GENOMICS

Tuning in to flower power

Researchers generated a high-resolution snapshot of the epigenome of *Arabidopsis thaliana* by constructing and integrating the methylome, transcriptome and small RNAome using next-generation sequencing.

In plants, DNA methylation is used to fine-tune gene expression and to ward off transposon invaders. To better understand how DNA methylation is regulated in *Arabidopsis*, Joe Ecker's team at the Salk Institute set out to chart its methylome.

The researchers first used immunoprecipitation and subsequent hybridization to whole-genome tiling microarrays to identify methylated sequences, but this approach generated maps of insufficient resolution. A breakthrough came in the form of a precommercial release of a new sequencing instrument, the Illumina Genetic Analyzer. As Ryan Lister, a postdoc in Ecker's lab and co–lead author on the paper, explains, "Joe's lab was an early-access site, so this allowed us to start looking at the location of DNA methylation at single-base-pair resolution—a very large improvement over the 500 bp–1 kb resolution from microarrays."

Coupling bisulfite conversion—a treatment that converts only nonmethylated cytosines to uracils—to this new high-throughput sequencing technology, Ecker's team determined the *Arabidopsis* methylome at single-base-pair resolution. Comparing the maps, the researchers noticed that bisulfite sequencing was a lot better than the microarray approach at finding methylated sites regardless of their density or sequence context. In particular, they identified methylation sites in low-complexity transposon-rich regions, which are difficult to detect by microarray-based techniques because of cross-hybridization.

To mine the methylome, Ecker's team with Julian Tonti-Filippini from the University of Western Australia modified a user-friendly web-based visualization tool, AnnoJ, originally created to analyze proteomics datasets. "We developed a browser [that] allows you to view [the genome] from the megabase down to the single-nucleotide level," says Lister. This tool revealed certain patterns on the genome scale, such as the clustering of hypermethylation sites in promoters and 3′ untranslated regions.

To delve more deeply into the impact of methylation on gene expression, Ecker's team turned their sequencing efforts to the transcriptome, looking at RNAs in plants with and without DNA methylation and demethylation enzymes. For this direct RNA sequencing approach, called mRNAseq, they devised a sample preparation method that retained both polyadenylated and non-polyadenylated transcripts. "We opted to go for a [locked nucleic acid]–based depletion method of ribosomal RNA [as] longer transcripts can be more evenly represented," says Lister. They found that hundreds of genes were switched on and transposons were activated in DNA methylase–deficient plants, demonstrating the critical role of DNA methyltransferase in gene expression and in genome defense.

A third level of epigenetic regulation exists in plants: small RNAs (15–30 nucleotides) control DNA methylation by directing the enzymes to their target sequences. To investigate this process in more detail, Ecker's team sequenced the 'small RNAome'. Again, they used the Illumina technology and aligned the small RNAs to their complementary sequences in the methylome using AnnoJ. Approximately 85% of the small RNAs co-localized, suggesting an important role for small RNAs in DNA methylation.

The *Arabidopsis* epigenome can be viewed in AnnoJ at the Ecker lab website (http://neomorph.salk.edu/epigenome/). The data are also available in the US National Center for Biotechnology Information Gene Expression Omnibus database.

What is next for Ecker's team? Plants adapt to their environment. Some plants are more resistant to pathogens and some are better adapted to different climates. By sequencing the epigenomes in geographically distinct *Arabidopsis* plants, the researchers hope to determine whether or not such differences have an epigenetic component.

No doubt others will soon plug into this flower power by charting mammalian epigenomes using the methods developed by Ecker's team. Insight into stem cell self-renewal and cancer mechanisms could be just around the corner.

Michelle Pflumm

RESEARCH PAPERS
Lister, R. *et al.* Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* **133**, 523–536 (2008).

PROTEOMICS

MAPPING THE *ARABIDOPSIS* PROTEOME

Using mass spectrometry–based proteomics methods, Baerenfaller and colleagues report protein 'catalogs' for cells, seeds and, at different developmental stages, roots, leaves, flowers and fruit (siliques). The researchers found chloroplast and photosynthetic proteins enriched in leaves, intracellular protein transport and oxidative stress response proteins in roots, and heat- and water-sensitive and development-specific proteins in seeds. The identified proteins represent nearly 50% of those predicted from gene models. The data are available in the publicly accessible proteomics identifications database (PRIDE) and can be viewed at their website (http://www.atproteome.ethz.ch/). Happy mining. *MP*

RESEARCH PAPERS
Baerenfaller, K. *et al.* Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. *Science*, published online 24 April 2008.