Additive inheritance of histone modifications in Arabidopsis thaliana intra-specific hybrids

Ali M. Banaei Moghaddam, François Roudier, Michael Seifert, Caroline Bérard, Marie-Laure Martin-Magniette, Raheleh Karimi Ashtiyani, Andreas Houben, Vincent Colot

the plant journal

SEB Society for Experimental Biology

The Plant Journal (2011) **67**, 691–700 doi: 10.1111/j.1365-313X.2011.04628.x

Additive inheritance of histone modifications in *Arabidopsis thaliana* intra-specific hybrids

Ali M. Banaei Moghaddam[1], Francois Roudier[2,†], Michael Seifert[1,†], Caroline Bérard[3], Marie-Laure M. Magniette[3], Raheleh Karimi Ashtiyani[1], Andreas Houben[1], Vincent Colot[2] and Michael F. Mette[1,*]

[1]*Leibniz Institute of Plant Genetics and Crop Plant Research, Corrensstraße 3, 06466 Gatersleben, Germany,*
[2]*Institut de Biologie de l'Ecole Normale Supérieure (IBENS), Centre National de la Recherche Scientifique (CNRS) UMR8197, Institut National de la Santé et de la Recherche Médicale (INSERM) U1024, Paris, France, and*
[3]*AgroParisTech, Institut National de la Recherche Agronomique (INRA) UMR518, Paris, France*

Received 14 March 2011; revised 29 April 2011; accepted 4 May 2011; published online 11 July 2011.
*For correspondence (fax +49 39482 5137; e-mail mette@ipk-gatersleben.de).
†These authors contributed equally to this work.

SUMMARY

Plant genomes are earmarked with defined patterns of chromatin marks. Little is known about the stability of these epigenomes when related, but distinct genomes are brought together by intra-species hybridization. *Arabidopsis thaliana* accessions and their reciprocal hybrids were used as a model system to investigate the dynamics of histone modification patterns. The genome-wide distribution of histone modifications H3K4me2 and H3K27me3 in the inbred parental accessions Col-0, C24 and Cvi and their hybrid offspring was compared by chromatin immunoprecipitation in combination with genome tiling array hybridization. The analysis revealed that, in addition to DNA sequence polymorphisms, chromatin modification variations exist among accessions of *A. thaliana*. The range of these variations was higher for H3K27me3 (typically a repressive mark) than for H3K4me2 (typically an active mark). H3K4me2 and H3K27me3 were rather stable in response to intra-species hybridization, with mainly additive inheritance in hybrid offspring. In conclusion, intra-species hybridization does not result in gross changes to chromatin modifications.

Keywords: *Arabidopsis thaliana*, epigenome, heterosis, histone methylation, intra-specific hybrids, ChIP on chip.

INTRODUCTION

Extensive studies of DNA methylation and histone modifications in *Arabidopsis thaliana* (Turck *et al.*, 2007; Zhang *et al.*, 2007, 2009; Cokus *et al.*, 2008) and rice (*Oryza sativa*) (He *et al.*, 2010; Zemach *et al.*, 2010) have revealed that plant genomes are earmarked by well-defined patterns of chromatin marks. In the context of their transcriptional activity or inactivity, particular sequence classes such as genes or repeat elements are preferentially associated with distinct patterns of DNA methylation and histone modifications (Roudier *et al.*, 2009; Teixeira and Colot, 2010).

The combining of related but distinct genomes with their respective patterns of chromatin marks in inter-species hybridization and allopolyploid formation often results in changes to chromatin marks. In hybrids of various Arabidopsis species, one parental set of ribosomal RNA genes was shown to be silenced within a few generations, but could be re-activated by interfering with either DNA methylation or histone deacetylation, suggesting a pivotal role for chromatin modification in the regulation of expression of orthologous genes (Lee and Chen, 2001; Lawrence *et al.*, 2004). Gene expression studies in synthetic allopolyploid Arabidopsis (Comai, 2000) and cotton (*Gossypium hirsutum*) (Brubaker *et al.*, 1999) revealed that gene silencing occurs during the first or second generation after hybridization. However, other studies in allopolyploid *Spartina anglica*, *Brassica juncea* and cotton showed that the activity of parental genomes remained unchanged (Axelsson *et al.*, 2000; Liu *et al.*, 2001; Baumel *et al.*, 2002). Induction of DNA methylation changes after hybridization was reported for synthetic *Cucumis* allopolyploids (Chen and Chen, 2008). Similarly, in experimentally synthesized *Brassica napus* (Xu *et al.*, 2009) and Arabidopsis allopolyploids (Madlung *et al.*, 2002), 7 and 8%, respectively, of the tested DNA sites showed changes in cytosine methylation status in comparison with their respective diploid progenitors.

Less is known about the stability or dynamics of chromatin modifications in response to intra-species hybridization. Limited differential DNA methylation (approximately 1% gain or loss) in comparison with the respective progenitors was found for intra-species hybrids of rice (Xiong *et al.*, 1999) and cotton (Zhao *et al.*, 2008). In two rice cultivars from different sub-species, *Oryza sativa japonica* and *O. sativa indica*, variation of DNA methylation, and, to a lower extent, variation of histone modifications H3K4me3 and H3K27me3, was observed between parental lines. In reciprocal hybrids of the two rice sub-species, distinct non-additive patterns of chromatin marks were observed. The level of changes after hybridization was higher for DNA methylation, and both histone modifications were mainly inherited additively in the hybrids (He *et al.*, 2010). Inbred accessions of *A. thaliana* also display substantial DNA methylation variation between each other (Vaughn *et al.*, 2007). Investigation of the DNA methylation pattern in two different accessions of *A. thaliana* and their reciprocal F_1 hybrid progeny showed that DNA methylation polymorphisms are mostly inherited additively, with only limited changes after hybridization (Zhang *et al.*, 2008; Banaei Moghaddam *et al.*, 2010; Groszmann *et al.*, 2011).

Here, we determined whether intra-species crosses between inbred lines lead to changes in chromatin marks other than DNA methylation using accessions of *A. thaliana* as a model. We selected histone H3 dimethylated at lysine 4 (H3K4me2) and histone H3 trimethylated at lysine 27 (H3K27me3) as contrasting histone H3 modifications marks (Fuchs *et al.*, 2006; Kouzarides, 2007; Roudier *et al.*, 2009). Histone H3K4me2 was chosen as a general euchromatic mark that is absent from silent repeat elements, and H3K27me3 was chosen as a euchromatic mark that is mostly associated with genes repressed by polycomb repressive complex 2 (Schubert *et al.*, 2006; Turck *et al.*, 2007; Zhang *et al.*, 2007). A genome-wide 'ChIP on chip' analysis of H3K4me2 distribution in *A. thaliana* indicated that 6% of sequences are targeted by this mark (Zhang *et al.*, 2009). Of these target regions, 93% were genes, with particular enrichment of H3K4me2 in the promoter and 5′ end of transcribed regions, and only 1.3% of the target regions were transposable elements (TEs). In contrast, histone H3K27me3 is associated with silent genes distributed in euchromatic regions that are subject to tissue-specific or developmentally regulated expression (Turck *et al.*, 2007; Zhang *et al.*, 2007). Genome-wide analyses revealed that at least 15–20% of *A. thaliana* genes are targeted by this histone mark and show tissue-specific expression patterns. H3K27me3-marked domains are largely coincident with the entire transcribed region of genes (Turck *et al.*, 2007; Zhang *et al.*, 2007).

To study the stability of histone modification patterns in response to intra-species hybridization of various inbred accessions of *A. thaliana* (Col-0, C24 and Cvi), we performed chromatin immunoprecipitation in combination with genome tiling array hybridization (ChIP on chip) analyses. Changes in the H3K27me3 and H3K4me2 distribution between Col-0, Cvi and C24 were identified, with a greater range of variations for H3K27me3 than H3K4me2. H3K4me2 and H3K27me3 were rather stable after intra-species hybridization, with additive inheritance in Col-0 × Cvi and Col-0 × C24 F_1 hybrid offspring. Changes in the distribution of histone modifications after hybridization were detected in 346 genes for H3K4me2 in Col-0 × Cvi progeny, and in 1233 and 876 genes for H3K27me3 in Col-0 × Cvi and Col-0 × C24 progeny, respectively. However, these changes were rather random and were not associated with particular sequence categories.

RESULTS

Genomic sequence variation among *A. thaliana* accessions resides mainly in transposable elements

Comparative genomic hybridization (CGH) experiments were performed using Arabidopsis whole-genome tiling NimbleGen arrays to identify differences in the genomic sequences of *A. thaliana* accessions Cvi and C24 compared to the reference accession Col-0. The CGH analysis, which detects copy number variation and sequence polymorphisms, was a necessary prerequisite for the comparison of histone modification patterns between Cvi, C24 and Col-0. Based on this information, it was possible to distinguish whether differential hybridization signals of labelled DNA derived from immunoprecipitated chromatin from the various accessions were due to differences in histone modifications rather than variation in the DNA sequence. In total, 6.0 and 5.5% of tiles showed significant CGH polymorphisms

Table 1 CGH analysis for Col-0 versus C24 and Col-0 versus Cvi

	Col-0 versus Cvi	Col-0 versus C24
Percentage of tiles showing CGH polymorphism	6.0	5.5
Percentage of tiles with lower copy number in C24 or Cvi (total size in kb)	5.3 (5395)	5.2 (5221)
Number of tiles per domain	2–369	2–372
Mean size of domain (kb)	3.3	3.9
90% of domains had a size of less than (kb)	6	8
Size of largest domain (kb)	40	57
Percentage of tiles with higher copy number in C24 or Cvi (total size in kb)	0.67 (631)	0.35 (308)
Number of tiles per domain	2–169	2–54
Mean size of domain (kb)	3.0	1.9
90% of domains had a size of less than (kb)	6.5	4
Size of largest domain (kb)	26	9

for Col-0 versus Cvi and Col-0 versus C24, respectively (Table 1). Most of the CGH polymorphic tiles indicated a decrease in copy number of the corresponding sequence in C24 and Cvi compared to Col-0.

Further analysis focused on CGH polymorphic domains containing two or more consecutive CGH polymorphic tiles (Table 1). Of the tiles present in CGH polymorphic domains identified between Col-0 and Cvi, as well between Col-0 and C24, 93% were identical. However, copy number varied in some cases in opposite directions between accessions, with an increase in Col-0 versus Cvi and a decrease in Col-0 versus C24, and vice versa. Annotation of tiles within CGH polymorphic domains indicated that copy number variation mainly affects TEs, while genic regions and 5′ and 3′ untranslated regions (UTRs) are more conserved between the analysed accessions (Figure 1). Ontology categorization of the genes (excluding TEs) associated with CGH polymorphic domains according to the Munich Information Center for Protein Sequences (MIPS) Functional Catalogue (Ruepp *et al.*, 2004) indicated an excess of genes with functions in signal transduction (Figure 2, category 30) or cell defence (Figure 2, category 32). As annotated TEs were excluded prior to ontology analysis, they are absent from the gene ontology data (Figure 2, category 38). In addition, unclassified proteins (Figure 2, category 99) were common.

In summary, the CGH data revealed sequence polymorphisms among the analysed accessions of *A. thaliana.* These polymorphisms occurred in all types of sequences, but TEs and genes involved in signal transduction and cell defence were over-represented.

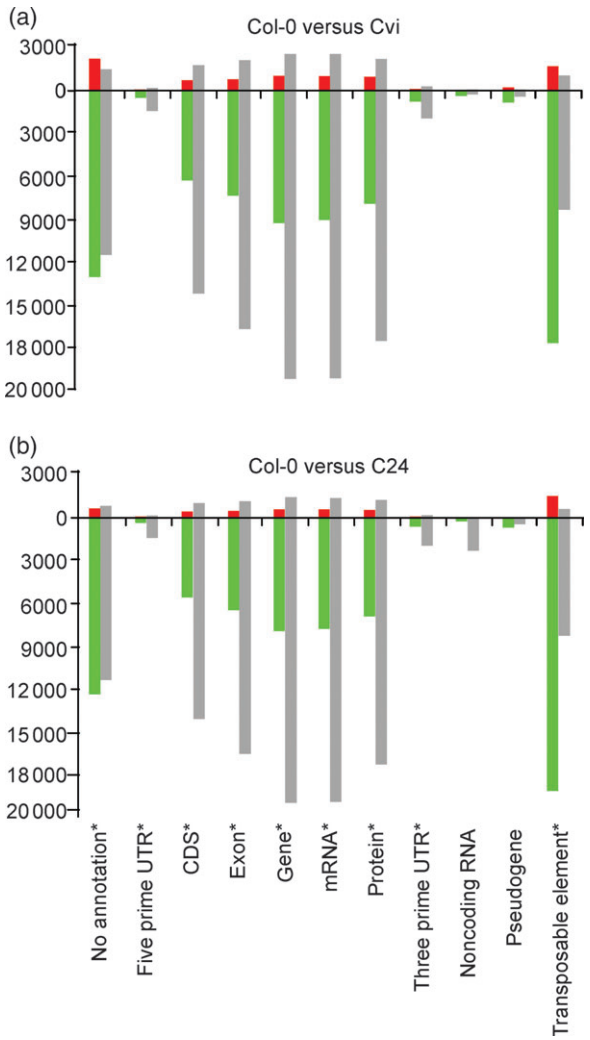

Figure 1. Classification of CGH polymorphic tiles based on their sequence types.
CGH polymorphic tiles between *A. thaliana* accessions (a) Col-0 versus Cvi and (b) Col-0 versus C24, were classified based on annotation (TAIR8) of their underlying sequences. Red and green bars represent tiles with higher and lower copy numbers in Cvi (a) and C24 (b) relative to Col-0, respectively. Grey bars indicate controls by random counts. Cases of significant deviation between CGH data (green or red bars) and random counts (grey bars) as indicated by a *P* value <0.01 are indicated by asterisks.

Genome-wide patterns of histone modifications H3K4me2 and H3K27me3 show variation among *A. thaliana* accessions

Next, we performed ChIP on chip analysis using Arabidopsis whole-genome tiling NimbleGen arrays to determine the genome-wide distribution of H3K4me2 in accessions Col-0 and Cvi and in Col-0 × Cvi F_1 hybrids, and of H3K27me3 in Col-0, Cvi and C24, and in Col-0 × Cvi and Col-0 × C24 F_1 hybrids. The quality of ChIP assays was confirmed by quantitative PCR using primers specific for sequences known to be associated with H3K4me2 or H3K27me3 (Figure S1) from previous studies (Turck *et al.*, 2007; Zhang *et al.*, 2007, 2009).

Tiles showing polymorphic hybridization signals in CGH were excluded from the ChIP on chip analysis to avoid the influence of sequence differences among accessions. Tiles that were found to be significantly associated with a given modification in any of the analysed genotypes were assigned to histone modification domains of at least three successive tiles that correspond to a region of at least 0.3 kb (Table S1). Given the mean size of chromatin fragments (0.8 kb) and the resolution provided by the NimbleGen microarrays (165 bp) used in the ChIP on chip experiments, tiles that could not be assigned to such domains were not considered for further analysis. Subsequently, histone modification domains were categorized according to the TAIR8 annotation into genes or TEs. In the case of large domains, it may be that one domain simultaneously harbours genic sequences and TEs.

First the distribution of histone modifications was compared between parental accessions Col-0 and Cvi, as well as Col-0 and C24 (Figure S2). H3K4me2, a classical euchromatic histone mark, was associated with domains ranging in length from 314 bp to 31.2 and 21.9 kb in Col-0 and Cvi, respectively (Table 2 and Figure S3). Overall, 94% of these

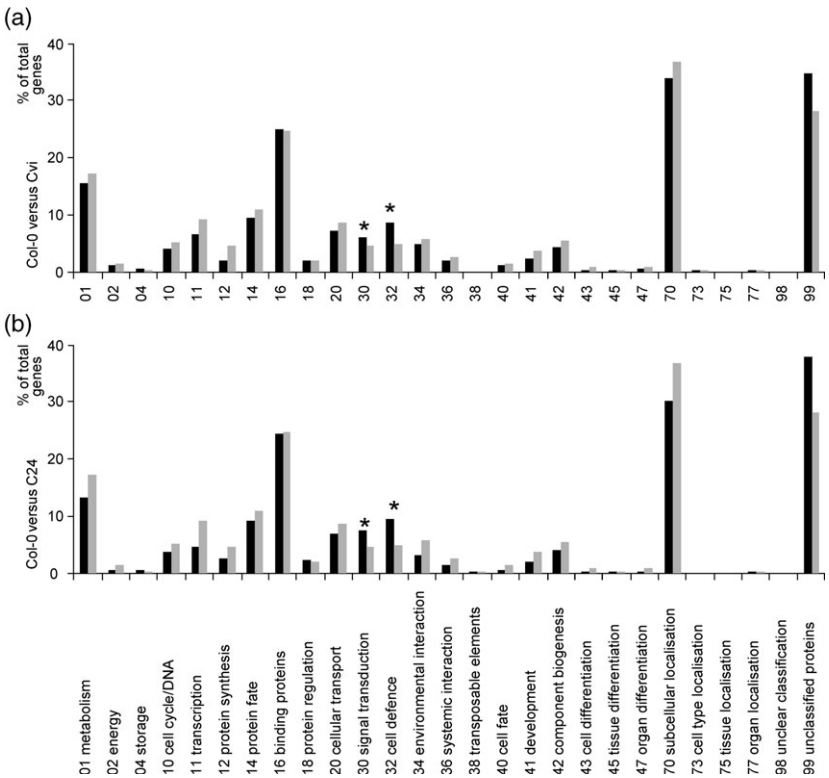

Figure 2. Ontology classification of genes localized in CGH polymorphic regions.
Genic sequences (according to TAIR8) that were found to be localized in CGH polymorphic regions indentified between (a) Col-0 and Cvi and (b) Col-0 and C24 were classified into gene ontology groups according to the MIPS Functional Catalogue (Ruepp *et al.*, 2004) to identify potentially over-represented categories. Black bars indicate the frequencies of CGH polymorphic genic sequences; as a control, grey bars indicate the frequency of randomly selected genic sequences in the various gene ontology classes as a percentage of the total number of genes in each class. Asterisks indicate cases of significant deviation between the frequencies of polymorphic and randomly selected sequences in a given category with a P value <0.001.

domains coincided with genes (Table S2) and 9% with TE sequences (Table S3). Intersection analysis of H3K4me2-marked domains in Col-0 and Cvi revealed that 87% of them were common between both accessions. These common domains coincided with 93% (21 908) of the genes and 60% (1526) of the TE sequences associated with H3K4me2 (Figure 3). The length of the 13% of H3K4me2-marked regions that differed between the two accessions (Table 3) ranged from 314 bp to 6.7 kb, and corresponded to 4.6% (1053) and 2.6% (581) of the genes and 26% (537) and 23.9% (479) of the TE sequences in Col-0 and Cvi, respectively (Table 3). Overall, a larger proportion of TEs than genes showed H3K4me2 polymorphisms. However, the absolute numbers of genes and TEs showing differential association with H3K4me2 between Col-0 and Cvi were similar, as fewer TEs than genes are associated with this histone modification.

Genes that were differentially associated with H3K4me2 in either of the parental lines were randomly distributed in various gene ontology groups according to the MIPS Functional Catalogue (Figure S4) (Ruepp *et al.*, 2004) and the DAVID tool (Table S4) (Huang *et al.*, 2009). Similar results were obtained for the distribution of H3K4me2 in and between Col-0 and C24 accessions analysed using a chromosome 4 tiling array (data not shown) (Turck *et al.*, 2007).

H3K27me3, a histone mark associated with the repression of genes, was found over domains ranging from 315 bp to 26.7 kb in length (Table 2 and Figure S3) in Col-0, Cvi and C24. In Col-0, 67% of these domains coincided with genes and 36% with TEs. Similar results were obtained for Cvi and

Table 2 H3K4me2- and H3K27me3-associated domains in Col-0, Cvi and C24

Experiment[a]	Genotype	Modification	Associated tiles (from total of 717 235)	Number of domains	Domain size (kb)			Annotation	
					Maximum	Minimum	Mean	Genes	TEs
A	Col-0	H3K4me2	287 285	21 539	31.2	0.3	2.2	22 961	2063
	Cvi	H3K4me2	281 000	20 972	21.9	0.3	2.2	22 489	2005
B	Col-0	H3K27me3	140 035	11 182	22.7	0.3	2.1	9125	4950
	Cvi	H3K27me3	138 791	12 585	20	0.3	1.9	9326	5357
C	Col-0	H3K27me3	147 318	9834	26.7	0.3	2.5	9427	4704
	C24	H3K27me3	125 947	10 068	22.5	0.3	2.1	8118	4471

[a]Chromatin preparations for experiments A and B were performed in parallel but the one for experiment C was performed independently.

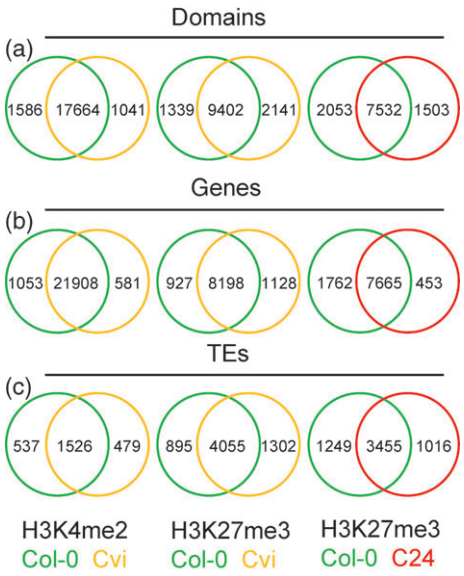

Figure 3. Intersection analysis of domains, genes and TEs associated with H3K4me2 or H3K27me3 in three *A. thaliana* accessions.
The sets of (a) domains, (b) genes and (c) TEs that were found to be associated with H3K27me3 or H3K4me2 in ChIP on chip experiments using whole-genome NimbleGen tiling arrays in *A. thaliana* accessions Col-0 (green circles), Cvi (yellow circles) and C24 (red circles) were subjected to intersection analysis to determine common elements (intersection area of two circles) and unique elements (areas outside intersections). Due to the analysis method, the sums of domain numbers do not necessarily correspond to the total numbers of domains in Table 2.

C24 (Figure 3, Table 2, and Tables S2 and S3). Intersection analysis of H3K27me3-marked domains revealed that 87% of them were common between Col-0 and Cvi. These common domains coincided with 80% (8198) of the genes and 58% (4055) of the TEs marked by H3K27me3 (Figure 3). Compared to H3K4me2, fewer genes and more TEs were associated with H3K27me3 in Col-0 and Cvi (Figure 3). Some H3K27me3 domains were unique to one of the analysed accessions (Table 3). The length of these H3K27me3 polymorphic domains ranged from 314 bp to 6.6 kb in Col-0 and 5.7 kb in Cvi. These domains coincided with 9% (927) and 11% (1128) of the genes, and 17% (895) and 25% (1302) of the TEs associated with H3K27me3, respectively. Consistent results were obtained for comparison of H3K27me3 polymorphic domains between Col-0 and C24 (Figure 3 and Table 3). Thus, in general, more TEs coincided with H3K27me3 than with H3K4me2 polymorphic domains. Similar numbers of genes and TEs were differentially associated with H3K27me3 in Col-0 compared to Cvi and Col-0 compared to C24. These genes were randomly distributed in various gene ontology groups (Figure S4 and Table S4).

Comparison of genes associated with H3K27me3 in the reference accession Col-0 with those found in previous studies (Turck *et al.*, 2007; Zhang *et al.*, 2007) revealed approximately 80% overlap, indicating good reproducibility despite the differences in Plant materials and Experimental Procedures (Figure S5). In addition, our study identified H3K27me3-associated genes that were not found in previous studies.

Taken together, our findings indicate that H3K4me2 and H3K27me3 distribution patterns are highly conserved in different A. *thaliana* accessions, with few local variations in a random manner regardless of TEs and gene ontology.

Intra-species hybridization mediates limited alterations in H3K4me2 and H3K27me3 distribution patterns

We next investigated whether intra-species hybridization could generate alternative distribution patterns of H3K4me2 in hybrid offspring between Col-0 and Cvi (Figure 4a–c, Figure S2, and Tables S1–S3). In order to focus on the most relevant cases, the analysis was performed on domains that were associated with H3K4me2 exclusively in both parental accessions, but not in F_1 hybrid progeny (Figure 4d, category IV), or in the F_1 hybrid progeny, but not in the parental accessions (Figure 4d, category III). The few hybridization-responsive domains identified (3% of all H3K4me2-associated domains) corresponded to 346 genes and 226 TEs (1.5 and 8% of H3K4me2-associated genes and TEs, respectively) and their length ranged from 0.3 to 1.6 kb (Table 4). In comparison to genes, TEs were over-represented among hybridization-responsive H3K4me2 domains. The hybridization-responsive H3K4me2-associated genes were randomly distributed between various ontology groups and gene families (Figure S6 and Table S4). Similar results were

Table 3 H3K4me2 and H3K27me3 polymorphic domains in Col-0 versus Cvi and Col-0 versus C24

Experiment[a]	Specific for accession	Modific ation	Number of domain	Domain size (kb)			Annotation	
				Maximum	Minimum	Mean	Genes	TEs
A	Col-0	H3K4me2	1585	6.6	0.3	0.67	1053	537
	Cvi	H3K4me2	1041	4.8	0.3	0.72	581	479
B	Col-0	H3K27me3	1339	4.6	0.3	0.72	927	895
	Cvi	H3K27me3	2141	5.7	0.3	0.81	1128	1302
C	Col-0	H3K27me3	2053	7.4	0.3	0.9	1762	1249
	C24	H3K27me3	1503	6.7	0.3	0.77	453	1016

[a]Chromatin preparations for experiments A and B were performed in parallel but the one for experiment C was performed independently.

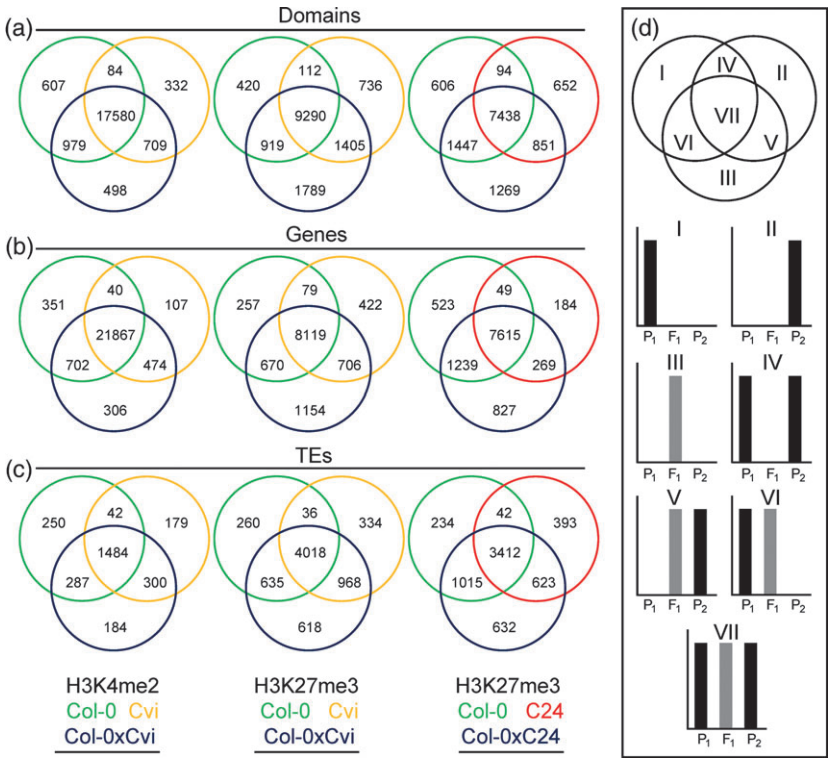

Figure 4. Intersection analysis of domains, genes and TEs associated with H3K4me2 or H3K27me3 in *A. thaliana* accessions and their F_1 hybrids.
(a–c) The sets of (a) domains, (b) genes and (c) TEs that were found to be associated with H3K27me3 or H3K4me2 in ChIP on chip experiments using whole-genome NimbleGen tiling arrays in *A. thaliana* parental accessions Col-0 (green circles), Cvi (yellow circles) and C24 (red circles), and their F_1 hybrids Col-0 × Cvi, and Col-0 × C24, respectively (blue circles), were subjected to intersection analysis.
(d) The seven areas of each three-circle intersection diagram define seven categories of elements. Categories I and II comprise cases in which a histone mark is detected in one of the parental accessions (P_1 or P_2, respectively), but not in the other parental accession or the F_1 hybrid. Category III comprises cases in which a histone mark is detected in neither parental accession, but is present in the F_1 hybrid. Category IV comprises cases in which a histone mark is detected in both parental accessions, but not in the F_1 hybrid. Categories V and VI comprise cases in which a histone mark is detected in one of the parental accessions (P_1 or P_2, respectively) and in the F_1 hybrid, but not in the other parental accession. Category VII comprises cases in which a histone mark is detected in both parental accessions and the F_1 hybrid. Categories III and IV indicate changes in histone marks in F_1 hybrids in comparison with the parental accessions, while the other categories either indicate no change (category VII) or are not conclusive with regard to changes (categories I, II, V and VI).

Table 4 H3K4me2 and H3K27me3 hybridization-responsive domains in Col-0 × Cvi and Col-0 × C24

			Hybridization-responsive domains (Figure 4, categories III and IV)					
				Domain size (kb)			Annotation	
Experiment[a]	Crosses	Modifications	Number of domains	Maximum	Minimum	Mean	Genes	TEs
A	Col-0 × Cvi	H3K4me2	582	1.6	0.3	0.6	346	226
B	Col-0 × Cvi	H3K27me3	1901	2.2	0.3	0.6	1233	654
C	Col-0 × C24	H3K27me3	1363	2.4	0.3	0.6	876	674

[a]Chromatin preparations for experiment A and B were performed in parallel but the one for experiment C was performed independently.

found for Col-0 × C24 F_1 hybrids in ChIP on chip analysis using the chromosome 4 tiling array (data not shown).

Whole-genome comparison (Figure S2) and intersection analysis of H3K27me3-marked domains, genes and TE sequences involved were also performed for Col-0, Cvi and Col-0 × Cvi (Figure 4). Hybridization-responsive domains (13% of all H3K27me3-associated domains; Figure 4d, categories III and IV) corresponded to 1233 genes and 654 TEs (11 and 10% of total H3K27me3-associated genes and TEs, respectively), and ranged in size from 0.3 to 2.2 kb (Table 4). No over-representation of either TEs or genes was detected. The hybridization-responsive H3K27me3-marked genes were randomly distributed between ontology groups and gene families (Figure S6 and Table S4). Similar

H3K27me3 dynamics were found in Col-0 × C24 hybrids (Figure 4a–c and Tables S1–S3).

In conclusion, the genome-wide distribution of H3K4me2 is rather stable and that of H3K27me3 is slightly more dynamic in response to intra-species hybridization. TEs as well as genes showed changes in both histone modifications. No obvious over-representation of particular gene ontology groups or families was found. The changes in the two marks analysed happened largely independently. For very few genes, histone modifications changed simultaneously in response to hybridization (Figure S7).

DISCUSSION

By combining comparative genomic hybridization and epigenomic profiling, we have shown that H3K4me2 and H3K27me3 distribution patterns are overall very similar in various *Arabidopsis thaliana* accessions, and remain largely unchanged in their F_1 progeny.

Sequence polymorphisms between *A. thaliana* accessions are mainly found in transposable elements and genes undergoing rapid evolution

Our CGH analysis of the genomes of Cvi and C24 relative to Col-0 identified regions showing DNA sequence polymorphism between these accessions of *A. thaliana*. The number of polymorphic regions identified by CGH was slightly higher for Cvi than for C24 in comparison with Col-0. This is consistent with a previous analysis based on single nucleotide polymorphisms, which indicated that Cvi is more divergent from Col-0 than C24 is (Schmid *et al.*, 2003; Clark *et al.*, 2007).

As TE sequences change more rapidly than genes (Kazazian, 2004), sequence differences between *A. thaliana* accessions are not randomly distributed, and TEs differ more than other parts of the genome (Figure 1). A similar non-random distribution of sequence polymorphisms was previously found by CGH analysis between *A. thaliana* accessions (Clark *et al.*, 2007) and between maize (*Zea mays*) inbred lines B73 and Mo17 (Springer *et al.*, 2009). In maize, approximately 70% of polymorphic regions were located in intergenic regions.

Among the CGH polymorphic regions, gene ontology analysis indicated over-representation of functions associated with signal transduction and cell defence (Figure 2). Copy number variation between accessions of *A. thaliana* has previously been reported for loci involved in plant disease resistance (Noel *et al.*, 1999). Furthermore, consistent with our findings, a comparison of 20 accessions of *A. thaliana* by microarray-based whole-genome re-sequencing revealed that members of defence-related families such as nucleotide-binding leucine-rich repeat genes and receptor-like kinase genes are over-represented among genes that are affected by sequence polymorphisms (Clark *et al.*, 2007). The enhanced rate of sequence variation indicates rapid evolution of these gene families.

Fewer CGH polymorphic tiles indicated an increase in copy number in C24 and Cvi compared with those that indicated a decrease in copy number in both accessions (Figure 1). This could be due to the fact that the tiling arrays used in this study are based on the Col-0 reference sequence. Thus, sequences that solely exist in Col-0 but not in C24 or Cvi are classified as absent from C24 or Cvi, whereas sequences that solely exist in C24 or Cvi but not in Col-0 are not detected. Therefore, our CGH results probably under-estimate the level of sequence polymorphism between *A. thaliana* accessions. CGH-identified polymorphic sequences were excluded from the ChIP on chip analysis, as it was impossible to distinguish whether differences in signal intensities between the accessions were due to differences in the DNA sequence or the histone modification status. Thus, the possibility cannot be excluded that some fast-evolving genomic regions that were removed from analysis showed differential chromatin marking between accessions.

Histone modification patterns are conserved between accessions of *A. thaliana*

Several previous studies have detected differential DNA methylation in various *A. thaliana* accessions (Vaughn *et al.*, 2007; Zhang *et al.*, 2008; Banaei Moghaddam *et al.*, 2010). Not only sequence conservation, but also DNA methylation patterns, were found to be more similar between Col-0 and C24 than between Col-0 and Cvi (Vaughn *et al.*, 2007). Our ChIP on chip data extend these observations to histone H3K4me2 and H3K27me3 distribution patterns. These differ between accessions of *A. thaliana*, consistent with the situation observed in cultivars of rice (He *et al.*, 2010).

Among domains with different histone modification status between accessions, the mean length was <1 kb for both histone marks, while the mean length of the conserved H3K4me2- and H3K27me3-marked domains exceeded 2 kb. Thus, polymorphic histone modification domains are restricted to rather small regions. Despite their shorter length, these differentially marked domains may be associated with locus-specific differential regulation in the various accessions. For instance, TE sequences next to genes can affect their transcriptional regulation through deposition of repressive chromatin modifications. A TE in an intron of *FLOWERING LOCUS C* (*FLC*) in *A. thaliana* accession Landsberg *erecta* causes transcriptional inactivation of this locus, and consequently earlier flowering of L*er* in comparison to Col-0 (Liu *et al.*, 2004). TEs also alter the expression of adjacent genes in wheat (*Triticum aestivum*) (Kashkush *et al.*, 2003). It remains to be determined whether the regions differentially marked by the repressive H3K27me3 or active H3K4me2 modifications in Cvi or C24 in comparison to Col-0 also show differential expression between accessions.

Consistently, we found that polymorphisms in H3K27me3 (typically a repressive mark) were associated with both TEs and genes. In contrast, polymorphisms in H3K4me2 (typically an active mark) were mainly restricted to genes. This agrees well with genome-wide high-resolution analyses of histone modifications in *Saccharomyces cerevisiae* and mammals, which detected H3K4me2 in regions undergoing transcription across the body (Pokholok *et al.*, 2005) and in the vicinity of active genes (Bernstein *et al.*, 2005). Histone modification polymorphisms in genes were not associated with particular gene ontology classes.

The mechanisms by which histone modification polymorphisms are maintained over generations are not clear (Saze, 2008). On the one hand, heritable maintenance of particular chromatin states cannot be excluded. On the other hand, chromatin modifications may alter as consequences of changes in transcriptional activity. In *A. thaliana*, H3K27me3 is deposited by polycomb repressive complex 2 (Schubert *et al.*, 2006), and has been found to be associated with approximately 4400 genes, many of which are differentially expressed during development (Zhang *et al.*, 2007). As accessions may have evolved specific developmental programs (Chen, 2010), some of the H3K27me3 polymorphisms observed in our study could correspond to differences in gene expression patterns. Indeed, gene activity and H3K4 and H3K27 methylation levels were correlated in different rice cultivars (He *et al.*, 2010).

The observed similarities and differences suggest a role for chromatin modifications in addition to that of DNA sequence polymorphisms in the diversity of various accessions of *A. thaliana*.

Inheritance of H3K27me3 and H3K4me2 distribution patterns in hybrid offspring is additive and only to a small extent responsive to intra-specific hybridization

Comparison of H3K4me2 and H3K27me3 distribution patterns between hybrid offspring and parental inbred lines revealed limited changes in intra-specific hybrids. These results are in agreement with a previous study on *A. thaliana* accessions and their intra-specific hybrids regarding the distribution of histone methylation marks at the microscopic level (Banaei Moghaddam *et al.*, 2010), as well as a study that compared histone methylation patterns in hybrids between rice cultivars (He *et al.*, 2010). Similarly, inheritance of DNA methylation polymorphisms has been shown to be additive (Zhang *et al.*, 2008; Banaei Moghaddam *et al.*, 2010).

We conclude that intra-specific hybridization in *A. thaliana* does not result in global epigenomic rearrangements. More investigations are required to analyse whether this conclusion is also valid for other chromatin modifications. More generally, it would be interesting to determine whether heritable variations in epigenomic patterns between inbred lines contribute to hybrid performance in addition to sequence polymorphisms.

EXPERIMENTAL PROCEDURES

Plant materials

A. thaliana accessions Col-0, C24 and Cvi and their reciprocal F_1 hybrid offspring Col-0 × C24, C24 × Col-0, Col-0 × Cvi and Cvi × Col-0 were used (Banaei Moghaddam *et al.*, 2010). Approximately 400 seeds of each sample were surface-sterilized and cultured in liquid medium, and grown for 10 days under controlled conditions with 16 h light per day (light intensity of approximately 100 μE), 22°C day temperature and 18°C night temperature (Lippman *et al.*, 2004).

DNA extraction, chromatin immunoprecipitation and array hybridization

Plant genomic DNA used for CGH was extracted using the Qiagen DNeasy plant DNA extraction system (http://www.qiagen.com/) according to manufacturer's instructions. ChIP assays were performed essentially as described previously (Gendrel *et al.*, 2005) using anti-H3K4me2 (07-030) and H3K27me3 (07-449) antibodies from Upstate/Millipore (http://www.millipore.com). Each experiment was performed in two biological replicates.

DNA recovered after immunoprecipitation (IP) and directly from input Col-0 chromatin (INPUT), or genomic DNA extracted from the various genotypes, was amplified, differentially labelled and co-hybridized in dye-swap experiments to correct for dye biases, as previously described (Lippman *et al.*, 2004; Turck *et al.*, 2007) for the chromosome 4 tiling microarray, or according to the manufacturer's instructions for the whole-genome tiling arrays (Roche NimbleGen, http://www.nimblegen.com). The Arabidopsis chromosome 4 tiling microarray comprised 21 800 printed features, with a mean size of 1 kb, covering the main part of chromosome 4. The heterochromatic knob on the short arm and several megabases of pericentromeric heterochromatin are included, and account for 16% of the 18.6 Mb covered by the array (Martienssen et al., 2005). Details of array design and production are described by Vaughn *et al.* (2007). This platform has been deposited to the Gene Expression Omnibus (GEO) under accession number GPL10172. The whole-genome tiling microarray consists of 50–75 nt tiles, with a mean spacing of 165 nt, that are distributed across the entire genome sequence (TAIR7) without repeat masking. These tiles have a mean melting temperature of 74°C, and 88% of them match a unique position in the genome. This custom design was split into two arrays of 360 718 tiles each, using every other tile in each array (GEO accessions GPL10919 and GPL10920).

Comparative genomic hybridization (CGH) analysis

CGH experiments were performed for Col-0 versus Cvi and Col-0 versus C24. Data obtained using Arabidopsis whole-genome tiling NimbleGen arrays were analysed using hidden Markov models (Seifert *et al.*, 2009). A fully connected three-state hidden Markov model with state-specific Gaussian emission densities was adapted to each CGH experiment using a Bayesian Baum–Welch algorithm. Decoding of the status of each tile (deleted, unchanged or amplified) was performed using the Viterbi algorithm. Groups of contiguous tiles with log ratios that were significant different from zero for each set of compared accessions were interpreted as representing regions of copy number variation. For interpretation of CGH data, only contiguous CGH polymorphic tiles that were consistently found in both directions of dye-swap experiments were included.

ChIP on chip analysis

Raw hybridization data obtained with the chromosome 4 tiling array were normalized as described previously (Turck *et al.*, 2007),

and an ANOVA model was applied to whole-genome data to remove technical biases. Normalized data were analysed using the ChIPmix method (Martin-Magniette *et al.*, 2008), which was adapted to handle multiple biological replicates simultaneously. This method is based on a mixture model of regressions, the parameters of which are estimated using an expectation-maximization (EM) algorithm. For each tile, a posterior probability, defined as the probability of enrichment given the log(INPUT) and log(IP) intensities, is used to classify the tile into the normal or enriched class. A false-positive risk is determined by defining the probability of obtaining a posterior probability at least as extreme as the one that is actually observed when the tile is normal. False-positive risks are then adjusted by the Benjamini–Hochberg procedure, and tiles for which the adjusted false-positive risk is lower than 0.01 are considered enriched. Previously published data (Turck *et al.*, 2007; Vaughn *et al.*, 2007) were re-analysed using the same procedure. Neighbouring enriched tiles are combined into domains, requiring minimal runs of 1.6 kb or 300 bp and allowing maximal gaps of 800 or 200 bp for chromosome 4 or whole-genome data, respectively. Enriched but isolated tiles were not considered for further analyses.

Scatter plots and Pearson correlation analyses revealed higher correlation between data sets for histone modification polymorphic tiles among biological replicates for one accession than between datasets of different accessions (Figure S8). ChIP assays were validated by quantitative PCR for sequences known to be associated or not associated with the histone modification of interest (Figure S1 and Table S5) (Turck *et al.*, 2007; Zhang *et al.*, 2007, 2009). For subsequent analysis of ChIP on chip data for Col-0, Cvi and C24 and their intra-specific hybrids, all CGH polymorphic tiles (including singletons and those tiles that were detected only in one of two CGH technical replicates) were excluded.

Data availability and computational analyses

Raw and processed data have been deposited to the National Center for Biotechnology Information Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession GSE24836, and to CATdb (http://urgv.evry.inra.fr/CATdb) (Samson *et al.*, 2004; Gagnot *et al.*, 2008). In addition, array data and genome annotations are available for visualization at http://epigara.biologie.ens.fr/cgi-bin/gbrowse/a2e/. Gene ontology and gene family analysis were performed using the MIPS Functional Catalogue (http://mips.helmholtz-muenchen.de/proj/funcatDB/) (Ruepp *et al.*, 2004) and the DAVID tool (http://david.abcc.ncifcrf.gov/) (Huang *et al.*, 2009).

ACKNOWLEDGEMENTS

This work was supported by Deutsche Forschungsgemeinschaft grant HO 1779/7-1/2 to M.F.M. and A.H. within the framework program SPP 1149 'Heterosis in Plants'. Bioinformatics analyses by M.S. were supported by the Ministry of Culture of the State of Saxony Anhalt, Germany (XP3624HP/0606T) and the Deutscher Akademischer Austauschdienst (PROCOPE 50748812). We would like to thank O. Weiß for excellent technical assistance, and I. Schubert, M. Strickert and S. Scholten (Department of Biology, University of Hamburg) for discussions and helpful comments on the manuscript. We also thank E. Duvernois-Berthet for data visualization on Genome Browser (GB).

SUPPORTING INFORMATION

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

Figure S1. Confirmation of ChIP preparations by quantitative PCR of reference sequences.

Figure S2. Sample comparison of H3K4me2 and H3K27me3 patterns.

Figure S3. Size distribution of histone modification domains.

Figure S4. Ontology classification of histone modification polymorphic genes.

Figure S5. H3K27me3-associated genes in this study in comparison with previous analyses.

Figure S6. Ontology classification of hybridization-responsive genes.

Figure S7. Intersection analysis of genes for which H3K4me2 or H3K27me3 changed after hybridization.

Figure S8. Reproducibility of histone modification data for H3K4me2 and H3K27me3.

Table S1. H3K4me2 and H3K27me3 domains in parental accessions and F_1 hybrid offspring.

Table S2. Genes associated with H3K4me2 and H3K27me3 in parental accessions and F_1 hybrid offspring.

Table S3. TEs associated with H3K4me2 and H3K27me3 in parental accessions and F_1 hybrid offspring.

Table S4. Gene ontology and family analysis using the DAVID tool.

Table S5. Protocol for quantitative PCR of ChIP reference sequences.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

REFERENCES

Axelsson, T., Bowman, C.M., Sharpe, A.G., Lydiate, D.J. and Lagercrantz, U. (2000) Amphidiploid *Brassica juncea* contains conserved progenitor genomes. *Genome*, **43**, 679–688.

Banaei Moghaddam, A.M., Fuchs, J., Czauderna, T., Houben, A. and Mette, M.F. (2010) Intraspecific hybrids of *Arabidopsis thaliana* revealed no gross alterations in endopolyploidy, DNA methylation, histone modifications and transcript levels. *Theor. Appl. Genet.* **120**, 215–226.

Baumel, A., Ainouche, M., Kalendar, R. and Schulman, A.H. (2002) Retrotransposons and genomic stability in populations of the young allopolyploid species *Spartina anglica* C.E. Hubbard (Poaceae). *Mol. Biol. Evol.* **19**, 1218–1227.

Bernstein, B.E., Kamal, M., Lindblad-Toh, K. *et al.* (2005) Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*, **120**, 169–181.

Brubaker, C., Brown, A., Stewart, J., Kilby, M. and Grace, J. (1999) Production of fertile hybrid germplasm with diploid Australian *Gossypium* species for cotton improvement. *Euphytica*, **108**, 199–213.

Chen, Z.J. (2010) Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci.* **15**, 57–71.

Chen, L. and Chen, J. (2008) Changes of cytosine methylation induced by wide hybridization and allopolyploidy in *Cucumis. Genome*, **51**, 789–799.

Clark, R.M., Schweikert, G., Toomajian, C. *et al.* (2007) Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana. Science*, **317**, 338–342.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M. and Jacobsen, S.E. (2008) Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature*, **452**, 215–219.

Comai, L. (2000) Genetic and epigenetic interactions in allopolyploid plants. *Plant Mol. Biol.* **43**, 387–399.

Fuchs, J., Demidov, D., Houben, A. and Schubert, I. (2006) Chromosomal histone modification patterns – from conservation to diversity. *Trends Plant Sci.* **11**, 199–208.

Gagnot, S., Tamby, J.P., Martin-Magniette, M.L., Bitton, F., Taconnat, L., Balzergue, S., Aubourg, S., Renou, J.P., Lecharny, A. and Brunaud, V. (2008) CATdb: a public access to Arabidopsis transcriptome data from the URGV-CATMA platform. *Nucleic Acids Res.* **36**, D986–D990.

Gendrel, A.V., Lippman, Z., Martienssen, R. and Colot, V. (2005) Profiling histone modification patterns in plants using genomic tiling microarrays. *Nat. Methods*, **2**, 213–218.

Groszmann, M., Greaves, I.K., Albertyn, Z.I., Scofield, G.N., Peacock, W.J. and Dennis, E.S. (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.

He, G., Zhu, X., Elling, A.A. *et al.* (2010) Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell*, **22**, 17–33.

Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57.

Kashkush, K., Feldman, M. and Levy, A.A. (2003) Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat. Genet.* **33**, 102–106.

Kazazian, H.H. Jr (2004) Mobile elements: drivers of genome evolution. *Science*, **303**, 1626–1632.

Kouzarides, T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.

Lawrence, R.J., Earley, K., Pontes, O., Silva, M., Chen, Z.J., Neves, N., Viegas, W. and Pikaard, C.S. (2004) A concerted DNA methylation/histone methylation switch regulates rRNA gene dosage control and nucleolar dominance. *Mol. Cell*, **13**, 599–609.

Lee, H.S. and Chen, Z.J. (2001) Protein-coding genes are epigenetically regulated in Arabidopsis polyploids. *Proc. Natl Acad. Sci. USA*, **98**, 6753–6758.

Lippman, Z., Gendrel, A.V., Black, M. *et al.* (2004) Role of transposable elements in heterochromatin and epigenetic control. *Nature*, **430**, 471–476.

Liu, B., Brubaker, C.L., Mergeai, G., Cronn, R.C. and Wendel, J.F. (2001) Polyploid formation in cotton is not accompanied by rapid genomic changes. *Genome*, **44**, 321–330.

Liu, J., He, Y., Amasino, R. and Chen, X. (2004) siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis. *Genes Dev.* **18**, 2873–2878.

Madlung, A., Masuelli, R.W., Watson, B., Reynolds, S.H., Davison, J. and Comai, L. (2002) Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic Arabidopsis allotetraploids. *Plant Physiol.* **129**, 733–746.

Martienssen, R.A., Doerge, R.W. and Colot, V. (2005) Epigenomic mapping in Arabidopsis using tiling microarrays. *Chromosome Res.* **13**, 299–308.

Martin-Magniette, M.L., Mary-Huard, T., Bérard, C. and Robin, S. (2008) ChIPmix: mixture model of regressions for two-color ChIP-chip analysis. *Bioinformatics*, **24**, i81–i86.

Noel, L., Moores, T.L., van Der Biezen, E.A., Parniske, M., Daniels, M.J., Parker, J.E. and Jones, J.D. (1999) Pronounced intraspecific haplotype divergence at the RPP5 complex disease resistance locus of Arabidopsis. *Plant Cell*, **11**, 2099–2112.

Pokholok, D.K., Harbison, C.T., Levine, S. *et al.* (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*, **122**, 517–527.

Roudier, F., Teixeira, F.K. and Colot, V. (2009) Chromatin indexing in Arabidopsis: an epigenomic tale of tails and more. *Trends Genet.* **25**, 511–517.

Ruepp, A., Zollner, A., Maier, D. *et al.* (2004) The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res.* **32**, 5539–5545.

Samson, F., Brunaud, V., Duchêne, S., De Oliveira, Y., Caboche, M., Lecharny, A. and Aubourg, S. (2004) FLAGdb++ : a database for the functional analysis of the Arabidopsis genome. *Nucleic Acids Res.* **32**, D347–D350.

Saze, H. (2008) Epigenetic memory transmission through mitosis and meiosis in plants. *Semin. Cell Dev. Biol.* **19**, 527–536.

Schmid, K.J., Sorensen, T.R., Stracke, R., Torjek, O., Altmann, T., Mitchell-Olds, T. and Weisshaar, B. (2003) Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in *Arabidopsis thaliana*. *Genome Res.* **13**, 1250–1257.

Schubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T. and Goodrich, J. (2006) Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *EMBO J.* **25**, 4638–4649.

Seifert, M., Banaei, A., Keilwagen, J., Mette, M.F., Houben, A., Roudier, F., Colot, V., Grosse, I. and Strickert, M. (2009) Array-based genome comparison of Arabidopsis ecotypes using hidden Markov models. In *Proceedings of Biosignals: Second International Conference on Bio-inspired Systems and Signal Processing* (Encarnacao, P. and Veloso, A., eds). Setúbal: INSTICC press, pp. 3–11. http://dig.ipk-gatersleben.de/HMMs/ACGH/ACGH.html [accessed 23 April 2009].

Springer, N.M., Ying, K., Fu, Y. *et al.* (2009) Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. *PLoS Genet.* **5**, e1000734.

Teixeira, F.K. and Colot, V. (2010) Repeat elements and the Arabidopsis DNA methylation landscape. *Heredity*, **105**, 14–23.

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R.A., Coupland, G. and Colot, V. (2007) Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS Genet.* **3**, e86.

Vaughn, M.W., Tanurdzic, M., Lippman, Z. *et al.* (2007) Epigenetic natural variation in *Arabidopsis thaliana*. *PLoS Biol.* **5**, e174.

Xiong, L.Z., Xu, C.G., Saghai Maroof, M.A. and Zhang, Q. (1999) Patterns of cytosine methylation in an elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification polymorphism technique. *Mol. Gen. Genet.* **261**, 439–446.

Xu, Y., Zhong, L., Wu, X., Fang, X. and Wang, J. (2009) Rapid alterations of gene expression and cytosine methylation in newly synthesized *Brassica napus* allopolyploids. *Planta*, **229**, 471–483.

Zemach, A., McDaniel, I.E., Silva, P. and Zilberman, D. (2010) Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science*, **328**, 916–919.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y.V., Pellegrini, M., Goodrich, J. and Jacobsen, S.E. (2007) Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. *PLoS Biol.* **5**, e129.

Zhang, X., Shiu, S., Cal, A. and Borevitz, J.O. (2008) Global analysis of genetic, epigenetic and transcriptional polymorphisms in *Arabidopsis thaliana* using whole genome tiling arrays. *PLoS Genet.* **4**, e1000032.

Zhang, X., Bernatavichute, Y.V., Cokus, S., Pellegrini, M. and Jacobsen, S.E. (2009) Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol.* **10**, R62.

Zhao, Y., Yu, S., Xing, C., Fan, S. and Song, M. (2008) Analysis of DNA methylation in cotton hybrids and their parents. *Mol. Biol.* **42**, 169–178.

