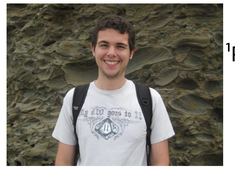
Transgenerational Epigenetic Instability as a Source of Novel Methylation Variants



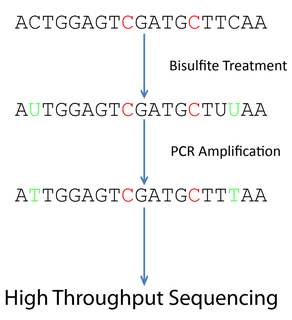
Matthew D. Schultz[1,2,3], Robert J. Schmitz[1,2], Mathew G. Lewsey[1,2], Ronan C. O'Malley[2], Mark A. Urich[1,2], Ondrej Libiger[4], Nicholas J. Schork[4], Joseph R. Ecker[1,2,5]

[1]Plant Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA. [2]Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA.
[3]Bioinformatics Program, University of California at San Diego, La Jolla, CA 92093, USA.
[4]The Scripps Translational Science Institute and the Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA 92037, USA.
[5]Howard Hughes Medical Institute, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA.

Motivation and Background

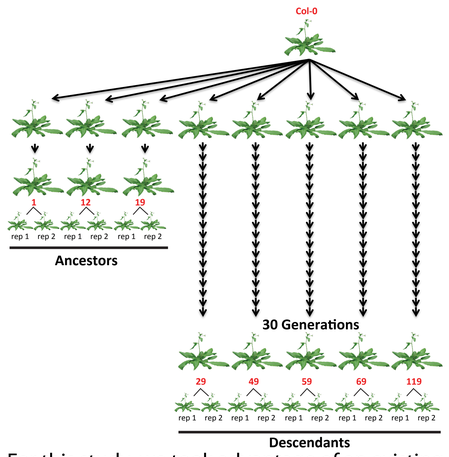
Every cell that makes up an individual has the exact same genome and yet that individual is made up of a whole host of different cells. To achieve this variety of cell types, organisms utilize a variety of chemical modifications to DNA, which can be thought of as dials that selectively turn up or down how much of that gene's product is produced. For example, some of the genes that are turned on in heart cells will not be turned on in liver cells and vice versa. Consequently, when a heart cell divides, it needs to send a signal to its daughter cells about which genes should be on and off. The term epigenetics is used to describe the inheritance of these kinds of non-genetic modifications. An open question is whether or not changes in these modifications can vary over time like mutations in DNA. Undoubtedly, if one picked any of these modifications and compared them across two individuals there would be many differences. What makes this problem challenging is disentangling those differences from genetic and environmental effects that could have also caused them. To overcome these challenges in an attempt to answer this fundamental question, we turned to the model plant *Arabidopsis thaliana* and DNA sequencing technology.

Methylome Sequencing



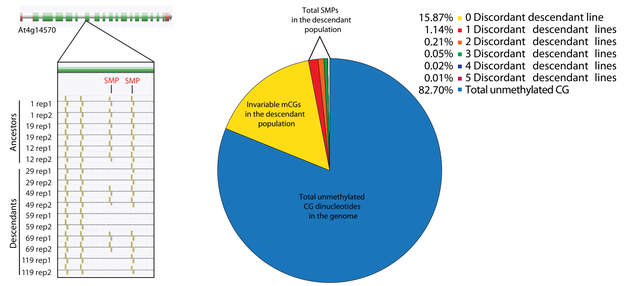
The epigenetic modification that we chose to focus on in this study is DNA methylation, which is the addition of a methyl group to cytosines in DNA. To measure this modification, we take advantage of a chemical known as bisulfite, which will convert unmethylated cytosines to thymines and leave methylated cytosines (red Cs above) unconverted. After this conversion, any remaining cytosines detected by a DNA sequencing platform should be methylated. We refer to the sum total of these methylated cytosines as the methylome.

Mutation Accumulation Population



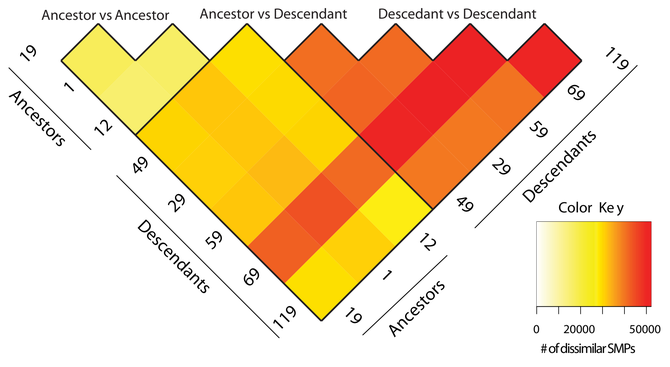
For this study, we took advantage of an existing population of plants that had been highly inbred and is consequently very genetically similar. We sequenced the methylomes of 5 individuals that were 31 generations removed from the original founder of this population and 3 individuals that were only 3 generations removed from the founder. To account for technical variability, we also sequenced a biological replicate of each of these samples which were highly similar.

Single Methylation Polymorphisms



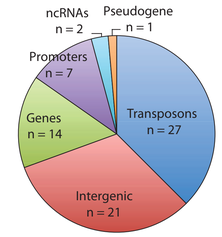
With these methylome data, we were able to ask how different the descendants were from the original founder. We called these differences single methylation polymorphisms (SMPs). Although most of the cytosines in all of our individuals were either completely methylated or unmethylated, a fraction of them varied between individuals. In the left panel above, gold bars represent sites of methylation.

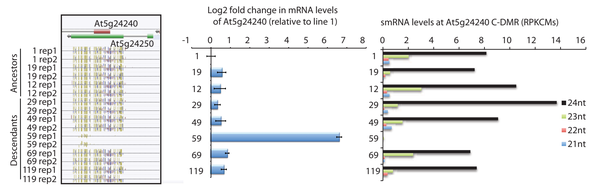
SMPs Diverge Over Time



When we compared the SMPs from different individuals we found a striking pattern. The more generations that separated two individuals the more the methylation patterns at SMP sites differed. Above, dark red indicates more differences.

Differentially Methylated Regions





We expanded our search to look for differentially methylated regions and found 72 regions. Interestingly, 14 of these regions fell within protein-coding genes, and we looked at the correlation between methylation, expression, and small RNAs, which are known to help direct DNA methylation. As can be seen above, the absence of methylation correlates with the expression of this gene and with a loss of 24-nucleotide small RNAs. This result supports the idea that these epigenetic changes can have an effect on gene production and potentially an individual.

Conclusions

In the design of this experiment, we have tried to account for genetic and environmental variability that could have led to the changes we observed. *Arabidopsis thaliana* offered us a unique opportunity to show convincing evidence that it is possible for these modifications to vary across generations independent of genetic changes. Furthermore, we were able to find examples of these changes that had an effect on the transcriptional output of some genes. Although no major phenotypic changes have ever been observed in this population of plants, our results leave open the intriguing possibility that these changes could be acted upon by evolution.