

# **The epigenome and beyond: Are there alternatives to primary DNA and RNA sequence that organize and transmit heritable information over generations and would alter our view of evolution and adaptation?**

Butch Brodie<sup>1</sup>, Brian Gregory<sup>2</sup>, Damon Lisch<sup>3</sup>, Nicole C. Riddle<sup>4</sup>

<sup>1</sup> Department of Biology, University of Virginia, Charlottesville, VA

<sup>2</sup> Department of Biology, University of Pennsylvania, Philadelphia, PA

<sup>3</sup> Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN

<sup>4</sup> Department of Biology, The University of Alabama at Birmingham, Birmingham, AL

## **Summary**

Heritability is at the core of our understanding of both adaptation and evolution. Most analysis of heritability assumes that changes in DNA sequence are the foundation of heritable changes in phenotype. However, given our improved understanding of the importance of epigenetics, it is possible that heritable changes in phenotypes not dictated by DNA sequence also have significant effects on both adaptation and evolution. Given known connections between epigenetic variation and environmental variation, it is important to determine the extent to which the environment, including both abiotic and biotic interactions, can impact and perturb epigenetic pathways both within individual organisms and in subsequent generations. Transgenerational epigenetic effects are particularly important because they are most likely to be shaped by natural selection and other evolutionary forces. Furthermore, how epigenetic pathways are employed and regulated differs among organisms, but we have little knowledge of this variation across the diversity of life. Addressing these questions will have a significant impact on our understanding of heritability, evolution, and adaptation and might lead to improved, more predictable models of phenotypic evolution.

## **Epigenetics Defined**

For the purpose of this article, we broadly define epigenetics as heritable changes in phenotype in the absence of changes in primary DNA sequence and in the absence of the initial trigger for those changes. Although specific biochemical modifications such as DNA methylation and histone modification are known to be associated with epigenetic changes in gene expression, we are agonistic with respect to mechanism. Epigenetic changes are distinctive because they are heritable, labile, and potentially influenced by the environment. Because of these qualities, epigenetic changes generate patterns of inheritance that are not necessarily amenable to traditional Mendelian genetic analysis and thus will require distinctive strategies.

## **Why now?**

While epigenetics began as a study of oddities in a few model organisms, we now have numerous well-studied examples of epigenetic inheritance in a variety of species, and there are

many examples of unexplained forms of heritability across the tree of life that need to be explored further (<https://www.ncbi.nlm.nih.gov/pubmed/24384572>). Technological innovations such as next generation sequencing methods, genome-wide epigenome assays, and CRISPR/dCas9-based methods for epigenome editing have enabled researchers to explore epigenetic phenomena at an unprecedented scale and depth. We are now at a point where we can begin to determine the extent to which the heritability of large numbers of traits in multiple species are dependent on known epigenetic mechanisms, such as DNA and histone modification. Thus, it is now possible to obtain reliable information about the extent of epigenetic variation in known pathways within individuals, between individuals and populations, as well as between species. Knowledge of the level and type of epigenetic variation is the first step in understanding the impact of this variability on phenotypes.

While DNA and histone modifications are an attractive target of study, analysis of epigenetic phenomena should not be restricted to known mechanisms. It is tempting to simply apply known technologies such as bisulfite or ChIP (chromatin immunoprecipitation) sequencing on a much larger scale, but such a focus on well understood mechanisms carries the risk of missing important aspects of epigenetic inheritance. There is ample evidence from yeast, for instance, that prions can mediate heritable changes in the absence of changes in DNA sequence that impact host fitness, and it is worth noting that until very recently, little was known about the importance of small RNAs, a class of molecules that now sit at the center of our understanding of epigenetic regulation. Therefore, it would be beneficial to take a broader and more integrated view, remaining open to the existence of additional mechanisms mediating epigenetic inheritance. Furthermore, this approach might be extended to behavioral or cultural evolution, where information is transferred also between generations, without the involvement of changes in DNA sequence. Overall, the goal is not just to find more examples of what we know, but also to uncover phenomena that may involve novel mechanisms.

Heritable epigenetic changes are particularly difficult to assess for the same reasons that they are interesting: they are both transmissible and unstable. For this reason, analysis of instance in which DNA sequence is largely invariant will be invaluable (<https://link.springer.com/article/10.1007/s12224-017-9308-x>). Additionally, there is still a large amount of missing heritability across the myriad genome-wide association studies (GWAS) performed for both plant and animal traits (Solving the missing heritability problem) (<https://www.ncbi.nlm.nih.gov/pubmed/22035733>). This unexplained variance could be due to the relative instability of traits that are driven by epigenetic mechanisms. Focusing on the heritability of these GWAS tested traits could also provide a resource for identifying novel epigenetic mechanisms of inheritance. It is also likely that the plant and animal breeding communities have encountered instances of unexplainable phenotypic variation, some of which is likely due to relatively unstable epigenetic variance. Based on past experience, a potentially fruitful phenotype on which to initially focus would be variation in flower morphology and color in various plants. The pathways that influence these phenotypes are well understood, and breeders have been selecting for variation, some of which is unstable, for centuries.

## **New technologies enable experiments designed to test the relationship between epigenetics and evolution**

Over the last ten years, methods development and technical innovations made a plethora of long-standing research questions amenable to study. One of these questions is the evolutionary significance of phenotypic variation under epigenetic control. To address this question, the development of low cost, high-throughput sequencing methods has been essential. Combining these next generation sequencing approaches with standard assays used to interrogate the epigenome and epitranscriptome has produced genome-wide profiles from diverse sets of epigenetic/epitranscriptomic marks in numerous cell types in many species. These methods allow us to assay numerous heritable histone, DNA and RNA modifications, and higher level chromatin structure can be profiled by a variety of methods that assay accessibility to nucleases (DNase-seq), transposases (ATAC-seq), and other processes (FAIRE-seq, salt extraction, etc). These methods have delivered a view of the epigenome in unprecedented detail and complexity.

Many of these methods are applicable across a wide range of species and thus, the data available for scientists to estimate the extent of variation among epigenome features within and between organisms as well as within and between populations and species has increased exponentially. With the recent foray of epigenome profiling methods into the scale of single cells, we anticipate that the amount of data available will increase by orders of magnitude in the near future. This increase in capacity will open up additional areas of inquiry, but will also increase the challenges associated with handling the increased data complexity. Thus, computational methods and tools to efficiently and appropriately handle these data will become increasingly important. While computational modeling *per se* is not new, the methods to process highly dimensional data and model the complex interactions between genetics, epigenetics, and environment have become available only recently. Tools to integrate these data types and to partition phenotypic variance components are still being developed and refined. In addition, the integration of these tools into evolutionary models will require further development of computational methods in multi-disciplinary collaborations between evolutionary biologists, geneticists, molecular biologists, and computational biologists.

Targeted manipulation of features within the epigenome that are now available will make it possible to directly test hypotheses (<https://www.ncbi.nlm.nih.gov/pubmed/29852072>). ZNF (zinc finger) nucleases, TALEN (Transcription activator-like effector nucleases), and CRISPR/dCas9 systems can be used to specifically target regions in the genome of model systems for epigenetic manipulation. With these methods, it is possible to experimentally assess the phenotypic effects of specific modifications. Follow-up experiment utilizing artificial selection then can test directly if this epigenetic information and the associated phenotypes can be acted on by selection and serve as basis for evolutionary change and adaptation. In particular the CRISPR/dCas9-based approaches likely will be transferable from model systems into other species, thus allowing for hypotheses regarding the epigenetics-evolution relationship to be tested across different branches of the tree of life. These approaches present exciting new possibilities, as we might explore if we can purposefully employ epigenome engineering to drive evolution, at least in the short term.

## **Challenges and opportunities**

Our understanding of epigenetic mechanisms of inheritance is incomplete. Our understanding of even the most well studied phenomena is largely correlative, often anecdotal, and restricted to a relatively small number of model organisms. We know very little about the impact that these epigenetic mechanisms have on most organisms, although there is certainly evidence that there is a great deal of variation with respect to specific mechanisms among species. DNA methylation, for instance, is essential for proper epigenetic regulation in many species, but is nearly absent in many other species (ref). We know very little about how selection operates on these pathways/mechanisms and how changes in these pathways mediates changes in responses to selection on the phenotypes they affect.

Another challenge currently limiting our understanding of epigenetic inheritance is that epigenetic information is dynamic. Time is an essential component of epigenetic regulation because the causes of epigenetic changes are manifest both in individual organisms as they develop and in their progeny. But we know very little about the dynamics of these epigenetic changes, either spatially or temporally. There are, for instance, multiple epigenomes/epitranscriptomes even within an individual, with different cell types/tissues having distinct epigenomes and epitranscriptomes. In addition, patterns of epigenetic modification change as tissues differentiate or experience different environmental conditions. Thus, we need to carefully consider at what scale (e.g. cells, tissue, etc.) and at what time we need to sample to understand the inheritance of the epigenome between parents and offspring.

Interestingly, the best and most ubiquitous examples of stable epigenetic inheritance are linked to silencing of transposable elements (TEs), which make up a substantial proportion of most eukaryotic genomes. The bulk of these examples involve selection in favor of silencing of TEs rather than regulation of gene expression. Thus, if one is interested in traits that affect phenotype, there is a substantial signal to noise problem. It will be important to recognize distinctions between selectively neutral epigenetic variation arising from transposon silencing and epigenetic variation that has a direct effect on fitness.

To the extent that epigenetic differs from genetic inheritance, we will need to transform our thinking about predictive modeling. For each phenotype, models will have to include the contributions of both epigenetic and genetic factors that give rise to that phenotype. Overall, these barriers provide important areas of inquiry in order to drive this important research direction forward.

Finally, it is important to note that epigenetic variation not linked to known modifications is likely to have been ignored; it may well be that re-examination of data from a wide variety of experiments in fields as diverse as breeding, ecology and medicine could yield valuable insights into hidden sources of variation.

### **Moving forward**

Transgenerational transmission of information is the basis of evolutionary change, yet existing conceptualizations consider only direct inheritance of variation in DNA. If whole organism

phenotypes are comprised of a mosaic of characters that are inherited through distinct mechanisms, evolutionary theory will have to determine how to simultaneously include different processes to predict change. A mechanistic understanding of how these indirect pathways function will dictate how we determine their impact on evolutionary processes. The inclusion of a new focus on both biotic and abiotic environmental factors will require a change in the overall scale at which we are currently analyzing these phenomena. This question will require the integration of approaches and expertise from a number of biological disciplines in order to truly assess the full importance and functionality of these extra-chromosomal pathways and mechanisms of inheritance.

One key problem will be recognizing patterns of inheritance that are distinct from those observed that result from only genetic variation. These patterns will likely be considerably more stochastic and contingent. Machine learning can be used to look for patterns of inheritance that do not match normal genetic inheritance patterns.

In order to determine the scale and magnitude of how these mechanisms affect trait inheritance across scales and populations, the proposed work will involve interactions between evolutionary biologists, geneticists, molecular biologists, ecologists, computational biologists, and systems biologists. They also need to involve collaborations between biologists working in different phyla.

Integration: what are the roles for each biological field in this research question:

- *Computational biologists* will be needed because the approaches that are being widely adopted and developed as well as those that are needed for addressing this question tend to be data-intensive and produce numerous large-scale datasets. *Transmission geneticists* are very good at finding new epigenetic phenomena, while *population geneticists* are very good at finding variation in allele (and epiallele) frequencies. In order to understand the overall impact of epigenetic variation, both in populations and over time, it will be important for these specialists to work closely with both ecologists and evolutionary biologists.
- *Molecular biologists* will be needed because the methodologies and approaches that need to be developed will require molecular genetic techniques and expertise. Thus, molecular biologists will continue to be at the forefront in providing the methodological framework to drive our understanding of this important research question and to uncover novel mechanisms of epigenetic inheritance.
- *Ecologists* will reveal the bidirectional connections between phenotype and the environment. Environmental conditions have the potential to drive changes in extra-genetic inheritance, while simultaneously acting as the background in which resulting phenotypic variation is tested by selection.
- *Evolutionary biologists* will be required to understand how epigenetic inheritance determines long-term change of populations and species and to develop the predictive framework that incorporates novel pathways of inheritance into models of transgenerational change and response to environmental challenges.

## **Conclusion**

Exploring the diversity of epigenetic inheritance could reveal novel mechanisms that explain examples of transmission that cannot be understood by primary sequence-based mechanisms. Understanding how different pathways determine the inheritance of each part of a larger phenotypic output could lead to independent manipulation of components. For instance, the regulation of disease resistance genes in plants involves a complex relationship between genetic and epigenetic mechanisms (<https://www.ncbi.nlm.nih.gov/pubmed/28154240>). Epigenetic modes of inheritance could be especially important for understanding social and cultural modes of inheritance that cannot be explained under current molecular mechanisms of inheritance. Complex social traits including collective behavior, empathic learning, and social network structure that exist only in groups of many individuals are likely to be influenced by epigenetic processes. Inheritance that operates outside of DNA sequence transmission opens the potential for horizontal and vertical pathways of transmission both within and among groups and species, which would fundamentally shift the rules of evolutionary change. Ultimately, uncovering new mechanisms by which biological inheritance takes place would transform our understanding of how organisms respond to environmental changes and selection, both artificial and natural.