Molecular Epigenetics in Evolution and Development

James J. Lewis

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Arts

In

Philosophy

Richard Burian

William Fitzpatrick

William Hopkins

Chris Haufe

June 25, 2010

Blacksburg, VA

Keywords: Evolution, Development, Evo-Devo, Epigenetics

Molecular Epigenetics in Evolution and Development

James J. Lewis

ABSTRACT

The dominantly held view in evolutionary theory focuses on gradual or punctuated change, primarily via natural selection, as the mechanism by which novel traits arise and evolution occurs. Noticeably absent from this portrayal of evolution is mention of the conservation of general characteristics, such as homologous morphological features or conserved nucleotide sequences, commonly observed across even distantly related groups at both the molecular and organismal levels. This raises at least the following questions: a) How does the evolution of conserved traits fit into an evolutionary theory that emphasizes change? b) What components of an evolving system provide the capacity for adaptation in spite of this apparent conservation of general traits? And c) How do these components affect the evolution of lineages? Here I suggest that heritable traits such as DNA methylation and histone modifications provide one place to look when addressing these questions. Current quantitative and population genetic models reflect the dominant view of evolution described above, and act as the foundation for both formal and informal descriptions and predictions of evolutionary change. Using results from recent work in molecular epigenetics, I consider the evolutionary implications for these traits, and show how current models of evolution fail to accurately capture this influence. In doing so, I also address some of the philosophical implications for how we conceptualize evolution, and what potential changes might be necessary for a more complete theory.

Table of Contents:

Chapter 1: Introduction. Epigenetic Systems and Evolutionary Biology

Despite the amazing variation among living organisms, it is conservation rather than difference which appears to be the norm. Evolution is often portrayed as being a process where organisms change to fit altering environmental conditions, with an emphasis on more successful groups changing to fit their environment while those less successful die off. Instead a review of many cellular and morphological characters indicates that it is most often a slight modification of a highly conserved process or trait that proves most beneficial—the "hopeful monster" (Goldschmidt 1940) is a tantalizing idea for many of us, but has little opportunity for success with numerous, almost insurmountable, obstacles in front of it .Thus the following questions are, in my opinion, perhaps the most important for a deep understanding of how evolution occurs: In what ways are these processes and traits conserved, and how do slight changes result in the wide degree of diversity found in the natural world. (Notice I said *slight* change, not gradual.)

Let us consider first cellular traits and processes, the conservation of which occurs at several levels (Gerhart & Kirschner 1997). First, nucleotide sequences can remain very similar between groups of organisms. Identity in nucleotide sequence ensures identical primary, secondary, and tertiary protein structures (assuming that environmental conditions are also preserved), though quaternary structure is not necessarily identical. If nucleotide sequence is different, the next level of conservation is amino acid sequence. Due to the redundancies in the genetic code, nucleotide sequence can change without altering the amino acids produced. Because protein conformation is often decided by number of specialized regions in the primary structure, similarity in protein conformations can also be conserved in spite of differing amino acid sequences, though there will be less similarity than is found in redundant nucleotide sequences. Finally, when looking for conservation of cellular processes, we can look to the function of the protein or protein complex. Even when protein molecules differ greatly in their structures, they can sometimes retain a similar function. This is perhaps the least successful form of conservation, for protein structure is intricately tied to function, but in cases where delicate bio-chemical processes are not necessary, a certain degree of dissimilarity can occur even when the function remains the same.

Though I have emphasized conservation of cellular and molecular processes and traits, similarities arise at higher degrees of organization as well (Wagner 1989). At the morphological level, homologs to the human arm and hand bones are an often cited example of how basic structures or plans are conserved between groups of organisms with slight modifications for adaptation to local environments and niches. But how do we get such great diversity if many or most processes and traits are conserved over large groups of organisms? This very issue was raised in the debate over punctuated equilibrium versus gradualism, and continues to be worth considering even though that particular topic has since been mostly resolved. In this and the last chapters, where I will return to this issue, I refer to this as the conservation problem, which I state explicitly as: "How does conservation of traits and processes in large phyletic groups fit into their evolutionary history, and what initiates evolution of phenotypic form in spite of or because of this conservation?" Answering this question requires more than the basics of evolutionary theory—we must also understand development. A detailed map of the

developmental processes found in nature is not necessary at this point. Rather, a few specific details indicate the correct direction for addressing this question, all of which relate to gene expression.

Gene expression at the molecular level is often described to students in introductory genetics courses, but rarely is any thought given to how and when *specific* genes are expressed from an organismal standpoint. Two key aspects to understanding the role of gene expression in phenotypic development are 1) the standard state of expression, and 2) the standard time of expression. Every cell in an organism contains the exact same set of DNA (excluding for the moment mutations in somatic cells during mitosis). This is true even in those cells that have specialized to perform specific functions dependent upon location and environment. As this suggests, the standard state for any non-housekeeping gene (e.g., not actin genes) is non-expression, and for most of the nucleotide sequences found in any given cell, expression rarely, if ever, occurs. Expression of these few genes also tends to occur in a condition dependent manner, where biochemical signals initiate expression of specific genes under the specific circumstances in which these signals arise. Taken together, the standard state and time of expression for genes results in liver cells in the liver, kidney cells in the kidney, and increased production of growth hormone during periods of relative youth, all of which occur as they should.

What does this mean for understanding evolution through conservation? It means that small changes to existing processes and traits can alter the phenotype to varying degrees by changing either the standard state of expression or the time of expression for a gene. Obviously, genes also change, and new alleles appear frequently. This is the basic evolutionary theory which I have already suggested is necessary, but insufficient, for explaining evolution. But an alternative to attempting to "fix something that is not broken" is to use it in another way. That is, expressing a normally dormant gene, or doing so earlier or later than normal (heterochrony), can accommodate the (sometimes necessary) conservation of a process or trait while initiating a potentially viable new phenotype.

My goal here is to consider the significance of one small part of the answer to the conservation problem for evolutionary theory. A purely genetic study of evolutionary history will by its very nature miss anything of evolutionary significance not based on nucleotide sequence alone. Thus here I consider the importance of epigenetic regulation in evolutionary biology. In particular I examine two mechanisms for epigenetic regulation of gene expression– DNA methylation and histone modification. As an intermediary between molecular genetics and higher level phenotypic and morphological traits, molecular epigenetics has the potential to act as a bridge between the genetic based studies of evolution that were the foundation for the modern evolutionary synthesis and the embryological, anatomical, and comparative studies which have returned to a place of importance with the advent of evolutionary developmental biology.

The issues with which I deal in this thesis fall into two overlapping categories: applied philosophy of biology and conceptual issues in evolutionary theory. The goal of any philosophical approach to biological systems (with varying ontological properties), as I see it, is

to help guide scientific studies and models by addressing the conceptual issues that provide the assumptions upon which other research is based, or models created. The application of philosophical inquiry to a specific set of biological systems, namely epigenetic systems, frees up the opportunity to approach this goal from a position based on recent research into relatively new features of partially unknown, but potentially great, biological significance in a way that addresses standard conceptual models. It also provides some suggestions for the development of the conceptual framework underlying a rapidly advancing field in evolutionary and developmental biology (as well as medicine and other related fields).

On a more concrete level, the topics I will address here are familiar to many, if not all, philosophers of biology. The development of the argument requires extensive biological knowledge, making it easy to lose sight of the traditional questions in the philosophy of biology. Thus, I will address the following philosophical debates (though some additional topics will be covered implicitly): The central dogma of molecular biology (Crick 1958, 1970), which has been the subject of numerous revisions, criticisms, and studies, both from biology (Hunter 1999, Mattick 2003, Morange 2006, 2008) and philosophy (Torres 1999, Thieffry & Sarkar 1998, Rosenberg 2006); the nature of selection and adaptation in evolutionary theory, which have received much attention in the philosophy of biology, where issues such as the replicator view of selection (Dawkins 1984), the levels of selection (Brandon 1982), and the units of selection (Sober 1984, Lewontin 1970) are frequently debated; and reductionism in evolutionary theory (Schaffner 1984, Nagel 1984, Ruse 1984). Finally, I will also address the relationship between macro- and microevolution, and the more recent suggestions that a new evolutionary synthesis is required (Pigliucci 2007). Recent work on epigenetic systems can be applied to many more issues in the philosophy of biology, but it will not be possible to address them here.

The primary molecular epigenetic mechanisms I will discuss (DNA methylation and histone modification) are not the only ones that could be considered: for example small interfering RNAs (siRNAs), DNA methylation and histone modifications are relatively well studied and provide important examples to support the claims I make about the importance of epigenetics for evolution and for philosophy of biology. These mechanisms tend to work together and thus provide an excellent set of model characters for discussion. But it suffices to focus on the two mechanisms I have chosen, for the others only add to and reinforce the morals that I will draw. I now provide a basic account of DNA methylation and histone modification necessary for understanding their evolutionary significance.

As should be familiar to readers of this thesis, in eukaryotic cells DNA is found primarily in the nucleus as a double stranded string of adenine, guanine, cytosine, and thymine (A, G, C, T) occurring in pairs (A-T or G-C) ordered as rungs of a ladder between two opposed strands of deoxyribonucleic acid, wound tightly around histone proteins and bundled into chromosomes, which are then attached to a homologous chromosome in diploid organisms. During periods of gene expression, the tightly bound DNA sequence is loosened in the region where transcription will occur, RNA polymerases bind to the promoter sequence, transcribe DNA sequence to RNA sequence, which may subsequently be altered and translated into an amino acid chain.

Bacterial and other prokaryotic genetics differ significantly from this picture, and the epigenetic phenomena involved are correspondingly distinct. I limit my discussion for the most part to eukaryotes, though epigenetics in the prokaryotes is well worth attention of its own elsewhere.

In eukaryotes, epigenetic modifications yield important additions at the molecular level to this simplified view of nucleotide sequence and protein synthesis by introducing heritable controls over expression of genes, or, more generally, of transcription of DNA and handling of DNA products (see Gilbert & Epel 2008 for an excellent overview). DNA methylation is the process where C-G pairs are methylated at the cytosine nucleotide. DNA methyltransferases identify the appropriate binding site along the nucleotide sequence, then attach a methyl group (-CH_3) to the 5-carbon of the cytosine molecule. Though a single methyl group will generally have little effect on the region where it is present, heavily methylated regions of DNA are significantly less likely to be transcribed than unmethylated sites (Holliday and Ho 2002). Where heavy methylation exists, the RNA polymerases that normally bind to promoter sequences just upstream of the start codon in a gene are unable to attach to the DNA strand because the methyl groups interfere with the binding apparatus. This results in the inhibition of transcription, and thus translation and protein synthesis (*ibid*.).

Returning to the view that the standard state for most genes is non-expression, this means that changes in methylation patterns act by altering the inhibition or activation activities for these sites. Many genes have operator sequences that accept proteins that, when they are available, attach to the operator sequence of the gene and either inhibit or stimulate binding of RNA polymerase. (The *lac* gene operon in bacteria is a common example of this.) Operons are found primarily in prokaryotes, though they are present in a number of eukaryotes as well. Enhancer and silencer sequences, however, perform a similar role in many eukaryotes by different mechanisms of action. Normally dormant genes are kept so through a variety of inhibition processes. Methylation of translation inhibitor sequences, such as silencer sequences, would result in the inhibition of the inhibition of the gene, and thus stimulate gene expression. Likewise, methylation of either the enhancer region or the promoter sequence results in non-expression of the gene by interfering with the expression activity which has overcome (usually) some form of silencing.

Histone modifications also act to alter transcription, though their mode of action is quite different. Core histone complexes are composed of two proteins that form histone subunits labeled H2A, H2B, H3 and H4. Thus, core histones are octomers around which DNA is tightly wound and then condensed into the familiar form of chromosomes. The linker histone proteins H1 and H5 promote binding of the DNA strand to the core histone complex. The histone proteins and DNA together are collectively referred to as chromatin. Every protein has a structure depending on its amino acid sequence and neighboring proteins with two terminal tails called the N-terminus and C-terminus (the N-terminus being the amino (-NH_2R) end which is first produced and the C-terminus being the carbon end (-COOH), where synthesis ends). Much like DNA and (some) RNA sequences, the terminal tails of histone proteins can receive a number of different chemical groups which alter the way they bind to DNA.

The main forms of histone modification are methylation, acetylation (-$COCH_3$), and phosphorylation (-PO), though others also exist. When these groups are added at specific sites to the protein tails, their conformation changes, thus altering their function (Zhang and Reinberg 2001). Histone H3, for example, often has three methyl groups attached to the lysine amino acid at position four and acetyl groups on the lysines at positions nine and fourteen, which are thought to be necessary for the start of transcription. Transcription initiation-enhancing activities by the histone modifications are due to the closeness of the histone-DNA bonds. More tightly bonded DNA is less available to polymerases for transcription, while more loose DNA is more easily bound by the polymerase binding sites and thus expressed more frequently. Just as the previous examples of histone modification result in transcription initiation, others inhibit transcription by stabilizing the histone-DNA bonds and restricting polymerase access (*ibid*.).

In addition to their separate activities, DNA methylation and histone modifications often work together to repress transcription (Martin and Zhang 2007). Besides directly blocking RNA polymerase binding, DNA methylation has also been found to recruit H1 histone proteins and histone acetylases to neighboring histones, both of which are capable of stabilizing the nucleosomes and further reducing the potential for transcription. Similarly, histone modification activity has been shown to recruit methyltransferases to nearby regions of DNA and increase the density of methylation in these sites. Thus both DNA methylation and histone modifications by themselves repress transcription potential for nearby DNA, but also work interactively to reinforce this repression activity.

The integration of developmental phenomena in general, and molecular epigenetic processes in particular, into evolutionary theory will require more than just a description of their activity in phenotype formation and their potential for heritability. Since the early twentieth century and the introduction of mathematical models to the study of biological systems, evolutionary biology has been heavily based upon the models of population genetics. Others, working primarily in plant and animal breeding programs, developed a second set of mathematical models for evolution based primarily on phenotypic characters, which resulted in the discipline of quantitative genetics. Both are based on the same fundamental principles of random assortment and segregation of chromosomes and the Mendelian concept of the gene, but each provides a slightly different characterization of the evolutionary processes of drift, selection, etc. Since their introduction to evolutionary biology, both population and quantitative genetics have been enhanced over time to include new concepts in molecular and evolutionary biology, such as molecular evolution and single nucleotide polymorphic alleles in population genetics, and quantitative trait loci in quantitative genetics. Nonetheless, the fundamental mathematics of each has remained more-or-less unchanged for almost a century, and the mathematical models which act as the paradigms for evolutionary biology must be addressed to proceed with the inclusion of developmental biology.

Though based on the same principles, population and quantitative genetics are not equal when considering the introduction of developmental phenomena into evolutionary biology. Population genetics is the study and modeling of genetic change over time due to the "forces" of evolution. It often considers only the most basic of scenarios, such as two or more alleles at a single locus, or two alleles at two loci, and focuses mainly on qualitative genes (genes which

have discrete forms of expression such as hair color). Furthermore, population genetics is just that, a study of the genetic structure of populations, with little concern for the phenotypic expression of these genes. It is instead assumed that the genes studied are expressed in some standard or normal fashion, and that the genetic basis of these traits is the most significant aspect of evolutionary change.

Quantitative genetics differs from population genetics most significantly in that it is primarily concerned with phenotypic traits, often a statistical distribution of such ranging over the whole population, which are the result of the expression of additive genetic effects with non-additive variance affecting the spread of the distribution. Though quantitative genetics also assumes that the most significant factor in evolution is the genetic basis of phenotypic traits, it models the genetic architecture underlying these traits by identifying the genes which affect the trait in question, then models effects of different combinations of alleles on the phenotype. Once a phenotypic value is given, other models describe the effects of selection, drift and so on, on the population or some subset of the population.

Correctly addressing the incompleteness of mathematical models for including developmental phenomena in evolution requires, first, choosing the appropriate models for review. Though population genetics is undoubtedly the primary method of modeling evolution, to attempt to point out the areas where these models fail would be futile. That is, population genetics is almost exclusively concerned with genetic change, while developmental phenomena are, by nature, phenotypic (with the exception of developmental genetics, which acts as a bit of a link between development and more standard genetics). Thus identifying the flaws of population genetics in this regard is both simple and impossible. To do so at length would be to create a straw man of the discipline, making it easy to attack, yet simultaneously rendering any claim irrelevant. Quantitative genetics, however, is almost perfectly suited to discussing phenotypic characters such as molecular epigenetic patterns, and it is this discipline which will be the focus of my critical inquiry.

My purpose here with regard to the mathematical models of evolutionary biology is not to "correct" them, but instead to identify their limitations in application to quantitative genetics, though I will in some cases make suggestions as to how these problems might be fixed. One problem for the project of identifying deficiencies in these models is that with the correct manipulation of their structure, they can be made to "fit" almost any data. This problem, known as the curve-fitting problem, is not a new problem to science, but nonetheless requires appropriate attention. Forcing a model to fit the data may be possible, but the goal of quantitative genetics is to model evolution, not to fit a data set. Thus, we must keep in mind the mechanisms and processes which underlie the various components of any model we use, and make certain that the components represent the actual processes involved as accurately as possible, rather than fitting our data sets well. I return to this issue in chapter 4 after I have discussed the deficiencies in the models in question.

In addition to the informal and formal components required to expand evolutionary theory to take fuller account of developmental and epigenetic processes, one final barrier needs to be addressed: reductionism. A number of well known philosophers of biology and biologists,

such as Rosenberg (1997) and Wolpert, have taken a reductionist stance in evolutionary biology. That is, they have suggested that everything about an organism can ultimately be reduced to the molecular level, and even to nucleotide sequence alone. This view then implies that the study of evolution through molecular genetic means alone is sufficient for a full understanding of the natural history of any organism or phyletic group. The remainder of this thesis will show why this position should be rejected. This is not, as such, the focus of my project, but it remains a necessary step in its completion. For the most part I will not address this issue explicitly since the implications of what I suggest in other areas are sufficient for the purpose, though I will return to this subject again in chapter 4, e.g., in discussing the curve-fitting problem.

In sum, the purpose of my project here is to consider the evolutionary implications of heritable molecular epigenetic processes and the impact they might have on the way we model evolution. Conservation of molecular, cellular, and organismal traits indicates that something beyond nucleotide substitution alone is behind evolution, and developmental phenomena are likely a large part of this missing component. In the following chapters I will address the formal and informal aspects of molecular epigenetics and how it impacts the quantitative genetic models of evolutionary theory in chapters 2 and 3. For pedagogical purposes, I will address heredity and heritability in chapter 2, saving the forces which act on variation within a population for chapter 3. In chapter 4 I will return to many of the issues raised in this introduction and provide a bit of a synthesis of the claims made in chapters 2 and 3.

Chapter 2: Epigenetic Heredity

I have divided evolution into two parts—heredity and forces acting on heritable information—in discussing the importance of epigenetic characters to evolutionary biology for two reasons. First, this division is pedagogically useful, since analysis of the effects of heritable information on the evolution of a group follows naturally from a discussion of its heredity. Secondly, and more importantly, I separate out heredity as a necessary component of any form of evolution, be it biological or not. Though evolution is often presented in a more complex form, the sole requirements for the evolution of any biological system are: 1) Heredity, 2) A source of variation, 3) Selection, and 4) Environmental stability or regularity (over varying amounts of time). These requirements need not be realized in the way we understand them today, with natural selection acting on a genetic heredity system. Rather, so long as we have inheritance of characters from one generation to the next, selection to reduce the least fit members of our population, and sufficient environmental stability that multiple generations can be produced, evolution can occur without necessitating specific forms of heredity or selection. This has become increasingly relevant as our understanding of the modes of selection in natural populations has grown to include numerous forms of selection, such as sexual and ecological selection.

This digression into the necessary and sufficient components of evolution highlights what may be the single greatest oversight in the growth of the modern synthesis during the mid-twentieth century: heredity does not have to be genetic, and there is little reason to believe otherwise, other than the application of Occam's razor. Despite the sufficiency of any form of heredity as a condition for evolution to occur, since the discovery of the double helix and the advent of molecular genetics, nucleotide sequences have been portrayed as the sole bearer of heritable information. The central dogma of molecular biology acts as the foundation of this belief. First formulated by Crick (1958) following the discovery of the double helix structure of DNA, the central dogma states that transfer of information in biological systems occurs unidirectionally, from nucleic acids to protein, and that information cannot be transferred from protein to protein or from protein to nucleic acid. Though this was later amended to include the special forms of information transfer found in RNA viruses, and retroviruses in particular, the basic structure of the dogma remains.

The importance of heredity in the evolutionary significance of epigenetic mechanisms cannot be understated. Most notably, the presence of heritable epigenetic information, by which I mean the physical instances of methylcytosine or histone modifications (similar to the nucleotide sequence information of a gene), could potentially undermine the central dogma, thus overthrowing a paradigm in molecular biology. For this to occur, however, epigenetic processes must be shown to be heritable. That is, the presence of epigenetic information and its ability to affect the phenotype must be shown to be heritable through non-genetic mechanisms.

This indicates the second most significant effect of the non-genetic heredity of epigenetic information—the loss of strict genetic reductionism in the biological sciences. Here I take strict genetic reductionism to suggest that all components of a biological system can be reduced to, or explained by, nucleotide sequence alone. Inheritance of epigenetic information through a genetic

intermediate would be of note, but could ultimately be expressed as a particular form of gene expression and thus be of evolutionary significance, but not sufficient for undermining the central dogma or reductionism. Thus it is important to distinguish between heredity in the quantitative genetic sense, where heredity is identified with transmission of phenotypic traits from parent to offspring, regardless of mechanism, and heredity in the molecular genetic sense, which requires a specific mechanism for the transfer of genetic information. Should non-genetic inheritance of epigenetic information be expressible in the latter sense, an expansion of evolutionary theory must follow.

In this chapter I will briefly discuss the different conceptions of heredity and heritability in evolutionary models, which necessitates a brief overview of the gene concept in several fields of evolutionary biology. I then consider how heritable epigenetic information best fits into evolutionary theory by considering quantitative genetic models of heritability and how heritable epigenetic information requires some alteration of these models. To retain the less formal significance of this information, I discuss both mechanisms of heredity for molecular epigenetic states and how they are induced through variation in environmental conditions.

Despite the often simplified view of heredity in evolutionary biology as the transfer of nucleotide sequence from parental to filial generations, understanding heredity depends upon the sub-discipline from which the researcher is working. That is, the concept of heredity changes depending on the field of study. Intrinsically tied to heredity is the gene concept, which describes the units of heredity and differs across disciplines. Similarly, the level of organization studied will alter the concept of heredity, where mechanisms and properties of heredity change between genetic and phenotypic, molecular, cellular, or organismal foci. Heredity is often portrayed as nucleotide sequence transfer by meiosis or mitosis. This is a molecular mechanism for heredity, but it does not describe the inheritance of traits or characters. To understand trait heredity requires understanding how traits are brought about in subsequent generations. A key claim of this thesis is that inheritance of nucleotide sequence is only one of several ways in which traits can be transferred from parental to filial generation. The basic focus of explaining heredity is not on nucleotide sequence transfer, but on explaining how correlations between traits in parents and the traits in their offspring are brought about.

Population and quantitative genetics began in the early 20th century and still reflect the biological knowledge of the time. Well prior to the discovery of the double helical structure of DNA and the development of molecular genetics, population and quantitative genetics inherited a gene concept based on phenotypic traits. The underlying structure of heredity was unknown at the time. Mendelian geneticists observed phenotypic variation and attributed this to genes, with alleles as the genetic variants on basic genetic forms. In population genetics, differences in qualitative traits were attributed to single (or very few) genes, where variation within these traits was due to new alleles for these specific genes. Note that phenotypic or chromosomal variation was necessary for positing new alleles. Quantitative genetics, being most concerned with quantitative characters with continuous (usually normal) distributions around a population mean, viewed genes as the hereditary units which provided additive variance to a trait. Again, observable phenotypic variation was necessary for positing the existence of new genes or alleles, though different combinations of alleles could produce the same phenotypic traits, and thus

complicated this matter. Since the rise of molecular genetics, both population and quantitative genetics have introduced some measures to capture the molecular basis of genes, such as quantitative trait loci (QTLs) or methods for determining the molecular genetic distance between groups, but on the whole they continue to adhere to modified versions of the pre-molecular gene concept when modeling heredity, selection, drift, etc.

Significantly different is the gene concept in molecular genetics. Here the gene is reduced to the nucleotide sequence found within a region of DNA, with adjustments to the exact sequence made to account for introns in eukaryotes (admittedly this leaves out accounts of regulatory regions and evolutionary molecular genetics to some extent). Often the function of the sequence is unknown, and putative genes are frequently used for the construction of molecular phylogenetic trees. This reductionist gene concept is also born out in positing genetic variants. While phenotypic variation was necessary for new alleles in population and quantitative genetics, single nucleotide polymorphisms provides a basis for recognizing alleles in molecular genetics. More than other disciplines, molecular genetics is particularly sensitive to the evolutionary distance between individuals or groups being studied. Single synonymous or neutral non-synonymous substitutions can act as the basis for new alleles in studies of closely related intra-specific groups, while allelic variants in studies of more distant groups often require a non-neutral amino acid change. In either case, however, non-molecular phenotypic variation is often less important for describing genes and alleles in molecular genetics, where amino acid substitutions often provide statistically insignificant phenotypic variation, but provide a method for tracking evolutionary trees.

Concepts of heredity for population, quantitative, and molecular genetics are similarly varied. Molecular genetics adopts the view of heredity given above, where in diploid organisms meiosis leads to chromosomal duplication, segregation, and the formation of gametes, which are then inherited from both parents after sexual reproduction. Population genetics assumes a similar Mendelian heredity, though it rarely addresses this subject in detail. Rather, population genetic models such as the equations describing equilibrium and linkage assume random segregation and assortment of chromosomes during reproduction, and identify equilibrium states (or deviations from expected values in the case of linkage disequilibrium) based on this assumption. Quantitative genetics take a different approach to heritability where it is the character or trait that is heritable, rather than the underlying genetic material. Unlike molecular and population genetics, heredity is not merely an assumption of other models, but is given a value based on the covariance between the mid-parent value and mean offspring value and the phenotypic variance of the mid-parent value. Under standard quantitative genetic assumptions, environmental variance is included, but equals 0, since the standard view holds that variance due to environmental conditions has a normal distribution with mean 0, and thus does not contribute to heredity. This reflects the initial subject of quantitative genetics--artificial breeding programs--where controlled environments are the norm.

For describing the uncontrolled setting of natural populations, environment induced phenotypic variance is much more significant in evolutionary quantitative genetics. Excluding concerns about epigenetic characters for the moment, the environment has been shown to induce highly significant phenotypic variation under a number of circumstances. For example, the

Nemoria anizonaria caterpillar mimics the flowers of oak trees after hatching in the spring, while those hatching in the summer resemble twigs. Even more extreme, should the colorful male *Thalassoma bifasciatum* fish die, one of the less colorful females in its cohort will develop testes and the accompanying phenotypic traits of a male (Gilbert and Epel 2008, chapter 1). More generally, we can divide this sort of phenotypic plasticity, or environment induced phenotypic change, into two types: reaction norms and polyphenisms. Though similar, the reaction norm indicates a continuous phenotypic range from potential genome expression, while polyphenisms are discrete phenotypes formed under differing environmental conditions. Reaction norms often include an environmental threshold, such as the horns of the male dung beetle, which only grow after a certain food threshold has been met, and then display a continuous range of lengths according to available food supplies beyond this threshold *(ibid.)*.

Whether due to a reaction norm or polyphenism, phenotypic variation in morphs from all of these examples display the impact the environment can have on phenotypic variation, and suggest that environmentally induced variation is often non-normal, or at least has a non-zero mean. Whether this requires its inclusion in quantitative heritability models requires one further observation: environments are inherited. Environments tend to be stable, but when they change, they often do so (relatively) drastically. Stable environmental conditions increase heredity of specific phenotypic traits, but reduce the stability of plasticity. With less selection in favor of rapid adaptability, any selective costs associated with plasticity for the trait(s) in question will, over time, eliminate plastic phenotypes if fluctuating environmental conditions are not experienced with sufficient frequency. That is, the greater the instability, the less heritable a phenotypic trait will be, while environmental fluctuations sufficient for inducing different polyphenic traits will favor or stabilize genetic support for a polyphenism for the relevant group of traits.

Fitting these pieces together with epigenetic characters requires a more formal study of heritability in quantitative genetics. Roff (1997) proposes three categories of heredity based on phenotypic similarity between offspring and parent, namely: 1) similarities due to environmental stability and environment-specific traits, 2) similarities due to parental phenotype which are transferred to offspring, and 3) similarities due to expression of inherited alleles. (1) and (2) could more properly be viewed as sub-types of environmental heredity, since, for example, both stable weather conditions and parental food choice are equally conditions of the environment during development. Regardless, Roff suggests that if we discard both forms of environmental heredity, heredity is reduced to similarities due to expression of inherited alleles. I will consider the importance of cases where these forms of environmental heredity are *not* discarded after a brief review of genetic heredity.

Heritability in the narrow sense in quantitative genetics is described as the slope of the relationship between the mid-parent value and the mean offspring value, written formally as: $h^2 = Cov_{OP}/V_{PAR}$. Here V_{PAR} is the phenotypic variance between mid-parent values, while Cov_{OP} is the covariance between mid-parent and mean offspring values. Using the simplest case of two additive alleles at a single locus without dominance while considering only genetic based heredity, in traits where males and female are similar, the covariance between mid-parent value and mean offspring value, as well as the variance between mid-parent values, will equal *pq*

(where p is the frequency of the first, or more frequent, allele and q is the frequency of the other allele), and the slope will remain 1 under equilibrium conditions. Including non-heritable environment influences that are usually modeled as an additional normally distributed value E, with mean 0 and variance V_E, the phenotypic variance will increase by V_E, while the mean offspring value will remain unchanged, giving a slope of $pq/(pq + V_E)$. Including non-heritable environmental variance forces the slope to be less than 1--decreasing the value for h^2 as V_E increases. One additional observation on genetic based heredity: Heritability changes according to allele frequencies, being greatest at p = q = .50 (Roff 1997, 25-27).

A more general notion of heritability, or heritability in the broad sense, identifies heritability with the "proportion of heritability attributable to both additive and nonadditive genetic variance" (Roff 1997, 30). The broad sense of heredity is often used to set limits on the value of heritability by including environment induced variance in the phenotypic value of parents and offspring, thus expanding the value of heritability according to genetic heredity to accommodate potential environmental influences and demonstrate the outer limits of heritability in the narrow sense. Here we can also suggest that the broad sense of heritability, even from a genetic position, must include environment induced phenotypic variance as an additive value, rather than the nonadditive component often attributed to such variance.

In circumstances where the environment will prove stable over a number of generations, and thus produce similar traits in plastic genomes that produce phenotypic variants in differing environments, environment induced variance will act more like an additive component. Rather than contributing to the breadth of the distribution of phenotypic values without altering the mean, under such circumstances the environment will contribute as an additive component by increasing or decreasing the phenotypic value of the organism in accordance to the effects of the environment. Returning to the case of horn growth in dung beetle, suppose for simplicity that the genetic component of phenotypic value is 0, such that the existence of any combination with the same number of positive of alleles for horn growth results in the potential for horn growth. Take positive phenotypic values as horn growth, and negative values as no horn growth. Because food availability is continuous rather than discrete like the presence or absence of certain alleles, we attribute a continuous scale to food supply, though we could simplify this somewhat such that below the threshold level any amount of food receives the value -1, while the continuous scale above the threshold point begins at 0 and increases as appropriate. Here any positive values for environment (food supply) induced variation will be added to the genotypic value to provide the phenotypic value. Where food supply is relatively stable, this additive environmental component would also be included in heritability calculations.

This sets the groundwork for including epigenetic mechanisms into quantitative genetic models of heritability. Like environment-induced variance in phenotypic value, epigenetic information is also capable of producing both additive and non-additive components to phenotypic value. Nonadditive variance from epigenetic characters would behave similarly to nonadditive environmentally induced variance, and could even be captured through inclusion in this component. Thus, here I am mostly concerned with an additive component to phenotypic value attributable to epigenetic characters. Addressing the additive component of epigenetic characters in heritability requires understanding the molecular basis of their heredity.

Here are some relevant (for my discussion) features of epigenetic mechanisms: 1) They can be divided into three (non-exhaustive) groups--DNA methylation, Histone modifications, and small interfering RNAs (siRNA). 2) Epigenetic characters essentially act as on-off switches or rheostats. These metaphors are particularly apt, as these traits tend to work much like the negation operator of logical switches, where the epigenetic character either negates (turns off) the expression of a gene, or enables the expression of a repressed region of DNA by binding to the suppressing site and hindering its expression (thus acting as a double negation) or as modulators of the intensity of expression when they regulate the rate or circumstances in which a messenger RNA is transcribed. This negation potential could also become extremely important in cases where additive regulatory effects are necessary for control (suppression or induction) of gene expression 3) Finally, DNA and histone modifications are enabled through a series of enzymes which can add and remove methyl or acetyl groups (depending on the enzymes) in a site-specific manner as various stages during development and the life of the organism. Should these switches be ultimately reduced to gene expression, we could simply allow for a more complex additive genetic component, perhaps by increasing the number of loci to include those for the epigenetic regulators. If epigenetic information is not independently heritable from one generation to the next, they would form a non-additive component to phenotypic value that would widen the distribution of values, but do little to change the mean. Thus it is important to demonstrate that these epigenetic characters are both irreducible to gene expression and heritable through means other than nucleotide transfer from parental to filial generations.

Molecular heredity of DNA methylation patterns can take two different paths--methylation patterns are inherited across cellular and organismal generations. Cellular heredity of methylation patterns is more easily recognized and explained than heredity across organismal generations, and relies on the mechanism of DNA replication. Though three initial hypotheses were suggested for how DNA replicates, a decisive experiment by Meselson and Stahl showed that DNA replicates semiconservatively by separating the two strands of the double helix, then using each of these as the pattern for a new copy of the opposite strand by adding the corresponding nucleotide to the opposing chain (Meselson and Stahl 1958). It has since been shown that when this occurs, methyl groups are retained in replication of the DNA molecule by a mechanism that allows the methylation enzyme DNA methyltransferase-1 (DNMT1) to recognize methylation sites. Essentially, where DNMT1 finds a hemimethylated site with the CG-pair of the template strand methylated, it completes the methylation pattern by adding a methyl group to the corresponding pair of the newly synthesized strand (Chan et al. 2004). In this way, DNMT1 can work its way down the synthesized strand and correctly reproduce the previous methylation patterns. This process, however, is not nearly as accurate as DNA replication, which has several “proofreading” units, so the error rate is considerable higher.

Though a strong commitment to the idea that acquired characters cannot be inherited by offspring has been the standard view in genetics and evolutionary theory since the rejection of "Lamarckism" and the development of the modern synthesis, DNA methylation is one of several forms of epigenetic character that is contrary to this commitment.

Non-genetic intergenerational heredity of these patterns requires that the methyl groups either be retained or replaced during gameto- and embryo-genesis through non-genetic means. This problem is compounded by the diversity among the ways in which methylation pattern are processed. E.g., in mice nearly all methyl groups are removed during embryogenesis, but in zebrafish, in contrast, they are retained (Macleod et al. 1999). Where methyl groups are retained during the processes of gamete formation and early embryonic development, their inheritance is quite simple. As they are never removed, methylation patterns are preserved by the same mechanism as standard cellular inheritance of methylation patterns. For species undergoing mitotic, or asexual, reproduction, the same mechanism for inheriting methylation patterns in cellular division is responsible for intergenerational heredity of these patterns. In many mammals, however, it appears that the removal of methylation patterns during the early stages of embryonic development is crucial for the proper development of the young organism; in these cases their heritability requires additional mechanisms (Kafri et al. 1992).

While most methylation sites in mammals are wiped clean during two stages of gametogenesis and early development, a select few sites remain methylated as before and act as the basis for genomic imprinting (Macleod et al. 1999). How non-imprinted regions of mammalian DNA are re-methylated is as yet poorly understood. Circumstantial evidence for heritable methylation patterns in non-imprinted regions of mammalian DNA is easy to find, and studies have shown that methylation patterns are fairly well conserved from one generation to the next, despite several periods during which methyl groups are removed from the original DNA molecules. One minor way that some methyl groups might remain is that, while the male methylation patterns are directly removed from the parental DNA strands, the maternal methylation patterns are indirectly removed through lack of restoration (*ibid*.). Rather than removal through enzymatic activity, these groups essentially degrade through numerous rounds of DNA replication during which they are not restored. A more likely mechanism for reliable methylation pattern heredity is that de novo methylation enzymes somehow recognize the appropriate sites for methyl groups and attach them accordingly (Athanasiadou et al. 2010)**.** This is poorly understood however, and more work remains before we can identify exact mechanism(s) of inheritance in this case.

Inheritance of histone modifications through either mitosis or meiosis is also poorly understood, though several hypotheses exist (Martin and Zhang 2007). Just as DNA is replicated during cellular division, so too must the chromatin surrounding the DNA molecule be copied. One potential mechanism for heredity of histone modifications is a semi-conservative model similar to the semi-conservative replication of DNA molecules, though in this case the histone tetramers are divided into dimers, and the corresponding dimers are added to the template proteins to replicate the histone protein. This hypothesis has been dismissed by several studies as a general method of histone replication, but has been suggested as a good possibility under highly specific circumstances such as regions related to gene activity. A strongly supported alternative template hypothesis builds on the discovery that old histone proteins apparently segregate randomly during DNA replication, with new histone molecules being attached between old histones where needed. In this hypothesis, the old histone molecules act as a template for the newly synthesized histone proteins such that neighboring proteins will develop similar modifications as those present on the neighboring template proteins. The last posited mechanism

for histone modification replication relies on DNA methylation, where methylated regions of DNA would direct methylation enzymes to specific genes where they would be recruited to modify the histone protein modification. Though some support for the recruitment of methylation enzymes for histone modification exists, this model is not universally accepted, for a number of organisms possess histone methylation without the necessary corresponding DNA methylation (*ibid*.)**.**

Preservation of histone modifications through meiosis builds on mitotic inheritance. Though the mechanism for the preservation of histone modifications during mitotic division is still an important and poorly understood step, the remainder of meiotic heredity of histone modifications merely requires the preserved histone proteins transferred from parent to offspring. As gametic chromosomes are bound by chromatin, standard meiotic inheritance mechanisms will result in chromatin heredity.

I now return to examining the role of environment-induced variance in natural populations. As noted above, though nonadditive phenotypic variance, which expands the width of a phenotypic distribution, is the standard portrayal of environment induced variance, in some cases it might be more appropriate to include an additive component to phenotypic values attributable to the environment. In addition to the examples of plasticity in caterpillars, fish and beetles I described above, the environment can play an important role in molecular epigenetics. At least two forms of environmentally induced change in methylation patterns in mammals (described below) have been recently studied. Like the mechanism(s) of heredity, little has been done to better understand the role of the environment in histone modifications, and how this might alter phenotypic expression. None the less, the following examples of environmentally induced change in methylation patterns should at least indicate the possibility of an environmental impact on histone modifications.

In agouti mice, a mutation produces a distinct yellow fur phenotype and obesity. Although this trait is dominant (over the wild type *agouti*), the effects of the causative allele can be masked in offspring by the maternal diet. When the dam feeds on foods rich in methyl-donor supplements such as folate and choline, the offspring retain the wild type coloration and obesity is reduced (Gilbert and Epel 2008, chapter 2). A "kinky tail" trait, in which the tail is bent in several places, has also been attributed to methylation differences between wild type mice and kinky tailed mice, though this trait appears to be heritable from either parent and may not be as dependent on the parental diet. More importantly, these epigenetically induced traits are heritable--F2 mice of the methylated a*gouti* parent display the same phenotypic traits as the grandmother, despite cessation of methyl supplement feeding by the mother (*ibid*.).

Maternal behavior has also been shown to alter offspring methylation patterns in mice, which induces noticeable phenotypic variance. Stress response in mice has been associated with glucocorticoid receptors, with more receptors providing an increased capacity to decrease adrenal hormone production during stress management. The Egr1 transcription factor, which is part of a glucocorticoid gene enhancer, has been found to be methylated in all mice within one day of birth in rat pups. Pups which receive significant attention, such as grooming, during early post-birth development tend to lose this methylation pattern, enabling the Egr1 transcription

factor and allowing the enhancer region to produce additional glucocorticoid receptors in these mice (Gilbert and Epel 2008, chapter 2).

Examples of changes in methylation patterns due to environmental conditions do more than indicate the importance of heritable epigenetic characters--they demonstrate that environment induced variance can be more than a nonadditive component to heritability. Furthermore, the stability of the environment must be taken into account for models of heritability spanning numerous generations. The methylated *Agouti* mice displayed the phenotypic characters of the maternal ancestor for several generations. Glucocorticoid receptor production, however, seemed to require repetition of grooming behavior for each generation. Thus additive environmental variance could be due to methylation patterns transferred through the germ-line for several generations, or could be due to stable environmental conditions (such as continued care for subsequent generations of offspring). In either case, such variance does more than expand, compress or provide and interesting outlier for the distribution of phenotypic values. Rather, it acts as an important component in modeling the heritability of these traits. The true importance of this, along with other aspects of heritable epigenetic characters, will only become apparent once we consider how selection, drift, and other forces act on populations.

Chapter 3: Epigenetic Systems and the Mechanisms of Microevolution

I begin this chapter by discussing the forces acting on variation within and between populations and the sources of this variation—mutation and, to a lesser degree, migration. In classical population genetics, mutation acts as the sole source of new genetic variants. Broadly construed, mutation includes single nucleotide changes (transitions and transversions), insertions, deletions, translocations, inversions, etc. Mutation rate, described in terms of mutations per DNA duplication through mitosis or meiosis, gives the likelihood of a new allele arising through replication error or some other process (such as mismatch alignment or unbalanced recombination of chromosomes). Unlike the previous case of heritability, here we can assume that under most circumstances mutation will be modeled in quantitative genetics much as it would be in population genetics. The effects of epimutations, however, are quite significant for quantitative genetics, and will be discussed shortly.

Methylation patterns and histone modifications are just as capable of change as the more recognizable nucleotide sequences. Quite generally, epimutations in epigenetic patterns occur with a much greater frequency than in the nucleotide cases, but require significantly more mutations for noticeable effect. That is, often a single methyl group will not alter the function of a nucleotide sequence, and it is only in heavily methylated regions that change in transcription and translation occurs. The specifics of the molecular events related to this process are important for understanding how mutation in epigenetic patterns produces a corresponding change in phenotypic expression, and how this can alter the evolutionary history of a lineage (Holliday and Ho 2002).

As noted previously, remethylation of cytosine bases after DNA replication relies on the semi-conserved process of meiosis, where a double stranded DNA molecule is split and two new complimentary strands are created using the split strands as templates. When the appropriate cellular apparatus(es) find(s) a hemi-methylated region of DNA at a replication fork, they correct the differences between the two new segments of DNA by adding a new methyl group to the G-C pair corresponding to the mC-G pair. So if the cytosine of a CG pair is methylated on the template strand, the cytosine attached to the guanine of the template strand will be methylated, thus completing the methylation pattern at this site. (This also suggests some insight into why CG pairs or CnG [Cytosine-nucleotide-Guanine] triplets are often the site of methylation.) While the process of DNA replication through the use of a template strand is extremely accurate--due to several cellular mechanisms for DNA repair when a mismatch occurs--the process by which hemi-methylated regions receive additional methyl groups is much more prone to error (Holliday and Ho 2002). In fact, the mutation rate for methylcytosine transitions is 10 times that of other transitions (Holliday and Grigg 1993), which is one reason perhaps why methylation patterns tend to occur as a region of heavily methylated DNA rather than a series of individually methylated sites sparsely spaced throughout the genome. (Another reason is that a single base mismatch will often be repaired by the binding apparatus for transcription. In contrast, clumping of methylation is necessary to modulate transcription effectively, so isolated methylation changes do not typically activate any repair mechanism.)

The addition and subtraction of various chemical groups to histone proteins works in a similar way to the process of adding methyl groups to cytosine. A number of enzymes are recruited to attach or remove the appropriate chemical group. Repair of histone proteins to their unmodified state can also occur through the replacement of the entire nucleosome. In addition to this basic process, modified histone proteins can also recruit a second set of enzymes that promote transcriptional memory retention through mitosis: the polycomb proteins, for example, possess two types of activity. First, these proteins act as a histone methyltransferase and attach methyl groups to the histone tail of the H3 histone protein (this also involves removing any acetyl groups currently attached that might promote transcription), while the second set of polycomb proteins binds to this methylated region, prevents transcription, and recruits similar agents to neighboring nucleosomes to create a densely packed chromatin region that inhibits transcription (Gilbert and Epel 2008, appendix). The exact error rate of these processes is unknown, but due to evolutionary benefits described below, the fidelity of these processes is likely to remain relatively low in order to allow rapid transitions when the conditions reinforcing epigenetic control of repression, or a block of gene expression, are no longer present. Accordingly, these sorts of epigenetic control can produce significant change over evolutionary time only when the selective conditions are sufficient for the retention of these controls (i.e., selective conditions are both retained for a sufficient duration and remain focused on the same life cycle stage[s] for this duration). Additional complications arise due to the recruiting ability of both modified DNA and histone proteins, each of which can influence the modifications to the other (Zhang and Reinberg 2001). Thus, the rate at which histone modification occurs depends in part on the presence of DNA methylation patterns, and vice versa, leading to an increased likelihood of errors in regions where more modification activity--either methylation of cytosines or chemical modification of histone tails—occurs.. (This is not due to a higher probability of error, per se, but instead to the probability of a specific event occurring given an extremely high number of repetitions.)

The relatively high rate of error of these mechanisms as compared to the rate of mutation in nucleotide sequences has significant repercussions for gene expression. Intra-generational cellular heredity patterns have the potential for altered cellular fitness values due to changes in epigenetic patterns in modular regions of the organism (i.e. organ-specific cellular lineages have the potential to develop post-embryonic functional alterations due to changes in epigenetic patterns within ancestral cells and to respond rapidly to local selective regimes favoring particular DNA expression patterns). Epigenetic change in totipotent stem cells could increase this potential, but, since alterations of this sort this would affect numerous regions and bypass the efficiency and security of modularized organismal components it is likely that they would face greater constraints. Most changes in epigenetic patterns of this sort are likely to prove deleterious (resulting in cancerous growth, cell death, etc.), but just as with genetic mutations, the possibility of beneficial change in gene expression due to an epimutation increases with the frequency and number of opportunities for such change such that the probability of a beneficial epimutation occurring in a somatic cell lineage is very high over the lifetimes of all members of a population through multiple generations (this potential is only increased by the vastly greater number of somatic cells in an individual compared with germ cells when considering the potential for evolutionary change of this sort with regard to the more traditional change in germ-line nucleotide sequence). Though such changes would be absent from germ-line transfer of

nucleotide sequences (and thus be limited in their evolutionary potential from a genetic perspective), intra-generational changes combined with hereditary behavior patterns, such as the maternal grooming behavior in mouse and rat pups, and long-term change of gene expression in derived lineages of such an organism could lead to change in phenotypic population structure.

Epimutations of germ-line nucleotide sequence or histone molecules conveyed through the maternal germ cells have greater potential for evolutionary significance given a direct mechanism for transfer from parental to filial generation, though the likelihood that they will occur is lower due to the limited opportunity of the relatively low frequency of changes in epigenetic patterns. None the less, the same logic as above persists, where a large population taken over numerous generations will still have a moderately high probability of seeing germ-line epimutations. The result of these changes in epigenetic patterns will be similar to the somatic cell story, but have the potential to behave more like changes to epigenetic patterns on totipotent cells, as the initial cells that form the single celled embryo become the basis for all other cell formation and later specialization. For both somatic and germ-line cells, modification to epigenetic patterns on or around microRNA encoding regions of DNA (here I include most forms of regulatory RNA as well as tRNA and mRNA, etc.) could act as a second-order regulatory apparatus by regulating the expression of such transcripts and stopping or reducing their availability to gene transcripts, though more study of RNA regulatory sequences is required to evaluate their potential for epigenetic influence.

Just as mutations in epigenetic patterns will result in the introduction of new phenotypic traits into a population, migration of individuals carrying specific epialleles will introduce extant epigenetic patterns to new populations, or alter the frequency of such patterns should they already be present. From an evolutionary quantitative genetic perspective, little will be different regarding epigene flow from the standard population genetic concepts of immigration and emigration, though the shift toward the meta-population phenotypic mean corresponding to random migration will not be borne out in genetic traits, as the genetic basis of these traits does not necessarily correlate with the epigenetic patterns regulating them. That is, the frequency of an epigenetic pattern correlated with an allele can increase within a population without a corresponding change in allele frequencies. The result of such change would then be a shift in the population phenotypic mean towards the meta-population phenotypic mean due to a change in the additive non-genetic variance component of the mean phenotypic value equation I have described above.

I turn now to the forces of evolution often considered to act most strongly on populations--selection and drift. Often considered to be the "strong" force in evolution, natural selection easily proves to be the more significant cause of evolutionary change (or stasis) when population conditions are optimal for continued existence. When conditions are very unfavorable, however, drift frequently becomes an extremely important factor in understanding the evolution of biological groups at numerous levels. Taken together, drift and selection are sometimes considered to be the basis of all evolutionary change above the population level, though this is traditionally held to be the case only when sufficient variation due to mutations exists.

The basic theory of natural selection, as described by Darwin, implies that there are evolutionarily important limits on population size, and that one reason for the effectiveness of selection is that not all organisms in a population can survive. By the end of the 18th century, Malthus had written that (human) populations will, under most conditions, quickly grow to outstrip the available resources and thus will be forced to remain within resource limits. Darwin and (later) Wallace generalized this premise and took it to its logical conclusion in the theory of evolution by means of natural selection, arguing that those members of a population most able to survive and reproduce will be those most likely to bear offspring and (to put it somewhat anachronistically) retain their hereditary material within the population. Over time, the distinctive alleles of the least fit members of society will be reduced in frequency until they are present in a very small fraction of the population; unless the selection is very strong (or the effective population size very small), less fit alleles will remain within a population at very low levels rather than being completely eliminated.

Though this view of natural selection has been modified somewhat since Darwin (especially since the development of genetic theory), the basic structure remains the same, with the main exceptions being that selection can now be divided into three primary forms and a more formal conception of selection has since been devised. More specifically, selection is now seen as either stabilizing (selection for the continued existence of a trait within a population), disruptive (selection favoring two distinct traits without an intermediate), or directional (favoring a general shift in trait distribution in a population structure from a less fit trait toward a more fit trait). Since I will be primarily concerned with changes in the phenotypic mean of a population, I will restrict my discussion to directional selection as the other two will not often affect this value (though this does not mean it will not later become important to understand better the relationship between these other forms of selection and epigenetic patterns).

In quantitative genetics, directional selection is described with the equation

$$R = h^2 S$$

where R is the response to selection, S is the selective differential and V_p the phenotypic variance of the trait. Using hard selection, where any member of a population that exceeds a selective threshold will survive, the selective differential can be taken in term of the phenotypic standard deviation such that

$$S = i\sqrt{V_p}$$

and *i,* the selectional intensity, is

$$i = \frac{S}{\sqrt{V_p}} = \frac{z}{p}$$

Here *z* is the threshold point for selection and *p* is the proportion of the population selected (equations from Roff 1997). Thus the selection for a give trait over a single generation is equal to

the heritability of that trait multiplied by the product of the selective threshold divided by the proportion selected and the square root of the phenotypic variance. The strength of selection can vary significantly depending on the values for these components--low heritability will obviously produce minimal selective effects, since as we approach zero so does the selective response. High heritability, however, can still be moderated if the threshold point is low and the proportion of organisms that are above (or below) the threshold is high. Similarly, a more genetically homogeneous population will also less responsive to selection even for highly heritable traits.

Returning to an earlier point, the importance of understanding the heritability of a trait--including the epigenetic factors involved--is demonstrated in the above equations. Of the variables given, heritability has the greatest potential by itself to affect the strength of selection, so accurately describing heritability is crucial to predicting or explaining the effects of selection on trait distribution. Thus including the influence of heritable components due mainly to environmental factor, such as heritable epigenetic patterns, in our definition of heritability can significantly alter our model of selection. While heritability decreases as the value of V_E increases under the standard models, we actually see an increase in h^2 if some of this variance is due to heritable environmental factors.

Less formally, epigenetic patterns such as methylation of a gene of interest affecting a given trait have the potential to alter the selective pressure on that trait by either affecting the heritability of that trait, or by changing the proportion of the population that resides above the threshold for selection. That is, in addition to altering the heritability of a trait, a heritable methylation pattern could also mitigate the impact of slightly deleterious alleles, with the result that more members of a population meet the selective threshold. A side-effect of this process would be a slower rate of decline of slightly deleterious alleles and increased difficulty in removing them by selection. It is possible that the same could occur for highly deleterious alleles, if they were completely inhibited, but this is not very likely since the rapid rate of change in epigenetic regulation makes it unlikely that such a system would remain stable over a long term. Given the relatively fast turnover of epialleles, such deleterious genes would not likely survive long.

The effect on selection through change in heritability is in many ways the least significant effect of epigenetic patterns on evolution by natural selection. To be sure, it has the potential to be the most formally important point, as it changes our understanding of how we can model selection and thus alters our predictive and explanatory tools. But the functional potential of epigenetic patterns for reacting to selection pressure is perhaps more relevant from a population or species perspective. Here I suggest epigenetic patterns bear the potential to act on such groups in the following (likely non-exhaustive) ways: 1) Epigenetic patterns provide additional targets of selection, at multiple levels, 2) Epigenetic patterns have the potential to act as mediators of rapid response to shifting selection pressures, and 3) Epigenetic patterns have the potential to generate novel traits.

As described above, epigenetic changes are brought about at many locations within the organism, and by a variety of mechanisms. Furthermore, while individual epigenetic modifications can act as transcriptional regulatory devices on their own, in some cases they can

also recruit modifying enzymes to other viable sites for epigenetic modifications. Thus this regulatory change can be activated through either pathway, and can be potentially altered by the other.

This also suggests the potential for varying degrees of effect, with far greater import than the "on/off" switches for genes or gene cascades that one often encounters as the basic metaphor for development in developmental genetics. Varying degrees of transcriptional regulation controlled interactions between the two epigenetic mechanisms of transcription-inhibiting mechanisms already discussed and in combination with other mechanisms, such as control of translation by microRNAs, not discussed here). Obviously complete absence of epigenetic patterns would be a default state where transcription could occur as it normally would in the absence of these regulatory systems. Both the mechanisms of epigenetic change on which we have focused here have similar effects on transcription, but the specific kinds of control they provide and the extent of the DNA regions whose transcription they regulate differ considerably. While methylation of DNA cytosines requires the addition of numerous methyl groups to stretches of nucleotide sequence to be able to affect the expression of a single segment of DNA, histone changes can increase or decrease the transcriptional potential for larger regions of DNA. Thus one mechanism more readily allows for selective regulatory action on individual sites, while the other can more readily blanket a region (potentially containing a number of genes). A second pathway appears when we consider the cell-lineage, and sometimes individual cell, specific nature of epigenetic modifications. While genetic change is distributed equally to all cells within the organism, changes in (for example) methylation patterns can be limited to specific locations. Should expression patterns be altered through epigenetic change in specific regions of specific tissues, localized effects could be generated, or a gradient could be formed.

This suggests the second functional capacity of epigenetic modifications--not only can they work at different levels of selection and in different mechanical ways, they are also (relatively) easily reversed, particularly when compared with the difficulty of reversing most genomic changes. This is not to say that other, easily reversible, mechanisms for genomic modification haven't also evolved, but rather that this is one way of obtaining reversible change. Functionally, epigenetic changes provide a rapid and reversible response to selection pressures and a more labile response to frequently shifting environmental conditions than genetic changes with similar effects. Reversible change is a definite advantage in rugged adaptive landscapes, and even more so when rapid shifts in environment (and thus in the structure of the adaptive landscape) occur as is often the case where polyphenisms have been evolved. As noted by Wright, there is a significant problem with crossing from one adaptive peak to another if the landscape shifts or the optimal point is located on another peak that is difficult to access by genetic mutation. Reversible change to selection pressures can facilitate "exploratory" modifications, with the potential for more permanent genetic change following a successful "search".

A second case where reversible change is potentially beneficial concerns repetitive shifts in environmental conditions, such as wetness conditions. This sort of situation has produced heterophylly in some aquatic plants, where they have developed a plastic response to the environment depending on whether they are submerged, partially submerged, or fully exposed to

the air (Schlichting and Pigliucci 1998). Epigenetic controls enable Leaf structure to change to fit immediate environmental circumstances while maximizing potential for photosynthesis in each circumstance. Plasticity is generally treated as a genetic trait, but the potential for epigenetic-induced plasticity is there. As with the other case above, the reversible nature of epigenetic patterns mean that expression patterns could be changed under specific environmental conditions, and reversed if these conditions are not met later. Again, this also need not require a permanent plastic response, and can facilitate "exploration" of the adaptive landscape. Finally, this sort of response to repetitive environmental shifts could also lead to a more permanent genetic basis for plasticity, or could be integrated into a genetic pathway.

Another way in which epigenetic patterns could be selected for, or result in adaptation to selection pressure, is through the genesis of novel traits. This is by far the most evolutionarily significant result of epigenetic change, and also the least likely to occur. If it were to occur, however, the potential for macroevolutionary change in only a few generations suggests the importance of this potential. Such changes have often been considered as "hopeful monsters" by evolutionary biologists, but recent studies of the evolution of developmental genes in drosophila suggests that such monsters are not impossible (Lohr et al. 2001), though obviously they would occur rarely and generally be unsuccessful.

Nonetheless, it should now be clear that modification of developmental gene expression by epigenetic change could produce new morphological traits that differ greatly from previous phenotypes. While the amino acid structure of fully expressed developmental genes is important, it is often the timing of expression and location within the embryo that is most significant. Cell-specific or region-specific changes in expression patterns could alter morphogen gradients and result in an altered developmental pathway. A change in time of expression would act similarly. Most of these changes would, as noted above, be deleterious, but should any be beneficial the outcome is not merely a new variant within a population, but an organism with the potential for producing offspring that yield an entirely new species while retaining genetic compatibility with other members of its population and the ability to produce back-crosses that keep the offspring, with their modifications, within the population. (This solves one of the problems of a genetic hopeful monster.) Again, beneficial adaptation through epigenetic change could be assimilated in later generations to a more permanent genetic state.

After natural selection, drift is generally the most important factor in evolutionary change (though this has been questioned), and in specific circumstances such as small population sizes or bottlenecks, drift can be even more influential. Drift is essentially a change in allele frequency within a population due to random chance or sampling error. It is especially prominent in a small population, in which the removal of any single individual's genotype through chance alone would result in a shift in allele frequencies. In the simplest case of two alleles at a single locus, drift is the result of randomly combining all alleles from the population for this locus, then selecting a certain proportion of them as representative of the next generation. Selection has no bearing on this result since the pairs are randomly selected. In large populations drift will have little effect on allele frequency. Once population size becomes small, however, the odds of randomly eliminating more of one allele than another increase greatly and drift can have a significant influence on a population. If, for example, population size was reduced to only two

heterozygous individuals of opposite sex, there would be a 50% chance of homozygosity in any offspring from this pair.

Drift is a special case with regard to epigenetic patterns. Under most circumstances these patterns will not significantly alter the effects of drift on the organism, nor will drift act significantly differently on these patterns than drift would act on nucleotide sequences (with exceptions). Because methyl patterns and histone modifications are found on both sets of parental chromosomes, drift will occur as expected, with the exception of time- and location-specific patterns, such as the methylation-induced maternal effects found in mammals. In this exceptional case there is a single line of transmission, similar to that of mitochondria, and drift is less likely to play an important role, if any role at all.

As with genetic drift, populations experiencing bottlenecks, or founding populations of only a few individuals are likely to have a significantly reduced number of epiallelic patterns—potentially leading to fixation of a single epigenetic pattern. However, drift is less likely to behave similarly in the long term for epigenetic patterns as it does for genetic conditions. While population genetics has the neutral theory as a model for the long-term effects of drift in normal-sized populations ('normal' is relative to the species), it is unlikely that a similar theory will apply to epigenetic patterns. Since histone modifications tend to be inherently discrete, limited to a small number, and evolutionarily unstable (Kubicek et al. 2006), it is highly unlikely that any pattern of histone modifications will endure long enough for long-term effects of drift in the absence of positive selection favoring that pattern. In cases in which histone modifications prove sufficiently beneficial, it is likely that more permanent genetic modifications would be incorporated into that region and that the selective benefit would overpower any effect drift had on the locus.

The same situation is not quite as likely for methylation patterns, since cytosine methylation is only effective as a transcription inhibitor when large numbers of cytosines have been methylated within a single exon or genic region. As noted before, methylation patterns have a mutation rate similar to that of nucleotide mutations (Holliday and Ho 2002), but this is understood as an epimutation resulting in the presence or absence of the pattern, rather than individual methyl groups which are removed or added at a much greater frequency. While it is still too early to know, the effectiveness of transcription inhibition by methylation patterns offers the potential for limited effects due to long-term drift. If this inhibition effect is a threshold trait, then we can think of the pattern itself as discrete, and as with histone modifications it is unlikely that drift will have any effect over long periods of time. To clarify this point, since the rate of gain or loss of the entire pattern is roughly equal to that of mutations in a gene, the amount of time any single methylation pattern is extant will be limited. If methylation patterns are truly an "on/off" switch, it is probably the case that, due to the large number of generations required for drift to act on a population to any significant degree, the methylation pattern will not remain present for a duration sufficiently long for drift to effect major changes in population structure. If, however, there is a gradient in the effectiveness of these patterns on transcriptional regulation--with increased methylation promoting greater inhibition of DNA-binding, the potential for drift remains. In such a case, numerous epialleles could exist at each site, limited only by the range at

which the effectiveness of inhibition given a specific number of methyl groups is no longer distinguishable.

Chapter 4: Conceptual Issues for Evolution and Development

The basic view of epigenetic systems I have developed here suggests the potential these systems have for inducing evolutionary change and affecting its direction. I have covered the molecular and functional properties of these systems, so I now return to the focus of this project--the philosophical and theoretical issues raised by epigenetic systems and the roles they play in evolutionary and developmental biology. The focus of the previous chapters was primarily microevolutionary in perspective, and more technical than conceptual. With this in mind, I turn from a biologically-centered discussion to one which integrates philosophical work from both philosophy and biology.

More specifically, I now focus on three main topics of interest: Crick's Central Dogma of molecular biology, the units and levels of selection, and the issue whether microevolutionary forces are sufficient to account for macroevolutionary events in evolution and development. As I will suggest, the failure to complete the Modern Synthesis program suggests the need to discard or refine the modern evolutionary synthesis in favor of a more complete synthesis specifically incorporating epigenetic mechanisms of evolutionary innovation. I conclude by briefly reviewing the discussions in this and the preceding chapters and addressing the anti-reductionist implications deriving from these arguments. In closing, I will briefly mention a few potential ways for epigenetic research to move forward, not just in the biomedical fields, but also in the study of evolutionary epigenetics.

The Central Dogma of Molecular Biology

First formulated by Francis Crick in (Crick 1958), but later revised and clarified (Crick 1970), the *Central Dogma of Molecular Biology* (hereafter just "Central Dogma") has become one of the most important tenets of molecular biology and evolutionary theory. As intended by Crick, the Central Dogma is a purely negative claim about the flow of *sequence information* in biological systems. Crick (1970) explains that the Dogma is the result of analyzing the nine potential pathways of transfer of sequence information, then determining which of those are expected *not* to be actual methods of information transfer. Crick divided nine potential pathways into three types. Type (i) consisted of sequence information transfer from DNA to DNA, DNA to RNA, RNA to RNA, and RNA to proteins. Type (ii) transfers were limited to RNA to DNA and DNA to proteins, while type (iii) consisted of the remaining possibilities, sequence information transfer from protein to protein, protein to DNA and protein the RNA. Of these, on the basis of the evidence available at the time, Crick considered type (i) to be well-supported, type (ii) transfers to be rare if they occurred at all, and type (iii) transfers not to occur. Crick (1970) notes that type (ii) transfers were not theoretically unlikely, but had not been seen, while type (iii) transfers were faced with problems due to the chemical reactions required.

In his 1970 paper, Crick also discusses several mistakes about the Central Dogma in the then-current literature, one of which is of particular interest here. He suggests that several authors have confused his negative claim about sequence information transfer with the separate positive claim found in the sequence hypothesis, according to which there is a transfer (though in

some cases his meaning is better understood as a mapping) from nucleotide sequences to amino acid sequences. Thus, according to Crick, the Central Dogma claims only that one type of information transfer *doesn't* occur (and another is unlikely), and makes no claims about how sequence information is transferred. He is also careful to point out that when he discusses the transfer of sequence information, he is only referring to the residue by residue replacement of one sequence string by another according to a set of defined alphabets. He excludes all conformational information from this information, noting in particular that he is reducing proteins to amino acid sequence only under the belief (at the time) that protein structure was determined primarily by amino acid sequence alone. Though he does not discuss three dimensional nucleotide sequence structure, he implicitly suggests the same for the cases of DNA and RNA.

Here I will examine two issues about the Central Dogma. First, I will review some of the changes that subsequent biological findings impose on Crick's version of the Central Dogma, focusing on how epigenetic systems may pose such serious problems for it that it should be abandoned rather than continuing to add exceptions to it in special circumstances. Second, and perhaps more important, I will review how a strong interpretation of the Central Dogma (significantly different than Crick intended) has influenced molecular biology and genetics. This interpretation makes some of the mistakes Crick specifically treated as misconstruals in his 1970 article, but it rests on common conceptions amongst biologists about the nature of information transfer in biological systems and thus deserves attention as the currently orthodox version of the central dogma.

I will not elaborate further on the general structure of Crick's version of the Central Dogma. It is worth noting, however, that prior to the work in epigenetic systems I am addressing here a number of 'exceptions' were made to the Central Dogma. The primary exception, one recognized in principle by Crick in his 1970 article, refers to the existence of retroviruses in general, and the enzyme reverse transcriptase in particular. Though all viruses require a living host to provide the materials necessary for their replication (this being one of the primary reasons viruses are generally not considered to be "alive"), most viruses use the cellular processes and products of their host to assemble materials for their reproduction. Viruses using the reverse transcriptase replication pathway, however, transcribe their RNA sequence information to DNA sequence information, then integrate the resulting DNA sequence into the host genome so that the host transcribes the RNA virus genome in a manner causing the host cell to manufacture the components of virus particles, which then undergo a form of self-assembly within the host cell. This process has major evolutionary and medical significance not covered here. Thus we find that there is a direct transfer of sequence information from RNA to DNA, which Crick initially suggested was in-principle possible, but unlikely to occur when he categorized it as a type (ii) form of information transfer. Additional exceptions made to the Central Dogma do not strictly apply to Crick's version, so I leave those untouched for the moment.

Returning now to epigenetic systems, the implications of these systems for Crick's version of the Central Dogma are difficult to define clearly. The primary concern here is whether or not the traditional alphabets of the Central Dogma, and molecular genetics in general for that matter, are the correct nucleotide and amino acid alphabets. That is, we may ask what the

ontological status of 5-methylcytosine, or an acetylated lysine, is. Traditionally methylcytosine has been lumped with unmethylated cytosine in the standard four base DNA alphabet. Some researchers, however, have suggested that 5-methylcytosine be considered a new base (e.g. Lister and Ecker 2009), a position supported by some of its properties that would challenge the epigenetic status of methylation patterns by making the conversion of cytosine into methylcytosine a change in nucleotide sequence. Methylcytosine is, after all, structurally different than cytosine; it also possesses different chemical properties, e.g., different binding affinities. Thus methylcytosine differs from cytosine in its functional characteristics since the change in binding affinities to other molecules is the mechanism by which multiple methylation sites, physically close to one another, inhibit transcription.

One reason for suggesting that methylcytosine is *not* a new base comes from a primarily evolutionary conceptual approach. Given the traditional view that methylation patterns are non-heritable post-zygotic modifications (against which I argue here), it is easy to suggest that the heritable nucleotide structures are unmethylated cytosines, and thus only four bases exist. Even accepting heritability, in organisms such as mammals, where most methyl groups are removed from most methylcytosines (except those associated with genomic imprinting, which is epigenetically induced differential gene expression in early embryonic development favoring the maternal genes) and then replaced during embryogenesis or early embryonic development (Macleod et al. 1999, Kafri et al. 1992), research findings suggest at least some reason for accepting the evolutionary perspective of four bases. However, the apparent retention of methylcytosines throughout embryogenesis and development in other organisms, such as zebra fish (Macleod et al. 1999), favors the opposite position from an evolutionary perspective, so this issue cannot be treated as settled. Perhaps the strongest argument in favor of a four base alphabet from an evolutionary perspective is the relatively short duration (in evolutionary time) of methylation patterns. Even though they appear to be highly heritable in some cases, the rate at which such patterns are lost and gained in most scenarios tends to suggest retaining a four letter alphabet for evolutionary analysis (though this too perhaps breaks down when considering genomic imprinting, for which there may be extremely high positive selection pressures).

Ultimately it is unlikely that any single classification of methylcytosine will be accepted within all disciplines. In cases where methylcytosine *is* accepted as a new base, we clearly have a deviation from Crick's version of the Central Dogma. Though protein to nucleotide transfer of information in by alteration of methylation patterns does not yield nucleotide-by-nucleotide transfer of information, we nonetheless see that enzymes recruited to specific CpG rich regions of DNA alter the nucleotide sequence information contained within these regions since the enzymes transform cytosines to methylcytosines. If the discipline of interest does not accept methylcytosine as a new base in the nucleotide alphabet, there seems to be no new exception to Crick's strict conception of the Central Dogma.

While I have focused on methylation patterns as a specific subject for deviation from the claims of the central dogma, the protein-protein interactions of histone modification face all of the same problems raised for methylation patterns, and also yield similar conclusions. Just as methylation changes the chemical and functional properties of cytosine, so the addition of chemical groups to amino acids changes the properties of the histones. These changes often alter

the conformation of the proteins in such a way that new functional properties emerge for the entire chain. Since the functional properties of the segment of DNA wrapped around the histone molecule are at least partially dependent upon the functional properties of the histone, changes in histone conformation are distributed onto the corresponding region of DNA in the form of new informational properties. Thus, in spite of the differences in the molecular details, histone alterations, like methylation of cytosines, can alter the functional properties (and hence the informational properties) of DNA. This potentially qualifies as protein-protein transfer of information.

It is also important, however, to direct attention to the standard conception of the Central Dogma, which deviates from that laid out by Crick in significant ways. This common conception retains at least one mistaken portrayal of Crick's Central Dogma, addressed in his 1970 clarification. That is, we find that a number of textbooks describe the Central Dogma as stating, "genetic information passes from DNA to protein in a one-way information pathway... [and that] genotype codes for phenotype, but phenotype cannot code for genotype (Pierce 2006, 280)", or "...the flow of genetic information in cells is therefore from DNA to RNA to protein (Alberts et al. 2004, 229)." Thus we clearly see that, despite Crick's own admonitions to the contrary, the common conception of the Central Dogma still contains the positive statements that Crick embedded into the sequence hypothesis. That is, the current conception of the Central Dogma includes Crick's negative claim (with exceptions) *and* the positive claim about the mapping from genetic information onto protein.

For present purposes, however, the implicit statement contained in the phrase "*genetic information*" is more important. In both articles, Crick specifies that for the purposes of his Central Dogma, he treats protein structure (including conformation) and function as if they were determined by amino acid sequence alone. He admitted that this claim was not properly justified, but held that for practical purposes it was close enough to true in normal physiological circumstances that it was a useful heuristic step unless and until further study raised problems with it. That this latter aspect of Crick's position is incorrect is not under debate--we know that environmental factors play just as much of a role in protein conformation at the tertiary and quaternary levels as does amino acid sequence.

However, this view of "genetic information" also suggests that it is not sequence information alone that is being transferred from DNA to RNA to protein, but also the functional information contained within the string of initial amino acids. In my own view, a protein-encoding gene is the region of DNA that is transcribed to mature mRNA plus any adjoining sequences necessary for this activity. Note that this means that any adjoining promoter or untranslated regions belong to the gene, unlike any trans- or cis-acting sites, though these regulatory regions may be necessary for the transcription or expression of that region. Likewise, a single region of eukaryotic DNA can contain a number of distinct genes in this view, since alternate splicing of transcribed regions can produce many distinct variants as a final mRNA product, though there is often a considerable amount of shared genetic material between variants. This view of genes deviates from the standard usage of the term only in the way it handles variant transcripts. Using this conception of genetic information (in so far as we are concerned with genes), we find that the genetic information transferred upward according to the common

perception of the Central Dogma is more than merely sequence information, as, in the cellular context, it also contains the functional information (splicing sites, number of hydrogen bonds, ability to form disulfide bonds, etc.) that ultimately produces specific amino acid sequences with a given conformation. That is, we can distinguish between the sequence information in a nucleotide string as the alphabet character and order, and genetic information in a nucleotide string as the distributed information contained within the entire string such that the information contained within a single nucleotide is conditional upon the additional nucleotides around it.

Though epigenetic systems are a problem for Crick's version of the Central Dogma only in the potential cases where modified nucleotides and amino acids are a new addition to the respective alphabets, these systems pose a much greater threat to the common conception of the Central Dogma. As functional changes to the nucleotide regions they are either embedded in (methylation patterns) or bound to (histone modifications), epigenetic systems clearly alter the transcriptional activity of this region, and thus also the functional information and its ability to be transferred therein. The functional properties of histone modifications have been suggested as sufficiently strong to justify a "histone code," which outlines the chemical and functional interactions between chromatin and DNA based on a histone alphabet containing the various modifying chemical groups and their potential locations on the histone tails.

Clearly epigenetic systems pose a significant problem for the common conception of the Central Dogma. These mechanisms provide a novel (to the Central Dogma) method of information transfer that threatens to overthrow the Central Dogma as conventionally interpreted, and to undermine even the weaker (but more reasonable) version of the Central Dogma presented by Crick to a certain extent. If the potential for epigenetic control of transcriptional regulation is realized, a major exception to the conventional interpretation will force a revaluation of its use in molecular and evolutionary biology.

Epigenetics and the Units and Levels of Selection

Though frequently debated in both the biological and philosophical literatures for some time, the question whether selection at the microevolutionary level can provide a sufficient explanation of macroevolution and major evolutionary transitions continues to be a topic of interest. I have already discussed the impact of epigenetic systems on the action of natural selection and how we model it. I now suggest that our conceptual understanding of selection as a force is also affected. Following Brandon (1984), I distinguish here between the levels and units of selection. For present purposes we can take the unit of selection as the object on which selection acts (that is, the entity which has a fitness with respect to its environment, and whose trait heritability is directly the result of this fitness) and the level of selection as the level (gene, genome, protein, cell, etc.) at which variations in traits contribute to different fitnesses. Less abstractly, Mayr (1963) and Brandon (1984) both argue that the phenotype is the primary unit of selection (though perhaps with reservations in some instances). This can be seen briefly by noting that fitness is based on the differential fitness of organisms within a population. Though allelic differences of individual genes may play a significant role in fitness differences, no gene is heritable without the accommodating cellular structures necessary for replication and preservation of nucleotide sequences, and is only successful in light of the cumulative fitness

properties of the entire phenotype. That is, genes cannot replicate themselves, a fact that distinguishes them from organisms capable of replication or reproduction. The levels of selection, however, can be viewed differently according to different selection pressures and the specific traits these pressures target. In this section, I argue that epigenetic systems blur the boundaries of both the unit and the levels of selection.

Though Brandon's specific account of what qualifies as a level of selection (Brandon 1982) has been questioned, the intuitive concept behind his position is attractive. Without digressing into a detailed exposition, Brandon's account suggests that selection occurs at a specific level if and only if a) entities at that level reproduce differentially, and b) if adaptedness values (expected fitness values) of entities at this level screen off adaptedness values of entities at all other levels from the reproductive values of entities at the level in question. The concept of screening off is essentially as follows: If an event E is explained by A such that B becomes statistically irrelevant, but not vice versa, A is a better explainer for event E than is B (Brandon 1982).

Here I suggest that a more lenient conception of the criteria for a level of selection could be that: a) objects or entities at the level of interest exhibit differential replication or reproduction, and b) that reproductive or replicative values of entities at the level of interest are dependent of adaptedness values of entities at the same level, but are statistically independent of adaptedness values of entities at lower levels (and thus only depend on the adaptedness values of entities at the level of interest). Thus, genes constitute a level of selection if and only if they replicate (since genes don't reproduce) differentially (i.e., their alleles have different likelihoods of replication) and the reproductive success of an allele is independent of the success of any entity at lower levels. Here I suggest that once we introduce epigenetic systems, in the cases in which epigenetic systems affect the behavior or expression of genetic material, the relevant genetic or genomic units cannot constitute a discrete level of selection.

The argument for this is fairly straight forward, and goes as follows: If we accept my definition of the requirements for a given level to be a level of selection, we accept that the reproductive success of entities at that level are independent of the adaptedness of entities at lower levels. Furthermore, we accept that the reproductive success of these entities is dependent upon the adaptedness of entities at the same and higher levels. We have seen, however, that with histone modifications and methylation patterns, co-recruitment of enzymes associated with these systems occurs, and thus, for cases involving the appropriate epigenetic systems, any adaptedness value for one level must be tied to the adaptedness value of entities at another. That is, since methylcytosine patterns can recruit histone modifying enzymes, and histone modifications can induce cytosine methylation, each is at least partially dependent upon the adaptedness of the other. Since these systems occur in numerous locations at the genomic and protein levels, I suggest that any level of selection must be higher than the protein level where these systems occur. Since both epigenetic systems are direct regulators of the intermediate, transcript level, we can eliminate the possibility of RNA transcripts as a level of selection as well. I thus suggest that the potential levels of selection are constrained by these epigenetic systems to any level above that of proteins, and thus effectively become cell-level and higher.

One additional point worth making here is that what I have suggested about the levels of selection only constrains what can be considered a discrete level of selection, but says nothing about what objects are the focus of selection pressures. That is, a biological entity can be the focus of selection pressures without there having to be selection *at the level of those entities*. We can instead simply accept that entities at different levels whose reproductive success is at least partially dependent on the success of others are foci for selection pressure, rather than each one requiring an independent level of selection.

The case for blurring the units of selection is in some ways simpler, and in others more difficult than that for denying the independence of various levels of selection. The difficulty arises from requiring that evolutionary change not depend on teleological processes, while nonetheless requiring that selection prepare organisms and other units of selection to cope with future states. The issue can be addressed by arguing that the ability to evolve is itself an evolvable feature (Pigliucci 2008). Evolvability, I suggest, entails non-conscious pattern recognition response system based on repeated adaptive successes with regard to cyclically or irregularly recurring environmental conditions within an appropriate time scale for effective adaptation. Thus evolvability is the potential to adapt to new conditions in ways that not only succeed in the new conditions, but also retain the potential for adaptive advantage in other conditions matching those of the previous history of the lineage (perhaps by returning to previous trait states). I maintain that epigenetic systems can provide both adaptedness to current conditions and the potential to adapt easily to environmental conditions already experienced by the lineage. There can be little question that epigenetic change is a mechanism that allows adaptive change to current environmental conditions that also retains the ability to restore the traits that were successful in coping with recently-encountered conditions. The ability to lose or gain the previously adaptive traits provides the potential to adapt to cyclical changes and environmental excursions frequently experienced by the lineage, as the ability to lose a currently-adaptive modification is often as important to survival when environmental conditions revert to a previous state as the ability to produce new adaptations.

Before cries of teleology are heard, this process can be explained in a way that does not require the presence of some sort of foresight or directedness by the evolutionary process. Consider two lineages, one which adapts to an environment by means of a permanent (or almost permanent) genetic change, while the other adapts through transcriptional regulation by way of epigenetic systems. Should the environment revert to the previous state, in which both lineages were highly fit, the lineage able to also revert to a previous state would be more successful than the lineage required to re-adapt through an extended process of selection pressure and genetic change according to the standard models of microevolution laid out in the modern synthesis (more on this later). That is, the lineage that adapted through easily reversible epigenetic change is at a significant advantage over the other, and is likely to increase in frequency within the population while the other decreases. When a series of such environmental changes occur sufficiently frequently within selectively significant intervals of time, those lineages able to respond epigenetically to the changing environments, as well as their current conditions, will be favored. In effect, this will be because they have the ability, acquired over a number of generations, to adapt to previous patterns of environmental change on a suitable time scale. In

absence of better terminology, one might say that they have evolved a non-teleological "expectation" of future patterns.

I am not claiming that epigenetic systems are the only way this sort of evolvability can arise, but that they provide one set of structures well suited to produce evolvability. They also effectively shift the unit of selection away from the individual organism onto the lineage. Any adaptation to "expected" future conditions based on previous patterns is no longer selection on the individual, but selection on the lineage itself. There will still likely remain some cases where the individual is the true unit of selection, such as development of resistance to synthetic pesticides which have little or no natural equivalent in the evolutionary history of most organisms. In many cases, however, we will find that the lineage is the proper unit of selection, and thus the ability to delimit a precise single unit of selection becomes blurred. On a case by case basis, this ability remains, but from an abstract conceptual perspective we can no longer accept either as *the* unit of selection. It is beyond the scope of my project here, but this also has important implications for the nature of fitness and about the entity (or entities) that may properly be said to be the bearers of fitness in evolutionary theory.

Macroevolution and a New Evolutionary Synthesis

The evolutionary synthesis of the early-mid twentieth century was perhaps the greatest advance in evolutionary theory to occur since Darwin's theory of evolution through natural selection. It essentially combined the Darwinian theory of evolution with Mendelian genetics using mathematical models (primarily population genetic models) of mutation, gene flow, drift, and natural selection. Though complex in its entirety, here I focus on two primary conclusions of the modern synthesis: gradualism and the sufficiency of microevolutionary mechanisms to account for macroevolution.

The modern synthesis suggested that evolution occurred through gradual changes in allele frequencies over time, claiming that after sufficient genetic divergence had occurred, speciation was the result. The gradualist theory of evolutionary change was later questioned by Eldredge and Gould (1972), who posited a punctuated equilibrium model where evolutionary change occurred in quick in bursts in evolutionary time, followed by periods of relative stability. Debate over the impact of punctuated equilibria on the modern synthesis was intense for a time, but eventually both were accepted as likely to describe evolutionary histories that arose in different situations. Thus, for example, genuinely panmictic oceanic organisms with very large numbers of broadcast larvae are unlikely to undergo punctuated speciation, while organisms with small relict populations or high levels of inbreeding are far more likely to do so if they are not driven to extinction. Other than this objection, however, after the modern synthesis was widely accepted, most evolutionary biologists adopted a strongly gradualist viewpoint, and it is still considered the "standard" view of evolutionary change in most (stable) environments.

A second significant component of the modern synthesis was the view that macroevolution could be fully explained by microevolutionary mechanisms. That is, while there were notable quantitative differences between clear examples of microevolution and macroevolution, the mechanisms shaping macroevolutionary change are the same as those

shaping microevolution. Under this view, macroevolution became merely an extension of microevolution; as such macroevolutionary change was simply the accumulated result of microevolutionary mechanisms acting on populations.

I conclude my discussion of the implications of epigenetic systems and developmental biology for evolutionary theory by suggesting that macroevolution can no longer be considered to be solely the result of the microevolutionary mechanisms that were taken into account during the Modern Synthesis, acting gradually over long periods of time. Epigenetic systems are likely to enforce significant conceptual change by supplementing or displacing those mechanisms, although a number of additional developmental phenomena will probably also play an important role in this regard. The epigenetic mechanisms on which we have focused already suffice to break the connection between evolutionary change and change of gene frequency. A number of scientists working in evolution and development have already suggested that we abandon, or replace, significant portions of the Modern Synthesis in favor of a new theory of evolution that incorporates developmental processes and mechanisms (and likely ecology as well) (Wagner and Laubichler 2004, Pigliucci 2007; 2008). The importance of epigenetic systems in deviations from the modern synthesis view of macroevolutionary change is not in the processes of change, but in the mechanisms which produce these processes. Thus, I first discuss the processes that are emerging as significant to understanding macroevolution, then return to epigenetic systems as potential mechanisms.

A key reason for distinguishing macroevolutionary from microevolutionary change is the presence of key innovations and the genesis of morphological novelty in a lineage that promotes deviation from similar biological groups. These processes distinguish macroevolutionary from microevolutionary change in homologous features, and often are the method by which "large" evolutionary jumps occur while nonetheless conserving the high levels of genetic homology often present in related taxa. To better understand this process, I turn to the concept of homology in development, and then address key innovations and morphological novelty.

In studying evolution, determining which features are ancestral and which are derived is key to understanding the evolution of one taxon from another. This feat is often obscured over long distances by homoplasy (the independent origin of a specific trait in two taxa, also referred to as convergent evolution), or by the loss of a suite of traits in a derived group, thus masking its identity as a derived taxon. When a single trait is shared uniquely by a group of taxa related by descent, this trait is considered homologous. More specifically, I follow Wagner (2007) in accepting that homology requires "historical continuity through inheritance with modification." This has the consequence that homologous traits need to be distinguished from those with multiple copies within or between populations not related by descent. For example, two individual organisms with the same allele for a gene acquired by convergent evolution are not said to share homologous genes, i.e., genes acquired by descent from a common ancestor, while two organisms with *different* alleles acquired by descent from a common ancestor do, according to the above definition, possess homologous genes. This distinction might seem trivial, but it points to an important distinction between extant morphological features within a population and morphological features from separate populations that share a common historical source. It is the evolutionary relationship in the latter case that points to the importance of homology, for we find

in homologous traits a difference in adaptive significance between two divergent lineages that produces two distinct taxa. Furthermore, it is often the case that modification of homologous features results in key innovations (to be discussed hereafter) important in distinguishing the evolution of major taxa.

Understanding the origin of morphological novelties and innovations is key to understanding evolutionary change, and is a distinctly developmental problem. Novelty in evolution is "the origination of new phenotypic characters in organismal evolution (Muller 2002, 827). Key innovation, then, is the use of a novel trait to invade new ecological niches previously inaccessible through adaptive change. In general, whether the novelty is a key innovation or not, we can utilize the definition of homology given above to suggest that novelty requires a qualitative difference in homologous features, rather than merely a quantitative difference (*Ibid*., 828). This of course, excludes the independent and spontaneous origin of new morphological features, but since the sudden appearance of a complex trait without any "historical continuity" is highly unlikely, I submit this as an acceptable description of the source of key innovations and morphological novelties. That is, even something like paralogous or orthologous relationships, where the traits in question are the result of gene duplications (paralogous) and subsequent evolution, or just evolution of a gene without duplication (orthologous), retain a "historical continuity".

The importance of the claim that changes in homologous features drives macroevolution is not to be underestimated. A prime example concerns the scales on butterfly wings. Previously suggested to be homologous to the sensory bristles of the peripheral nervous system, individually colored wing scales are clearly an example of key innovation driving macroevolutionary change (Galant et al 1998). With individual scales bearing a single color aggregating into a mosaic color pattern, the wing scales of *Lepidoptera* demonstrate macroevolutionary change in morphological features, and in the potential for new behavior patterns, defense mechanisms and ecological roles. Recent study of the relationship between sensory bristles and wing scales indicates that the *ASH1* gene has been co-opted, and is expressed during both the early development of both sensory bristles and wing scales (*Ibid*.). This suggests that there is indeed a homologous relationship between the two morphological traits, though they exhibit distinct features, and that the resulting key innovation in the morphological novelty forming from the co-option and adaptation of the sensory bristles has played a significant role in the macroevolution of the *Lepidoptera*.

The above example of the origin of morphological novelty and a key innovation through change in homologous traits also raises a second set of mechanisms through which macroevolutionary change can occur: change in time (heterochrony) or location of gene expression (heterotopy). Perhaps the best example of this producing macroevolutionary change is the radiation of the arthropod taxa from the onychophora. A study of the evolution of the *Hox* gene set in both groups within this clade suggests that the entire *Hox* family found in the arthropod lineages was present prior to the separation of the arthropods from the onychophorans and that the differences between the two groups with regard to segment formation and identity are the result of differential deployment of the *Hox* gene family. The differential expression of the trunk *Hox* genes along the anterior-posterior axis and the arthropod-specific repression of

distal-less are of particular importance. Thus it is likely that no major changes of nucleotide sequence in the *Hox* gene family was necessary for the separation of the arthropod and onychophoran lineages from their common ancestor, but rather only changes in timing and location of their expression (Grenier et al. 1997).

As I suggested above, the mechanisms that yield morphological novelties and key innovations can occur rapidly (over geological or evolutionary time) without necessarily being reduced to the microevolutionary mechanisms of the modern synthesis. For example, a (hypothetical) change in segregation of maternal transcripts during early development due to temporary epigenetic regulation of embryonic proteins necessary for development (such as disheveled, which appears associated with inhibiting GSK-3 opposite the point of sperm entry in *Xenopus*, thus allowing β-catenin to remain stable and promote dorsal axis identity) producing an altered phenotype cannot be explained solely through reference to the microevolutionary mechanisms described in the modern synthesis. The mechanisms by which morphological novelty and key innovation arise are diverse, but it now seems likely that they often result from alteration of transcriptional regulation in gene regulatory networks. Though numerous methods of transcriptional regulation exist, I suggest that epigenetic systems bear the potential to be a highly efficient mechanism for such activity in macroevolutionary change. Synthetic theorists could, of course, accept all of this, but that would be part of a revised synthesis I am calling for.

The high turnover potential of epigenetic systems provides the opportunity for "exploratory" transcriptional regulation in gene regulatory networks, while avoiding genetic 'hopeful monsters'. Granting Mayr's arguments that hopeful monsters have no chance of evolutionary success, this has the virtue of bypassing all the population-genetic difficulties associated with this concept. That is, we find that Goldschmidt (1940) is likely correct in his supposition that macroevolutionary change can occur, but that Mayr is correct that such change is not likely to occur through the standard genetic pathways of microevolution found in the modern synthesis. By regulating transcription of developmental genes in both time and location within the embryo, pattern formation and morphogenesis can be altered in ways that yield phenotypes with qualitative differences without requiring the sorts of genetic changes that bar successful interbreeding or reproduction.

Successful evolutionary change due to epigenetic regulation of transcription can later be integrated into a more permanent genetic structure through genetic assimilation of developmental modifications. The modern notion of genetic assimilation differs from the Wadingtonian version in that it is centered around reaction norms in a given environment. In the case that an organism is adapted to an environment X, but possesses the potential for a range of developmental variation if the environment changes to Y, the organism (or, rather, its lineage) is capable of survival in the new environment, so long as the reaction norm encompasses phenotypes fit for environment Y. Though the standard view of the reaction norm is one of quantitative rather than qualitative differences (which is probably correct in most cases), there is no reason why such differences cannot extend beyond simple differences in measurement. Furthermore, given appropriate reaction norms, non-genetic regulatory mechanisms are able to generate novel phenotypes, though the applicability of this ability to macroevolution is in need of further research. Finally, where environment Y becomes stable, genetic assimilation of the novel

phenotypic characteristics successfully expressed in environment Y can yield a genetic basis of the phenotype adapted to a new environment (Pigliucci et al. 2006; Newman and Muller 2000).

I have now suggested that, contrary to the tenets of the modern synthesis, macroevolution cannot be attributed to cumulative microevolutionary change achieved by the genetic mechanisms encompassed within the modern synthesis. Furthermore, I have suggested processes by which rapid macroevolutionary change can occur, and mechanisms by which these processes can be initiated. This, in turn, leads me back to the modern evolutionary synthesis and the call for a new evolutionary theory that incorporates developmental phenomena into its theoretical structure. This is not a call to abandon the modern evolutionary synthesis entirely, but instead to allow for new discoveries, such as those mentioned above, to be included as pathways for evolutionary change, while discarding those portions proven inadequate by more recent research.

Pigliucci (2008) suggests that evolvability, or the ability and potential for evolutionary change, is itself an evolvable trait, making an "extended evolutionary synthesis" which incorporates this notion a necessity. He further suggests (2007) that the modern synthesis is a theory of genes, while an extended synthesis must be a theory of genes, and of forms. Here I add that a conceptual foundation for a new evolutionary synthesis will be driven by research such as described above, and that epigenetic systems have the potential to become a significant source of information regarding the processes and mechanisms by which new evolutionary pathways have emerged. That is, in moving beyond the basic mechanisms of microevolutionary change described in the modern synthesis, we must take into account both the interactions between these microevolutionary mechanisms and epigenetic systems (as I have described previously) *and* the potential for epigenetic systems to effect the rapid macroevolutionary change seen throughout nature.

Conclusion

I began this project by questioning how evolutionary change could occur in spite of the obvious conservation seen at numerous levels and instances across the natural world. I specialized this issue to a specific set of systems that, I suggest, provide part of the answer to this question, and considered how they might interact with current evolutionary models. The first two thirds of this thesis were devoted to a description of the epigenetic systems of DNA methylation and histone modifications, followed by a fairly basic, though important review of how these systems impact microevolutionary change and the models we have created for these changes. This review led to the suggestion that evolutionary processes at the microevolutionary level are much more complex than previously described, and that we must be able to account for these new complexities if we wish to provide a full account of evolutionary change.

In this chapter I have shifted to a macroevolutionary perspective and focused on conceptual problems relevant to a number of philosophical issues that arise when considering the role of epigenetic systems in evolutionary change and our conceptual approach to evolutionary theory. We find that, as conventionally understood, the Central Dogma has yet another exception, and a potentially devastating one. We also find that the units and levels of selection previously suggested are brought into question. Perhaps most importantly, we have diagnosed

some major reasons for the insufficiency of the modern evolutionary synthesis as it currently exists and indicated a starting point for their resolution in a new "extended evolutionary synthesis," the construction of which is likely to take a number of years.

While I did not mention it explicitly in the discussions above, the argument shows that, once again, reductionism in evolutionary theory has proven problematic. The reasons for this derive from several areas of discussion above, but essentially go like this: a) Separation of the levels of selection, together with the argument that selection at the level of the phenotype drives most genetic selection (so selection generally "acts" at cellular and higher levels), suggests that we cannot reduce selection solely to the genetic level in cases where epigenetic systems are present and acting in a regulatory capacity. b) When microevolution within a lineage is taken as proposed in the modern synthesis, an understanding of macroevolutionary change is not derivable from strictly microevolutionary genetic analyses limited to the resources of the Modern Synthesis. Thus reductionism of a different sort is also discarded. c) Finally, we can no longer suggest that the phenotype is reducible to the genotype in any strong or complete sense. As the description of the epigenetic systems discussed in this thesis have shown, there are components of development that cannot be understood as genetically caused, for they depend on epigenetic mechanisms that alter the effects of the genetic information contained in the genes in ways that cannot be understood – or caused – by the genes alone.

Though I maintain that I have made a strong case for the impact of epigenetic systems on evolutionary change, much of what I have suggested here concerns the potential impact of these systems. To understand their full importance, and the role they might play in future models of evolution at both the microevolutionary and macroevolutionary levels, a significant amount of research remains necessary. Though it is difficult to suggest concrete avenues of research at this early stage, the most fruitful pursuit is likely to be in identifying model organisms through study of environmental conditions. If what I have suggested above is correct, a good way to begin is by looking for organisms whose life cycle or circumstances place them in shifting environments or expose them frequently to high levels of chemical waste. Comparative studies between related taxa from these areas could also provide some indication of the evolutionary significance of epigenetic mechanisms for these model systems.

Though biomedical fields are driving epigenetic research forward at a tremendous rate, specifically evolutionary research and experiments will also become necessary. Some aspects of this research will be fairly routine. Population- and species-wide comparative epigenomic studies are relatively accessible for methylation patterns through methyl-specific PCR or by use of methylation profiling kits. Histone modifications can be detected with antibodies and western blotting or several other experiments. More problematic will be an understanding of the long term evolutionary significance of epigenetic systems, as the relatively rapid decay rate of epigenetic patterns suggests that a single snapshot of a lineage will miss most of the significant epigenetic phenomena. One possible way to begin is by probing the significance of epigenetic systems over a number of generations by experimental evolution protocols and then shifting to mathematical and computational modeling, though this will not guarantee an understanding of epigenetic systems *in situ*.

Conceptual analysis of empirical research, such as my project here, has the potential to drive the study of evolutionary biology in (at least) two primary ways. First, experimental study will ultimately be the deciding factor in changes to how we conceive of evolution and development, but conceptual analyses can direct these studies towards potentially fruitful experiments. There is no question that the study of evolution is dependent upon empirical data—but while detailed research into biological organisms or systems is important, it is often a few well-chosen experiments that provide the missing information necessary to connect separate bodies of work into a more cohesive theoretical system. This is what conceptual analyses can provide for the future study of evolutionary biology. It is becoming more and more apparent that, as noted above, various subjects previously left out of the modern synthesis need to be integrated into a modified evolutionary theory. I hope that my work here, along with other conceptual studies of recent empirical findings, will help drive the correct lines of research necessary for making these connections.

We also find, however, that for the present, conceptual analysis of empirical research in the biological sciences often fills a void left by the rapid advancement of new technologies and methods for research in the life sciences. The body of knowledge associated with many areas of research in evolutionary biology alone is rapidly increasing, at least partly due to new techniques involving high-throughput sequencing and computational biology. With this rapid acquisition of data, we find missing a correspondingly rapid acquisition of *knowledge*. While not dismissing the efforts of scientists working in these fields, in many cases the body of information being generated is quickly outpacing the ability of researchers to analyze and integrate into a conceptual frame work. In this we see the second primary benefit of solely conceptual analyses in the study of evolution. Thus, my project here is intended not only to suggest future lines of research, but also to begin the process of integrating current research into our conceptual foundations, while simultaneously presenting a broad body of knowledge to those working in related fields in a way that opens further avenues of research, both conceptual and empirical.

Bibliography

Abrams, Marshall. 2007. How do Natural Selection and Random Drift Interact. *Philosophy of Science* 74: 666-79.

Alberts, Bruce, Bray, Dennis, Hopkin, Karen, Johnson, Alexander, Lewis, Julian, Raff, Martin, Roberts, Keith & Walter, Peter. 2004. *Essential Cell Biology*. New York, NY: Garland Science.

Athanasiadou, Rodoniki, de Sousa, Dina, Myant, Kevin, Merusi, Cara, Stancheva, Irena & Bird, Adrian. 2010. Targeting of De Novo DNA Methylation Throughout the Oct-4Gene Regulatory Region in Differentiating Embryonic Stem Cells. PLoS One. 5(4): e9937. doi:10.1371/journal.pone.0009937

Ariew, Andre and Richard C. Lewontin. 2004. The Confusions of Fitness. *British Journal for the Philosophy of Science* 55: 347-63.

Ariew, Andre. 2008. Population Thinking. In *The Oxford Handbook of Philosophy of Biology*, ed. Michael Ruse. Oxford: Oxford University Press.

Bistort, O., R. L. Christina and Massimo Pigliucci. 2008. Epigenetics for Ecologists. *Ecology Letters* 11: 106-11.

Brandon, Robert N. 1982. The Levels of Selection. *Proceedings of the Biennial Meeting of the Philosophy of Science Association*. Vol. 1: 315-323.

Brandon, Robert N. 1999. The Units of Selection Revisited: The Modules of Selection. *Biology and Philosophy* 14: 167-80.

Brandon, Robert N. 2005. The Difference Between Drift and Selection: A Reply to Millstein. *Biology and Philosophy* 20: 153-70.

Brandon, Robert and H. Frederik Nijhout. 2006. The Empirical Nonequivalence of Genic and Genotypic Models of Selection: A (Decisive) Refutation of Genic Selectionism and Pluralistic Genic Selectionism. *Philosophy of Science* 73: 277-97.

Brenner, Sydney. 1981. Genes and Development. In *Cellular Controls in Differentiation*, eds. C. W. Lloyd and D. A. Rees, 3-7. London and New York: Harcourt Brace Johanovich.

Brenner, Sydney. 1997. Biological Computation. In *The Limits of Reductionism in Biology*, eds. G. R. Bock and J. A. Goode, 106-16. Chicherster: John Wiley.

Carey, C. C., K. F. Gorman and S. Rutherford. 2006. Modularity and Intrinsic Evolvability of Hsp90-Buffered Change. PLoS ONE 1(1): e76. Doi: 10.1371/journal.pone.0000076.

Chan, Simon W.-L., Zilberman, Daniel, Xie, Zhixin, Johansen, Lisa K., Carrington, James C. & Jacobsen, Stephen E. 2004. RNA Silencing Genes Control de Novo DNA Methylation. *Science*. 303: 1336.

Conner, Jeffrey K. and Daniel L. Hartl. 2004. *A Primer of Ecological Genetics*. Sunderland, MA: Sinauer Associates, Inc.

Crick, Francis. 1958. On Protein Synthesis. *Symp. Soc. Exp. Biol. XII*: 139-163.

Crick, Francis. 1970. Central Dogma of Molecular Biology. *Nature.* 227: 561-563

Dawkins, Richard. 1984. Replicator Selection and the Extended Phenotype. In *Conceptual Issues in Evolutionary Biology: An Anthology*, ed. Elliott Sober. 125-41. Cambridge, MA: MIT Press.

Dietrich, Michael R. 2006. Three Perspectives on Neutrality and Drift in Molecular Evolution. *Philosophy of Science* 73: 666-77.

Eldredge, Niles & Gould, Stephen J. Punctuated Equilibria: An alternative to phyletic gradualism. 1972. In Models in Paleobiology, ed. Schopf, T.J.M. 82-115. San Francisco, CA: Freeman, Cooper & Co.

Frost-Arnold, Greg. 2004. How to be an Anti-Reductionist About Developmental Biology: Response to Laubichler and Wagner. *Biology and Philosophy* 19: 75-91.

Galant, R., Skeath, J.B., Paddock, S., Lewis D.L., & Carroll S.B. 1998. Expression pattern of a butterfly *achaete-scute* homolog reveals the homology of butterfly wing scales and insect sensory bristles. *Current Biology*. 8: 807–813

Gerhart, John & Kirschner, Marc. 1997. *Cells, Embryos, and Evolution: Towards a Cellular and Developmental Understanding of Phenotypic Variation and Evolutionary Adaptability.* Malden, MA: Blackwell Science, Inc.

Gilbert, Scott F. & Epel, David. 2008. *Ecological Developmental Biology: Integrating Epigenetics, Medicine, and Evolution*. Sunderland, MA: Sinauer Associates.

Gillespie, John H. 2004. *Population Genetics: A Concise Guide*. Baltimore and London: The Johns Hopkins University Press.

Goldschmidt, R. 1940. *Material Basis of Evolution.* New Haven, CT: Yale University Press.

Grenier J.K., Garber, T.L., Warren, R., Whitington, P.M. and Carroll, S.B. 1997. Evolution of the entire arthropod Hox gene set predated the origin and radiation of the onychophoran/arthropod clade. *Current Biology*. 7(8): 547-553.

Hartl, Daniel L. 2000. *A Primer of Population Genetics, 3rd Edition*. Sunderland, MA: Sinauer Associates, Inc.

Hoekstra, Hopi E. and Jerry A. Coyne. 2007. The Locus of Evolution: Evo Devo and the Genetics of Adaptation. *Evolution* 61 (5): 995-1016.

Holliday, R. & Grigg, G.W. 1993. DNA Methylation and Mutation. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*. 25 (1): 61-67.

Holliday, Robin & Ho, Thu. 2002. DNA methylation and epigenetic inheritance. *Methods*. 27 (2): 179-183.

Hunter, N. 1999. Prion diseases and the central dogma of molecular biology. *Trends Microbiol*.7 (7):265-6.

Jablonka, Eva and Marion J. Lamb. 1998. Epigenetic Inheritance in Evolution. *Journal of Evolutionary Biology* 11: 159-83.

Jablonka, Eva and Marion J. Lamb. 2005. *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. Cambridge, MA: The MIT Press.

Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H. & Razin, A. 1992. Developmental pattern of gene-specific DNA methylation in mouse embryo and gem line. *Genes & Development*. 6: 705-714.

Kubicek, S., Schotta, G., Lachner, M., Sengupta, R., Kohlmaier, A., Perez-Burgos, L., Linderson, Y., Martens, J.H., O'Sullivan, R.J., Fodor, B.D., Yonezawa, M., Peters, A.H. & Jenuwein, T. 2006. The role of histone modifications in epigenetic transitions during normal and perturbed development. *Ernst Schering Res Found Workshop*. 57: 1-27.

Laubichler, Manfred D. and Günter P. Wagner. 2001. How Molecular is Molecular Developmental Biology? A Reply to Alex Rosenberg's Reductionism Redux: Computing the Embryo. *Biology and Philosophy* 16: 53-68.

Laubichler, Manfred D. and Jane Maienschein. 2007. *From Embryology to Evo-Devo: A History of Developmental Evolution.* Cambridge, MA: The MIT Press.

Lewontin, Richard C. 1970. The Units of Selection. *Annual Reviews of Ecology and Systematics* 1: 1–18.

Lister, Ryan & Ecker, Joseph R. 2009. Finding the fifth base: Genome-wide sequencing of cytosine methylation. *Genome Research*. 19: 959-966.

Lohr, Ulrike, Yussa, Miyuki & Pick, Leslie. 2001. Drosphila *fushi tarazu*: a gene on the border of homeotic function. *Current Biology*. 11 (18): 1403-1412.

Macleod, Donald, Clark, Victoria H. & Bird, Adrian. 1999. Absence of genome-wide changes in DNA methylation during development of the zebrafish. *Nature Genetics*. 23: 139-140.

Martin, Cyrus & Zhang, Yi. 2007. Mechanisms of epigenetic inheritance. *Current Opinions in Cell Biology*. 19 (3): 266-72.

Mattick, John S. 2003. Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. *Bioessays*. 25 (10): 930-39.

Mayr, Ernst. 1963. *Animal Species and Evolution*. Cambridge, UK: Harvard University Press.

Meselson, M.S. and Stahl, F.W., 1958. The Replication of DNA. *Cold Spring Harbor Symposia for Quantitative Biology*. 23: 9-12.

Millstein, Roberta L. 2005. Selection Vs Drift: A Response to Brandon's Reply. *Biology and Philosophy* 20: 171-75.

Millstein, Roberta L. 2006. Discussion of "Four Case Studies on Chance in Evolution": Philosophical Themes and Questions. *Philosophy of Science* 73: 678-87.

Millstein, Roberta L. 2007a. Hsp90-Induced Evolution: Adaptationist, Neutralist, and Developmentalist Scenarios. *Biological Theory* 2 (4): 376-86.

Millstein, Roberta L. 2007b. Population Genetics. In *Cambridge Companion to the Philosophy of Biology*, eds. D. Hull and M. Ruse, 22-43. Cambridge: Cambridge University Press.

Milutinovic, Snezana, Brown, Shelley E., Zhuang, Qianli & Szyf, Moshe. 2004. DNA Methyltransferase 1 Knock Down Induces Gene Expression by a Mechanism Independent of DNA Methylation and Histone Deacetylation. *Journal of Biological Chemistry*. 279: 27915-27927.

Morange, Michel. 2006. The Protein Side of the Central Dogma: Permanence and Change. *History and Philosophy of the Life Sciences* 28: 513-24.

Morange, Michel. 2008. What History Tells Us XIII. Fifty Years of the Central Dogma. *Journal of BioSciences* 33 (2): 171-5.

Müller, Gerd B. 1990. Developmental Mechanisms at the Origin of Morphological Novelty: A Side-effect Hypothesis. In *Evolutionary Innovations*, ed. Matthew H. Nitecki, 99-130. Chicago: University of Chicago Press.

Müller, Gerd B., and Günter P. Wagner. 1996. Homology, Hox Genes, and Developmental Integration. *American Zoologist* 36: 4-13.

Müller, Gerd B. 2002. Novelty and Key Innovations. In *Encyclopedia of Evolution*, ed. Mark D. Pagel. Oxford: Oxford University Press.

Müller, Gerd B., Stuart A. Newman and Werner Callebaut, eds. 2003. *Origination of Organismal Form: Beyond the Gene in Development and Evolutionary Biology*. Cambridge, MA: MIT Press.

Müller, Gerd B. and Günter P. Wagner. 2003a. Innovation. In *Keywords and Concepts in Evolutionary Developmental Biology*, eds. Brian K. Hall and Wendy M. Olson. Cambridge, MA: Harvard University Press.

Müller, Gerd B. and Günter P. Wagner. 2003b. Epigenesis and Epigenetics. In *Keywords and Concepts in Evolutionary Developmental Biology*, eds. Brian K. Hall and Wendy M. Olson. Cambridge, MA: Harvard University Press.

Nagel, Ernest. 1984. The Standpoint of Organismic Biology. In *Conceptual Issues in Evolutionary Biology: An Anthology*, ed. Elliott Sober. 125-41. Cambridge, MA: MIT Press.

Newman, Stuart A. & Müller, Gerd B. 2000. Epigenetic mechanisms of character origination. *Journal of Experimental Zoology*. 288 (4): 304-317.

Nijhout, H. Frederik. 1990. Metaphors and the Role of Genes in Development. *BioEssays* 12: 441-46.

Nijhout, H. Frederik. 1994. Developmental perspectives on evolution of butterfly mimicry. *BioScience* 44: 148-57.

Nijhout, H. Frederik, and S. M. Paulsen. 1997. Developmental Models and Phylogenic Characters. *American Naturalist* 149: 394-405.

Nijhout, H. Frederik, and Douglas J. Emlen. 1998. Competition among Body Parts in the Development and Evolution of Insect Morphology. *Proceedings of the National Academy of Sciences, USA* 95: 3685-89.

Nijhout, H. Frederik. 2003. Development and Evolution of Adaptive Polyphenisms. *Evolution and Development* 5: 9-18.

Pierce, Benjamin A. 2006 *Genetics: A Conceptual Approach.* New York, NY: Freeman & Company.

Pigliucci, Massimo. 2001. *Phenotypic Plasticity: Beyond Nature and Nurture*. Baltimore and London: The Johns Hopkins University Press.

Pigliucci, Massimo, Murren, Courtney J. & Schlichting Carl D. 2006. Phenotypic plasticity and evolution by genetic assimilation. *Journal of Experimental Biology.* 209: 2362-2367.

Pigliucci, Massimo. 2007. Do we need an extended evolutionary synthesis? *Evolution*. 61 (12): 2743-49.

Pigliucci, Massimo. 2008. Is evolvability evolvable? *Nature Reviews Genetics*. 9: 75-82.

Richardson, Robert C. 2000. The Organism in Development. *Philosophy of Science* 67: S312-21.

Richardson, Robert C. 2003. Adaptationism, Adaptation, and Optimality. *Biology and Philosophy* 18: 695-713.

Richardson, Robert C. 2006. Chance and the Patterns of Drift: A Natural Experiment. *Philosophy of Science* 73: 642-54.

Robert, Jason Scott. 2004. *Embryology, Epigenesis, and Evolution: Taking Development Seriously*. Cambridge: Cambridge University Press.

Roff, Derek A. 1997. *Evolutionary Quantitative Genetics*. New York, NY: Chapman & Hall.

Rosenberg, Alex. 1997. Reductionism Redux: Computing the Embryo. *Biology and Philosophy* 12 (4): 445-70.

Rosenberg, Alex. 2006. Is Epigenetic Inheritance a Counterexample to the Central Dogma? *History and Philosophy of the Life Sciences* 28: 549-66.

Ruden, Douglas M., Mark D. Garfinkel, Vincent E. Sollars, Xiangyi Lu. 2003. Waddington's Widget: Hsp90 and the Inheritance of Acquired Characters. *Seminars in Cell and Developmental Biology* 14: 301-10.

Ruse, Michael. 1984. Reduction in Genetics. In *Conceptual Issues in Evolutionary Biology: An Anthology*, ed. Elliott Sober. 125-41. Cambridge, MA: MIT Press.

Rutherford, S. L. and S. Lindquist. 1998. Hsp90 as a Capacitor for Morphological Evolution. *Nature* 396: 336-42.

Rutherford, Suzanne L. 2003. Between Genotype and Phenotype: Protein Chaperones and Evolvability. *Nature Reviews* 4: 263-74.

Schaffner, Kenneth. 1984. Reduction in Biology: Prospects and Problems. In *Conceptual Issues in Evolutionary Biology: An Anthology*, ed. Elliott Sober. 125-41. Cambridge, MA: MIT Press.

Schlichting, Carl D. and Massimo Pigliucci. 1998. *Phenotypic Evolution: A Reaction Norm Perspective*. Sunderland, MA: Sinauer Associates, Inc., Publishers.

Schlosser, Gerhard, and Günter P. Wagner, eds. 2004. *Modularity in Development and Evolution.* Chicago: Chicago University Press.

Schwenk, Kurt, and Günter P. Wagner. 2004. The Relativism of Constraints on Phenotypic Evolution. In *Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes*, eds. Massimo Pigliucci and Katherine Preston, 390-408. New York: Oxford University Press.

Skipper, Robert A. 2006. Stochastic Evolutionary Dynamics: Drift Versus Draft. *Philosophy of Science* 73: 655-65.

Sober, Elliott. 1984. Holism, Individualism, and the Units of Selection. In *Conceptual Issues in Evolutionary Biology: An Anthology*, ed. Elliott Sober. 125-41. Cambridge, MA: MIT Press.

Sollars, V., X. Lu, L. Xiao, X. Wang, M. D. Garfinkel and D. M. Ruden. 2003. Evidence for an Epigenetic Mechanism by which Hsp90 acts as a Capacitor for Morphological Evolution. *Nature Genetics* 33: 70-4.

Stotz, Karola. 2006. Molecular Epigenesis: Distributed Specificity as a Break in the Central Dogma. *History and Philosophy of the Life Sciences* 28: 533-48.

Strasser, Bruno J. 2006. A World in One Dimension: Linus Pauling, Francis Crick and the Central Dogma of Molecular Biology. *History and Philosophy of the Life Sciences*. 28: 491-512.

Suzuki, Yuichiro and H. Frederik Nijhout. 2006. Evolution of a Polyphenism by Genetic Accommodation. *Science* 311: 650-2.

Thieffry, Denis and Sarkar, Sahotra. 1998. Forty years under the central dogma. *Trends in Biochemical Sciences*. 23: 312-316

Torres J. M. 1999. On the Falsification of the *Central Dogma* and the *De novo* Synthesis of Molecular Species. *Philosophia Naturalis*. 36:1-18.

Wagner, Günter P. 1988a. The Influence of Variation and Development on the Rate of Multivariate Phenotypic Evolution. *Journal of Evolutionary Biology* 1: 45-66.

Wagner, Günter P. 1988b. The Significance of Developmental Constraints for Phenotypic Evolution by Natural Selection. In *Population Genetics and Evolution*, ed. Gerdien de Jong, 222-29. Berlin and Heidelberg: Springer.

Wagner, Günter P. 1989. The Origin of Morphological Characters and the Biological Basis of Homology. *Evolution*. 43 (6): 1157-71.

Wagner, Günter P. and B.Y. Misof. 1993. How can a Character be Developmentally Constrained Despite Variation in Developmental Pathways? *Journal of Evolutionary Biology.* 6: 449-55.

Wagner, Günter P. and Lee Altenberg. 1996. Complex Adaptations and the Evolution of Evolvability. *Evolution* 50 (3): 967-76.

Wagner, Günter P., Ginger Booth and Homayoun Bagheri-Chaichian. 1997. A Population Genetic Theory of Canalization. *Evolution* 51 (2): 329-47.

Wagner, Günter P., Chi-Hua Chiu and Thomas F. Hansen. 1999. Is Hsp90 a Regulator of Evolvability? *Journal of Experimental Zoology (Molecular and Developmental Evolution)* 258B (2): 116-8.

Wagner, Günter P., Chi-Hua Chiu and Manfred D. Laubichler. 2000. Developmental Evolution as a Mechanistic Science: The Inference from Developmental Mechanisms to Evolutionary Processes. *American Zoologist* 40 (5): 819-31.

Wagner, Günter P. 2000. What is the Promise of Developmental Evolution? Part I: Why is Developmental Biology Necessary to Explain Evolutionary Innovations? *Journal of Experimental Zoology (Molecular and Developmental Evolution)* 288: 95-8.

Wagner, Günter P. 2001. What is the Promise of Developmental Evolution? Part II: A Causal Explanation of Evolutionary Innovations May Be Impossible. *Journal of Experimental Zoology (Molecular and Developmental Evolution)* 291: 305-9.

Wagner, Günter P., and Gerd B. Müller. 2002. Evolutionary Innovations Overcome Ancestral Constraints: A Re-examination of Character Evolution in Male Sepsid Flies (Dipteral *Sepsidae*). *Evolution and Development* 4 (1-6).

Wagner, Günter P. and Hans C.E. Larsson. 2003. What is the Promise of Developmental Evolution? Part III: The Crucible of Developmental Evolution. *Journal of Experimental Zoology (Molecular and Developmental Evolution)* 300B (1): 1-4.

Wagner, Günter P., and Manfred D. Laubichler. 2004. Rupert Riedl and the Re-Synthesis of Evolutionary and Developmental Biology: Body Plans and Evolvability. *Journal of Experimental Zoology (Molecular and Developmental Evolution)* 302B: 92-104.

Wagner, Günter P. 2005. How Wide and How Deep is the Divide Between Population Genetics and Developmental Evolution? *Biology and Philosophy* 22: 145-53.

Wagner, Günter P. 2007. The developmental genetics of homology. *Nature Review Genetics*. 8: 473-479.

Walsh, Denis M., Tim Lewens and Andre Ariew. 2002. The Trials of Life: Natural Selection and Random Drift. *Philosophy of Science* 69: 452-73.

West-Eberhard, Mary Jane. 2003. *Developmental Plasticity and Evolution.* Oxford and New York: Oxford University Press.

Zhang, Yi & Reinberg, Danny. 2001. Transcription regulation by histone modification: interplay between different covalent modifications of the core histone tails. *Genes & Development*. 15: 2343-2360.