

Review Article

Genetics: Polymorphisms, Epigenetics, and Something In Between

Keith A. Maggert

Department of Biology, Texas A&M University, College Station, TX 77843, USA

Correspondence should be addressed to Keith A. Maggert, kmaggert@tamu.edu

Received 22 July 2011; Accepted 20 September 2011

Academic Editor: Victoria H. Meller

Copyright © 2012 Keith A. Maggert. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

At its broadest sense, to say that a phenotype is epigenetic suggests that it occurs without changes in DNA sequence, yet is heritable through cell division and occasionally from one organismal generation to the next. Since gene regulatory changes are oftentimes in response to environmental stimuli and may be retained in descendent cells, there is a growing expectation that one's experiences may have consequence for subsequent generations and thus impact evolution by decoupling a selectable phenotype from its underlying heritable genotype. But the risk of this overbroad use of "epigenetic" is a conflation of genuine cases of heritable non-sequence genetic information with trivial modes of gene regulation. A look at the term "epigenetic" and some problems with its increasing prevalence argues for a more reserved and precise set of defining characteristics. Additionally, questions arising about how we define the "sequence independence" aspect of epigenetic inheritance suggest a form of genome evolution resulting from induced polymorphisms at repeated loci (e.g., the rDNA or heterochromatin).

1. Epigenetics and Evolution

The importance of sequence polymorphisms in evolution is fundamental and irrefutable. The contribution of epigenetic gene regulation is considerably less well established. In this perspective, I will not attempt to summarize all the studies that have contributed to our current understanding of epigenetics; instead, I will thread together a handful of salient studies, taken particularly but not exclusively from *Drosophila* research, to illuminate how common and consequent "epigenetic" gene regulation may result from induced polymorphism. Inclusion of induced polymorphism in the panoply of epigenetic gene regulatory mechanisms may force us to reconsider our definitions, but is in accord with current and historic uses of "epigenetics," and may provide a new mechanism to understand how stable changes in gene expression can be established and maintained.

To understand the role of epigenetics in evolution, it is necessary to consider definitions of both evolution and epigenetics. For the purpose of any discussion linking the two, "evolution" must expand to include the change of frequency of *phenotypic variants* irrespective of underlying allelic variants. This is a mild departure from a sequence-centric view of changes in allele frequencies in evolving

populations, but is ironically more aligned with the original use of "epigenetic" to describe the abstract processes that produce a phenotype from a genotype prior to the elucidation of the central dogma, gene regulation, and developmental genetics. Now, "epigenetics" are instances of changes in gene regulation that do not correspond to underlying changes in nucleotide sequence. What one means by "changes in nucleotide sequence" is worth dwelling on, which I will do later. In general, changes in nucleotide sequence are "polymorphisms" although it is common to see them called "genetic" in order to contrast them with "epigenetic." However, this is a misuse, and genetics is the study of inheritance and variation whatever their cause; polymorphism and epigenetics are subsets of genetics, and as I hope to convince you, they are neither exclusive nor exhaustive (Figure 1).

Understanding the joint contributions to evolution of polymorphism and epigenetics, particularly the latter, requires understanding the difference between them. This difference is profound since while polymorphisms are thought to be characterized by random, permanent, and well-understood changes to genetic information, epigenetic gene regulation is more volatile and hence has come to

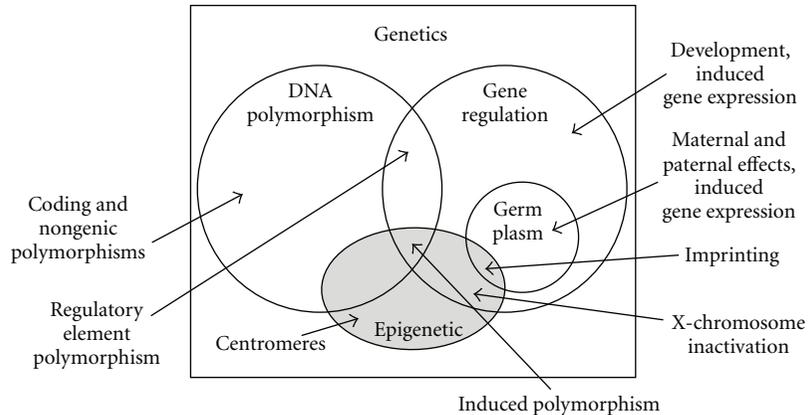


FIGURE 1: Relationships within genetics: random sequence polymorphisms, epigenetics, gene regulatory mechanisms, and induced polymorphisms.

include induced and reversible alterations in heritable traits. This raises the popular view (however, unfair) of Jean-Baptiste de Lamarck, that our own actions or experiences may come to bear on our offspring. Lamarck envisioned that an organism evolves by passing its experience to offspring. On the surface, the inheritance of acquired characteristics was consistent with slow change in species over time. It was not until Weismann articulated the difficulty in a giraffe's neck discussing its experience with a giraffe's sperm that a Lamarckian mechanism of evolution was cast aside. The resurgence of Lamarckian models of evolution has recently occurred for a number of reasons. First, there are clear examples of inheritance of information outside of DNA sequence, which has opened the possibility of experience affecting gene expression and such changes in expression being transmitted to offspring. Second, not only is this hypothetical model possible, but it is heretical and provocative, and thus exciting. Third, perhaps many of us feel more than a little guilty in heaping ridicule on an otherwise superb scientist who happened to be wrong.

2. What Is/Are Epigenetics?

A clear, concise, and comprehensive definition of epigenetics is tricky to articulate, not because it is difficult *per se*, but because the term has seen an expansion over the last decade and has started to include things that are arguably not epigenetic. To clarify the situation, Youngson and Whitelaw gave a cogent description of the difference between two types of "epigenetics": transmissible changes in expression (which they called "transgenerational epigenetic effects") and transmissible chromosome modifications ("transgenerational epigenetic inheritance") [1, 2]. They were attempting to separate two very different sets of phenomena that are both described as epigenetic. Many cases of "epigenetics" in recent literature fall into the former category and are not epigenetic at all, but rather are examples of germ cell gene regulation. To be meaningfully distinct from simple "transcription factor → enhancer → promoter → expression" forms of gene regulation, epigenetic phenomena must display three characteristics: they must manifest as (1) heritable

genetic changes that (2) are associated with chromosomes but (3) are not based on DNA sequence. These are criteria that should not be abandoned, but should be evaluated.

The second characteristic is important because it is the essence of epigenetic inheritance. Why? Because if epigenetics did not require chromosome association, every genetic pathway that included a positive-feedback loop would be epigenetic. Female-specific sex-lethal splicing in *Drosophila* to form more active sex-lethal splicing factor would be considered epigenetic. Bacterial expression of LacY, the lactose permease, increasing sensitivity to further exposure to lactose would be considered epigenetic. Autophosphorylation of CaMKII upon witnessing a calcium spike would be considered epigenetic. Suppressor-of-Hairless-induced expression of Notch, the Suppressor-of-Hairless activator, would be considered epigenetic. In short, just about every genetic network could be considered epigenetic, and "epigenetics" would not differ in any meaningful way from "gene regulation." Since proteins, lipids, RNAs, metabolic intermediates, and even toxins are passed through cell division in the cytoplasm, it is trivial to say that their effects are "inherited," and it is wrong to conclude that cells retaining consequences of their antecedents' experiences are necessarily epigenetic.

Without requiring chromosomal inheritance of epigenetic phenomena, expression in the germ line would be sufficient to demarcate any genetic pathway as epigenetic, which would serve merely to rename those genes expressed during the creation of eggs and sperm. It should not be surprising that such networks might span multiple organismal generations; after all it is the mother's genetics (and experiences) that create the egg and the father's genetics (and experiences) that create the sperm; alteration in these processes will certainly result in alterations to the next generation. Mammalian biology aggravates the issue even more, since late-term pregnancies can involve three concentric organisms: by the end of gestation, female mammals contain half-genomes of all their potential grandchildren; the oocytes housing those pronuclei are filled with gene products created by their mothers from the nutritional environment provided by their grandmother. Many cases called epigenetic are instead this form of transgenerational gene regulation.

To discriminate broad concepts of “memory” or “potentiation” in gene regulation from specific epigenetic inheritance, it is necessary to show that the epigenetic factors altering gene activity map specifically to the chromosomal locus being regulated. Experimentally, *epigenetic gene regulation is demonstrated when DNA violates the law of mass action*: two identical sequences can act in different ways despite identity in their sequences and in the proteins that bind to them. Conceptually, if a “naive” DNA was introduced into the system, would it behave as do the existing DNAs? If so, it is not epigenetic. Practically, this is most easily shown by showing that identical pieces of DNA (homologs, duplications, transgenes, individuals of repeated gene arrays, etc.) possess different behaviors. This has been shown for centromere identity [3, 4], genomic imprinting in mammals [5], plants [6–8], fungi [9, 10], nematodes [11], and insects [12–14]; it is this requirement that many examples of “epigenetics” do not test. The strong connection between epigenetics and chromatin structure has only contributed to a conflation of these terms. It is not unusual to find the term “epigenetic” associated with studies that merely show changes in histone modifications of a gene, perhaps even acetylation, with no experiments that test for heritability, sequence polymorphism, or chromosome association.

3. Does Epigenetics Exist?

Changes in nucleotide sequence resulting in phenotypic variants are clear, established, and the very foundation of the neo-Darwinian synthesis that married Darwin’s theories of variation and selection with Mendel’s rules of inheritance. What was, and remains, magical about epigenetics is that substantial variation may be seen with no evident underlying changes in nucleic acid sequence and as such changes are relatively unstable. What first drew attention to epigenetic inheritance was the different behavior of identical genomes, in the variegation as a result of cosuppression which inactivates duplicated gene copies in plants, heterochromatin-induced position effects of *Drosophila* [15], or somatic mosaicism due to X chromosome inactivation in female mammals. These differences in phenotypes would not be surprising if they were due to differences in DNA sequence.

But how carefully have we tested for sequence identity in these cases? Imagine a hypothetical situation. What if creating a centromere required an enzyme (centromerase?) to cut the DNA and insert a specific sequence necessary and sufficient to establish centromere activity? What if cases of neocentromeres were cases of rare random expression and activity of centromerase? What if loss of centromeric activity in dicentric Robertsonian fusion chromosomes was evidence of the reversibility of centromerase? The hypothetical existence of centromerase is unnecessary, to be sure, given what we know about centromeric histones and chromatin structure, but it is illustrative that in many cases specific induced polymorphism is not even considered. We have a mindset that random mutation is the only mechanism allowed to alter DNA sequence, and therefore that rapid, induced, and reversible changes to chromosome behavior must occur without changes in sequence. But this assumes

clearer lines in defining “sequence” than really exist, and it ignores many well-established observations.

Consider mating type switching in *Schizosaccharomyces pombe*. Switching occurs when a silent cassette of information from a “storage” locus is transferred to the active mating-type locus [16–18]. The mechanism of switching requires a mark, likely a break or ribonucleotide on one strand [19]. Tracing the ancestry of this strand has revealed that the altered strand comes from the switched locus in the previous generation. The result is that switching is limited in frequency and direction. A ribonucleotide in a chain of deoxyribonucleic acid is indeed a surprising way to carry information on a chromosome, but nonetheless it is genetic: it is heritable and consequent. And most surprising, it is inducible.

Consider also genomic imprinting in mammals. Is genomic imprinting really epigenetic? Although perhaps the most accepted form of epigenetics, it may be argued that it is not, for trivial nomenclatorial reasons: do you count 5-methylcytosine as cytosine, or as a fifth base that merely has an additional requirement for incorporation (a replication-coupled DNA methyltransferase)? While your answer may reveal something about your philosophy, it has impact on how we think of epigenetic mechanisms. If we count 5-methylcytosine as a fifth base, then the maternally and paternally derived alleles of genomically imprinted genes are indeed polymorphic. Can we also count dehydroxylation or deglycosylation as a polymorphism? Considering these cases of induced polymorphism would exclude both *S. pombe* mating type switching and imprinting at the *Medea* locus (where cytosine methylation induces a strand break on one homolog, alleviating it from silencing) as epigenetic. And why not? Selenocysteine is an amino acid even though a ribosome requires an extensive elaborated system to incorporate it [20, 21]. Methylcytosine is chemically and genetically distinct from cytosine; it merely requires an extensive elaborated system to incorporate it. A nicked DNA strand is again chemically and genetically distinct. It is a fun argument to make but seems overly contrived and unnecessary, and probably a little bizarre. It is not that we need to remove these cases from the list of epigenetics, but rather that we must consider what we mean by “sequence” when using it as the key criterion discriminating “epigenetics” from “polymorphisms.” There is a lot of landscape in that gray area.

4. Something In Between

Understanding how and why we define “sequence” and “epigenetic” is important when categorizing modes of gene regulation. But such considerations also reveal insight into how these phenomena might interact and lead us to consider how important induced polymorphism could be in evolution. The above examples—*Medea*, mating type, and imprinting—are all cases of induced polymorphism which result in changes in genetic activity of the sequence. The fact that they are “sequence independent” is an artifact of our ACGT-sequence bias. Still, it seems doubtful that these handful of examples would by themselves upset our views of

evolution. First, such modes of epigenetic gene regulation are apparently uncommon. It is estimated that there are perhaps hundreds of imprinted loci in mammals, and as few as one in plants. Second, they are not cases of presence/absence of genetic pathways, but rather expression biases of different alleles, and so siblings do not differ markedly because of this mode of regulation; imprinted genes are essentially haploid and so are not much different than sex-linked genes in terms of evolution. Third, they are reset after one generation. It is therefore difficult to imagine that these forms of epigenetic inheritance drive evolution in profound or novel ways.

Are there examples of induced polymorphism that are widespread, consequent, and long-lived and might therefore affect genome evolution?

Almost one-half of the genomes of many popular metazoa are highly polymorphic, but those polymorphisms go unnoticed in genome-wide association, quantitative trait loci, and population genetics studies. This portion, the heterochromatin—alpha and beta repeats, transposable elements, satellites, repetitive sequence, and so forth, all typically linked to centromeres—are not amenable to our modern approaches to genomics. Heterochromatin can comprise hundreds, thousands, and even millions of copies of simple (e.g., AATAT, AAGAG, and AAGAGAG) repeats [22–26]; hence they cannot be easily cloned, sequenced, or assembled using the techniques directed at whole-genome sequencing. In fact, the definition of “whole” has been altered to ignore this half of the genome [27–29]. Quantifying repeat copy number is cumbersome and imprecise, and stumbling upon rare sequence polymorphisms in otherwise homogenous blocks of satellite DNA is lucky [30]. It is therefore difficult to estimate the degree of differences or rate of polymorphism in this substantial portion of the genome.

Heterochromatin was first described by Emil Heitz in the 1920s and 1930s. At the time, its discriminating feature was heteropycnotic staining, which is still arguably the best definition. Subsequently, it was discovered that heterochromatin is generally late replicating, repressive for gene expression, and enriched in specific modifications of the DNA and the histones that package it although there are exceptions to all of these features [15, 31, 32]. What is agreed is that heterochromatin forms easily on highly-repetitive sequence and exists as a complex with heterochromatin proteins (e.g., histone methyltransferases, HP1, and possibly RNAs). Genetic and mutational manipulations that alter the amount of repetitive sequence or protein components demonstrate a natural balance between the sequence and protein components in forming heterochromatin [33–37]. Excess sequence compromises heterochromatin formation elsewhere by competing for limited heterochromatin proteins. Increases or decreases in heterochromatin proteins increase or decrease the ease of forming heterochromatin or increase or decrease the amount of sequence that can be packaged as heterochromatin.

Malik and Henikoff described their view of a specific example of an evolutionary balance at the centromeric chromatin (or “centrochromatin”) [38–40]. They envision a coevolution of sequence expansion and DNA binding by the centromeric histone Cid. Excess centromeric DNA is bound

by Cid, and changes in Cid binding (or expression) result in altered centromeric sequence. This may be an example of a broader mechanism of expansion and contraction limited (or promulgated) by the characteristics of DNA-binding proteins that stabilize repetitive sequence. The mix of multiple polymorphic simple repeats in the genome [25, 26, 41, 42] may be stabilized by a mix of dedicated or overlapping heterochromatin proteins [43–48]. The balance between the sequences and proteins that together form heterochromatin is expected to be important because the protein components of heterochromatin play double duty as general transcriptional regulators [49, 50]. Genes shift between “heterochromatin-like” and “euchromatin-like” as they shift between silent and expressed during development or as a response to environmental stimuli. Mutations in the genes that encode these protein components often act dominantly, suggesting that their dose matters [34, 36]. One can easily imagine a three-way balance between heterochromatic sequence, heterochromatin proteins, and euchromatic gene regulatory mechanisms. This predicts that copy number polymorphisms of heterochromatin-forming sequence might impact gene regulation throughout the genome.

It has been very difficult to test whether copy number polymorphisms are consequential because there are few molecular-genetic tools that allow manipulation of copy number in otherwise isogenic backgrounds. We know from classic studies in *Drosophila*, where the DNA and protein components of heterochromatin are easily manipulated, that the amount of heterochromatic sequence in a cell dramatically affects sensitized variegating genes [33, 51, 52]. At an extreme, multiple supernumerary heterochromatic chromosomes are lethal [53]. Although the reason remains unclear, one can imagine such a disruption in sequence-to-protein balance to cause massive misregulation of many genes. Y chromosomes captured from wild populations vary in their ability to affect heterochromatin-induced position effect variegation and euchromatic gene expression elsewhere in the genome [54–56], yet have very few protein-encoding genes [57–59], strongly suggesting that heterochromatin polymorphisms, perhaps copy number polymorphisms, affect gene expression throughout the genome. Our work has induced copy number variation in one repeat, the ribosomal DNA (rDNA) [60]. The rDNA has precedent for housing-induced phenotypic variation in plants [61, 62], but without being able to induce changes at the rDNA, it had been difficult to test this phenomenon further. In flies, however, induced copy number variation has consequences for heterochromatin-induced position effect variegation and gene expression across the genome [63, 64]. These variations in gene regulation overlap with those seen from isolated natural Y chromosomes [54, 64], suggesting a significant portion of natural variance in rDNA repeat copy number [65, 66] may contribute to phenotypic variance in natural populations. Equally importantly, much of the variance that maps to the Y chromosome does not map to rDNA, suggesting that most phenotypic variance maps to other sequences on the Y, perhaps to the other repeats that are less experimentally manipulable.

Natural variation in repeat sequence copy number may play a role in evolution, but the uniquely dynamic biology of the rDNA implies the more exciting possibility. Changes in copy number may be induced and inherited.

The rDNA contains interspersed active and inactive rRNA genes and thus contains characteristics of both euchromatin and heterochromatin in some cells. The physical manifestation of the tremendous expression and processing of the rRNAs is the nucleolus. The stability of these long stretches of direct repeats in the nucleus is likely due to the heterochromatic packaging of a subset of the repeats. Peng and Karpen observed multiple nucleoli in postmitotic cells of animals carrying mutations of suppressor-of-variegation genes, which encode the protein components of heterochromatin and regulate the rDNA [67]. Their results suggest that repeat sequence not packaged as heterochromatin is unstable and prone to damage/repair or intrachromosomal recombination. Our experiments showed that mutation of suppressor-of-variegation genes resulted in destabilization and reduction of rDNA copy number through mitosis [63]. We further quantified loss in the soma and also showed that loss was seen through the germline, resulting in a permanent decrease of rDNA copy number in a population after exposure to a mutation that disrupts heterochromatin formation. We more recently showed that mutation of a repressor of rDNA expression (CCCTC-Factor, or CTCF) also destabilizes the rDNA, resulting in permanent loss [68]. These results are consistent with heterochromatin-like silencing stabilizing repeated DNA sequence, and a balance between repeat sequences and the protein components that regulate them.

In *Drosophila*, the ribosomal DNA is a compelling compartment because its dynamism is unmatched. It is the most highly expressed set of genes in the genome [69], coordinates the activity of all three polymerases, shrinks naturally through the formation of extrachromosomal circles [70], can repair itself through meiotic magnification or somatic pseudomagnification [71–75], and can compensate its output through alteration of elongation rate and possibly initiation rate [76, 77]. It possesses side-by-side copies that possess heterochromatic and euchromatic chromatin structures [76, 78–80]. As the central body in protein synthetic capacity, it is also responsive to nutritional status, sensitive to toxins and drugs, and susceptible to instability by alterations of gene products necessary for its regulation [81–87]. Altering regulation of the rDNA through mutation or drug treatment affects not only rRNA output, but also stability [88–90]. Alteration of the activity of a protein component of heterochromatin might therefore affect the copy number of the sequence to which it binds.

Dynamism (of rDNA) and balance (of heterochromatic sequence and proteins) establishes a situation of heterochromatin homeostasis (Figure 2). Sequences are protected from loss by packaging as heterochromatin. Loss of protein (or reduced protein activity, arrow “a”) would destabilize repeat DNA (white state) and result in loss, reestablishing an equilibrium (arrow “b” to the gray state). Similarly, excess sequence would revert through loss if there is not sufficient protein to package it for stability. But excess protein is not

without consequence, since any heterochromatin protein not bound in constitutive sequence would alter gene expression throughout the genome (dark gray state), favoring either reduced protein expression/activity (arrow “c”) or expansion of repeat sequence (arrow “d”) to reestablish balance (light gray states). On the whole, the instability of repeat sequence and the consequence of excess heterochromatin proteins creates multiple states that balance the factors and naturally drives the number of repeat sequences and protein expression to equilibrate. Of course, any external factors that influence heterochromatin protein activity would be expected to result in induced and heritable changes in repetitive DNA copy number. The rDNA is particularly sensitive to induced copy number polymorphism, since it is affected by nutritional status throughout the lifetime of an organism and rDNA copy number exists in excess of what is required for translational demands, allowing some plasticity in copy number without being unduly disadvantageous.

On the surface, induced copy number polymorphism is similar to epigenetic modification (particularly if one cannot easily sequence and assemble repetitious DNAs), and the ability of repeat sequences to change in copy number relatively easily adds the degree of volatility common in epigenetic gene regulation. Unlike many forms of transgenerational gene regulatory effects, induced copy number polymorphisms are linked to chromosomes, and thus are both heritable and selectable. Unlike epigenetic regulation of imprinted or inactivated chromosomes, induced copy number polymorphisms can be inherited over multiple generations. But like both transgenerational and epigenetic effects, the role of induced polymorphism is only beginning to be considered in evolution. Such investigation will likely be done in simple organisms, such as *Drosophila*, that have relatively simple rDNA architecture [91, 92]. By contrast, humans have multiple rDNA arrays which change in size frequently [93], and the complex regulation that renders some arrays active and others inactive means it may be some time before we understand how rDNA polymorphisms and rDNA instability [94] contribute to phenotypic variance in human population or to disease etiology.

5. Is the rDNA Special?

Induced polymorphism of rDNA copy number offers a convenient mechanism by which changes may be inherited although the same objections apply here as they do for the environmentally induced changes in gene expression that craned Lamarck’s neck: how is the germline affected? In the case of induced polymorphism, germ cells may be more, not less, sensitive to induced alterations in heterochromatin composition, for three reasons. First, in many cases, gene expression is limited in these cell types. Perdurance of heterochromatin proteins, or the presence of ample gene product to endure fluctuation in gene activity, may be less in these cell types. Second, at least in males, the genome is stripped of most somatic chromatin components in favor of packaging proteins and polyamines. This may increase the sensitivity of such chromosomes to DNA rearrangements or specifically mark some regions for hypervariability. Third,

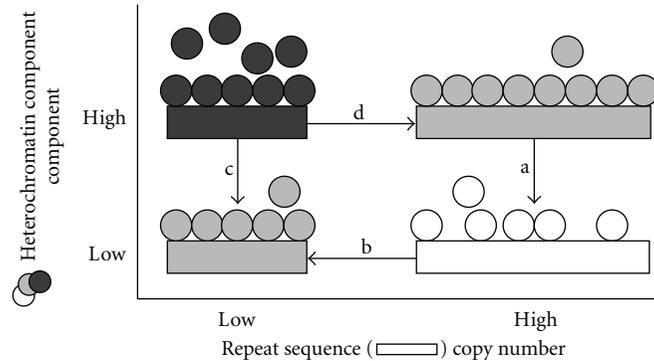


FIGURE 2: An illustration of a balance between heterochromatin sequences and heterochromatin components (e.g., proteins or RNAs). Repetitious heterochromatin-forming sequences (*rectangles*) are normally in balance with the proteins that bind them (*circles*), package them as heterochromatin, and thereby stabilize them (conditions in *gray*). Since these factors are used to regulate expression of euchromatic genes, the balance must accommodate “excess” factors for that purpose (denoted as circles apart from rectangles). If the expression or activity of proteins is reduced (a), repetitious sequence is exposed, destabilized, and lost through damage-repair, recombination, or extrachromosomal circle formation (b), until a new balance is established. Excess protein has gene regulatory consequence throughout the genome and presses to reestablish balance by altering expression level or activity (c) or perhaps through repeat expansion (d).

germ cells naturally undergo recombination at a high rate. It is well established that changes in microsatellite and rDNA copy number occur in meiosis, while the same sequences are relatively stable in mitosis. The challenge is to understand what identifies a gene as “sensitive to rDNA copy number,” because it would be those genes selected for phenotypic variation in response to rDNA copy number changes.

We do not yet understand whether repeated sequences are different from “nonexpressed” sequences in ability to be induced to change, but we do know from mutational and molecular analyses that “heterochromatin” is not monolithic and is more accurately thought of as multiple “colors” [95]. Mutations may affect one chroma of heterochromatin and not another [96]. The five enumerated chromas significantly expand our understanding of chromatin structure, but even those five are likely still a simplification caused by our failure to resolve more subtle differences. Cumbersome work has detected alterations of repeat sequence copy number in few studies, suggesting that this may be a very widespread form of genetic variation [66, 97, 98]. Peng and Karpen showed an increase in DNA damage repair foci in the heterochromatin of suppressor-of-variegation mutants in diploid cells [99, 100]. They did not identify the sequences that were being repaired, but the number and distribution of repair foci in the nuclei indicated that it was not clustered (i.e., limited to the rDNA arrays). This observation suggests that the heterochromatin formed on simple repeats (and not just the highly-expressed rDNA) also is stabilized by packaging as heterochromatin. As our understanding of what heterochromatin is, and as tools become available to probe it in more surgical ways, we may begin to unravel complex interactions between types of heterochromatin as they struggle to keep each other in check or ally to fend off common enemies.

The term “epigenetics” may retain its strict definitions of chromosome-bound nonsequence-based genetic information, or it may be expanded to include induced mutation

or gene regulatory networks that impact subsequent generations. In the end, all forms of regulation are genetic, and so are salient in understanding how complex, pleiotropic, and epistatic genetic interactions conspire to create phenotypes. However one defines epigenetics, its legacy is that we cannot understand the comprehensive synthesis of forces that drive a genome’s evolution without understanding how all the alleles within that genome are regulated.

References

- [1] N. A. Youngson and E. Whitelaw, “Transgenerational epigenetic effects,” *Annual Review of Genomics and Human Genetics*, vol. 9, pp. 233–257, 2008.
- [2] S. Chong, N. A. Youngson, and E. Whitelaw, “Heritable germline epimutation is not the same as transgenerational epigenetic inheritance,” *Nature Genetics*, vol. 39, no. 5, pp. 574–575, 2007.
- [3] K. A. Maggert and G. H. Karpen, “The activation of a neocentromere in *Drosophila* requires proximity to an endogenous centromere,” *Genetics*, vol. 158, no. 4, pp. 1615–1628, 2001.
- [4] K. A. Maggert and G. H. Karpen, “Acquisition and metastability of centromere identity and function: sequence analysis of a human neocentromere,” *Genome Research*, vol. 10, no. 6, pp. 725–728, 2000.
- [5] A. C. Ferguson-Smith, “Genomic imprinting: the emergence of an epigenetic paradigm,” *Nature Reviews Genetics*, vol. 12, no. 8, pp. 565–575, 2011.
- [6] M. Gehring, J. H. Huh, T. F. Hsieh et al., “DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation,” *Cell*, vol. 124, no. 3, pp. 495–506, 2006.
- [7] P. Wolff, I. Weinhofer, J. Seguin et al., “High-Resolution analysis of parent-of-origin allelic expression in the Arabidopsis endosperm,” *PLoS Genetics*, vol. 7, no. 6, Article ID e1002126, 2011.
- [8] M. Luo, J. M. Taylor, A. Spriggs et al., “A Genome-Wide survey of imprinted genes in rice seeds reveals imprinting

- primarily occurs in the endosperm," *PLoS Genetics*, vol. 7, no. 6, Article ID e1002125, 2011.
- [9] A. Kaykov and B. Arcangioli, "A programmed strand-specific and modified nick in *S. pombe* constitutes a novel type of chromosomal imprint," *Current Biology*, vol. 14, no. 21, pp. 1924–1928, 2004.
 - [10] S. Vengrova and J. Z. Dalggaard, "RNase-sensitive DNA modification(s) initiates *S. pombe* mating-type switching," *Genes and Development*, vol. 18, no. 7, pp. 794–804, 2004.
 - [11] J. K. Arico, D. J. Katz, J. van der Vlag, and W. G. Kelly, "Epigenetic patterns maintained in early *Caenorhabditis elegans* embryos can be established by gene activity in the parental germ cells," *PLoS Genetics*, vol. 7, no. 6, Article ID e1001391, 2011.
 - [12] K. A. Maggert and K. G. Golic, "The Y chromosome of *Drosophila melanogaster* exhibits chromosome-wide imprinting," *Genetics*, vol. 162, no. 3, pp. 1245–1258, 2002.
 - [13] V. K. Lloyd, D. A. Sinclair, and T. A. Grigliatti, "Genomic imprinting and position-effect variegation in *Drosophila melanogaster*," *Genetics*, vol. 151, no. 4, pp. 1503–1516, 1999.
 - [14] B. S. Haller and R. C. Woodruff, "Varied expression of a Y-linked P[W+] insert due to imprinting in *Drosophila melanogaster*," *Genome*, vol. 43, no. 2, pp. 285–292, 2000.
 - [15] K. S. Weiler and B. T. Wakimoto, "Heterochromatin and gene expression in *Drosophila*," *Annual Review of Genetics*, vol. 29, pp. 577–605, 1995.
 - [16] J. I. Nakayama, A. J. S. Klar, and S. I. S. Grewal, "A chromodomain protein, Swi6, perform imprinting functions in fission yeast during mitosis and meiosis," *Cell*, vol. 101, no. 3, pp. 307–317, 2000.
 - [17] J. Z. Dalggaard and A. J. S. Klar, "Orientation of DNA replication establishes mating-type switching pattern in *S. pombe*," *Nature*, vol. 400, no. 6740, pp. 181–184, 1999.
 - [18] A. J. S. Klar, "Regulation of fission yeast mating-type interconversion by chromosome imprinting," *Development*, vol. 110, pp. 3–8, 1990.
 - [19] S. Vengrova, J. Z. Dalggaard, B. Arcangioli, and A. Kaykov, "The Schizosaccharomyces pombe imprint—Nick or ribonucleotide(s)," *Current Biology*, vol. 15, no. 9, pp. R326–R327, 2005.
 - [20] T. C. Stadtman, "Selenocysteine," *Annual Review of Biochemistry*, vol. 65, pp. 83–100, 1996.
 - [21] R. Longtin, "A forgotten debate: is selenocysteine the 21st amino acid?" *Journal of the National Cancer Institute*, vol. 96, no. 7, pp. 504–505, 2004.
 - [22] S. Pimpinelli, M. Gatti, and A. De Marco, "Evidence for heterogeneity in heterochromatin of *Drosophila melanogaster*," *Nature*, vol. 256, no. 5515, pp. 335–337, 1975.
 - [23] S. Pimpinelli, G. Santini, and M. Gatti, "Characterization of *Drosophila* heterochromatin. II. C and N banding," *Chromosoma*, vol. 57, no. 4, pp. 377–386, 1976.
 - [24] M. Gatti, S. Pimpinelli, and G. Santini, "Characterization of *Drosophila* chromatin. I. Staining and decondensation with Hoechst 33258 and quinacrine," *Chromosoma*, vol. 57, no. 4, pp. 351–375, 1976.
 - [25] S. Bonaccorsi and A. Lohe, "Fine mapping of satellite DNA sequences along the Y chromosome of *Drosophila melanogaster*: relationships between satellite sequences and fertility factors," *Genetics*, vol. 129, no. 1, pp. 177–189, 1991.
 - [26] A. R. Lohe, A. J. Hilliker, and P. A. Roberts, "Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*," *Genetics*, vol. 134, no. 4, pp. 1149–1174, 1993.
 - [27] E. S. Lander, L. M. Linton, B. Birren et al., "Initial sequencing and analysis of the human genome," *Nature*, vol. 409, no. 6822, pp. 860–921, 2001.
 - [28] J. C. Venter, M. D. Adams, E. W. Myers et al., "The sequence of the human genome," *Science*, vol. 291, no. 5507, pp. 1304–1351, 2001.
 - [29] M. D. Adams, S. E. Celniker, R. A. Holt et al., "The genome sequence of *Drosophila melanogaster*," *Science*, vol. 287, no. 5461, pp. 2185–2195, 2000.
 - [30] X. Sun, J. Wahlstrom, and G. Karpen, "Molecular structure of a functional *Drosophila* centromere," *Cell*, vol. 91, no. 7, pp. 1007–1019, 1997.
 - [31] J. C. Yasuhara and B. T. Wakimoto, "Molecular landscape of modified histones in *Drosophila* heterochromatic genes and euchromatin-heterochromatin transition zones," *PLoS Genetics*, vol. 4, no. 1, p. e16, 2008.
 - [32] K. Weiler and B. Wakimoto, "Suppression of heterochromatic gene variegation can be used to distinguish and characterize E(var) genes potentially important for chromosome structure in *Drosophila melanogaster*," *Molecular Genetics and Genomics*, vol. 266, no. 6, pp. 922–932, 2001.
 - [33] J. B. Spofford and R. DeSalle, "Nucleolus organizer-suppressed position-effect variegation in *Drosophila melanogaster*," *Genetical Research*, vol. 57, no. 3, pp. 245–255, 1991.
 - [34] G. Reuter and P. Spierer, "Position effect variegation and chromatin proteins," *BioEssays*, vol. 14, no. 9, pp. 605–612, 1992.
 - [35] G. Wustmann, J. Szidonya, H. Taubert, and G. Reuter, "The genetics of position - effect variegation modifying loci in *Drosophila melanogaster*," *Molecular and General Genetics*, vol. 217, no. 2-3, pp. 520–527, 1989.
 - [36] G. Reuter, W. Werner, and H. J. Hoffmann, "Mutants affecting position-effect heterochromatinization in *Drosophila melanogaster*," *Chromosoma*, vol. 85, no. 4, pp. 539–551, 1982.
 - [37] G. Reuter and I. Wolff, "Isolation of dominant suppressor mutations for position-effect variegation in *Drosophila melanogaster*," *MGG Molecular & General Genetics*, vol. 182, no. 3, pp. 516–519, 1981.
 - [38] H. S. Malik and S. Henikoff, "Adaptive evolution of Cid, a centromere-specific histone in *Drosophila*," *Genetics*, vol. 157, no. 3, pp. 1293–1298, 2001.
 - [39] B. A. Sullivan and G. H. Karpen, "Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin," *Nature Structural and Molecular Biology*, vol. 11, no. 11, pp. 1076–1083, 2004.
 - [40] H. S. Malik and S. Henikoff, "Conflict begets complexity: the evolution of centromeres," *Current Opinion in Genetics and Development*, vol. 12, no. 6, pp. 711–718, 2002.
 - [41] A. R. Lohe and P. A. Roberts, "Evolution of DNA in heterochromatin: the *Drosophila melanogaster* sibling species subgroup as a resource," *Genetica*, vol. 109, no. 1-2, pp. 125–130, 2000.
 - [42] A. R. Lohe and D. L. Brutlag, "Multiplicity of satellite DNA sequences in *Drosophila melanogaster*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 83, no. 3, pp. 696–700, 1986.
 - [43] E. B. Tchoubrieva and J. B. Gibson, "Conserved (CT)n.(GA)n repeats in the non-coding regions at the Gpdh locus are binding sites for the GAGA factor in *Drosophila melanogaster* and its sibling species," *Genetica*, vol. 121, no. 1, pp. 55–63, 2004.

- [44] J. W. Raff, R. Kellum, and B. Alberts, "The *Drosophila* GAGA transcription factor is associated with specific regions of heterochromatin throughout the cell cycle," *EMBO Journal*, vol. 13, no. 24, pp. 5977–5983, 1994.
- [45] L. Perrin, O. Demakova, L. Fanti et al., "Dynamics of the sub-nuclear distribution of Modulo and the regulation of position-effect variegation by nucleolus in *Drosophila*," *Journal of Cell Science*, vol. 111, no. 18, pp. 2753–2761, 1998.
- [46] C. Monod, N. Aulner, O. Cuvier, and E. Käs, "Modification of position-effect variegation by competition for binding to *Drosophila* satellites," *EMBO Reports*, vol. 3, no. 8, pp. 747–752, 2002.
- [47] N. Aulner, C. Monod, G. Mandicourt et al., "The AT-hook protein D1 is essential for *Drosophila melanogaster* development and is implicated in position-effect variegation," *Molecular and Cellular Biology*, vol. 22, no. 4, pp. 1218–1232, 2002.
- [48] F. Girard, B. Bello, U. K. laemml, and W. J. Gehring, "In vivo analysis of scaffold-associated regions in *Drosophila*: a synthetic high-affinity SAR binding protein suppresses position effect variegation," *EMBO Journal*, vol. 17, no. 7, pp. 2079–2085, 1998.
- [49] L. Fanti and S. Pimpinelli, "HP1: a functionally multifaceted protein," *Current Opinion in Genetics and Development*, vol. 18, no. 2, pp. 169–174, 2008.
- [50] G. Schotta, A. Ebert, and G. Reuter, "SU(VAR)3-9 is a conserved key function in heterochromatic gene silencing," *Genetica*, vol. 117, no. 2-3, pp. 149–158, 2003.
- [51] J. W. Gowen and E. H. Gay, "Chromosome constitution and behavior in eversporting and mottling in *Drosophila melanogaster*," *Genetics*, vol. 19, no. 3, pp. 189–208, 1934.
- [52] P. Dimitri and C. Pisano, "Position effect variegation in *Drosophila melanogaster*: relationship between suppression effect and the amount of Y chromosome," *Genetics*, vol. 122, no. 4, pp. 793–800, 1989.
- [53] M. Ashburner, K. G. Golic, and R. S. Hawley, *Drosophila: A Laboratory Handbook*, Cold Spring Harbor Press, New York, NY, USA, 2nd edition, 2005.
- [54] B. Lemos, L. O. Araripe, and D. L. Hartl, "Polymorphic Y chromosomes harbor cryptic variation with manifold functional consequences," *Science*, vol. 319, no. 5859, pp. 91–93, 2008.
- [55] B. Lemos, A. T. Branco, and D. L. Hartl, "Epigenetic effects of polymorphic Y chromosomes modulate chromatin components, immune response, and sexual conflict," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 36, pp. 15826–15831, 2010.
- [56] P. P. Jiang, D. L. Hartl, and B. Lemos, "Y not a dead end: epistatic interactions between Y-linked regulatory polymorphisms and genetic background affect global gene expression in *Drosophila melanogaster*," *Genetics*, vol. 186, no. 1, pp. 109–118, 2010.
- [57] D. L. Lindsley and G. G. Zimm, *The Genome of Drosophila Melanogaster*, Academic Press, San Diego, Calif, USA, 1992.
- [58] A. B. Carvalho, B. A. Dobo, M. D. Vibranovski, and A. G. Clark, "Identification of five new genes on the Y chromosome of *Drosophila melanogaster*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 23, pp. 13225–13230, 2001.
- [59] M. D. Vibranovski, L. B. Koerich, and A. B. Carvalho, "Two new Y-linked genes in *Drosophila melanogaster*," *Genetics*, vol. 179, no. 4, pp. 2325–2327, 2008.
- [60] S. Paredes and K. A. Maggert, "Expression of I-CreI endonuclease generates deletions within the rDNA of *Drosophila*," *Genetics*, vol. 181, no. 4, pp. 1661–1671, 2009.
- [61] C. A. Cullis, "Mechanisms and control of rapid genomic changes in flax," *Annals of Botany*, vol. 95, no. 1, pp. 201–206, 2005.
- [62] R. G. Schneeberger and C. A. Cullis, "Specific DNA alterations associated with the environmental induction of heritable changes in flax," *Genetics*, vol. 128, no. 3, pp. 619–630, 1991.
- [63] S. Paredes and K. A. Maggert, "Ribosomal DNA contributes to global chromatin regulation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 42, pp. 17829–17834, 2009.
- [64] S. Paredes, A. T. Branco, D. L. Hartl, K. A. Maggert, and B. Lemos, "Ribosomal dna deletions modulate genome-wide gene expression: "rDNA-sensitive" genes and natural variation," *PLoS Genetics*, vol. 7, no. 4, Article ID e1001376, 2011.
- [65] E. M. S. Lyckegaard and A. G. Clark, "Ribosomal DNA and Stellate gene copy number variation on the Y chromosome of *Drosophila melanogaster*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 86, no. 6, pp. 1944–1948, 1989.
- [66] E. O. Long and I. B. Dawid, "Repeated genes in eukaryotes," *Annual Review of Biochemistry*, vol. 49, pp. 727–764, 1980.
- [67] J. C. Peng and G. H. Karpen, "H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability," *Nature Cell Biology*, vol. 9, no. 1, pp. 25–35, 2007.
- [68] P. A. Guerrero and K. A. Maggert, "The CCCTC-binding factor (CTCF) of *Drosophila* contributes to the regulation of the ribosomal DNA and nucleolar stability," *PLoS ONE*, vol. 6, no. 1, 2011.
- [69] J. R. Warner, "The economics of ribosome biosynthesis in yeast," *Trends in Biochemical Sciences*, vol. 24, no. 11, pp. 437–440, 1999.
- [70] S. Cohen, K. Yacobi, and D. Segal, "Extrachromosomal circular DNA of tandemly repeated genomic sequences in *Drosophila*," *Genome Research*, vol. 13, no. 6 A, pp. 1133–1145, 2003.
- [71] K. D. Tartof, "Regulation of ribosomal RNA gene multiplicity in *Drosophila melanogaster*," *Genetics*, vol. 73, no. 1, pp. 57–71, 1973.
- [72] K. D. Tartof, "Unequal mitotic sister chromatid exchange as the mechanism of ribosomal RNA gene magnification," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 71, no. 4, pp. 1272–1276, 1974.
- [73] R. S. Hawley and K. D. Tartof, "A two-stage model for the control of rDNA magnification," *Genetics*, vol. 109, no. 4, pp. 691–700, 1985.
- [74] R. Terracol and N. Prud'Homme, "Differential elimination of rDNA genes in bobbed mutants of *Drosophila melanogaster*," *Molecular and Cellular Biology*, vol. 6, no. 4, pp. 1023–1031, 1986.
- [75] R. Terracol, Y. Iturbide, and N. Prud'Homme, "Partial reversion at the bobbed locus of *Drosophila melanogaster*," *Biology of the Cell*, vol. 68, no. 1–3, pp. 65–71, 1990.
- [76] B. McStay and I. Grummt, "The epigenetics of rRNA genes: from molecular to chromosome biology," *Annual Review of Cell and Developmental Biology*, vol. 24, pp. 131–157, 2008.
- [77] D. A. Schneider, A. Michel, M. L. Sikes et al., "Transcription elongation by RNA polymerase I is linked to efficient rRNA

- processing and ribosome assembly,” *Molecular Cell*, vol. 26, no. 2, pp. 217–229, 2007.
- [78] M. Jamrich and O. L. Miller, “The rare transcripts of interrupted rRNA genes in *Drosophila melanogaster* are processed or degraded during synthesis,” *EMBO Journal*, vol. 3, no. 7, pp. 1541–1545, 1984.
- [79] O. L. Miller and B. R. Beatty, “Visualization of nucleolar genes,” *Science*, vol. 164, no. 3882, pp. 955–957, 1969.
- [80] D. G. Eickbush, J. Ye, X. Zhang, W. D. Burke, and T. H. Eickbush, “Epigenetic regulation of retrotransposons within the nucleolus of *Drosophila*,” *Molecular and Cellular Biology*, vol. 28, no. 20, pp. 6452–6461, 2008.
- [81] A. Murayama, K. Ohmori, A. Fujimura et al., “Epigenetic control of rDNA loci in response to intracellular energy status,” *Cell*, vol. 133, no. 4, pp. 627–639, 2008.
- [82] L. Li, B. A. Edgar, and S. S. Grewal, “Nutritional control of gene expression in *Drosophila* larvae via TOR, Myc and a novel cis-regulatory element,” *BMC Cell Biology*, vol. 11, article no. 7, 2010.
- [83] S. S. Grewal, L. Li, A. Orian, R. N. Eisenman, and B. A. Edgar, “Myc-dependent regulation of ribosomal RNA synthesis during *Drosophila* development,” *Nature Cell Biology*, vol. 7, no. 3, pp. 295–302, 2005.
- [84] S. S. Grewal, J. R. Evans, and B. A. Edgar, “*Drosophila* TIF-IA is required for ribosome synthesis and cell growth and is regulated by the TOR pathway,” *Journal of Cell Biology*, vol. 179, no. 6, pp. 1105–1113, 2007.
- [85] C. M. Koh, B. Gurel, S. Sutcliffe et al., “Alterations in nucleolar structure and gene expression programs in prostatic neoplasia are driven by the MYC oncogene,” *American Journal of Pathology*, vol. 178, no. 4, pp. 1824–1834, 2011.
- [86] T. Kumazawa, K. Nishimura, T. Kuroda et al., “Novel nucleolar pathway connecting intracellular energy status with p53 activation,” *Journal of Biological Chemistry*, vol. 286, no. 23, pp. 20861–20869, 2011.
- [87] S. Boulon, B. J. Westman, S. Hutten, F. M. Boisvert, and A. I. Lamond, “The Nucleolus under Stress,” *Molecular Cell*, vol. 40, no. 2, pp. 216–227, 2010.
- [88] P. Oberdoerffer, S. Michan, M. McVay et al., “SIRT1 redistribution on chromatin promotes genomic stability but alters gene expression during aging,” *Cell*, vol. 135, no. 5, pp. 907–918, 2008.
- [89] C. E. Fritze, K. Verschuere, R. Strich, and R. E. Esposito, “Direct evidence for SIR2 modulation of chromatin structure in yeast rDNA,” *EMBO Journal*, vol. 16, no. 21, pp. 6495–6509, 1997.
- [90] C. W. Ha and W. K. Huh, “Rapamycin increases rDNA stability by enhancing association of Sir2 with rDNA in *Saccharomyces cerevisiae*,” *Nucleic Acids Research*, vol. 39, no. 4, pp. 1336–1350, 2011.
- [91] M. Pellogrini, J. Manning, and N. Davidson, “Sequence arrangement of the rDNA of *Drosophila melanogaster*,” *Cell*, vol. 10, no. 2, pp. 213–224, 1977.
- [92] P. K. Wellauer and I. B. Dawid, “The structural organization of ribosomal DNA in *Drosophila melanogaster*,” *Cell*, vol. 10, no. 2, pp. 193–212, 1977.
- [93] D. M. Stults, M. W. Killen, H. H. Pierce, and A. J. Pierce, “Genomic architecture and inheritance of human ribosomal RNA gene clusters,” *Genome Research*, vol. 18, no. 1, pp. 13–18, 2008.
- [94] M. Derenzini, L. Montanaro, and D. Treré, “What the nucleolus says to a tumour pathologist,” *Histopathology*, vol. 54, no. 6, pp. 753–762, 2009.
- [95] G. J. Filion, J. G. van Bommel, U. Braunschweig et al., “Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells,” *Cell*, vol. 143, no. 2, pp. 212–224, 2010.
- [96] S. Phalke, O. Nickel, D. Walluscheck, F. Hortig, M. C. Onorati, and G. Reuter, “Retrotransposon silencing and telomere integrity in somatic cells of *Drosophila* depends on the cytosine-5 methyltransferase DNMT2,” *Nature Genetics*, vol. 41, no. 6, pp. 696–702, 2009.
- [97] G. C. Kuhn et al., “The 1.688 repetitive DNA of *Drosophila*: concerted evolution at different genomic scales and association with genes,” *Molecular Biology and Evolution*. In press.
- [98] J. Davison, A. Tyagi, and L. Comai, “Large-scale polymorphism of heterochromatic repeats in the DNA of *Arabidopsis thaliana*,” *BMC Plant Biology*, vol. 7, article 44, 2007.
- [99] J. C. Peng and G. H. Karpen, “Epigenetic regulation of heterochromatic DNA stability,” *Current Opinion in Genetics and Development*, vol. 18, no. 2, pp. 204–211, 2008.
- [100] J. C. Peng and G. H. Karpen, “Heterochromatic genome stability requires regulators of histone H3 K9 methylation,” *PLoS Genetics*, vol. 5, no. 3, Article ID e1000435e, 2009.