*Chromosome Research* **13**: 299–308, 2005.



# Epigenomic mapping in *Arabidopsis* using tiling microarrays

Robert A. Martienssen[1]*, R. W. Doerge[2] & Vincent Colot[3]
[1]*Cold Spring Harbor Laboratory, Cold Spring Harbor NY11724, USA; E-mail: martiens@cshl.edu;* [2]*Department of Statistics, Purdue University, West Lafayette IN 47907, USA;* [3]*Unité de Recherche en Génomique Végétale (URGV), INRA/CNRS/UEVE, 2 Rue Gaston Crémieux, 91057 Evry Cedex, France*
**Correspondence*

*Key words: Arabidopsis*, epigenetics, heterochromatin, histone, methylation, tiling microarrays

## Abstract

In addition to genetic information, chromosomes transmit epigenetic information from cell to cell during division, and sometimes from generation to generation. While genetic information is encoded directly in the DNA sequence, epigenetic information is not, although it is usually associated with specific chromosomal regions. Epigenetic modifications in plants include cytosine methylation as well as modification of histones and other chromosomal proteins. Small interfering RNA play major roles in targeting these modifications to specific regions. Genomic tiling microarrays are powerful tools for analysing epigenetic information, and we review their application in building epigenomic maps in the model plant, *Arabidopsis*.

## Introduction

Epigenetic modifications are superimposed on the DNA sequence to form a 'second code' (Jenuwein & Allis 2001). This additional information is inherited from cell to cell but can be reprogrammed without changing the underlying nucleotide sequence. Such modifications allow the inheritance (or the 'memory') of gene expression (Ringrose & Paro 2004), as well as chromosomal properties such as replication, cohesion, condensation and kinetochore function (Karpen & Allshire 1997, Harvey *et al.* 2002, McNairn & Gilbert 2003). Epigenetic modifications not only play a major role in the maintenance of differentiated cells during development, but also influence gene expression more broadly. Examples include monoallelic gene expression such as imprinting (Reik & Walter 2001), dosage compensation (Gilfillan *et al.* 2004) and nucleolar dominance (Lawrence & Pikaard 2004). Epigenetic modifications can also provide environmental memory, exemplified in plants by the vernalization response, in which histone modifications are programmed in the cold, and then propagated by polycomb group and other chromatin proteins (Henderson *et al.* 2003, Amasino 2004).

The mechanism by which epigenetic modifications are aligned with the DNA sequence is not understood but, at least in some cases, it appears to involve short interfering RNA (Lippman & Martienssen 2004, Matzke *et al.* 2004, Gendrel & Colot 2005). Repetitive sequences, viruses and transgenes give rise to transcripts that are targeted by RNA interference and, at least in fission yeast, short interfering RNA guides the chromatin modification apparatus to related sequences (Volpe *et al.* 2002, Verdel *et al.* 2004, Motamedi *et al.* 2004). Non-coding RNA has been implicated in many epigenetic phenomena, including X inactivation, imprinting and dosage compensation (Sleutels & Barlow 2002, Gilfillan *et al.* 2004, Morey & Avner 2004) and is likely to play a major role in aligning genetic and epigenetic information.

Empirically, mapping epigenetic modifications is a critical step in reading this 'second code' (van Steensel & Henikoff 2003, Fazzari & Greally 2004). Cytologically, mapping is achieved using indirect immunofluorescence with antibodies specific for each modification. Modified chromosomal regions can then be aligned with the genetic map by DNA-FISH with genetically mapped probes (Fransz *et al.* 2003). With the advent of genome sequencing, it has become possible to map genome modifications at high resolution: modifications associated with single nucleosomes, and in some cases with single nucleotides, can be determined. This has led to the concept of the 'epigenome': a genome-wide map of such modifications characteristic of individual strains, cell types, or tissues, which is sufficiently reproducible and detailed to yield relevant biological information. The *Arabidopsis* genome represents an excellent system for epigenomic profiling. This is because of the comprehensive genome sequence, the relatively small amount of repetitive DNA, and the extensive collection of mutants in chromatin remodelling, histone and DNA modification (Bender 2004, Lippman & Martienssen 2004, Matzke *et al.* 2004, Tariq & Paszkowski 2004).

As an example of epigenomic profiling, we have profiled heterochromatin from *Arabidopsis* chromosome 4 (Lippman *et al.* 2004). Analysis of segmental duplications revealed that *Arabidopsis* heterochromatin is derived from euchromatin by insertion of transposable elements (TEs) and related tandem repeats. We used genomic tiling microarrays to map heterochromatic modifications, such as DNA methylation and histone H3 lysine-9 di-methylation (H3K9me2), and we found that these marks are restricted to TEs (Gendrel *et al.* 2002, Lippman *et al.* 2004). TEs are distinguished from genes by the chromatin remodelling ATPase DECREASE IN DNA METHYLATION 1 (DDM1). Small interfering RNAs (siRNAs) correspond to TEs and repeats, providing a basis for their distinction (Llave *et al.* 2002, Lippman *et al.* 2004). Genes are mostly insulated from the silencing effects of heterochromatin but TEs can control genes epigenetically when inserted within them. The euchromatic imprinted gene *FWA* resembles such genes in that its promoter is provided by a TE which contains tandem repeats associated with siRNA, and is silenced epigenetically by DDM1 and the DNA methyltransferase MET1. Thus TEs and related repeats define heterochromatin and probably play major regulatory roles in repeat-rich genomes (Lippman *et al.* 2004). Here we review the methods, applications and limitations of genomic tiling microarrays for profiling the 'epigenome' in *Arabidopsis*. We also review recent progress in using these methods as well as their implications for epigenetic mechanisms of gene regulation.

## Genomic tiling microarrays

A variety of microarray technologies are available for genome-wide profiling of epigenetic modifications (Buck & Lieb 2004, Mantripragada *et al.* 2004, Mockler & Ecker 2005). These differ in the sensitivity and specificity of the probes (the individual DNA sequences arrayed on the microarray), as well as in their coverage of the genome. Primer design can be used to avoid repetitive sequences if unambiguous signals are desired. However, some of the most important epigenetic modifications are found in repeated sequences, so that representation of at least one copy of each repeat is recommended in any design. Transposable elements (TEs) and repeats contribute to important epigenetic mechanisms, such as chromosome organization and heterochromatic silencing (Lippman *et al.* 2004).

The simplest arrays are spotted microarrays in which PCR products are amplified using sequential primers taken from the genome sequence and then printed robotically on glass slides (Bowtell & Sambrook 2004). Each PCR product must be validated by gel electrophoresis, and amplification serves as quality control to ensure that the primers are correct. Such quality control is much more difficult for spotted (synthesized) oligonucleotides. On the other hand, spotted oligonucleotides allow greater specificity, particularly in plant genomes where gene families cross-hybridize in coding regions. Oligonucleotide lengths are much shorter (50–70 nt) than PCR products (typically $>$200 bp), allowing selection from unique genic and intergenic regions. Oligonucleotides are synthesized first and then printed on glass slides so that feature density is comparable to that achieved with PCR products (26–40 000 features per slide). As a result, single slide printed oligonucleotide arrays can only offer a fragmented representation of large genomic

regions but, unlike with PCR product arrays, hybridization is strand specific. Agilent oligonucleotide arrays are printed using inkjet technology, and genomic hybridization methods have been developed for the human genome (The Arabidopsis Genome Initiative 2000).

High-density oligonucleotide probes synthesized *in situ* are now available commercially from a number of distributors. Affymetrix uses the same photolithographic technology used to manufacture electronic microchips with millions of features per slide. Sequential masks control the addition of each nucleotide, and oligos are typically short (25 nt). Although this reduces sensitivity, signals from unamplified genomic DNA targets are sufficiently strong to be detected in genomes the size of *Arabidopsis* (Borevitz *et al*. 2003). With careful optimization, it is possible to detect SNPs (single nucleotide polymorphisms) by differential hybridization to these short oligos. This has allowed massively parallel mapping for positional cloning and other applications (Borevitz *et al*. 2003). A prototype genomic tiling array has been used for transcriptional profiling in *Arabidopsis*, and comprised 3 series, offset by 8 nt, of consecutive and non-overlapping 25-nt probes that cover the genome on both strands (more than 25 million features). This required an entire wafer, or several slides per hybridization, and was prohibitively expensive but extremely effective. Early in 2005, a tiling array will be made commercially available from Affymetrix comprising one 25-nt probe per 35 bp of *Arabidopsis* genomic sequence. Nimblegen (www.nimblegen.com) uses programmable HDTV mirrors to manufacture arrays by maskless photolithography and can achieve very high densities (400 000 features per slide) of long oligonucleotides, up to 70 nt. Longer probes permit more sensitive hybridization and, in the case of printed and Nimblegen arrays, design 'on-the-fly' allows optimization of oligo sequences from one hybridization to the next (Lucito *et al*. 2003, Sebat *et al*. 2004). Agilent will soon have comparable feature densities, allowing the entire *Arabidopsis* genome to be represented by 60- or 70-nt oligos located every 200 bp. Programmable oligonucleotide arrays offer superior performance but cannot currently be used to hybridize short targets (such as siRNA) unless even higher density tiling arrays become available. Cost is also a factor and, even when expensive commercial platforms are used for data collection, home-made spotted arrays can often be very useful to test labelling protocols beforehand.

## Applications

We have developed a tiling microarray in *Arabidopsis* using printed PCR products (Lippman *et al*. 2004). Sequential 1-kb fragments from *Arabidopsis* chromosome 4 were amplified and then spotted onto glass slides. Initially, only a 1.5-Mb region of chromosome 4 was 'tiled' in this way (Lippman *et al*. 2004) but we have now printed arrays representing the whole chromosome (unpublished data). Probe selection was accomplished by BLAST analysis of sequential 100-bp windows of sequence against the whole genome to discriminate between unique and repeated DNA. However, repeats were still amplified (from BAC clones) and printed on the array in order to interrogate repetitive as well as unique genomic sequences. Since the *Arabidopsis* genome has relatively little repetitive DNA, most features are unique (Figure 1). In addition, most *Arabidopsis* repeats are low copy (Figure 1), enabling cross-hybridization to be readily resolved by PCR validation with specific primer pairs in a majority of cases. Moreover, we have found that, except in rare instances, hybridization provided an accurate measure of expression, as well as DNA and histone methylation patterns across the 1.5-Mb region represented on the array (Lippman *et al*. 2004). We will use results obtained with these arrays to illustrate a variety of applications in epigenomic profiling described below.

### *Comparative genome hybridization*

Comparative genome hybridization has been widely used to detect deletions and insertions in several genomes (Mantripragada *et al*. 2004) including *Arabidopsis* (Borevitz *et al*. 2003). We have successfully detected deletions and amplifications in *Arabidopsis* by shearing the DNA from two different strains, labelling with random primers and hybridizing to our tiling arrays (Figure 2). Because the microarrays were constructed based on the sequence of the Columbia landrace, we hybridized them with total genomic DNA from both Columbia and Landsberg. We

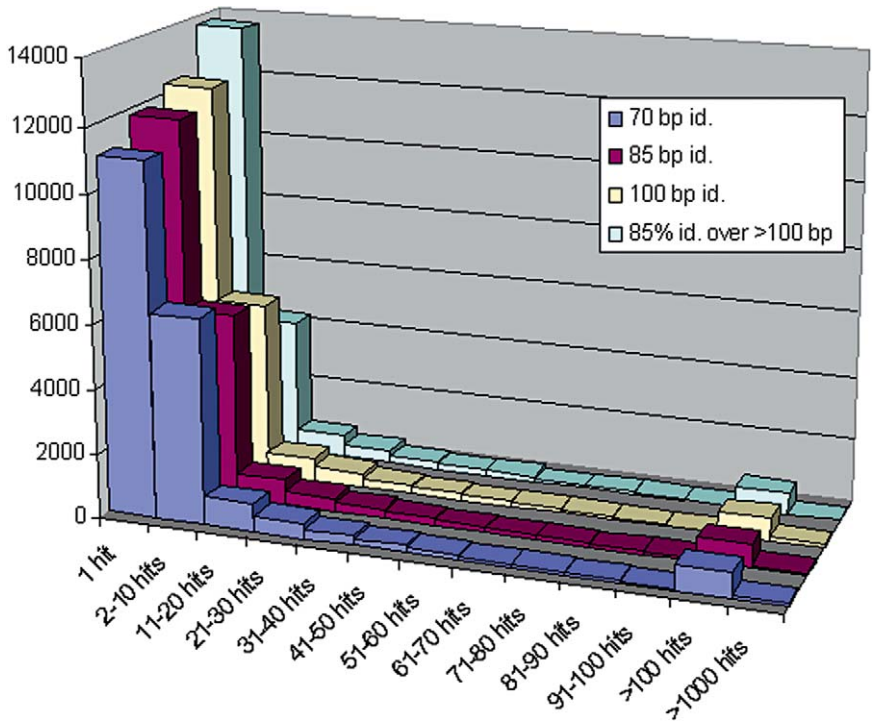
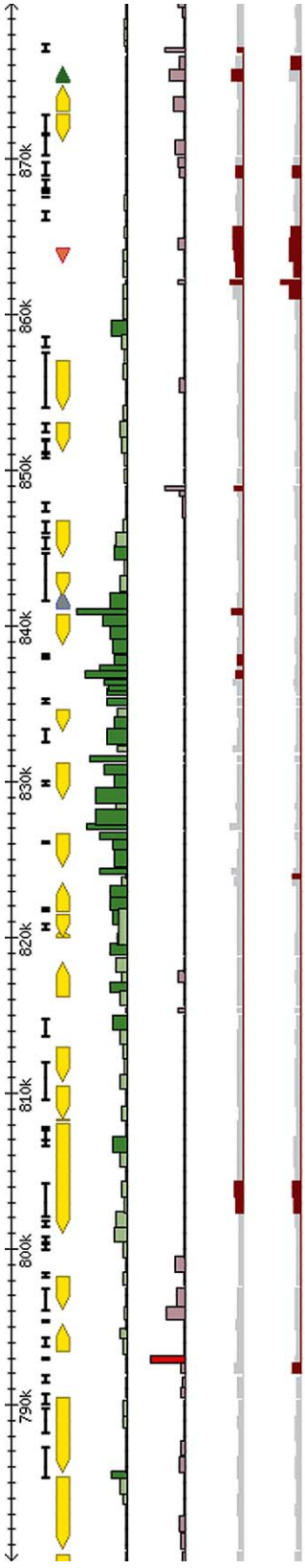



*Figure 1.* Number of BLASTN hits for the >21000 features that constitute the *Arabidopsis* chromosome 4 tiling array. Each feature was compared to the whole genome sequence using BLASTN. Over 50% of features are unique, and most repeat sequences have less than 20 hits across the genome.

chose Landsberg because shotgun reads covering approximately 70% of the genome are available in Genbank (AGI 2000). Figure 2 shows an example of a large deletion detected by comparative genome hybridization using our microarrays. Relative hybridization was calculated using a linear model and significantly depleted or enriched features were detected. Regions significantly depleted in Landsberg are highlighted in green, while those enriched in Landsberg are highlighted in red (Figure 2). In Landsberg, 9 CHP-class Zinc-finger genes are missing from this region of chromosome 4. Furthermore, this large region (26 kb) is also missing from the shotgun sequencing reads, which supports our conclusion that it has been deleted. We subjected this region to methylation profiling (Figure 2) as described below.

*DNA methylation profiling*

Several methods have been developed to dectect genomic DNA methylation on microarrays (Shi *et al.* 2003, van Steensel & Henikoff 2003, Lippman *et al.* 2004). These methods depend on either: (1) pre-digestion of the DNA with methylation sensing restriction enzymes, or (2) pretreatment with bisulphite, which converts unmethylated cytosine residues to uracil. Bisulphite treatment has the advantage that individual cytosine residues can be detected by hybridization with short oligo arrays, such as Affymetrix. In contrast, predigestion of DNA with restriction enzymes can only measure the density of methylation at any one feature. However, bisulphite-treated DNA is amplified unevenly during labelling, requiring primers specific for a given region in many cases. Furthermore, when treated molecules are cloned and sequenced, methylation in many regions of the genome are found to be highly variable, so that each cytosine residue might only be methylated in 50% of the individual molecules. In extreme cases, this methylation would not be detected in genomic DNA. By constrast, enzymatic digestion still occurs regardless of which cytosines are methylated, resulting in strong differential signals.




We use an enzymatic methodology for methylation detection (Lippman *et al.* 2004). In our method, genomic DNA from *Arabidopsis* was sheared to a constant size and then divided into two equal portions. One portion was digested with McrBC and then both portions were size-fractionated on agarose gels. DNA greater than 1 kb was purified, amplified and labelled with different fluorochromes for hybridization on microarrays (Figure 3). The digested sample is depleted of methylated DNA sequences so that the ratio of hybridization intensities between digested and undigested samples is a measure of DNA methylation. These ratios are calculated, taking all sources of variation into account, and plotted onto the genome using Genome Browser (Stein *et al.* 2002). Those ratios that are deemed to be significantly greater than the average euchromatic value are highlighted. An example of a region of methylated DNA from chromosome 4 is shown in Figure 3. Almost all of the methylation is restricted to transposable elements but occasionally genes are also methylated at the 3′ end (Figure 3). Methylation of genes differs genetically from methylation of TEs in that TE methylation depends very strongly on the chromatin remodelling gene *DECREASE IN DNA METHYLATION 1*, but methylation of genes does not. Methylation of both classes of sequences depends on the DNA methyltransferase MET1 (Lippman *et al.* 2004).

### *Chromatin immunoprecipitation*

Association of a DNA sequence with a given protein can be determined by chromatin immunoprecipitation (ChIP). Chromatin is first cross-linked to chromosomal DNA by fixing cells and tissues and then sheared to a small size (300–1000 bp) and immunoprecipitated using

*Figure 2.* Comparative genome hybridization (CGH) in *Arabidopsis*. A 100-kb region from the short arm of chromosome 4 is shown. Genes (yellow), DNA transposons (red) and LTR retrotransposons (green) are annotated as ORFs, along with shotgun sequencing reads from Landsberg (⊢). Deletions from Ler were detected as significantly increased CGH ratios (green) while reduced ratios (pink) indicate increases in copy number. DNA methylation of Columbia and Landsberg, respectively, is shown in the last two tracks (brown). The deleted region from Landsberg (green) corresponds to missing sequences in the Cereon Ler shotgun reads. The fourth annotated gene has a methylated exon near the 3′ end.

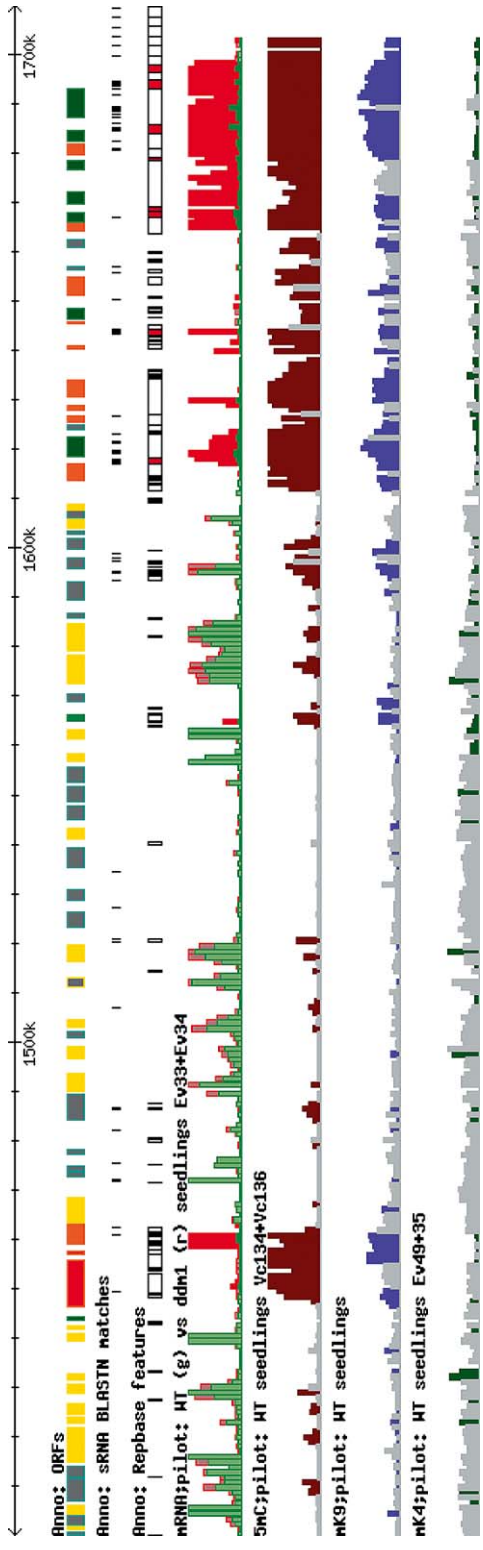



*Figure 3.* Epigenomic profiling of *Arabidopsis* chromosome 4. Epigenomic microarray profiling data from a 300-kb region from the short arm of chromosome 4 is displayed using Genome Browser. Genes (yellow), DNA transposons (red) and LTR retrotransposons (green) are annotated as Open reading frames, while repeats from Repbase are indicated in a separate track. Matches to a database of small RNA sequences are indicated also. Hybridization of mRNA from WT (green) and ddm1 mutant (red) is shown (top track), along with DNA methylation (brown), H3K9me2 (blue) and H3K4me2 (green) profiling. In each case, features that have significant levels of methylation relative to euchromatic sequences are highlighted. Euchromatic levels are indicated in grey.

antibodies raised against chromosomal proteins. After reversing the crosslink, bound DNA is eluted, amplified and labelled for hybridization. Sheared chromatin (before immunoprecipitation) is amplified and labelled as a control. 'ChIP chip' has been widely used in conjunction with genomic microarrays to profile histone modifications and other chromatin proteins along chromosomal DNA. Examples from yeast and from mammalian cells have demonstrated the resolution and power of this technique (Horak & Snyder 2002, Kurdistani & Grunstein 2003, Buck & Lieb 2004).

Thus far, we have profiled two histone modifications in *Arabidopsis*, namely histone H3 lysine-9 dimethylation (H3K9me2) and histone H3 lysine-4 dimethylation (H3K4me2), which are associated with heterochromatin and euchromatin, respectively (Jenuwein & Allis 2001). Examples of profiles are shown in Figure 3. We found that the vast majority of H3K9me2 was associated with TEs while most of the H3K4me2 was associated with genes. This distinction was erased in the chromatin remodelling mutant *ddm1* (Gendrel *et al*. 2002, Lippman et al. 2004). By displaying the data in a browser format, we were able to align regions of the chromosome that are associated with specific histone modifications, such as matches to small RNA, with sequence annotations and with each other. We found a close correlation between small RNA matches, H3K9me2 and DNA methylation. This was further investigated using unsupervised hierarchical clustering and found to be focused on TEs (Lippman *et al*. 2004). The interdependence of these modifications has been demonstrated genetically (Bender 2004) and are mechanistically linked in *Arabidopsis* as they are in fission yeast and *Drosophila*, which have little or no DNA methyation (Lippman & Martienssen 2004, Matzke & Birchler 2005).

## Statistical analysis and data display

Assessing the significance of epigenomic profiling data presents a number of challenges. Epigenomic profiling involves substantial experimental variation, and controls need to be carefully assigned prior to normalization. For example, some histone modifications are not found in certain chromosomal regions at all, and therefore should not be used as controls since they would unfairly bias the findings. We have found the use of a linear model (Craig *et al*. 2003) to estimate experimental effects, determine the level of variability, and detect features undergoing statistically significant changes in fluorescence intensity between the conditions investigated by each experiment to be extremely powerful. The linear model procedure partitions the sources of variation such that global and feature-specific array and dye (technical) effects are removed, creating a corrected signal for each feature. This signal is then used to assess changes in fluorescence intensity for each feature solely due to differences between the conditions of interest. In the case of DNA methylation the average ratio of hybridization intensities for unmethylated DNA and total DNA is calculated for each feature in each dye-swap experiment (see below). Unmethylated features give a ratio of close to 1.0, whereas methylated features give a ratio greater than 1.0. Similarly, for chromatin immunoprecipitation, the ratio of immunoselected and total DNA is calculated for each feature. In this case, the ratio found in euchromatin is arbitrarily set to 1.0 in both mutant and wild-type, as recovery of immunoselected DNA is always much less than 100%. This is the equivalent of using euchromatic genes such as actin as a control. Statistical tests are used to detect features undergoing significant changes, based on control of both the family-wise error rate (FWER) and the false discovery rate (FDR). This statistical analysis allows widely differing profiles to be compared such that the resulting data are highly reproducible. In one study that included biological replicates, there was a very good overlap (>90%) in the results obtained from the two comparisons (Lippman *et al*. 2004).

Data from dye-swap experiments were analysed using an analysis of variance (ANOVA) model (Kerr *et al*. 2000). The following linear model was employed to partition the sources of variation:

$$\mathrm{Y_{ijklm}} = \mu + \mathrm{A_i} + \mathrm{D_j} + \mathrm{T_k} + \mathrm{G_l} + \mathrm{AG_{il}} + \mathrm{DG_{jl}} + \mathrm{TG_{kl}} + \epsilon_{\mathrm{ijklm}}$$

where Y is the overall mean log-intensity, A, D, T and G represent the array, dye, treatment and gene main effects, AG, DG and TG are respectively array by gene, dye by gene and treatment by gene interaction terms, and $\epsilon$ is random error.

Observations with a background-corrected intensity of less than or equal to zero were set to be 1. The data were transformed using the natural logarithm prior to producing the $Y_{ijklm}$ values used in the analysis. For each of the features represented on the array, the hypotheses (Black & Doerge 2002) tested were:

$$H_0 : (T_1 + TG_{1g}) - (T_2 + TG_{2g}) = 0$$
$$H_1 : (T_1 + TG_{1g}) - (T_2 + TG_{2g}) \neq 0$$

with rejection of the null hypothesis indicating a statistically significant change in fluorescence intensity. To accommodate the large number of hypothesis tests being made and to provide some level of error rate control, significance was assessed using both a family-wise error rate (FWER) and a false discovery rate (FDR). A step-down multiple comparisons procedure (Holm 1979) was used to control the FWER below $\alpha = 0.01$, while a step-up procedure (Benjamini & Hochberg 1995) was used to control the FDR below $\alpha = 0.01$. For the purposes of these experiments, the hypotheses were assumed to be independent.

As outlined above, tiling arrays can be used to measure different epigenetic characteristics and one obvious challenge is how different measurements can be combined and displayed. For this purpose, a custom MySQL database (http://www.mysql.com) was designed to import normalized array data, which are then displayed using an adapted version of Generic Genome Browser (Stein *et al*. 2002). Additional features, such as siRNA sequence homologies and locations of genetrap transposon insertions (http://genetrap.cshl.org) were added to the *Arabidopsis* genome annotation by stringent BLASTN (Altschul *et al*. 1997) against complete genome sequence. Repeat and TE sequences were annotated by CENSOR analysis (Jurka *et al*. 1996) with version 8.4 of RepBase (Jurka 2000) using normal stringency. By aligning data from a variety of profiling experiments in this way, inferences based on quantitative and qualitative assessments can be readily drawn by simple computation or even visual inspection.

## Future directions

Genomic tiling arrays have been used for a wide variety of additional applications. Small RNA profiling is a recent example. Small RNA can be ligated and amplified using fluorescently labelled primers, and hybridized to microarrays (Miska *et al*. 2004). While the purity of small RNA preparations and background hybridization remains a problem, small RNA profiling in *Arabidopsis* would be very informative because the majority of small RNA are siRNA, derived from relatively long repetitive sequences, rather than miRNA derived from individual precursor genes. The profile of small RNAs from repetitive sequences for example, differs from wild-type in mutants in RNA interference (Xie *et al*. 2004) as well as in mutants in DNA and histone modification where it appears to affect inheritance of transposon activation, indicating association with the chromosome (Lippman *et al*. 2003).

DNA replication can also be detected using tiling microarrays. Synchronized *Drosophila* cells, for example, can be pulse-labelled with BrdU to detect replicated DNA. BrdU-labelled DNA can then be immunoprecipitated with specific antibodies and hybridized to genomic tiling micoarrays (MacAlpine *et al*. 2004). By labelling at different time-points following cell-cycle arrest and release, it is possible to track early and late origins of replication. These can be correlated with other epigenetic modifications, such as histone modification and DNA methylation, not to mention binding of DNA replication complexes via ChIP (MacAlpine *et al*. 2004).

Finally, heritability will be a key issue in profiling the *Arabidopsis* epigenome. Epigenetic modifications are often inherited in plants leading to natural variation and a high frequency of 'epimutations': heritable changes in gene expression that are not due to changes in DNA sequence (Kakutani *et al*. 1999, Stokes *et al*. 2002). While the prevalence of epimutations is unknown, they can be maintained over hundreds of generations and may play an important role in somaclonal variation, as well as population genetics and evolution (Riddle & Richards 2002). In *Arabidopsis*, it is possible to assess the heritability of individual epigenetic modifications by identifying polymorphisms and then following their segregation in mapping populations. Heritability *in cis* indicates a stable epigenetic modification, while heritability *in trans* might identify candidates for regulatory loci (Riddle & Richards 2002).

## Acknowledgements

We thank members of our laboratories, especially Zach Lippman, Matt Vaughn, Mik Black, Hongmei Jiang, Anne-Valérie Gendrel, Séverine Gagnot and Sébastien Aubourg for sharing their data and for comments on the manuscript. This work was supported by grants from the National Institutes of Health and the National Science Foundation (to RM and RWD) and by grants from Genopole and the CNRS (to VC).

## References

Altschul SF, Madden TL, Schaffer AA *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Amasino R (2004) Vernalization, competence, and the epigenetic memory of winter. *Plant Cell* **16**: 2553–2559.

Bender J (2004) Chromatin-based silencing mechanisms. *Curr Opin Plant Biol* **7**: 521–526.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B* **57**: 289–300.

Black MA, Doerge RW (2002) Calculation of the minimum number of replicate spots required for detection of significant gene expression fold change in microarray experiments. *Bioinformatics* **18**: 1609–1616.

Borevitz JO, Liang D, Plouffe D *et al.* (2003) Large-scale identification of single-feature polymorphisms in complex genomes. *Genome Res* **13**: 513–523.

Bowtell D, Sambrook J (2004) *Microarrays*. Cold Spring Harbor Press.

Buck MJ, Lieb JD (2004) ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics* **83**: 349–360.

Craig BA, Black MA, Doerge RW (2003) Gene expression data: the technology and statistical analysis. *J Agricult Biol Environ Stat* **8**: 1–28.

Fazzari MJ, Greally JM (2004) Epigenomics: beyond CpG islands. *Nat Rev Genet* **5**: 446–455.

Fransz P, Soppe W, Schubert I (2003) Heterochromatin in interphase nuclei of *Arabidopsis thaliana*. *Chromosome Res* **11**: 227–240.

Gendrel AV, Colot V (2005) *Arabidopsis* epigenetics: when RNA meets chromatin. *Curr Opin Plant Biol* (in press).

Gendrel AV, Lippman Z, Yordan C *et al.* (2002) Dependence of heterochromatic histone H3 methylation patterns on the *Arabidopsis* gene *DDM1*. *Science* **297**: 1871–1873.

Gilfillan GD, Dahlsveen IK, Becker PB (2004) Lifting a chromosome: dosage compensation in *Drosophila melanogaster*. *FEBS Lett* **567**: 8–14.

Harvey SH, Krien MJ, O'Connell MJ (2002) Structural maintenance of chromosomes (SMC) proteins, a family of conserved ATPases. *Genome Biol* **3**: REVIEWS3003.

Henderson IR, Shindo C, Dean C (2003) The need for winter in the switch to flowering. *Annu Rev Genet* **37**: 371–392.

Holm S (1979) A simple sequentially rejective multiple test procedure. *Scand J Stat* **6**: 65–70.

Horak CE, Snyder M (2002) ChIP-chip: a genomic approach for identifying transcription factor binding sites. *Meth Enzymol* **350**: 469–483.

Jenuwein T, Allis CD (2001) Translating the histone code. *Science* **293**: 1074–1080.

Jurka J (2000) Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet* **16**: 418–420.

Jurka J, Klonowski P, Dagman V *et al.* (1996) CENSOR—a program for identification and elimination of repetitive elements from DNA sequences. *Comput Chem* **20**: 119–121.

Kakutani T, Munakata K, Richards EJ *et al.* (1999) Meiotically and mitotically stable inheritance of DNA hypomethylation induced by ddm1 mutation of *Arabidopsis thaliana*. *Genetics* **151**: 831–838.

Karpen GH, Allshire RC (1997) The case for epigenetic effects on centromere identity and function. *Trends Genet* **13**: 489–496.

Kerr MK, Martin M, Churchill GA (2000) Analysis of variance for gene expression microarray data. *J Comput Biol* **7**: 819–837.

Kurdistani SK, Grunstein M (2003) *In vivo* protein–protein and protein–DNA crosslinking for genomewide binding microarray. *Methods* **31**: 90–95.

Lawrence RJ, Pikaard CS (2004) Chromatin turn ons and turn offs of ribosomal RNA genes. *Cell Cycle* **3**: 880–883.

Lippman Z, Martienssen R (2004) The role of RNA interference in heterochromatic silencing. *Nature* **431**: 364–370.

Lippman Z, May B, Yordan C *et al.* (2003) Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. *PLoS Biol* **1**: E67.

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

Llave C, Kasschau KD, Rector MA *et al.* (2002) Endogenous and silencing-associated small RNAs in plants. *Plant Cell* **14**: 1605–1619.

Lucito R, Healy J, Alexander J *et al.* (2003) Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. *Genome Res* **13**: 2291–2305.

MacAlpine DM, Rodriguez HK, Bell SP (2004) Coordination of replication and transcription along a *Drosophila* chromosome. *Genes Dev* **18**: 3094–3105.

Mantripragada KK, Buckley PG, de Stahl TD *et al.* (2004) Genomic microarrays in the spotlight. *Trends Genet* **20**: 87–94.

Matzke MA, Birchler JA (2005) RNAi-mediated pathways in the nucleus. *Nat Rev Genet* **6**: 24–35.

Matzke M, Aufsatz W, Kanno T *et al.* (2004) Genetic analysis of RNA-mediated transcriptional gene silencing. *Biochim Biophys Acta* **1677**: 129–141.

McNairn AJ, Gilbert DM (2003) Epigenomic replication: linking epigenetics to DNA replication. *Bioessays* **25**: 647–656.

Miska EA, Alvarez-Saavedra E, Townsend M *et al.* (2004) Microarray analysis of microRNA expression in the developing mammalian brain. *Genome Biol* **5**: R68.

Mockler TC, Ecker JR (2005) Applications of DNA tiling arrays for whole-genome analysis. *Genomics* **85**: 1–15.

Morey C, Avner P (2004) Employment opportunities for non-coding RNAs. *FEBS Lett* **567**: 27–34.

Motamedi MR, Verdel A, Colmenares SU *et al.* (2004) Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119**: 789–802.

Reik W, Walter J (2001) Genomic imprinting: parental influence on the genome. *Nat Rev Genet* **2**: 21–32.

Riddle NC, Richards EJ (2002) The control of natural variation in cytosine methylation in *Arabidopsis*. *Genetics* **162**: 355–363.

Ringrose L, Paro R (2004) Epigenetic regulation of cellular memory by the polycomb and trithorax group proteins. *Annu Rev Genet* **38**: 413–443.

Sebat J, Lakshmi B, Troge J *et al.* (2004) Large-scale copy number polymorphism in the human genome. *Science* **305**: 525–528.

Shi H, Wei SH, Leu YW *et al.* (2003) Triple analysis of the cancer epigenome: an integrated microarray system for assessing gene expression, DNA methylation, and histone acetylation. *Cancer Res* **63**: 2164–2171.

Sleutels F, Barlow DP (2002) The origins of genomic imprinting in mammals. *Adv Genet* **46**: 119–163.

Stein LD, Mungall C, Shu S *et al.* (2002) The generic genome browser: a building block for a model organism system database. *Genome Res* **12**: 1599–1610.

Stokes TL, Kunkel BN, Richards EJ (2002) Epigenetic variation in *Arabidopsis* disease resistance. *Genes Dev* **16**: 171–182.

Tariq M, Paszkowski J (2004) DNA and histone methylation in plants. *Trends Genet* **20**: 244–251.

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 796–815.

van Steensel B, Henikoff S (2003) Epigenomic profiling using microarrays. *Biotechniques* **35**: 346–350, 352–354, 356–357.

Verdel A, Jia S, Gerber S *et al.* (2004) RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**: 672–676.

Volpe TA, Kidner C, Hall IM *et al.* (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**: 1833–1837.

Xie Z, Johansen LK, Gustafson AM *et al.* (2004) Genetic and funtional diversification of small RNA pathways in plants. *PLoS Biol* **2**: 104.