TCdM event also occurs at an adjacent segment at At1g64790. We tracked At3g43340/50 and At1g64790 in the F2 generation and in progeny of the F1 hybrids (C24\*/Ler\*; asterisk denotes the segment derived from the F1 hybrid) backcrossed to the original unmethylated Ler parent or outcrossed to another accession unmethylated at this region of interest (*SI Appendix, Figs. S1 and S2*). This enabled us to determine whether the newly acquired  $m^C$  pattern on the F1 Ler\* segment is maintained in the following generation and is capable of promoting a new TCM event on the unmethylated chromosome in that generation.

The TCM-derived methylation pattern at the At3g43340/50 segment spans from Chr3: 15,285,707 to Chr3: 15,286,347, covering the central upstream sequence and extending into the first

intron of each of the two genes, which are transcribed in opposite directions (Fig. 1A). Within this region, C24 and Ler are differentiated by seven SNPs and one small insertion/deletion, enabling us to distinguish between the two parental segments in the hybrid and look for changes in  $m^C$  patterns. The DMR between parents and the  $m^C$  pattern of the hybrid TCM event may extend further into the genes, but because of the multicopy characteristics of both genes, we are not able to define the termini of the TCM event. The same DMR occurs between Columbia (Col) and Ler, and hybrids between these two accessions also demonstrate TCM at this segment (*SI Appendix, Fig. S3*).

In F2 plants derived from the C24/Ler F1 plants, the TCM-derived  $m^C$  pattern at At3g43340/50 is present in all progeny genotypes (C24\*/C24\*, C24\*/Ler\*, and Ler\*/Ler\*), indicating transgenerational transmission of the TCM-derived  $m^C$  pattern (Fig. 1B). In backcrosses of the F1 hybrid (C24\*/Ler\*) to the Ler parent, both heterozygous (C24\*/Ler) and homozygous (Ler\*/Ler) Ler progeny have an  $m^C$  pattern similar to the F1 (Fig. 1B and C and *SI Appendix, Fig. S4A*). Although in the Ler homozygotes the two Ler segments (Ler\* and Ler) cannot be differentiated by SNPs, the level of methylated cytosines approaches 100% for the CG and CHG contexts, well above the midparent value (MPV) and consistent with both Ler segments having the C24  $m^C$  pattern (Fig. 1B and C and *SI Appendix, Fig. S4A*). Digestion with the restriction enzyme McrBC followed by quantitative RT-PCR (qRT-PCR) analysis of the homozygous Ler individuals also showed high levels of  $m^C$ , consistent with both segments being methylated (*SI Appendix, Fig. S4B*). The



**Fig. 1.** Inheritance of  $m^C$  patterns at At3g43340/50. (A) Region in which TCM was observed, overlapping At3g43340 and At3g43350. Exons are indicated as grey and blue boxes (B) McrBC qRT-PCR of preanthesis floral buds from parents, F1 hybrids, F2s, Ler  $\times$  F1 backcrosses, and F1  $\times$  Cvi outcrosses. Parental and F1 columns are derived from three biological replicates, with each replicate containing  $\geq 4$  plants. Each F2 column represents the combined result from  $\geq 4$  individual plants. The Ler  $\times$  F1 (Ler/Ler combination) column is derived from six individuals, and the Ler  $\times$  F1 (C24/Ler combination) and Cvi  $\times$  F1 columns are derived from two individuals each. Column colors denote genotypes of plants. Black dotted line represents MPV. Error bars, SEM. (C) Bisulphite PCR of parents, F1 hybrids, and Ler  $\times$  F1 backcrosses. For parents and F1 hybrids, bisulphite results are derived from a pool of  $\geq 4$  plants. For backcrosses (Ler\*/Ler), six individuals were analyzed (*SI Appendix, Fig. S4*). Asterisk denotes a segment derived from F1, with individual plants labeled A and B. The box demonstrates how the individuals were derived. C24 and F1 graphs were modified from ref. 6. The  $m^C$  levels are represented by blue ( $m^C$ CG), red ( $m^C$ CHG), and green ( $m^C$ CHH) columns.

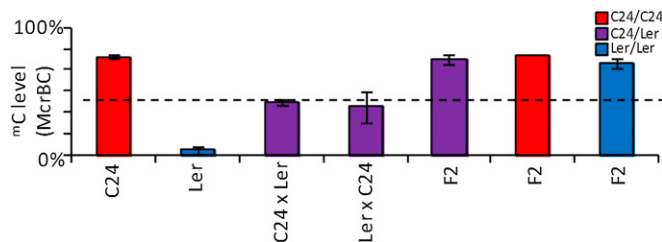
ability to maintain and initiate a new TCM event in the F1 × Ler backcrosses was observed in all Ler homozygous individuals studied (six plants; Fig. 1 *B* and *C* and *SI Appendix*, Fig. S4 *A* and *B*). Although we cannot rule out other cis or transacting factors mediating the TCM event, the fact that all six Ler\*/Ler individuals analyzed showed similar <sup>14</sup>C patterns at the locus suggests that the trans signal is linked to the TCM region. These data show that the C24 <sup>14</sup>C pattern transferred to the Ler\* segment by the original TCM event in the F1 is maintained in the Ler\* chromosome in the next generation, where it results in a new TCM event transferring the C24 <sup>14</sup>C pattern to the unmethylated Ler chromosome.

In a cross of the F1 hybrid to a Cape Verde Islands (Cvi) accession that, like Ler, lacks <sup>14</sup>C at the At3g43340/50 region (*SI Appendix*, Fig. S2), the Ler\* segment maintained its F1 <sup>14</sup>C pattern, whereas the Cvi segment remained unmethylated (*SI Appendix*, Fig. S5). In C24\*/Cvi progeny, there was a similar outcome, with the Cvi segment remaining unmethylated and C24\* maintaining its <sup>14</sup>C pattern (*SI Appendix*, Fig. S5). The Cvi region may have a property that prevents it from becoming methylated.

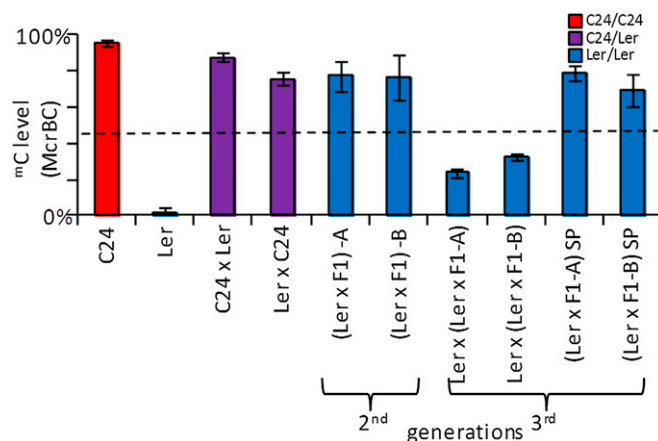
The TCM event at At3g43340/50 in F1 hybrids was present in preanthesis floral buds but not in the vegetative tissue of 15 days after sowing (DAS) seedlings (Figs. 1*B* and 2 and *SI Appendix*, Fig. S4*C*). In F2 progeny, the <sup>14</sup>C state of the Ler\* segment was present in both preanthesis floral buds and 15 DAS vegetative tissue (Figs. 1*B* and 2 and *SI Appendix*, Fig. S4*C*). The different results between the F1 and F2 generations imply the <sup>14</sup>C on the Ler\* segment in the F1 is not developmentally or tissue specific but may be the result of an <sup>14</sup>C-inducing signal that accumulates over time. Once the Ler\* segment has become methylated, it retains the <sup>14</sup>C pattern in the next generation.

To determine if the <sup>14</sup>C pattern on the Ler segment can be maintained into a third generation, we analyzed At3g43340/50 in progeny from selfed Ler\*/Ler second-generation backcrossed plants where cytosines within both segments were methylated (*SI Appendix*, Fig. S6*A*). We also examined third-generation plants in which the same methylated Ler\*/Ler backcross plants had been used in a second backcross to the Ler accession carrying the unmethylated Ler segment (*SI Appendix*, Fig. S6*A*). The third-generation plants, derived from selfing second-generation backcross plants, maintained the <sup>14</sup>C pattern and level. The third-generation plants derived from the second backcross to a Ler plant containing an unmethylated segment showed a reduction in <sup>14</sup>C levels and variation between individual progeny, with some showing low <sup>14</sup>C levels (Fig. 3 and *SI Appendix*, Fig. S6).

**TCM at At3g43340/50 Is Associated with Decreased mRNA Levels in the F1 and F2 Generations.** We examined the mRNA levels of At3g43340 and At3g43350 in the F1 and F2 generations. In both 15 DAS seedlings and in preanthesis floral buds, the unmethylated Ler parent had higher mRNA levels for each gene compared with the methylated C24 parent, suggesting that the <sup>14</sup>C at



**Fig. 2.** TCM-derived <sup>14</sup>C patterns at At3g43340/50 are present at different times in F1 and F2 15 DAS seedlings. Each column represents at least four individuals of each allelic combination. Column colors denote allelic combinations. Black dotted line represents MPV. Error bars, SEM.

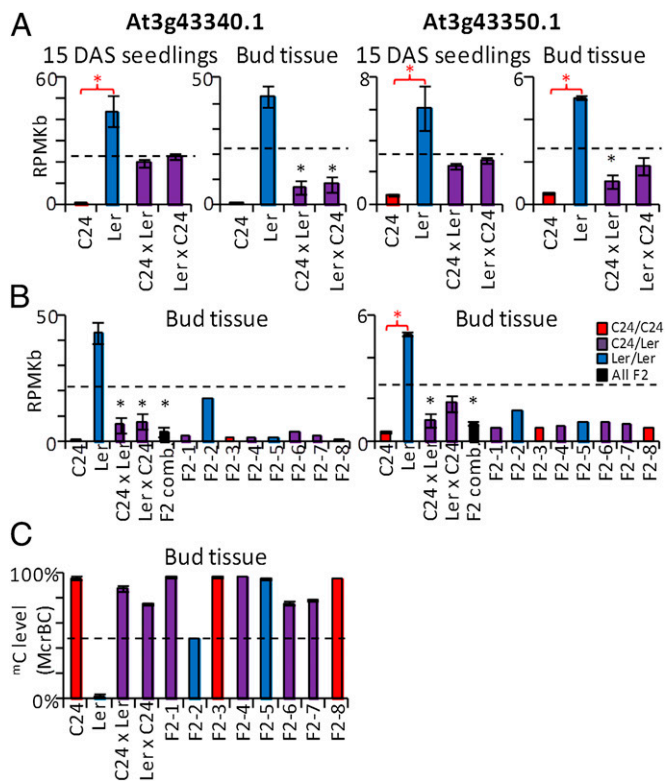


**Fig. 3.** The <sup>14</sup>C levels at At3g43340/50 in third-generation plants derived from two backcrosses or one backcross followed by a selfing (*SI Appendix*, Fig. S6*A*). McrBC qRT-PCR was used to analyze <sup>14</sup>C levels of preanthesis floral bud material from third-generation plants. Individuals (Ler × F1) -A and -B are presented on the graph, as they are the plants from which the third-generation plants are derived and represent the average level of <sup>14</sup>C of each individual from two independent McrBC digests. Results of third-generation plants are derived from a pool of individuals ( $n \geq 5$ ). SP, progeny of selfed plants. Column colors denote allelic constitution of plants. Black dotted line represents MPV. Error bars, SEM.

this region is associated with the decreased expression of both genes (Fig. 4 *A* and *C*). In F1 15 DAS seedlings, in which TCM has not yet occurred, the mRNA levels were at MPV, consistent with only the unmethylated Ler alleles being expressed. In preanthesis floral buds, where TCM has occurred, the hybrids show an expression level below MPV for both genes (Fig. 4*A*). In the F2 generation, we analyzed the mRNA and <sup>14</sup>C levels of both genes. In eight F2 individuals, which represented all possible genotype combinations (Fig. 4 *B* and *C*), seven had low mRNA levels and high levels of <sup>14</sup>C. One plant (F2-2) had higher mRNA levels coinciding with lower levels of <sup>14</sup>C (Fig. 4 *B* and *C*). These results suggest that these F1 TCM-derived <sup>14</sup>C patterns are associated with changes to mRNA levels in both the F1 and F2 plants.

**TCM/TCdM-Derived <sup>14</sup>C Patterns at At1g64790 Are Maintained in Selfed Progeny but Not in Backcrosses or Outcrosses.** In C24/Ler hybrids, TCM and TCdM regions are positioned close together in the first intron of At1g64790 (Fig. 5). The C24 and Ler segments can be differentiated by nine SNPs, three of which are in the TCM/TCdM region. Both the TCM and TCdM events occur on the Ler\* segment. The TCdM event involves the loss of CHH methylation, and the TCM event results in the gain of methylation of the only CG site and of several CHH sites, together with increases in CHG methylation levels. An SNP in the 24 nt siRNA reads from the TCM region showed that in the F1, in addition to those reads produced by the C24\* segment, the Ler\* segment produces 24 nt siRNAs (6). The <sup>14</sup>C at the TCM region appears to be dependent on the presence of 24 nt siRNAs, as C24 RNA polymerase IV mutants, which have reduced 24 nt siRNAs, lack CHH methylation and have decreased CHG methylation (20–21) (*SI Appendix*, Fig. S7).

The <sup>14</sup>C patterns observed in F2 plants indicated that TCM/TCdM-derived patterns found on the F1 Ler\* segment are inherited (Fig. 5), similar to the inheritance pattern seen at the At3g43340/50 segment. In an outcross of the F1 C24 × Ler hybrid to Col, which has an <sup>14</sup>C pattern at At1g64790 resembling the Ler parent (*SI Appendix*, Fig. S2), the C24\*/Col progeny have the same TCM- and TCdM-derived <sup>14</sup>C patterns as the C24 × Ler F1 hybrid (Fig. 5 and *SI Appendix*, Fig. S8). However, in the Ler\*/Col progeny, the Ler\* segment no longer has the TCM/



**Fig. 4.** TCM events at At3g43340/50 are associated with altered mRNA levels in F1 and F2 plants. (A) mRNA levels of parental and F1 hybrids for At3g43340 and At3g43350 in 15 DAS seedlings and preanthesis floral buds (Dataset S1, Table S1). (B) mRNA levels in buds of eight individual F2 plants (Dataset S1, Table S1). (C) MCRBC qRT-PCR of the same F2 plants used in the expression analysis. Column colors denote genotype. Black dotted line represents MPV. Error bars, SEM. Red and black asterisks indicate statistically significant differences ( $P \leq 0.05$ ) in expression levels between parents and between hybrids and MPV, respectively.

TCdM-derived  $m^C$  patterns, and has the original Ler parental  $m^C$  pattern except for the single CG methylation site gained from the F1 TCM event. In the Ler\*/Ler progeny of a Ler  $\times$  F1 backcross, the Ler\* genomic segment had the Ler parental  $m^C$  pattern and not the TCM/TCdM F1  $m^C$  pattern (Fig. 5). This result is consistent with the observations we made in individuals derived from a Col  $\times$  F1 outcross. This failure of inheritance in the outcross and backcross is similar to the loss of the TCM-derived  $m^C$  pattern at the At3g43340/50 segment in the second backcross to a Ler plant unmethylated at the At3g43340/50 segment. mRNA levels of At1g64790 are similar in C24 and Ler and do not change in response to the TCM and TCdM events (Dataset S1, Table S1).

#### TCM Events Are Associated with Changes to Histone H3K9 Acetylation.

Because histone modifications influence the presence of  $m^C$  (22), we analyzed key histone marks at At3g43340/50 and At1g64790. At At3g43340/50, H3K9ac showed an inverse correlation with  $m^C$  in the parental accessions, with Ler showing high H3K9ac levels and low  $m^C$  and C24 having low H3K9ac levels and high  $m^C$  (Fig. 6). In the hybrids, there was a significant decrease in H3K9ac to a level indicating that the Ler\* segment has lost its H3K9ac modifications. This demonstrates that increased  $m^C$  levels, due to TCM, can be accompanied by changes in histone epigenetic marks and changes in mRNA levels (Fig. 4). Concordant changes in  $m^C$  levels and histone marks may not always occur in TCM/TCdM regions. At At1g64790 we found a small but insignificant increase in the H3K9ac levels at the TCdM

region and only small changes for H3K9ac and H3K9me2 at the TCM region (SI Appendix, Fig. S9).

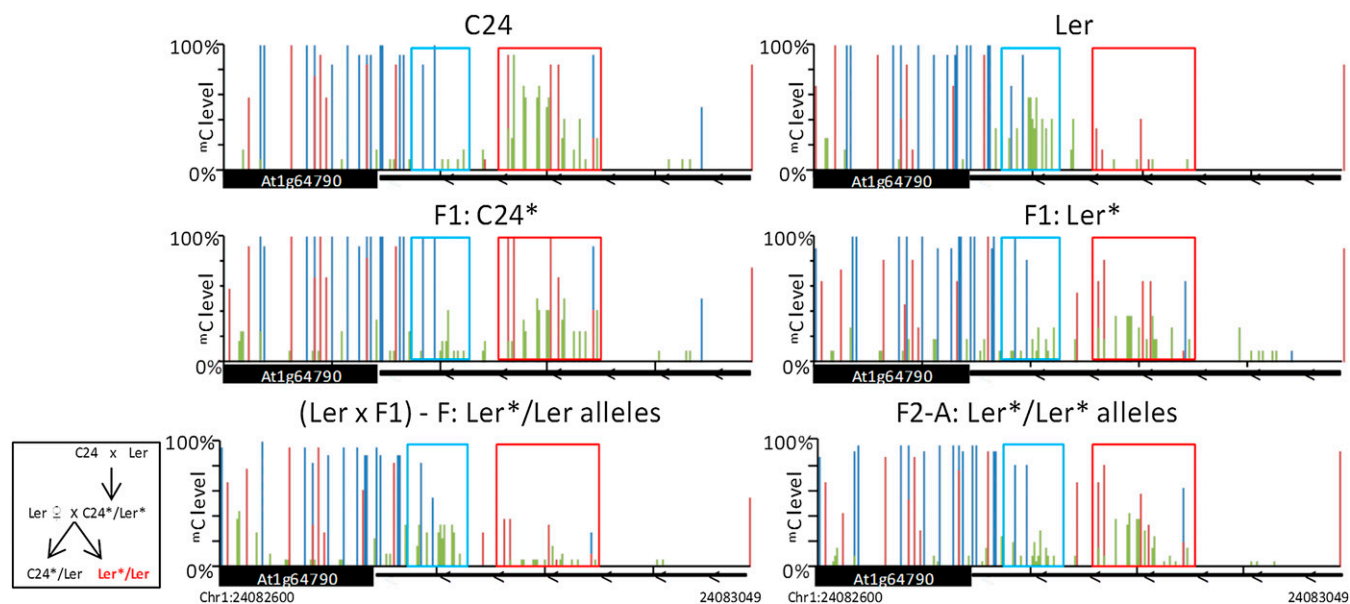
#### Discussion

The *Arabidopsis* C24  $\times$  Ler reciprocal hybrids exhibit a strong vigor phenotype even though the parents have similar genomes (23). Given the need for diversity between parents to generate heterosis, we have suggested that the differences between the two parental epigenomes contribute to the hybrid vigor phenotype (reviewed in refs. 4, 5).

In the F1 hybrids, 24 nt siRNAs show a marked reduction in frequency and coincide with a parallel reduction in  $m^C$  of homologous genomic regions (15) (reviewed in ref. 24). The reduced frequencies of 24 nt siRNAs occur at genomic segments that have different levels of 24 nt siRNAs in the two parents. At regions of the F1 genome where cytosines are differentially methylated between parents and where siRNAs are not reduced, increases in  $m^C$  levels occur (6, 7, 15). These processes result in the altered F1  $m^C$  patterns and are termed TCM and TCdM events. In TCM and TCdM events, the  $m^C$  pattern of one parental segment attains the same  $m^C$  pattern as the homologous chromosomal segment of the other parent. TCM and TCdM occur most frequently in genes and flanking regulatory regions and are sometimes associated with changes in mRNA levels that may contribute to the gene activity pattern determining heterosis (6, 7, 15). The altered  $m^C$  pattern in the genomic segment encompassing the upstream and 5' regions of At3g43340 and At3g43350 is associated with reduced mRNA levels of both genes in F1 and F2 plants. Alterations to  $m^C$  and 24 nt siRNA levels have also been reported in interspecific *Arabidopsis* hybrids and in hybrids of rice and maize (6–8, 15, 25–28).

A feature of F1 hybrid vigor is that the heterotic increase in yield, relative to the parents, is reduced in subsequent generations—a result assumed to be due to genetic segregation. The increased heterogeneity for biomass and seed yield traits in the F2 plants may also in part be due to changes in siRNA levels,  $m^C$  patterns, and histone modifications. We found that the new F1  $m^C$  patterns at At3g43340/50 and At1g64790 are inherited into the F2 regardless of the genotype combination (i.e., C24\*/Ler\* or Ler\*/Ler\*). At At1g64790, in backcrosses or outcrosses to plants carrying an unmethylated homologous segment, only progeny carrying the C24\*  $m^C$  segment have a TCM event. The F1  $m^C$  pattern of the Ler At1g64790 segment is only maintained in a subsequent generation if inherited together with a methylated homologous segment. The genomic segment loses the F1  $m^C$  pattern in the next generation if exposed to an unmethylated homologous segment. This shows that although the F1 C24\* and Ler\* segments have near-identical  $m^C$  patterns, there must be underlying differences between the two segments that enable only the C24\* At1g64790 segment to maintain the  $m^C$  pattern and also carry out TCM and TCdM events. No changes were observed in mRNA levels at At1g64790 in F1 hybrids, which may be due to the positioning of the TCM events in the intron of the gene, which unlike promoter  $m^C$  may have no impact on gene expression. Also, the TCM/TCdM event is restricted to a 200 bp region that is directly adjacent to many methylated CG and CHG sites present in both parental accessions. Changing the expression of At1g64790 may require the alteration of these cytosine methylated sites rather than just those of the TCM/TCdM region.

At the At3g43340/50 region, C24\* and Ler\* segments of the F1 hybrid can act as a TCM source in a backcross to unmethylated Ler segments for one generation. This shows that unlike At1g64790, the F1 At3g43340/50 Ler\* segment can maintain its F1  $m^C$  pattern when exposed to an unmethylated allele and is itself able to promote a TCM event. The outcross progeny of Ler\*/Cvi also show that  $m^C$  can be maintained by the Ler segment but does not necessarily result in a TCM event.



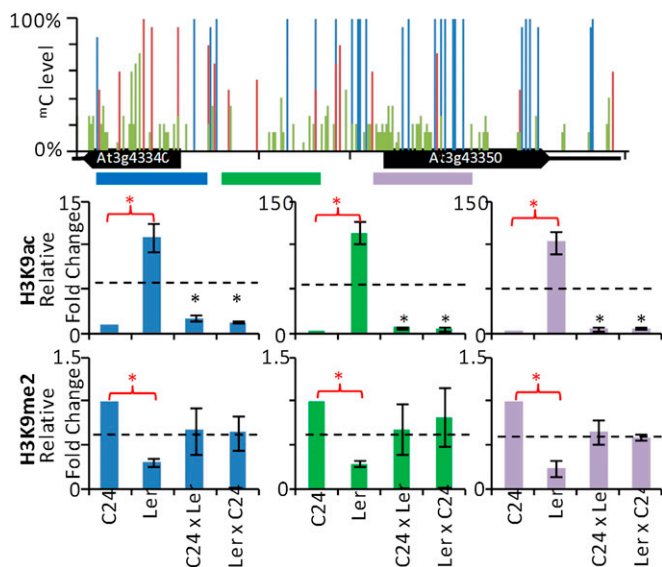
**Fig. 5.** TCM-derived <sup>14</sup>C patterns at At1g64790 in F1 and second-generation plants. Bisulphite PCR results of <sup>14</sup>C levels in parents, F1 hybrids, Ler × F1 backcross, and an F2 individual. For parents and F1 hybrids, bisulphite PCR results are derived from a pool of ≥4 plants. SNPs were used to differentiate alleles. Asterisk denotes a segment derived from F1, with individual plants labeled F and A. The box demonstrates how the individuals were derived. Regions that undergo TCdM (blue) and TCM (red) are highlighted in boxes. C24, Ler, and F1 graphs are modified from ref. 6. The <sup>14</sup>C levels are represented by blue (<sup>14</sup>CG), red (<sup>14</sup>CHG), and green (<sup>14</sup>CHH) columns. Error bars, SEM.

The difference in inheritance of <sup>14</sup>C between the At3g43340/50 and At1g64790 regions may be due to the strength of the TCM signal. This signal may involve 24 nt siRNAs, as they may be produced at only one chromosome segment but are capable of associating with both parental chromosome segments to establish <sup>14</sup>C patterns. Consistent with this, we have noted previously that TCM and TCdM events are associated with 24 nt siRNAs (6, 15), with levels of siRNA likely to be important for initiating these events (reviewed in refs. 4 and 5). The fact that Ler\* At3g43340/50 can maintain and also promote TCM in a backcross whereas Ler\* At1g64790 cannot suggests that the TCM signal is stronger in the Ler\* At3g43340/50 case. The TCM-derived <sup>14</sup>C region at At3g43340/50 is longer than that at At1g64790 (~6×), and contains many more methylated cytosine residues that may facilitate, through an increased production of 24 nt siRNAs, a stronger <sup>14</sup>C condition, enabling the Ler\* segment to maintain the methylation state for a generation longer than at At1g64790. Another possible factor is the juxtaposition of a number of transposable elements within At3g43350, with their chromatin state possibly being more favorable for perpetuating the TCM pattern.

Our observations suggest that TCM events may not occur in the hybrid nucleus immediately after fertilization. The TCM event at At3g43340/50 was found late in F1 development (in floral buds), coinciding with decreased mRNA levels. In F2 plants, the TCM-derived <sup>14</sup>C pattern was present in 15 DAS seedlings as well as in the floral buds. The timing of the TCM event may be dependent on the strength of the signal initiating TCM/TCdM and may depend on the concentration of 24 nt siRNAs in the segment. Such factors could vary between individual loci, as signatures of TCM events are recognizable earlier in development than observed for At3g43340/50 (7).

We have suggested that siRNAs from the donor segment are shared with the recipient segment (reviewed in ref. 4) and that this could dilute siRNA frequencies on a per-segment basis to levels insufficient to initiate <sup>14</sup>C on the recipient segment or to maintain methylation on the donor segment. Consistent with a dilution effect, the newly methylated cytosine residues of the Ler\* segment, at both loci, can lose its TCM-derived <sup>14</sup>C pattern

when exposed to an unmethylated segment through backcrossing (At1g64790 1× backcross; At3g43340/50 2× backcross). In



**Fig. 6.** TCM events correlate with decreases in H3K9ac levels. Chromatin immunoprecipitation (ChIP) analyses for H3K9ac and H3K9me2 levels at At3g43340/50 are from three independent ChIP experiments using qRT-PCR. The bisulphite graph from the C24 At3g43340/50 segment (Fig. 1) (6) is used as a reference and indicates regions analyzed by ChIP (blue, green, and purple bars shown underneath graph). For the ChIP graphs, columns represent the fold change relative to C24. H3K9ac levels at the TCM region were normalized against *ACT1N*, whereas H3K9me2 levels were normalized against *TA3*. The blue, green, and purple columns represent primer sets 1, 2, and 3, respectively, with primer set positions highlighted on the bisulphite PCR graph. The black dotted line represents MPV. Error bars, SEM. Red and black asterisks indicate statistically significant differences ( $P \leq 0.05$ ) in immunoprecipitation between parents and between hybrids and MPV, respectively.

contrast, this new <sup>14</sup>C pattern is maintained in F2 plants, which are homozygous for the methylated Ler\* segments.

The *trans*-chromosomal interactions changing <sup>14</sup>C patterns in F1 hybrids and subsequent generations have parallels to the interactions of alleles described in paramutation in maize (reviewed in refs. 29, 30). Both At3g43340/50 and the *booster1* locus in maize involve a donor segment suppressing a recipient segment through siRNA activity, <sup>14</sup>C, and histone modification (31–33). The late onset of TCM at At3g43340/50 is paralleled by the <sup>14</sup>C associated with the *B* locus, where the <sup>14</sup>C level increases up until the 10 leaf stage (33). Paramutation can be sequence specific, with the *B K55* and *B-peru* alleles being “immune” to paramutation (31), similar to our observations that the unmethylated At3g43340/50 locus of Cvi is not able to be involved in TCM. In the case of Cvi At3g43340/50, the lack of TCM could be due to sequence variation or to the chromatin state, which may include gene activity or unknown transacting factors (reviewed in refs. 4, 5). In the F1 hybrids, the level of <sup>14</sup>C induced by a TCM event varies between segments (6), as does the silencing ability of alleles that facilitate paramutation events in maize (reviewed in ref. 34). The ability to maintain the newly methylated state in subsequent generations varies in different cases of paramutation (reviewed in ref. 34), similar to the inheritance patterns at the two loci we have studied.

In our *Arabidopsis* hybrids, there are thousands of localized regions of TCM/TCdM (6). These events may be a common result of hybridization and could have implications for the transcriptome of the hybrid plant. Although paramutation was initially described at only a small number of loci in maize, recent

evidence indicates it occurs more frequently, with both <sup>14</sup>C and gene expression in maize hybrids being altered in many genes through paramutation-like events (35–37). The two loci examined here maintained the F1 <sup>14</sup>C patterns into the F2 generation, whereas other TCM events may show variable penetrance in subsequent generations. Further analysis will be needed to investigate how many TCM events are stably maintained and how many are absent from the next generation. Adding further to the complexity is that the timing of the TCM event in the F1 can differ from that in the F2. Given the dynamic patterns of epigenetic changes being uncovered in hybrids, it is possible that a proportion of the altered gene expression patterns may only occur in the F1 generation. Although, at this time, the importance of TCM/TCdM events in generating the heterotic phenotype in *Arabidopsis* is unknown, their frequency, and consequences for gene expression suggest such events could make significant contributions to hybrid vigor.

## Materials and Methods

Plant material, genotyping, and detailed experimental design are described in *SI Appendix, SI Materials and Methods*. Genomic DNA was bisulphite converted using the MethylEasy Xceed Kit (Human Genetic Signatures). Bisulphite PCR was carried out using unbiased bisulphite primers, allowing amplification of both converted and unconverted fragments (Dataset S1, Table S2). For McrBC digestion, 25–50 ng of genomic DNA were digested for 16 h at 37 °C. Results were obtained by qRT-PCR by comparing digested DNA with an undigested control. ChIP experiments were carried out as described in ref. 38. mRNA levels were ascertained from RNA-seq run on a Illumina HI-seq platform (Dataset S1, Table S1).

- Birchler JA, Yao H, Chudalayandi S, Vaiman D, Veitia RA (2010) Heterosis. *Plant Cell* 22(7):2105–2112.
- Shull GH (1948) What is “heterosis”? *Genetics* 33(5):439–446.
- Hollick JB, Chandler VL (1998) Epigenetic allelic states of a maize transcriptional regulatory locus exhibit overdominant gene action. *Genetics* 150(2):891–897.
- Groszmann M, et al. (2011) Epigenetics in plants-vernalisation and hybrid vigour. *Biochim Biophys Acta* 1809(8):427–437.
- Greaves IK, Groszmann M, Dennis ES, Peacock WJ (2012) Trans-chromosomal methylation. *Epigenetics* 7(8):800–805.
- Greaves IK, et al. (2012) Trans chromosomal methylation in *Arabidopsis* hybrids. *Proc Natl Acad Sci USA* 109(9):3570–3575.
- Shen H, et al. (2012) Genome-wide analysis of DNA methylation and gene expression changes in two *Arabidopsis* ecotypes and their reciprocal hybrids. *Plant Cell* 24(3):875–892.
- Chodavarapu RK, et al. (2012) Transcriptome and methylome interactions in rice hybrids. *Proc Natl Acad Sci USA* 109(30):12040–12045.
- Michalak P (2009) Epigenetic, transposon and small RNA determinants of hybrid dysfunctions. *Heredity (Edinb)* 102(1):45–50.
- Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* 11(3):204–220.
- Zhang XY, et al. (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* 126(6):1189–1201.
- Lister R, et al. (2008) Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* 133(3):523–536.
- Kurihara Y, et al. (2008) Identification of the candidate genes regulated by RNA-directed DNA methylation in *Arabidopsis*. *Biochem Biophys Res Commun* 376(3):553–557.
- Matzke M, Kanno T, Huettel B, Daxinger L, Matzke AJM (2007) Targets of RNA-directed DNA methylation. *Curr Opin Plant Biol* 10(5):512–519.
- 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(6):2617–2622.
- Jenuwein T, Allis CD (2001) Translating the histone code. *Science* 293(5532):1074–1080.
- Roudier F, Teixeira FK, Colot V (2009) Chromatin indexing in *Arabidopsis*: An epigenomic tale of tails and more. *Trends Genet* 25(11):511–517.
- Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S (2008) Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* 456(7218):125–129.
- Sasaki TKA, Kobayashi A, Saze H, Kakutani T (2012) RNAi-independent de novo DNA methylation revealed in *Arabidopsis* mutants of chromatin remodeling gene DDM1. *Plant J* 70(5):750–758.
- Hamilton A, Voinnet O, Chappell L, Baulcombe D (2002) Two classes of short interfering RNA in RNA silencing. *EMBO J* 21(17):4671–4679.
- Herr AJ, Jensen MB, Dalmay T, Baulcombe DC (2005) RNA polymerase IV directs silencing of endogenous DNA. *Science* 308(5718):118–120.
- Qian W, et al. (2012) A histone acetyltransferase regulates active DNA demethylation in *Arabidopsis*. *Science* 336(6087):1445–1448.
- Schneeberger K, et al. (2011) Reference-guided assembly of four diverse *Arabidopsis thaliana* genomes. *Proc Natl Acad Sci USA* 108(25):10249–10254.
- Groszmann M, Greaves IK, Fujimoto R, James Peacock W, Dennis ES (2013) The role of epigenetics in hybrid vigor. *Trends Genet* 29(12):684–690.
- 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(42):17835–17840.
- He GM, et al. (2010) Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell* 22(1):17–33.
- Barber WT, et al. (2012) Repeat associated small RNAs vary among parents and following hybridization in maize. *Proc Natl Acad Sci USA* 109(26):10444–10449.
- Li Y, Varala K, Moose SP, Hudson ME (2012) The inheritance pattern of 24 nt siRNA clusters in *Arabidopsis* hybrids is influenced by proximity to transposable elements. *PLoS ONE* 7(10):e47043, 10.1371/journal.pone.0047043.
- Chandler VL (2010) Paramutation’s properties and puzzles. *Science* 330(6004):628–629.
- Hollick JB (2012) Paramutation: A trans-homology interaction affecting heritable gene regulation. *Curr Opin Plant Biol* 15(5):536–543.
- Stam M, Belele C, Dorweiler JE, Chandler VL (2002) Differential chromatin structure within a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. *Genes Dev* 16(15):1906–1918.
- Nobuta K, et al. (2008) Distinct size distribution of endogenous siRNAs in maize: Evidence from deep sequencing in the mop1-1 mutant. *Proc Natl Acad Sci USA* 105(39):14958–14963.
- Haring M, et al. (2010) The role of DNA methylation, nucleosome occupancy and histone modifications in paramutation. *Plant J* 63(3):366–378.
- Pilu R (2011) Paramutation: Just a curiosity or fine tuning of gene expression in the next generation? *Curr Genomics* 12(4):298–306.
- Li L, et al. (2013) Mendelian and non-Mendelian regulation of gene expression in maize. *PLoS Genet* 9(1):e1003202, 10.1371/journal.pgen.1003202.
- Regulski M, et al. (2013) The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. *Genome Res* 23(10):1651–1662.
- Eichten SR, et al. (2013) Epigenetic and genetic influences on DNA methylation variation in maize populations. *Plant Cell* 25(8):2783–2797.
- Helliwell CA, Wood CC, Robertson M, James Peacock W, Dennis ES (2006) The *Arabidopsis* FLC protein interacts directly in vivo with SOC1 and FT chromatin and is part of a high-molecular-weight protein complex. *Plant J* 46(2):183–192.