Quantitative epigenetics: DNA sequence variation need not apply

Eric J. Richards1

Boyce Thompson Institute for Plant Research, Ithaca, New York 14853, USA

Two recent reports, including one by Reinders and colleagues (pp. 939–950) in the April 15, 2009, issue of *Genes & Development*, describe the construction of *Arabidopsis* recombinant inbred populations that maximize epigenetic rather than genetic variation. The distribution and behavior of phenotypic variation in these populations suggest that stable epialleles can control complex quantitative traits. However, stochastic epimutation and transposon movement in these populations present some unexpected technical hurdles to implementing quantitative epigenetic analysis.

Interest in the impact and potential evolutionary role of inherited epigenetic variation continues to grow (Jablonka and Lamb 1995; Rapp and Wendel 2005). Such variation involves information encoded in biochemical marks on the DNA or chromatin (e.g., differential histone modification states, histone variant distribution, and DNA methylation polymorphisms) that exist independently of strict control by cis- or trans-acting genetic variation (Richards 2006). Much of the work on stable transmissible epigenetic variation has exploited experimental models, such as mice, maize, and *Arabidopsis*, and most of these studies have examined either phenotypically silent epigenetic variation or macroepimutations that cause dramatic developmental phenotypes (Cubas et al. 1999; Riddle and Richards 2002; Vaughn et al. 2007; Youngson and Whitelaw 2008). Moving beyond experimental models, work on epigenetic variation in relation to cancer and other human diseases has sparked widespread interest (Gal-Yam et al. 2008; Whitelaw and Whitelaw 2008). However, the relevance of epigenetics is less obvious for the disciplines of evolutionary biology, ecology, and breeding, which are concerned primarily with complex polygenic traits and phenotypic variation in natural contexts (Bossdorf et al. 2008; Richards 2008).

Constructing recombinant inbred lines that parcel out epigenetic variation

Recent work from two groups—Reinders et al. (2009) in the April 15, 2009, issue of *Genes & Development* and Johannes et al. (2009)—promises to help bridge the gap between epigenetics in model systems and the approaches and tools commonly used by evolutionary biologists and breeders. Both reports describe the creation of a novel set of recombinant inbred lines in the flowering plant *Arabidopsis thaliana* constructed to maximize the segregation of epigenetic variation rather than genetic polymorphisms. The approach employed is simple and elegant [Fig. 1]. Both groups started with a strain carrying a mutation that disrupts proper epigenetic regulation, leading to the creation of epigenetic alleles through a loss of transcriptional silencing mixed with ectopic epigenetic silencing. In one case, a *met1* mutant defective in a CpG cytosine DNA methyltransferase was used [Kankel et al. 2003; Saze et al. 2003]. The other crossing scheme used a parent with a deficient *DDM1* nucleosome-remodeling ATPase gene, which is required to maintain cytosine methylation and histone H3 Lys 9 methylation marks characteristic of heterochromatic sequences [Jeddoloh et al. 1999; Lippman et al. 2004]. The mutant parent was crossed to an isogenic wild-type strain, and F2 or backcrossed progeny carrying only the wild-type *MET1* or *DDM1* allele were recovered as progenitors of multiple, independent recombinant inbred lines propagated through single-seed descent. The objective is the creation of recombinant inbred lines, each with a different patchwork of epialleles carried on chromosome blocks inherited from the mutant parent. As the mutant parent was isogenic with the wild-type inbred parent, no genetic variation should be segregating in the pedigrees. This strategy should maximize the epigenetic variation among lines, while completely eliminating genetic variation, hence, the term epiRILs—epigenetic recombinant inbred lines.

Perhaps the first question that springs to mind is why any epigenetic variation would persist in the epiRILs. One might expect that epigenetic variation generated in the *met1* or *ddm1* mutant parent would be reset after complementation of the mutation in the F1 generation and subsequent removal of the mutation by segregation. However, it has been long established that loci that lose
DNA methylation and epigenetic silencing in *ddm1* and *met1* mutants can be stably inherited in their altered epigenetic state through meiosis and transmitted for many generations after introduction into a wild-type background (Vongs et al. 1993; Kakutani et al. 1999). A breakthrough in the understanding of this phenomenon was achieved recently through the work of Colot and colleagues (Johannes et al. 2009), who demonstrated that loci targeted by small RNA (in this case, siRNA) undergo progressive remethylation over multiple generations after removal of the *ddm1* mutation, while loci without small RNA guides for de novo methylation and resilencing remain locked in their *ddm1*-induced epiallelic state (Teixeira et al. 2009). It is notable that even loci with matching small RNA species can, at some frequency, also escape resilencing (as shown in the epiRILs for the *fwa* epiallele affecting flowering time (Johannes et al. 2009; Reinders et al. 2009). In addition to the loci that lose epigenetic silencing marks in the mutants, other loci gain ectopic patches of heterochromatin and become inappropriately silenced in a manner resembling the CpG island methylation and silencing that occurs in human tumor genesis. The mechanisms at work are less clear but appear to involve off-target silencing by wayward heterochromatin modification machinery adrift without the underpinning of cytosine methylation marks on the DNA (Tariq et al. 2003; Mathieu et al. 2007). Inefficient or misregulated DNA demethylation in *met1* and *ddm1* mutants might also play a role in generating ectopically silenced epialleles (Mathieu et al. 2007). Once created, inappropriately silenced epialleles are also transmitted across generations in wild-type backgrounds, although often with less stability than their hypomethylated, unsilenced counterparts (Jacobsen et al. 2000).

**Phenotypic variation in the epiRILs**

Once generated, the epiRILs were used to address some key questions. Foremost among these questions was whether the epiRIL lines showed phenotypic variation, particularly for complex quantitative traits. Satisfyingly, the epiRIL lines differed in a number of traits, including flowering time, biomass, plant height, and response to pathogen attack (Johannes et al. 2009; Reinders et al.

---

**Figure 1.** Construction of epiRILs. The scheme for construction of conventional recombinant inbred lines is shown in the left column. A single chromosome pair in a diploid is illustrated. Two different strains (A and B) are intercrossed, and strain-specific polymorphisms (symbolized by yellow vs. blue) will segregate into the F_{2} (or backcrossed progeny). Propagation through single-seed descent generates lines that are fixed—homozygous at most loci. The center column of the figure depicts a model for the generation of epiRILs assuming stable inheritance of parental epigenetic states. The underlying genome sequence is the same in the wild-type and epimutator parents with the exception of the *met1* or *ddm1* mutation (red). The filled boxes that overlay the genome sequence illustrate the epigenetic states of the two parents. The black boxes indicate regions of epigenetic silencing marks, many of which are lost in the mutant parent. In addition, ectopic silencing marks (green boxes) are found in the mutant parent. Assuming faithful epigenetic inheritance of parental epigenetic marks, the resulting epiRIL lines will be composed of epigenetic haplotypes representing blocks derived from the parent chromosomes. The right column describes the sources of additional genetic and epigenetic variation observed in epiRILs, and the single chromosome pair in the inset illustrates the resulting deviations from the idealized scheme shown in the center column.
Moreover, the continuous distribution of the phenotypic values across the lines was consistent with polygenic traits rather than phenotypes controlled by a single locus of large effect. Such a result might also reflect increased variability within the individual lines caused by a breakdown of epigenetic regulation of gene expression and developmental programs. Data from both groups support the former interpretation, at least for some phenotypes. Reinders et al. [2009] showed that differences among epiRILs for bacterial pathogen resistance and salt sensitivity bred true (Reinders et al. 2009). Johannes et al. [2009] went further to show that variations for flowering time and plant height among different lines are higher than the variations observed among plants within individual lines; broad-sense heritability estimates for these phenotypes were on the order of 25%–30%. These findings support the hypothesis that heritable epiallelic variation can underlie polygenic traits important for fitness.

Another important question is whether phenotypic variation displayed by the epiRILs is fundamentally different than that seen in mutagenized populations or in traditional RILs made from genetically divergent parents. The two different epiRIL populations appear to behave differently in this regard. The met1-derived epiRIL population exhibited many unstable phenotypes, and ~30% of these epiRIL lines [28 of 96] were unable to be propagated to the F5 generation because of reduced fertility or developmental defects [Reinders et al. 2009]. In contrast, the ddm1-derived epiRIL populations displayed less phenotypic instability [V Colot pers. comm.], and >99% of the lines [505 of 509] were successfully advanced to the F5 generation [Johannes et al. 2009].

Epigenetic instability and epimutation in the epiRILs

A possible explanation for the phenotypic instability in the met1-derived epiRILs was uncovered when cytosine methylation patterns were examined in the lines. Cytosine methylation polymorphisms were segregating even in the F5 and F6 generations at a few selected loci examined [Reinders et al. 2009]. This finding indicates that not all epigenetic states are fixed, despite many generations of self-pollination. Further, a large number of nonparental cytosine methylation patterns were present in the epiRILs [Reinders et al. 2009]. Cytosine methylation alleles were detected using a whole-genome profiling technique combining bisulfite conversion of the target DNA (unmodified C → U, while 5mC is unconverted) and hybridization to a short-oligonucleotide tiling microarray [Reinders et al. 2008]. The profiling results for three epiRIL lines explode the simple model that the met1-derived epiRILs are a mosaic of parental epigenetic haplotypes [Fig. 1, center]. While the parental origin of chromosome segments could be discerned from the distribution and overall density of wild-type and met1-like cytosine methylation alleles, it was clear that wild-type methylation alleles were abundant on chromosome segments inherited from the mutant parent [Reinders et al. 2009]. RNA-directed de novo cytosine methylation could account for this pattern, equivalent to the cytosine remethylation process seen in ddm1-derived epiRILs [Teixeira et al. 2009]. More unexpected was the widespread occurrence of nonparental epialleles dispersed across the profiled epiRIL genomes irrespective of the boundaries between different parental chromosome segments [Reinders et al. 2009]. These nonparental alleles, from both ectopic and hypomethylation events, could not simply originate from the mutant parent as they are also found on chromosome segments derived from the wild-type parent. The continued generation and segregation of cytosine methylation polymorphism most likely reflect and parallel the creation, segregation, and instability of novel epigenetic alleles affecting development and plant fitness in the met1-derived epiRILs.

If the epigenetic instability observed in the met1-derived epiRILs is responsible for phenotypic instability, one would expect the more phenotypically stable ddm1-derived epiRILs to show less flux in epigenetic marks. Unfortunately, it is difficult to make a proper comparison because the cytosine methylation analysis of the ddm1-derived epiRILs focused on selected loci and used a method that detects the density of 5-methylcytosines, but does not distinguish cytosine methylation alleles with single-nucleotide sequence resolution [Johannes et al. 2009; Teixeira et al. 2009]. Consequently, it is not known whether nonparental cytosine methylation alleles are generated in the ddm1-derived epiRILs to an extent comparable with that seen in the met1-derived epiRILs.

What causes the nonparental epialleles to form in the met1-derived epiRILs? Might the union in the F1 generation between two drastically different epigenomes, albeit perched on the identical genomic sequence, lead to an explosive round of epigenetic variation? The epigenomic "shock" and restructuring invoked here is analogous to the type of genomic crisis and response that McClintock [1984] recognized in the mobilization of maize transposons triggered by DNA damage [for example, from a broken chromosome]. To evaluate properly the importance of epigenomic discordance as a genome shock, a baseline for epigenetic variation induced by genetic crosses needs to be established first. What happens to the cytosine methylation patterns when two identical wild-type strains carrying very similar epigenomes are crossed? Does epigenetic variation occur in all crosses, or is the abundance of nonparental cytosine methylation alleles created in hybrids dependent on the epigenetic (and genetic) divergence of the parents? How might epigenomic divergence be sensed? Some insight into potential mechanisms might be glimpsed from recent reports in plants and animals that suggest mobile or transmissible sets of small RNA can cross-reference the epigenetic state between different cells or between genomes in hybrids [Brennecke et al. 2008; Slotkin et al. 2009].

Transposon mobilization in the epiRILs

Regardless of the mechanisms at work, one might consider the generation of novel nonparental epigenetic
variation as a potential bonus for the purpose at hand, which is constructing populations that maximize epigenetic variation while minimizing genetic variation. However, standard linkage-based mapping approaches to identify important epialleles affecting complex traits using these epiRILs will be frustrated by the fact that the new epigenetic variation is often unstable and not sequestered in tidy blocks that can be traced unambiguously to one or the other parent. Confounding matters further, there is another type of variation generated in the epiRIL lines that undermines the premise of the lines’ construction. In both the met1- and ddm1-derived epiRILs, CACTA transposons are mobilized [Johannes et al. 2009; Reinders et al. 2009]. These DNA elements lose epigenetic silencing and are activated in the ddm1 mutant parents [Miura et al. 2001]. Once these elements shed their epigenetic silencing, they continue to jump in the different ddm1-derived epiRIL lineages [Johannes et al. 2009]. A less straightforward explanation must be proposed for CACTA element mobilization in the met1-derived epiRILs. CACTA elements are silent and immobile in wild-type plants and in met1 mutants [Reinders et al. 2009]—in the latter case, because complete release of epigenetic silencing requires a loss of both CpG and non-CpG methylation [Kato et al. 2003]. Despite being inactive in wild-type and met1 parents, CACTA elements begin to jump stochastically in the met1-derived epiRILs [Reinders et al. 2009]. The genome shock framework might also apply to this situation where dormant elements become activated only after crossing or later in the self-pollination regime as epigenetic instability persists. Transposon activation could also be a response to genomic damage more egregious than the sorting out of epiallelic variants. Cytosine hypomethylation in mammals is associated with an elevated frequency of deletions [Chen et al. 1998], and the true triggers of transposon mobilization could be DNA damage or DNA rearrangement intermediates that form as a consequence of epigenetic defects. Thus, it is important to consider whether other types of genetic variation, such as copy number variants, are also segregating in the epiRIL populations.

Disentangling the epigenetic and genetic components controlling phenotypic variation in the epiRIL populations will be a challenge [Johannes et al. 2008], but the same problem exists for any population, whether experimental or natural. The perspective and techniques employed most frequently in mapping phenotypic variation are focused solely on discovery and manipulation of genetic variation. Often, experimental populations are constructed to maximize genetic variation [e.g., irradiated or chemically mutagenized], but the molecular basis of phenotypic variation in natural populations or among varieties is usually unknown. In the case of the epiRIL lines, while it is clear that genetic mutation is occurring, these populations have been constructed to tip the balance toward epigenetic variation. Relatively few examples of populations and experimental designs to achieve this objective have been described in the literature [Rutherford and Henikoff 2003; Sollars et al. 2003; Garfinkel et al. 2004]. Verification that important phenotypic variation in the Arabidopsis epRILs is controlled by epigenetic variation comes from the demonstration that fwa epimutations are found in late-flowering lines from both epRIL populations [Johannes et al. 2009; Reinders et al. 2009]. Although the epRIL populations are more complex than originally anticipated, the expansion of epigenomic profiling techniques and the reduction in the costs of high-throughput genomic sequencing might facilitate the identification of the epigenetic and genetic basis of the complex trait variation in these populations without the reliance on linkage-based techniques. Defining and understanding the interplay between genetic and epigenetic variation represents one of the important challenges in epigenetics [Johannes et al. 2008], and the Arabidopsis epRIL studies lay the groundwork for tackling this challenge.

Acknowledgments

I thank Jeff Jeddeloh and Patrice Debois for comments on the manuscript. Research in E.J.R.’s laboratory is funded by grants from the National Science Foundation (MCB-0548597) and the National Institutes of Health (R01GM078256). I apologize to my colleagues whose work was not cited due to space limitations.

References

hypomethylation induced by ddm1 mutation of Arabidopsis thaliana. Genetics 151: 831–838.

Quantitative epigenetics
Quantitative epigenetics: DNA sequence variation need not apply

Eric J. Richards

Genes Dev. 2009, 23:23
Access the most recent version at doi:10.1101/gad.1824909

Related Content

Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes
Genes Dev. April, 2009 23: 939-950

References

This article cites 33 articles, 10 of which can be accessed free at:
http://genesdev.cshlp.org/content/23/14/1601.full.html#ref-list-1

Articles cited in:
http://genesdev.cshlp.org/content/23/14/1601.full.html#related-urls

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.