Published in final edited form as:

Curr Opin Plant Biol. 2011 April ; 14(2): 130–136. doi:10.1016/j.pbi.2011.03.010.

High-throughput approaches for plant epigenomic studies

Robert J Schmitz¹ and Xiaoyu Zhang²

¹Plant Biology and Genomic Analysis Laboratories, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

²Department of Plant Biology, University of Georgia, Athens, GA 30602, USA

Abstract

In plant cells, DNA is packaged into chromatin by wrapping around histone octamers. Pathways that lead to cytosine DNA methylation, posttranslational histone modifications and certain components of the RNA interfering (RNAi) pathway are critically important in modulating chromatin structure, thereby affecting many molecular processes that take place in a cell. Recent advances in microarray and high-throughput sequencing technologies have made it possible to study these pathways on a genome-wide scale. Results from such epigenomic studies are broadening our understanding of plant genomes and are also providing important clues regarding the mechanisms and functions of these pathways that can be further tested using genetic and biochemical approaches. This review focuses on the high-throughput approaches that have been successfully applied in plant epigenomic studies.

Introduction

DNA methylation, histone modifications and the siRNA pathway are responsible for modulating the chromatin environment in plant cells and, as a result, affect a myriad of biological processes. Numerous studies during the past decade have generated a vast knowledgebase of information regarding the components of each pathway that lead to these epigenomic modifications, the interactions between these pathways and their biological functions. It has also become clear that these pathways frequently target thousands of loci in plant genomes. In addition, loss-of-function mutations of genes in these pathways often exhibit locus-specific and sometimes opposite defects [1]. A number of recent studies have combined genomic approaches such as microarrays and high-throughput sequencing with biochemical methods previously used in locus-specific studies to better understand the mechanisms and functions of these pathways on a genome-wide scale. The combination of these methods has led to a new area of research, referred to as ‘epigenomics’. Several epigenomic approaches, such as those used for genome-wide profiling of small RNAs (sRNAs) and DNA methylation, were first developed in plants [2,3•,4•]. These approaches not only revolutionized plant epigenomic research, but also assisted in the development of similar approaches in animal and fungal systems [3•,5•]. The goal of this article is to provide an overview of these epigenomic approaches and their applications. The mechanistic and functional aspects of plant epigenetics are topics of other articles within this issue.

Methods for genome-wide profiling of DNA methylation

Three types of methods have been used to detect DNA methylation. The first involves digestion of genomic DNA by methylation-sensitive restriction enzymes such as *Msp* I (digestion blocked by CHG methylation; H = A, T or C) and *Hpa* II (blocked by CG and CHG methylation) [6]. This method was initially used to identify differentially methylated sites in wild-type and mutant *Arabidopsis* plants by microarray analysis of small DNA fragments (*i.e.* between unmethylated restriction sites) following digestion [7,8]. It is also possible to analyze a single DNA sample by comparing digestion efficiency by isoschizomer pairs with one enzyme being methylation-sensitive and the other insensitive. The methylation-specific enzymes McrBC, which preferentially digests methylated DNA, are also used to detect DNA methylation. Following McrBC digestion, longer DNA fragments mostly represent unmethylated DNA, which can be identified by microarray hybridization [9]. The effectiveness of these methods is limited by the intrinsic properties of the enzymes, such as the requirement for the presence of specific sites within suitable distances. Nevertheless, these methods led to several important findings in earlier studies, such as the differential methylation status of transposons, genes and intergenic sequences, as well as the unexpected presence of CG-methylation in the transcribed regions of plant genes with a pronounced 3' bias [7,9].

Second, methylated DNA fragments can be isolated by affinity purification using proteins that preferentially bind methylated DNA, or by immunoprecipitation using anti-³H antibodies (³H-CIP). The most widely used ³H-binding protein is the MBD domain of the human protein MeCP2, which binds methylated DNA but not unmethylated DNA at high salt concentrations [10]. It should be noted that MeCP2 only recognizes CG-methylated DNA, and its DNA binding affinity appears to be positively correlated with the number of methylated CG sites in a DNA fragment [11]. ³H-CIP does not appear to have these limitations, as a monoclonal antibody has been developed that recognizes ³H in all sequence contexts. Either microarray analysis or high-throughput sequencing can be used to reveal methylated DNA fragments isolated by ³H-CIP [3•,11,12]. This method was successfully applied to *Arabidopsis*, producing the first genome-wide maps of methylated regions at a resolution of a few dozen base pairs. However, in most cases, neither affinity purification nor ³H-CIP allows the determination of which particular cytosines within a genomic region are methylated.

The third method involves treating genomic DNA with sodium bisulfite, a process that converts unmethylated cytosines (but not methylated cytosines) to uracils. Following PCR amplification and sequencing of bisulfite-treated DNA, methylated cytosines can be identified as those that remain as cytosines whereas unmethylated cytosines appear as thymines [13,14]. This is the gold-standard method for detecting DNA methylation at specific loci. Bisulfite-treated DNA has also been analyzed by microarray hybridization, as methylated DNA retains the original sequence and therefore the ability to hybridize to the microarray [15]. However, the full potential of this method was only realized recently when bisulfite-treated DNA was subjected to shotgun sequencing (MethylC-Seq), revealing single-base resolution of cytosine methylation [3•,4•,16,17•,18•]. Compared to the other methods described above, MethylC-Seq offers two important advantages. First, it allows the detection of genome-wide DNA methylation patterns at single-nucleotide levels in all sequence contexts. This is particularly important in plants where three distinct DNA methyltransferases preferentially methylate cytosines in three different sequence contexts (*i.e.* MET1 at CG sites, CMT3 at CHG sites and DRM2 at CHH sites). MethylC-Seq results can therefore assist in the identification of the 'sites of action' of each individual methyltransferase, thus providing important information to address questions regarding the mechanisms and functions of DNA methylation. Second, deep sequencing allows the

quantitative measurement of the percentages of DNA fragments that are methylated at a particular cytosine. This effectively reflects the relative fractions of cells that contain DNA methylation at that cytosine, thus providing valuable information regarding cell-specific methylation that could function in regulating differential gene expression.

Methods for profiling histone modifications

Chromatin immunoprecipitation (ChIP) is the most widely used method for profiling histone modifications, as well as the binding sites of proteins that recognize modified histones or specific DNA sequences [9,19–27]. Plant chromatin can be fragmented by sonication after crosslinking, or digested using enzymes such as DNase I or micrococcal nuclease. Chromatin fragments containing certain histone modifications are isolated by ChIP using corresponding antibodies, and their genomic locations can then be determined by microarray analysis or deep sequencing of associated DNA (ChIP-chip and ChIP-Seq, respectively) [9,19–27]. Because histones are among the most conserved proteins in eukaryotes, commercially available antibodies developed for animals and fungi can also be used to study histone modifications in plants. The success of ChIP-chip or ChIP-Seq relies heavily on the antibodies used, and it is therefore desirable to assess antibody specificities in early stages of these experiments [28]. It should also be noted that a growing body of evidence has revealed the presence of histone modifications on very closely spaced amino acid residues, which are often functionally related and regulated developmentally or in a cell-cycle-specific manner [29]. The development of new antibodies that allow the detection of such combinatorial modifications will likely reveal previously unknown interactions in future epigenomic studies.

Methods for profiling RNAs

It is often of interest to generate gene expression profiles in epigenomic studies, such that the relationship and correlated changes between epigenomic states and transcriptional activities can be identified in a genome-wide manner. Detailed characterization of the cellular RNA pools has also uncovered novel intergenic transcripts, noncoding RNAs (ncRNAs), alternative splicing isoforms and antisense transcripts, thereby providing new opportunities for discovery-based research [4•,24,27,30–38]. In addition to established microarray-based methods, high-throughput RNA sequencing (RNA-Seq) is being used in an increasing number of studies (Figure 1). Two major types of libraries can be constructed for RNA-Seq: cDNA libraries and strand-specific RNA libraries. The advantage of cDNA library sequencing is the cost of library preparation and the ability to construct a library with limited starting material, and it is now possible to profile the transcriptome of a single cell [39]. The advantage of strand-specific RNA-Seq (also known as directional RNA-Seq) is the ability to determine the strandedness of the RNA, which allows the identification of novel ncRNAs and antisense transcripts [4•].

Plant cellular RNA pools also include several types of sRNAs, including small interfering RNAs (siRNAs), microRNAs (miRNAs), trans-acting small interfering RNAs (tasiRNAs) and natural antisense small interfering RNAs (nat-siRNAs). These sRNAs are highly abundant and play important roles in many biological processes, such as regulating gene expression, heterochromatic silencing, and antiviral defense. A major goal in sRNA studies is the detailed characterization of the sRNA populations accumulated in specific cells or tissues. The source and targets of sRNAs can then be identified to address questions regarding the mechanisms and functions of the sRNA pathways. There are currently two primary methods for isolation of sRNAs. First, size selection from total cellular RNA is the most popular method and has been extensively utilized to catalog sRNAs in different tissue types, mutants and accessions or subspecies [2,24,40,41]. Second, immunoprecipitation of

sRNA-binding proteins [42–44] has been employed, but it requires a few additional features, including a protein that strongly associates with sRNAs (such as ARGONAUTE proteins) and the ability to immunoprecipitate the protein of interest. Following ligation to RNA adaptors and reverse transcription, millions of sRNAs sequences can be revealed by RNA-Seq. In addition to the sequences and relative abundances, sequencing results also provide length and information about strand of origin, which can be used to infer the proteins that might be involved in generating and processing sRNA precursors.

Microarray and high-throughput sequencing platforms

Microarray platforms designed to interrogate gene expression levels contain features that either represent the entire transcripts (*e.g.* cDNA spotted arrays) or are synthesized to match the 3' portions of transcribed regions (*e.g.* oligo expression arrays). The effectiveness of these microarrays in epigenomic studies is limited as they only cover genomic regions that are annotated as being transcribed. In contrast, tiling microarrays contain features that are relatively evenly spaced to cover either a large genomic interval or the entire genome in an unbiased manner [11,12,45–47]. High-density microarray platforms have been developed to tile the entire genomes of *Arabidopsis* (on a single array) and rice (on a three-array set) at 35-bp resolution using 25-nt features. The high coverage and resolution, coupled with relatively low cost and the ease of data analysis, makes these tiling microarray platforms fairly attractive for epigenomic studies.

The first high-throughput sequencing method, massively parallel signature sequencing (MPSS), was developed over a decade ago [48]. Three distinct sequencing platforms have since been developed that offer longer reads and higher throughput. These include the 454 Genome Analyzer FLX (Roche), the HiSeq (Illumina) and the 5500xl SOLiD System (Life Technologies) (Table 1). Each sequencing platform has its own unique advantages, depending on the specific scientific inquiry under investigation. There are two key features that determine which sequencing platform is best suited for each experiment: the total number of sequenced reads output and the length of sequenced reads. In general, the 454 Genome Analyzer FLX sequencer is currently best suited for applications involving *de novo* genome and transcriptome assemblies because it produces longer reads. In contrast, HiSeq and SOLiD produce more reads and are therefore better suited for quantitative applications when reference genome sequences are available, such as gene expression and sRNA analysis, ChIP-Seq and DNA methylome analysis. It should be noted that even though each platform is best suited for specific experiments, anyone could be used for all the applications mentioned above (see Figure 2 for examples of datasets generated by high-throughput sequencing methods).

Both microarray-based and high-throughput sequencing methods have been successfully used in epigenomic studies, and the results are largely consistent. High-throughput sequencing offers more quantitative results at higher resolution, and is not confounded by several problems associated with microarray-based methods (such as cross-hybridization of similar sequences). With the continuing decline in cost, the increase in throughput and the maturation of analysis and data management tools, high-throughput sequencing has become more attractive for epigenomic studies.

Plant 'epigenomes'

The collective results from published epigenomic studies suggest that plant genomes are highly compartmentalized, with different components of the genome associated with distinct combinations of epigenomic modifications [49]. That is, transposable elements and other repetitive sequences are highly DNA-methylated in all three sequence contexts and are also targeted by the siRNA and histone lysine 9 dimethylation (H3K9me2) pathways [2,3•,4•],

9,21,50]. Developmentally repressed genes are enriched in H3K27me3 and bound by LIKE HETEROCHROMATIN PROTEIN 1 (LHP1) [19,22,23,51], whereas actively transcribed genes are enriched in a different set of histone modifications including acetylation, H3K4me, H3K36me and H3K9me3 [9,20,22,23,25], and many moderately to highly transcribed genes also contain heavy gene-body DNA methylation but only in the CG context [3•,4•,7,11,12]. In other words, constitutively silenced, developmentally repressed and actively transcribed regions are associated with largely nonoverlapping sets of modifications, indicating that these modifications likely play important roles in modulating transcriptional activities. Consistent with this notion, nonrepetitive intergenic regions that are transcriptionally inactive are generally devoid of these modifications. In addition to transcription, tantalizing new results have also implicated H3K27me1 in the activities of origins of DNA replication and associated DNA methylation with nucleosome positioning and histone variants [52–54]. It is conceivable that some of the pathways described above may also function in other molecular events that are affected by the local chromatin environment, such as recombination or DNA repair. The application of epigenomic approaches in combination with other methods, particularly those that allow cell-type-specific studies [55•], will undoubtedly lead to novel and exciting discoveries in the near future.

Acknowledgments

We apologize to researchers whose work could not be cited here due to space limitations. R.J.S. was supported by an NIH NRSA postdoctoral fellowship (F32-HG004830). X.Z. is supported by National Science Foundation grant 0960425.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Saze H, Kakutani T. Heritable epigenetic mutation of a transposon-flanked Arabidopsis gene due to lack of the chromatin-remodeling factor DDM1. *EMBO J*. 2007; 26:3641–3652. [PubMed: 17627280]
2. Lu C, Tej SS, Luo S, Haudenschild CD, Meyers BC, Green PJ. Elucidation of the small RNA component of the transcriptome. *Science*. 2005; 309:1567–1569. [PubMed: 16141074]
3. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature*. 2008; 452:215–219. [PubMed: 18278030] This study describes the use of high-throughput sequencing of bisulfite treated DNA from wild-type *Arabidopsis* and the DNA methyltransferase mutants *met1*, *cmt3* and *drm1/drm2*. The results represent the genome-wide DNA methylation map at single cytosine resolution. A proof of principle study was also performed using DNA from mouse germ-cell tissue.
4. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell*. 2008; 133:523–536. [PubMed: 18423832] This study describes the use of high-throughput sequencing method to study not only DNA methylation, but also small RNAs, mRNAs, ncRNAs and antisense transcripts in wild type and mutant *Arabidopsis* plants. The results represent a comprehensive analysis of the epigenome and transcriptome of *Arabidopsis*.
5. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, et al. Human DNA methylomes at base resolution show widespread epigenomic

- differences. *Nature*. 2009; 462:315–322. [PubMed: 19829295] This study describes the first DNA methylation map of the human genome at single cytosine resolution.
6. McClelland M, Nelson M. The effect of site-specific DNA methylation on restriction endonucleases and DNA modification methyltransferases—a review. *Gene*. 1988; 74:291–304. [PubMed: 2854811]
 7. Tran RK, Henikoff JG, Zilberman D, Ditt RF, Jacobsen SE, Henikoff S. DNA methylation profiling identifies CG methylation clusters in *Arabidopsis* genes. *Curr Biol*. 2005; 15:154–159. [PubMed: 15668172]
 8. Tran RK, Zilberman D, de Bustos C, Ditt RF, Henikoff JG, Lindroth AM, Delrow J, Boyle T, Kwong S, Bryson TD, et al. Chromatin and siRNA pathways cooperate to maintain DNA methylation of small transposable elements in *Arabidopsis*. *Genome Biol*. 2005; 6:R90. [PubMed: 16277745]
 9. Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD, et al. Role of transposable elements in heterochromatin and epigenetic control. *Nature*. 2004; 430:471–476. [PubMed: 15269773]
 10. Cross SH, Charlton JA, Nan X, Bird AP. Purification of CpG islands using a methylated DNA binding column. *Nat Genet*. 1994; 6:236–244. [PubMed: 8012384]
 11. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell*. 2006; 126:1189–1201. [PubMed: 16949657]
 12. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet*. 2007; 39:61–69. [PubMed: 17128275]
 13. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci USA*. 1992; 89:1827–1831. [PubMed: 1542678]
 14. Henderson IR, Chan SR, Cao X, Johnson L, Jacobsen SE. Accurate sodium bisulfite sequencing in plants. *Epigenetics*. 2010; 5:47–49. [PubMed: 20081358]
 15. Reinders J, Delucinge Vivier C, Theiler G, Chollet D, Descombes P, Paszkowski J. Genome-wide, high-resolution DNA methylation profiling using bisulfite-mediated cytosine conversion. *Genome Res*. 2008; 18:469–476. [PubMed: 18218979]
 16. Zemach A, Kim MY, Silva P, Rodrigues JA, Dotson B, Brooks MD, Zilberman D. Local DNA hypomethylation activates genes in rice endosperm. *Proc Natl Acad Sci USA*. 2010; 107:18729–18734. [PubMed: 20937895]
 17. Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science*. 2010; 328:916–919. [PubMed: 20395474] This study describes the DNA methylation patterns in several plant and animal species.
 18. Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME, et al. Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci USA*. 2010; 107:8689–8694. [PubMed: 20395551] This study together with Ref. [17•], describes a comparative analysis of DNA methylation in several plant and animal species.
 19. Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE. Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol*. 2007; 5:e129. [PubMed: 17439305]
 20. Zhang X, Bernatavichute YV, Cokus S, Pellegrini M, Jacobsen SE. Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol*. 2009; 10:R62. [PubMed: 19508735]
 21. Bernatavichute YV, Zhang X, Cokus S, Pellegrini M, Jacobsen SE. Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in *Arabidopsis thaliana*. *PLoS One*. 2008; 3:e3156. [PubMed: 18776934]
 22. Turck F, Roudier F, Farrona S, Martin-Magniette ML, Guillaume E, Buisine N, Gagnot S, Martienssen RA, Coupland G, Colot V. *Arabidopsis* TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS Genet*. 2007; 3:e86. [PubMed: 17542647]

23. Oh S, Park S, van Nocker S. Genic and global functions for Paf1C in chromatin modification and gene expression in Arabidopsis. *PLoS Genet.* 2008; 4:e1000077. [PubMed: 18725930]
24. He G, Zhu X, Elling AA, Chen L, Wang X, Guo L, Liang M, He H, Zhang H, Chen F, et al. Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell.* 2010; 22:17–33. [PubMed: 20086188]
25. Zhou J, Wang X, He K, Charron JB, Elling AA, Deng XW. Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in Arabidopsis reveals correlation between multiple histone marks and gene expression. *Plant Mol Biol.* 2010; 72:585–595. [PubMed: 20054610]
26. Charron JB, He H, Elling AA, Deng XW. Dynamic landscapes of four histone modifications during deetiolation in Arabidopsis. *Plant Cell.* 2009; 21:3732–3748. [PubMed: 20008096]
27. Elling AA, Deng XW. Next-generation sequencing reveals complex relationships between the epigenome and transcriptome in maize. *Plant Signal Behav.* 2009; 4:760–762. [PubMed: 19820310]
28. Egelhofer TA, Minoda A, Klugman S, Lee K, Kolasinska-Zwierz P, Alekseyenko AA, Cheung MS, Day DS, Gadel S, Gorchakov AA, et al. An assessment of histone-modification antibody quality. *Nat Struct Mol Biol.* 2010; 18:91–93. [PubMed: 21131980]
29. Gehani SS, Agrawal-Singh S, Dietrich N, Christophersen NS, Helin K, Hansen K. Polycomb group protein displacement and gene activation through MSK-dependent H3K27me3S28 phosphorylation. *Mol Cell.* 2010; 39:886–900. [PubMed: 20864036]
30. Filichkin SA, Priest HD, Givan SA, Shen R, Bryant DW, Fox SE, Wong WK, Mockler TC. Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Genome Res.* 2010; 20:45–58. [PubMed: 19858364]
31. German MA, Pillay M, Jeong DH, Hetawal A, Luo S, Janardhanan P, Kannan V, Rymarquis LA, Nobuta K, German R, et al. Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat Biotechnol.* 2008; 26:941–946. [PubMed: 18542052]
32. Gregory BD, O'Malley RC, Lister R, Urich MA, Tonti-Filippini J, Chen H, Millar AH, Ecker JR. A link between RNA metabolism and silencing affecting Arabidopsis development. *Dev Cell.* 2008; 14:854–866. [PubMed: 18486559]
33. Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, Kebrom TH, Provart N, Patel R, Myers CR, et al. The developmental dynamics of the maize leaf transcriptome. *Nat Genet.* 2010; 42:1060–1067. [PubMed: 21037569]
34. Mosher RA, Schwach F, Studholme D, Baulcombe DC. PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proc Natl Acad Sci USA.* 2008; 105:3145–3150. [PubMed: 18287047]
35. Nakano M, Nobuta K, Vemaraju K, Tej SS, Skogen JW, Meyers BC. Plant MPSS databases: signature-based transcriptional resources for analyses of mRNA and small RNA. *Nucleic Acids Res.* 2006; 34:D731–D735. [PubMed: 16381968]
36. Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ. Endogenous siRNA and miRNA targets identified by sequencing of the *Arabidopsis degradome*. *Curr Biol.* 2008; 18:758–762. [PubMed: 18472421]
37. Zheng Q, Ryvkin P, Li F, Dragomir I, Valladares O, Yang J, Cao K, Wang LS, Gregory BD. Genome-wide double-stranded RNA sequencing reveals the functional significance of base-paired RNAs in Arabidopsis. *PLoS Genet.* 2010; 6(9):e1001141. [PubMed: 20941385]
38. Li YF, Zheng Y, Addo-Quaye C, Zhang L, Saini A, Jagadeeswaran G, Axtell MJ, Zhang W, Sunkar R. Transcriptome-wide identification of microRNA targets in rice. *Plant J.* 2010; 62:742–759. [PubMed: 20202174]
39. Tang F, Barbacioru C, Nordman E, Li B, Xu N, Bashkurov VI, Lao K, Surani MA. RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat Protoc.* 2010; 5:516–535. [PubMed: 20203668]
40. Lu C, Kulkarni K, Souret FF, MuthuValliappan R, Tej SS, Poethig RS, Henderson IR, Jacobsen SE, Wang W, Green PJ, et al. MicroRNAs and other small RNAs enriched in the Arabidopsis RNA-dependent RNA polymerase-2 mutant. *Genome Res.* 2006; 16:1276–1288. [PubMed: 16954541]

41. Zhai J, Liu J, Liu B, Li P, Meyers BC, Chen X, Cao X. Small RNA-directed epigenetic natural variation in *Arabidopsis thaliana*. *PLoS Genet*. 2008; 4:e1000056. [PubMed: 18437202]
42. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, et al. Sorting of small RNAs into *Arabidopsis argonaute* complexes is directed by the 5' terminal nucleotide. *Cell*. 2008; 133:116–127. [PubMed: 18342361]
43. Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC. Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. *Cell*. 2008; 133:128–141. [PubMed: 18342362]
44. Havecker ER, Wallbridge LM, Hardcastle TJ, Bush MS, Kelly KA, Dunn RM, Schwach F, Doonan JH, Baulcombe DC. The *Arabidopsis* RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. *Plant Cell*. 2010; 22:321–334. [PubMed: 20173091]
45. Lippman Z, Martienssen R. The role of RNA interference in heterochromatic silencing. *Nature*. 2004; 431:364–370. [PubMed: 15372044]
46. Yamada K, Lim J, Dale JM, Chen H, Shinn P, Palm CJ, Southwick AM, Wu HC, Kim C, Nguyen M, et al. Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science*. 2003; 302:842–846. [PubMed: 14593172]
47. Li X, Wang X, He K, Ma Y, Su N, He H, Stolc V, Tongprasit W, Jin W, Jiang J, et al. High-resolution mapping of epigenetic modifications of the rice genome uncovers interplay between DNA methylation, histone methylation, and gene expression. *Plant Cell*. 2008; 20:259–276. [PubMed: 18263775]
48. Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, et al. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat Biotechnol*. 2000; 18:630–634. [PubMed: 10835600]
49. Zhang X. The epigenetic landscape of plants. *Science*. 2008; 320:489–492. [PubMed: 18436779]
50. Zhang X, Henderson IR, Lu C, Green PJ, Jacobsen SE. Role of RNA polymerase IV in plant small RNA metabolism. *Proc Natl Acad Sci USA*. 2007; 104:4536–4541. [PubMed: 17360559]
51. Zhang X, Germann S, Blus BJ, Khorasanizadeh S, Gaudin V, Jacobsen SE. The *Arabidopsis* LHP1 protein colocalizes with histone H3 Lys27 trimethylation. *Nat Struct Mol Biol*. 2007; 14:869–871. [PubMed: 17676062]
52. Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S. Histone H2A.Z. and DNA methylation are mutually antagonistic chromatin marks. *Nature*. 2008; 456:125–129. [PubMed: 18815594]
53. Chodavarapu RK, Feng S, Bernatavichute YV, Chen PY, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ, et al. Relationship between nucleosome positioning and DNA methylation. *Nature*. 2010; 466:388–392. [PubMed: 20512117]
54. Jacob Y, Stroud H, Leblanc C, Feng S, Zhuo L, Caro E, Hassel C, Gutierrez C, Michaels SD, Jacobsen SE. Regulation of heterochromatic DNA replication by histone H3 lysine 27 methyltransferases. *Nature*. 2010; 466:987–991. [PubMed: 20631708]
55. Deal RB, Henikoff S. A simple method for gene expression and chromatin profiling of individual cell types within a tissue. *Dev Cell*. 2010; 18:1030–1040. [PubMed: 20627084] This study describes a highly innovative and powerful method for chromatin isolation and epigenomic studies in specific cell types.
56. Boutrot F, Segonzac C, Chang KN, Qiao H, Ecker JR, Zipfel C, Rathjen JP. Direct transcriptional control of the *Arabidopsis* immune receptor FLS2 by the ethylene-dependent transcription factors EIN3 and EIL1. *Proc Natl Acad Sci USA*. 2010; 107:14502–14507. [PubMed: 20663954]

Figure 1.

A flowchart of some types of RNA-Seq library preparation workflows. After isolation of total RNA, a number of methods can be used for enrichment of target RNAs (a few examples are shown above). Enriched RNA is used for either conversion to cDNA or used in one of the many protocols that can be used to maintain the strand specificity of the RNA. Samples are fragmented, ligated to adapters and PCR amplified. Optional barcoding of samples can be performed during ligation or PCR amplification, depending on the protocol used.

Figure 2.

Selected examples of epigenomic datasets generated by high-throughput sequencing methods [4]. **(a)** An illustration of a transposon (gray rectangle) that is undergoing RNA-directed DNA methylation. mC = DNA methylation (gold lines = mCG, blue lines = mCHG and pink lines = mCHH). sRNAs = small RNAs (filled boxes = uniquely aligning small RNAs and shaded boxes = small RNAs that align to more than one genomic location). **(b)** An illustration of antisense RNAs identified by strand-specific RNA-Seq (green boxes = antisense reads originating from the Watson strand highlighted by a dashed red box (top) and maroon boxes = sense reads originating from the Crick strand (bottom)). **(c)** An illustration of DNA:protein interaction data from a ChIP-Seq experiment [56]. Enrichment of sequenced reads is indicated by a dashed red box. **(d)** An illustration of alternative splicing, highlighted by a dashed red box, revealed by strand-specific RNA-Seq.

Table 1

Comparison of major sequencing platforms.

Sequencing platforms	Aligned reads/run	Read length	Run time	Epigenomic applications
MPSS	<1,000,000	17–20 bp		RNA-Seq
454 FLX	>1,000,000	400 bp (Avg.)	10 h	RNA-Seq
SOLiD 5500xl	~1,000,000,000	25–75 bp	1–7 days	RNA-, ChIP-, MethylC-Seq
HiSeq 2000	~1,000,000,000	25–100 bp	1.5–8 days	RNA-, ChIP-, MethylC-Seq