The Metastable Genome

A Lamarckian Organ in a Darwinian World?

Ehud Lamm

In the context of a workshop celebrating the bicentennial of Lamarck's *Philosophie zoologique,* a paper focused on the genome may seem overly reductionistic and insensitive to the reciprocal interaction between the organism and its environment. My discussion of the dynamic nature of the genome will illustrate why overly zealous reduction fails even at the level of understanding the genome, and why the genome is a significant developmental unit.

This article is arranged around two general claims and a thought experiment.

I begin by suggesting that the genome should be studied as a developmental system, and that genes supervene on genomes (rather than the other way around). I move on to present a thought experiment that illustrates the implications a dynamic view of the genome has for central concepts in biology, in particular the information content of the genome, and the notion of responses to stress.

The Developing Genome

The genome, understood as the concrete physicochemical *system* carrying hereditary and developmental information—including the DNA, the non-DNA components of the chromosomes, and mechanisms related to them—develops during ontogeny, both as part of the regular cell cycle and through organized responses to various stimuli

(Lamm and Jablonka 2008). Thus, the genome and the set of interrelated developmental processes immediately associated with it should be considered a developmental system, not because it encapsulates the genes, nor because it is a physically localized entity, but rather because genomes exhibit recurring developmental processes. Moreover, these processes and related mechanisms (specifically, genomic epigenetic mechanisms or GEMs) share an evolutionary history, and in all likelihood evolved at least partly for their participation in these developmental processes, resulting in a system that reacts in a consistent way to developmental cues of various kinds, genetic cues among them (cf. Griffiths and Gray 1994).

It might seem that by talking about the genome as a delineated developmental system nested inside the developmental system constituting the organism, we implicitly privilege the genetic system in relation to other developmental resources. Individuating the genome as a developmental system, however, does not have to rely on the assumption that genes have a unique role in determining the organism. Developmental systems may be embedded inside one another (Griffiths and Gray 1994: 294–295), and thus considering the genome as a developmental system does not rule out the organism of which the genome is part, as well as more encompassing systems, from being developmental systems as well. These developmental systems can interact and affect the genome.

Viewing the genome as an integrated developmental system is used here as an explanatory strategy rather than an ontological claim. The rest of this chapter attempts to highlight some of the implications of this approach.

Genomes as Prior to Genes

My next explanatory move is to argue that genomes (as active, responsive systems) should be considered as conceptually prior to genes. This claim can be loosely summarized by the claim that *genes supervene on genomes*. A convenient biological summary is that *genes are manifestations of the physiology of genomes.*

The problems associated with the gene concept have been recognized by many (Pearson 2006). As Keller (2000) argues, the gene concept has in all likelihood run its course, and the gene as an explanatory concept combining both biological structure and biological function will have to be replaced with a different explanatory framework. The discussion below will make it apparent that some of the difficulties with the gene concept Keller and others have identified (for example, that genes may be coded in pieces that need to be combined, that the same sequence can result in different proteins due to alternative splicing, and the different ways in which gene regulation can span multiple genes) are the result of mechanisms that depend on genomic context. Indeed, it can be argued that the source of the difficulties in defining the gene and individuating genes is that the definition must acknowledge developmental and genomic context and mechanisms.

The basic strategy for substantiating the claim that genomes should be considered as prior to genes is to show that properties associated with genes are in fact properties of genomes. The classic gene concept articulated the gene as a unit of structure, function, mutation, and recombination. These properties are problematized in different ways, depending on how the relation between genes and genomes is elucidated. I suggest that there are three fundamental ways of interpreting the

gene/genome relation, which I call the Waddingtonian, the Mendelian, and the McClintock interpretations.

Dominance and recessivity are two generic properties often attributed to "genes." However, the degree of dominance of an allele is influenced, as Waddington emphasized, by the whole set of genes (Waddington 1961:62). Generic properties of genes are thus seen as the result of (*generic*) properties of the organism's developmental system. They cannot be properties of genes per se, but rather are properties of genes by virtue of genomic properties (indeed, properties of gene networks). It could be argued, however, that recessivity and dominance should indeed not be considered properties of genes per se but rather as properties of organismic development, and that the problems with attributing them to genes rather than to genomes do not generalize to other properties of genes, specifically nonfunctional ones.

The Mendelian interpretation of the relation between genes and genomes, in contrast, focuses on the developmental aspects of the genome, not those of the organism. If Mendelian inheritance is the result of chromosomal organization and developmental dynamics, then the generic properties of the genetic system are seen to be the result of the developmental behavior of the genomic system.

A third way of privileging genomes, which makes explicit the functional significance of their ontogenetic capabilities, is inspired by Barbara McClintock's arguments about genome responses to challenges (McClintock 1984). This suggests that the genome is a reactive system, and hence possibly homeostatic or even teleological.

The basic idea behind the claim that genomes should be considered as explanatorily prior to genes would be that properties of genes either originate from, or need to survive, genomic behavior. Genes, then, are not individuated entities apart from genome mechanics; they are passive, and do not control their own fate after they experience “random mutations.” Gene networks and genomes, in contrast, can respond to stimuli, reorganize, and adapt. If we observe “Darwinian genes” that are individuated and for the most part suffer random mutations passively, this should be attributed to the reaction of the genome—or lack of it.

The Dynamic Genome

It is well known that the genome is highly dynamic over various timescales and across wide intra- and interchromosomal distances. Taken together, the wide variety of genomic mechanisms, the timing and timescales of their operation, the scope of the changes, as well as their relationship to gene expression, and hence to development, demand a radical reevaluation of the genome as a static container of genetic information.

The following is a cursory look at some of the types of genomic maintenance mechanisms. It will illustrate some of the activity that is going on in eukaryotic genomes simultaneously with gene expression (transcription), and in addition to DNA replication, the two prominent functions typically attributed to genomes (for further details see appendix B of this volume).

DNA is wrapped around histone proteins, forming the nucleosomes that make up the chromatin. There are several histone variants (e.g., histones appearing only in

specific genomic regions), and histones may be in one out of several possible states. These affect the physical conformation of the chromatin, and hence gene expression. This is shown most clearly in the differences between the typically coding-gene rich, and not highly condensed, euchromatin, and gene-poor, highly condensed, heterochromatin. Some histone variants are deposited in a replication-independent manner, and may be evacuated and redeposited due to transcriptional events (Mito, Henikoff, and Henikoff 2007).

Histone modifications, and hence chromatin state, are maintained through cell division (Henikoff, Furuyama, and Ahmad 2004). DNA methylation, an additional epigenetic process which can lead to gene silencing, is also maintained when DNA is replicated. The supposedly gene-poor heterochromatic regions of the genome may have an important functional role enabling various genomic mechanisms (Grewal and Jia 2007; Zuckerkandl and Cavalli 2007). The state of the chromatin is correlated with recombination and DNA repair, and thus "mutation rate" (Prendergast, Campbell, Gilbert, Dunlop, Bickmore, and Semple 2007). More generally, there are mutational hotspots in the genome and regions of genomic instability, whose instability is the result of interaction with other, nonlocal, genomic elements (Aguilera and Gómez-González 2008).

Chromosomes may interact with neighboring chromosomes, and their location in the nucleus and the identity of their neighbors is not random (Fraser and Bickmore 2007). Centromeres and telomeres also turn out to be dynamic. Remarkably, the location of centromeres, which are pivotal for mitosis and meiosis, may be determined by non–DNA-sequence factors, and inherited via a chromatin-based inheritance

mechanism, rather than being defined solely by DNA sequence motifs (Henikoff, Ahmad, and Malik 2001). The end regions of linear chromosomes, the telomeres, are actively maintained either by the telomerase ribozyme or by other mechanisms.

RNA-mediated processes are involved in gene regulation by transcriptional repression as well as posttranscriptional gene silencing. They are also involved in genome rearrangement events (Lamm and Jablonka 2008), and RNAi and small RNAs may play a role in the maintenance of heterochromatin (Grewal and Elgin 2007).

It is worth including, in this picture of the dynamic genome, developmentally regulated genome rearrangements that operate on a different timescale than the processes mentioned above. Processes of this type are often induced by environmental and genomic stress conditions (Zufall, Robinson, and Katz 2005; Lamm and Jablonka 2008).

Each of the processes alluded to may or may not be active in different organisms, and each mechanism is the result of the evolution of several enzymes and structures. As could be expected, homologous genomic mechanisms have different functions in different organisms. The existence of such a range of mechanisms operating in the genome, at times necessarily simultaneously, suggests that they may occasionally interact and interfere with one another, whether directly or indirectly. Moreover, an enzyme may play a role in several genomic processes, serving different functions. All this leads to a complicated picture, both ontogenetically and evolutionarily.

The Metastable Genome

The notion of *metastability* may help uncover the conceptual implications of genomic dynamism. A system is said to be in a metastable state when it is in a delicate equilibrium state, likely to switch to a different attractor as a result of even small perturbations. In the case of the genome, the attractors I will consider consist of the configuration of the genome, the operating genomic mechanisms, and possibly the expression pattern of relevant genes. It should be emphasized that the systems I am talking about consume and accumulate energy, and thus the attractors are *not* simply deduced from considerations of minimal energy.

By focusing on metastability and the role of genomic mechanisms, I do not mean to discount various other forms of stability. Indeed, these form the context in which genomic metastability has to be understood. Metastability, however, highlights the fact that the genome is constantly dynamic and in constant danger of unintentional switching from one metastable state to another, and hence that the distance between attractors and the mechanisms maintaining their stability are sources of variation and are subject to evolutionary pressure. Thus, as a thought experiment, I am going to consider genome functioning as being dependent solely on the genomic system occupying and transitioning between attractor states that are metastable. The purpose of the metastable genome (MSG) thought experiment is to highlight the possible evolutionary consequences that might ensue if genomes are indeed metastable, or when their behavior approaches this extreme case. This allows me to explore the conceptual implications of the suggestion that the stability of many genomic states does not stem from structural stability or even simple dynamic stability, but is rather

metastability, actively maintained by GEMs, such as those responsible for histone replacement (Mito, Henikoff, and Henikoff 2007) and the RNAi GEMs involved in chromatin maintenance (Grewal and Elgin 2007).

The metastable states (or attractors) of the genome, which underlie the thought experiment, consist of the physical organization of the genome and the activation pattern (quantitative) of the GEMs. Attractors are not, I should stress, merely the expression level of genes in the transcriptome (cf. Bar-Yam, Harmon, and De Bivort 2009). They are, however, less than the complete cellular state.

To remain in the metastable states, in spite of ever present noise, active (energy-consuming) mechanisms are needed. These mechanisms, however, not only maintain stability; they may cause perturbations. This is one reason why the physical stability of DNA molecules is largely irrelevant to the question of genome stability. The MSG thought experiment helps elucidate the implications of current knowledge about the division of labor between the supposedly inert DNA molecules (but see Shapiro 2006) and their surrounding machinery.

The stage is now set to consider how this setup affects our understanding of key biological notions. I discuss two: responses to stress conditions and the notion of genomic information.

Stress Response

Various environmental stresses, such as temperature shocks and pathogen attacks, as well as genomic stress, such as hybridization and polyploidy, can cause anything from small-scale genomic changes to wide-ranging and repeatable genomic repatterning

(Lamm and Jablonka 2008). These genomic reorganization events may involve a variety of molecular mechanisms (e.g., RNAi) and activate transposition of transposable elements. Whether the ensuing genomic reorganization is an adaptive response that was selected for remains an open question. How would the MSG perspective account for such phenomena?

There is disagreement about whether similar stress conditions are frequent enough to lead to selection for stress responses. When the genome is thought of as being a metastable system, these issues take on a new shape. Stress, according to this perspective, would be any situation or event that disturbs precariously stable states. If a system is relatively stable, stress would be relatively infrequent, and hence stress responses would be less amenable to selection. If, in contrast, the system occupies metastable states most of the time, and their stability and the transition between them have to be actively maintained by homeostatic mechanisms (such as histone replacement, maintenance of small RNAs, etc.), these mechanisms will experience strong selection pressure. Stress, and consequently the stress response, are according to this perspective a matter of degree; normal conditions and stress are not a priori distinct, nor does stress occur only when some threshold is crossed. Rather, homeostatic mechanisms are constantly under pressure to retain the stability of the attractors, which is precarious.

Since stress is a matter of degree, stress responses will, inter alia, be selected (although this will not, of course, necessarily result in mechanisms that always manage to maintain the stable state or responses that always have adaptive value in

the environment). There is no principled distinction to be made between normally occurring selection of regulatory mechanisms and rare selection for stress response.

In the scenario we are considering, the very same mechanisms, and generally the same attractors, are involved in frequently and infrequently occurring conditions. Some states of the genome, which may exist in various external conditions, may, however, be more stable than others and exert less evolutionary influence on GEMs and genome organization. States of the genome whose stability is more likely to be affected by changes in the GEMs will exert a larger effect on the selection pressure of GEMs. Accordingly, the less stable functional states of the genome are likely to have a larger effect on the selection pressure of the GEMs than more stable states employing the same GEMs.

This perspective is in marked contrast to the standard account which focuses on frequency of “stress conditions” rather than stability of genomic states.

If stress is understood to be any situation or event that disturbs precariously stable states, it follows that stress may involve conditions that affect the functioning of the genomic mechanisms as well as conditions that affect the genomic substrate (i.e., chromatin) more directly. Additionally, selection may operate both on the dynamic homeostatic mechanisms *and* on the more stable genomic substrate. This is because changes to either can change the stable states, or attractors, of the system, and a change to either can alter the “distance” between attractors, a crucial factor in maintaining their developmental stability (homeorhesis) and the developmental trajectory of the system as whole. Exploring empirically whether selection can indeed

achieve similar results when variation in only one of these is permitted, is an intriguing possibility.

The observation that selection may operate both on the dynamic homeostatic mechanisms and on the more stable substrate of the genome, in order to establish the genomic attractors and the developmental trajectories through them, is important when considering coordinated or large-scale genomic reorganization in response to various stresses (see, for example, chapters 24 and 25 in this volume). Hybridization and polyploidy involve a change in the relatively inert "substrate" upon which the homeostatic mechanisms operate (i.e., the chromatin). Hybridization and polyploidy thus create new attractors (e.g., by creating new target sites for mechanisms or sinks that cause depletion of various enzymes) and possibly change distances between stable states. This may create new stable states "between" two existing states, making the transition fraught with the possibility of accidentally ending up in the new state.

Similar types of genomic outcomes resulting from different types of disruptions become easier to explain with the aid of the MSG framework: The new attractor depends not only on the direction of disruption and homeostatic mechanisms stabilizing the new state, but also on the topography of the attractor space. It is thus probable that different types of disruptions will end up leading to the same attractor. In the metastable scenario, most perturbations cause the system to move between attractors, making this possibility more likely.

It is critical to note that the effect of the physical organization of the genome, a fundamental factor shaping the attractor space, depends on the functional GEMs in each particular genome. One significant reason for this is that various perturbations to

the genomic substrate cause various repair mechanisms to be activated—and these affect the subsequent organization of the genome. For example, the distribution of homologous sequences in the genome is particularly important for genomes that repair double-stranded breaks using homologous recombination rather than nonhomologous end-joining (Argueso, Westmoreland, Mieczkowski, Gawel, Petes, and Resnick 2008; Scheifele and Boeke 2008). Naturally, these factors become more significant the more frequent the triggering events are (Zahradka, Slade, Bailone, Sommer, Averbeck, Petranovic, Lindner et al. 2006).

For understanding the response to stress it is therefore not enough merely to consider the frequency and repeatability of the stress conditions, since the organization of the system can affect the *effective repeatability* of stress (or conditions more generally); it is not the "external" conditions, but internal ones, that affect the response.

Genomic Information

Heredity, as well as developmental regulation, is often explained in terms of information. The MSG perspective has significant implications for what might be termed "genomic information." This can be thought of as analogous to the notion of genetic information, which is typically understood as referring to the coded information in the DNA sequence.

The functional effect of the genome depends on the attractor state which the genome occupies, since this determines which regions are expressible, which genomic processes can occur, where they can occur, and so on. The basic argument relating the

MSG scenario and genetic information is straightforward: In an MSG, the functional information *is* the attractor, and if these attractors depend on GEMs, then genomic information is dependent on GEMs. Ergo, epigenetic mechanisms *underlie* genetic information. This argument undermines the traditional account of epigenetic phenomena as an information or inheritance system operating "on top of" the genetic system.

The genome, according to this picture, is not only, as traditionally understood, a *message*, but also a *receiver* involved in the interpretation of messages (cf. Jablonka 2002). Putting this idea more concretely, the possible interpretations of genetic messages—and hence the set of possible messages that can be encoded—are constrained by the fact that the incorporation of a message (i.e., sequence) into the genome potentially affects the functioning of the genomic system.

This analysis leads to the following conclusions about genomic information: (1) The physical organization, and not only the nucleotide sequence, is relevant for assessing the information content, as it affects the GEMs and, hence, the identity of stable states and their effective proximity to one another. (2) Information is holistic because attractors are typically influenced by nonlocal events in the genome (e.g., a change in conformation can affect multiple loci). The holism is both structural and dynamical. The appearance of individuated elements, such as genes, calls for an evolutionary explanation. (3) The genetic information system and genome-related epigenetic phenomena should be understood as a single information system, not as separate information channels (Lamm 2009; cf. Jablonka 2002). The present argument does not rest on an in-principle objection to the possibility of distinguishing

inheritance channels in general, but rather provides a specific reason why the division between the genetic inheritance channel and genomic epigenetic inheritance should be rejected (cf. Griffiths and Knight 1998; Wimsatt and Griesemer 2007). This argument also means that the genomic inheritance system may inherit properties (or limitations) of the epigenetic inheritance system, leading, for example, to restrictions on the modularity of variation of genomic attractors (cf. Jablonka 2002). (4) Information content is relative to a specific environment (external or internal), and life strategy: Information depends on GEMs whose activation depends on environmental stimuli. For instance, the identity of the active GEMs may determine if noncoding DNA sequences have a functional role (Zahradka, Slade, Bailone, Sommer, Averbeck, Petranovic, Lindner et al. 2006).

Conclusions

The metastable genome suggests that the view of genomic homeostasis, inspired by McClintock's suggestions, is a productive way to conceptualize the genome. The genome, according to this picture, functions by virtue of being in (metastable) attractor states that consist of the configuration of genome, the operating genomic mechanisms, and the expression pattern of relevant genes. The attractors are maintained (and perturbed) by genomic mechanisms as well as by external forces. This has functional consequences, and affects and constrains genomic variation, resulting in what might be termed "genomic inherency" (cf. Müller and Newman 2005).

Understanding the transition between attractors (e.g., in development) requires attention to the topography of the attractor space, namely, to how close stable genomic states are to one another. The analysis may also need to appeal to the *ecology* of the genome, in the sense of considering the distribution of elements (e.g., transposable elements) rather than their specific location or exact copy number, since their distribution may affect the attractor space. Elements may be "useful without being indispensable" and "non-specific in their activity" (Darlington, quoted in Vanderlyn 1949). A set of stable genome states, nonidentical in sequence, epigenetic marks, or pattern of GEM activity, may nonetheless be functionally equivalent. These equivalence classes may be significant units of genomic analysis.

The systemic, reactive properties of genomes suggest that "neo-Darwinian genes" that are manifested in some, though not necessarily all, circumstances, and hereditary variation more generally, have to be explained by appealing to the reactive nature of the system and the propensities that result from its organization. Neo-Darwinian genetic elements, in this picture, are an outcome of the behavior of a reactive organized system. Mutation-inducing events should not be considered as random on the basis of a priori reasoning, since the type of randomness experienced by a system can be judged only with respect to its organization. Ironically, "Darwinian" systems biology may be "Lamarckian" in this sense, due to its attention to system-level properties.

Acknowledgments

The work presented in this chapter is part of the author's Ph.D. research, which was supported by the Cohn Institute for the History and Philosophy of Science and Ideas and the School of Philosophy at Tel Aviv University. I thank the participants of the Annual International Workshop on History and Philosophy of Science, Transformations of Lamarckism: 19th–21st Centuries, for their helpful comments, the editors of this volume, and Joel Velasco for comments on earlier drafts of this chapter.

References

Aguilera A, Gómez-González B. 2008. Genome instability: A mechanistic view of its causes and consequences. Nat Rev Genet. 9: 204–217.

Argueso JL, Westmoreland J, Mieczkowski PA, Gawel M, Petes TD, Resnick MA. 2008. Double-strand breaks associated with repetitive DNA can reshape the genome. Proc Natl Acad Sci USA. 105,33: 11845–11850.

Bar-Yam Y, Harmon D, De Bivort B. 2009. Systems biology: Attractors and democratic dynamics. Science 323,5917: 1016–1017.

Fraser P, Bickmore W. 2007. Nuclear organization of the genome and the potential for gene regulation. Nature 447: 413–417.

Grewal SIS, Elgin SCR. 2007. Transcription and RNA interference in the formation of heterochromatin. Nature 447: 399–406.

Grewal SIS, Jia S. 2007. Heterochromatin revisited. Nat Rev Genet. 8: 35–46.

Griffiths PE, Gray RD. 1994. Developmental systems and evolutionary explanation. J Philos. 91,6: 277–304.

Griffiths PE, Knight RD. 1998. What is the developmentalist challenge? Philos Sci. 65: 253–258.

Henikoff S, Ahmad K, Malik HS. 2001. The centromere paradox: Stable inheritance with rapidly evolving DNA. Science 293,5532: 1098–1102.

Henikoff S, Furuyama T, Ahmad K. 2004. Histone variants, nucleosome assembly and epigenetic inheritance. Trends Genet. 20,7: 320–326.

Jablonka E. 2002. Information: Its interpretation, its inheritance, and its sharing. Philos Sci. 69: 578–605.

Keller EF. The Century of the Gene. Cambridge, MA: Harvard University Press; 2000.

Lamm E. 2009. Conceptual and methodological biases in network models. Ann NY Acad Sci. 1178,4676: 291–304.

Lamm E, Jablonka E. 2008. The nurture of nature: Hereditary plasticity in evolution. Philos Psychol. 21,3: 305–319.

McClintock B. 1984. The significance of responses of the genome to challenge. Science 226: 792–801.

Mito Y, Henikoff JG, Henikoff S. 2007. Histone replacement marks the boundaries of cis-regulatory domains. Science 315,5817: 1408–1411.

Müller GB, Newman SA. 2005. The innovation triad: An EvoDevo agenda. J Exp Zoolog. B: Mol Dev Evol. 304B: 487–503.

Pearson H. 2006. What is a gene? Nature 441,7092: 399–401.

Prendergast JGD, Campbell H, Gilbert N, Dunlop MG, Bickmore WA, Semple CAM. 2007. Chromatin structure and evolution in the human genome. BMC Evol Biol. 7: 72.

Scheifele LZ, Boeke JD. 2008. From the shards of a shattered genome, diversity. Proc Natl Acad Sci USA. 105,33: 11593–11594.

Shapiro JA. 2006. Genome informatics: The role of DNA in cellular computations. Biol Theory 1,3: 288–301.

Vanderlyn L. 1949. The heterochromatin problem in cytogenetics as related to other branches of investigation. Bot Rev. 15,8: 507–582.

Waddington CH. The Nature of Life. London: Allen & Unwin; 1961.

Wimsatt WC, Griesemer JR. Reproducing entrenchments to scaffold culture: The central role of development in cultural evolution. In: Integrating Evolution and Development: From Theory to Practice. Sansom R, Brandon R, eds. Cambridge, MA: MIT Press; 2007:227–323.

Zahradka K, Slade D, Bailone A, Sommer S, Averbeck D, Petranovic M, Lindner AB et al. 2006. Reassembly of shattered chromosomes in Deinococcus radiodurans. Nature 443: 569–573.

Zuckerkandl E, Cavalli G. 2007. Combinatorial epigenetics, “junk DNA,” and the evolution of complex organisms. Gene 390: 232–242.

Zufall RA, Robinson T, Katz LA. 2005. Evolution of developmentally regulated genome rearrangements in eukaryotes. J Exp Zoolog. B: Mol Dev Evol. 304B: 448–455.