Research Statement: Chad E Niederhuth

What is the intersection between genetics, the epigenome, and phenotype? How does knowledge of the epigenome help us address larger questions of biology and the major challenges we face going forward? These are the questions I constantly ask myself in my research. I am broadly interested in the role of the epigenome in shaping the translation of genes into phenotype and how in turn the epigenome is shaped by both genetics and life history of an organism. These two over-arching themes are really two-sides of the same coin. One views DNA methylation, histone modifications, and other constituents of the epigenome as contributing causative factors to a trait. For example, gain or loss of chromatin marks at repetitive sequences and transposons can affect neighboring gene expression, leading to phenotypic variation. The other side of the coin views the epigenome itself as a trait; asking such questions as what are the underlying mechanisms and how might genetic variation result in epigenomic variation. Undergirding this is always the question what this means in the big picture.

Nature has done far more experiments, creating far more diversity, than any number of labs can do in a life time. My approach to answering these questions has been to let this natural diversity direct the questions I ask and the experiments I conduct. Using this philosophy, in my postdoctoral research I took a "comparative epigenomics" approach to examine DNA methylation of thirty four different flowering plant species (Figure 1) **(Niederhuth CE*, Bewick AJ*, et al in revision; *co-first authors)**. This included the first methylomes of many species, including economically important crops such as *Citrus clementina* (clementine), *Vitis vinifera* (grape), *Malus domestica* (apple), *Beta vulgaris* (sugar beet), and more. This revealed widespread diversity in DNA methylation; including differences in methylation distributions in grass species, potential effects of reproductive strategies, associations between gene duplication and methylation, and potential ways of inducing phenotypic variation through perturbing the epigenome.

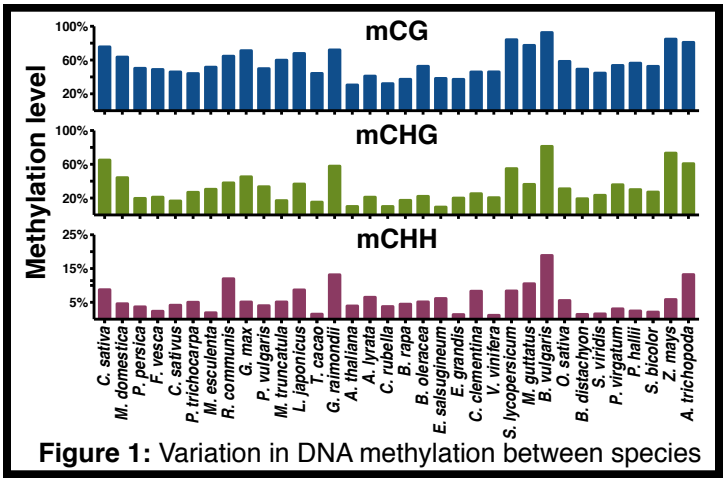

Figure 1: Variation in DNA methylation between species

The power of this approach has enabled me to generate new hypotheses and then go on to test these using both computational and experimental approaches. Two examples of published and ongoing research demonstrate this. First is the example of the evolution of gene-body methylation (gbM). This is a feature of plant genomes where methylation of CG di-nucleotide sites in gene bodies is associated with constitutive gene expression. Its function and origins have been a mystery since its discovery 10 years ago. I identified the first angiosperm species, *Eutrema salsugineum*, which lacks gbM. From this, we identified loss of CHROMOMETHYLASE 3 (CMT3) in *E. salsugineum* as the underlying cause of loss of gbM **(Bewick AJ*, Ji L*, Niederhuth CE*, Willing E-M*, et al PNAS, 2016; *co-first authors)**. Recently I have also used this approach to identify another methyltransferase, CMT2, as underlying unique features of CHH methylation in the grasses. I have been able to alter this pattern of methylation by transgenic expression of *Arabidopsis thaliana* CMT2 in the model grass *Brachypodium distachyon*. This has also resulted in phenotypic effects, showing that these discoveries can be exploited to generate novel traits.

In my lab I will take this research to the next level. I will continue to use comparative epigenomic approaches to identify key questions and guide me while incorporating new

sequencing technologies and methods to expand both the breadth and depth of my research. My work with CMT2, and the work of others, has also shown that epigenomes can be altered to create epigenetic diversity and potentially new phenotypes. These experimental approaches to "engineering" epigenomes can be used to test basic hypotheses of transcriptional regulation and to uncover hidden or cryptic genetic variation in genomes that was previously silenced. This will be an important complementary aspect to the other activities in my lab, both in testing hypotheses and in applying this knowledge to the larger issues we face. Below, I present brief descriptions of three major research areas and potential funding sources for each.

Genetic Basis of Epigenomic Diversity: (Target Funding Source: NSF MCB):

Observed variation in methylomes both within and between species (Figure 1) points towards genetic and mechanistic differences. Identification of the basis of this variation is important to understanding the epigenome interacts with other sources of variation to generate phenotypic diversity. DNA methylation is only part of the equation. A multitude of histone modifications, histone variants, and associated factors exist that all shape the epigenomic landscape. The natural diversity of these modifications have not been explored in the same way that DNA methylation has. Furthermore, many of these factors interact to define and maintain chromatin states. To more fully understand epigenomic diversity I will explore a range of epigenomic modifications using CHIP-seq and MethylC-seq. To complement this data, I will also identify open and closed chromatin regions using ATAC-seq and/or DNase-seq. This will be done for several species from across the angiosperms and for a diverse set of accessions within each species to identify both intra- and inter-species levels of variation.

It is clear that both cis- and trans-acting genetic factors underlie epigenomic variation (Figure 2). For example, a transposon insertion may act in cis to affect the chromatin state of a neighboring gene, while presence-absence variation of key enzymes in establishing the epigenome can have trans-acting effects at multiple loci. Technologies such as BioNano optical mapping and long-read sequencing technologies like PacBio or Oxford Nanopore will be used identify within species sources of variation that contribute to epigenomic variation. By combining these genomic and epigenomic data sets, the genetic basis of diversity for multiple epigenomic factors will be identified and integrated to provide a systems view of genetic and epigenomic interactions.

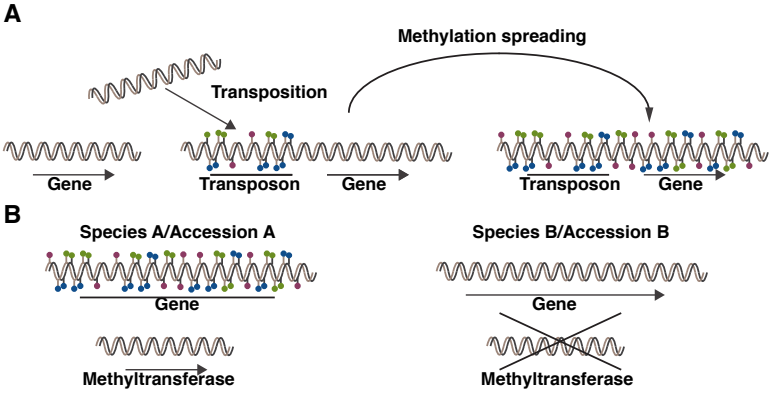

Figure 2: A) Cis-acting genetic variation from transposotion induces DNA methylation in a neighboring gene. **B)** In species/accession A, a functional copy of a methyltransferase leads to methylation of a gene in trans. In species/accession B, absence of the methyltransferase leads to a loss of DNA methylation at the corresponding loci.

The Epigenomics of Reproductive Modes (Target Funding Source: NSF PGRP):

Growing evidence in the field points to mechanisms of reinforcing chromatin states in germ-line cells, enabling the propagation of these chromatin states in the next generation. As a result of this reinforcement, random changes in chromatin state (epimutations), can occur and can be inherited through sexual reproduction. However, plants have evolved a multitude of different means of propagating offspring (Figure 3). In sexual reproduction there is variation in the degree of outcrossing and selfing amongst species. Some species have developed asexual means of reproduction either through seeds (apomixis) or by vegetative means, such as stolons

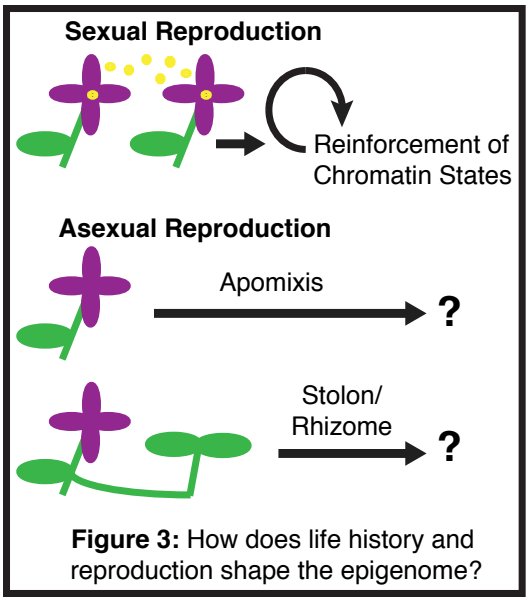

Figure 3: How does life history and reproduction shape the epigenome?

and rhizomes. In addition, humans propagate many species vegetatively through grafting or cuttings. These various reproductive strategies pose important, but unexplored questions regarding the epigenome. For instance, how are epigenomes shaped by these alternative reproductive strategies and how does this influence the evolution of the genome and phenotypes? How is the interaction of the epigenome and environment changed as a result? Can asexual means of reproduction be used to create, propagate, and utilize epigenetic variation?

In many species novel phenotypes arising from changes in chromatin state have been observed under tissue-culture conditions. These can also be induced through chemical demethylating agents and could provide novel phenotypes for crop improvement, but are rarely stable across sexual reproduction. Asexual reproductive strategies may provide a stable means of propagating these phenotypes. Asexual reproduction may also permit the inheritance of environmentally induced epimutations that confer adaptive advantages. Genomic and molecular resources are now available in species with multiple modes of reproduction, such as *Fragaria vesca* and *Boechera stricta*, that will allow me to address these questions. By identifying and tracing naturally occurring epigenomic differences between lines propagated by different reproductive means, the impact of reproductive mode on the epigenome will be assessed. Epigenetic variation will also be induced using chemical demethylating agents, generating populations of “epimutants” that can be further studied for stability under different reproductive modes. Finally, environmental stresses, such as heat or drought will be used to create environmentally-induced epigenomic variation which can be use to understand how the epigenome and reproductive strategy may contribute to adaptability.

Assessing Epigenetic Inheritance and Risk in Genetically and Epigenetically Modified Organisms : (Target Funding Source: USDA BRAG):

The future of epigenomics is not only in a mechanistic understanding, but also in the application of this knowledge. The ability to engineer the epigenome has already been demonstrated in human cells and such attempts are currently underway in plants. The development and use of epigenomic engineering methods will be a key part of my research program. However, there are many questions regarding these approaches that will need to be addressed, especially questions of inheritance and risk. Just as there has been and continue to be concerns about the safety of genetic engineering, there will be basic questions about the impacts of epigenome engineering. Concerns on the unintended epigenetic impacts of double stranded RNA was recently raised in discussions at the Cartagena Protocol, an international agreement governing the release of genetically modified organisms.

The goal of this work is to study the extent of induced epigenetic variation and create a framework for assessing the potential impacts of induced epialleles. This will be done both for genetic modification and epigenetic modifications. To assess the impact of genetic modification, I will examine the epigenomes of newly deregulated Papaya plants expressing viral coat genes targeting Papaya Ringspot Virus, kindly provided by Alan Chambers at the University of Florida. To assess intentional epigenetic modifications, I will study the effects of multiple epigenome modifying approaches such as chemically induced modifications from demethylating agents and targeted techniques that use CRISPR-CAS9 or TALE enzymes to target epigenome modifying enzymes. Not only will these studies provide a basis for assessing epigenetic modifications, but it will provide further insight into the behavior of epialleles and their inheritance.