Darwin, Gödel, Luria, Delbrück: Biomedical, Mathematical, and Metamathematical Perspectives on Attributes and Consequences of Random Somatic Mutations Subject to Selection

Neil S. Greenspan[1], Owen Han[1]

1 Case Western Reserve University

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.

Abstract

This perspective explores the ways in which random somatic mutation and selection can influence phenotypes that are biologically or clinically consequential. We also discuss the ways in which genetic research from recent decades has provided impetus to think about how *de novo* mutations in parental germ cells and somatic mutations in progeny might affect progeny phenotypes and heritability. The phenomenon of extended phenotype is characterized and discussed in terms of the impact it has on influencing phenotypes in ways that might affect heritability. We therefore propose revising the broad heritability equation to reflect these experimental advances and better convey the full range of factors that can exert influence over organismal phenotypes. Finally, we consider how an analogy between changes in axiom systems and mutational changes in genomes prompts consideration of a broader comparison between formal deductive systems and the structured genetic systems of cells and organisms.

Neil S. Greenspan[a,*] and **Owen Han**[b]

[a] *Department of Pathology, Case Western Reserve University*

[b] *Case Western Reserve University*

[*]Correspondence: nsg@case.edu

Keywords: somatic mutation, clonal evolution, heritability, extended phenotype, undecidability, consistency, context-dependence.

Introduction

A foundational notion of modern biological and biomedical research is that mutation is “random,” where random is defined in a relatively precise and technical sense. When the word “mutation” is preceded and characterized by “random,” at least in an evolutionary context, the standard intent to be conveyed is that the frequency of the particular nucleotide substitution is unrelated to the functional, and therefore, evolutionary consequences.

In this evolutionary context, “random” does not necessarily imply that for a point mutation (i.e., the replacement of an initial nucleotide at a particular position in the linear sequence of a gene by a different nucleotide) the “new” nucleotide at the position of interest has equal probabilities of being any of the three alternative nucleotides typically found in DNA. For example, if the original nucleotide was cytosine, a mutation could still be considered “random” in the conventional evolutionary sense noted above even if the post-mutation nucleotide, was preferentially one of the three possible alternatives: adenine, guanine, or thymine. The chemical non-randomness of the mutational mechanism does not preclude randomness with respect to functional consequences.

This alternative interpretation of “random” mutation as chemical randomness, however, is sometimes used even in publications devoted to science, like *Science*. For example, a recent news article in*Science* [1] includes the following passage:

> *“Nevertheless, researchers and citizen scientists from around the globe began to scan SARS-CoV-2 genome sequences deposited in the international GISAID database, looking for the kinds of mutations expected to be caused by molnupiravir.* ***Rather than inducing random changes in the virus’ RNA genome, the drug is more likely to cause specific nucleic acid substitutions, with guanine switching to adenine and cytosine to uracil.”*** *(emphasis added)*

Such a usage of “random mutation” is reasonable as long as it is clear that it is a different sense of the term from the one that is central in evolutionary theory. Unfortunately, the significance of instances of mutation for which all nucleotide substitutions are not equally probable can be misconstrued as violating the “dogma” of evolution. In fact, such mutations do not necessarily represent a significant challenge to evolutionary theory.

At this point, mathematicians focused on biology may be forgiven for heaving a sigh of relief. After all, their job is to turn ideas and words into equations and algorithms that can predict or simulate biological and biomedical realities. Randomness is often more mathematically tractable than non-randomness, especially when the non-randomness can involve multiple interacting elements or processes that generate outcomes that can be difficult to foresee or even fully account for in hindsight.

Below, we will discuss a series of topics for which the biology depends on mutation, presumed unless otherwise indicated to be random in the technical evolutionary sense delineated above, and for which mathematical treatment or mathematical concepts would be of value. These topics are: 1) the role of somatic mutations in influencing organismal phenotypes, 2) the implications of this phenomenon and others for the concept of heritability and the equations that are meant to convey the concept will be explored, and 3) the analogy between axiom systems and genomes.

With respect to the heritability equations, we propose that the broad heritability equation should be updated in light of information amassed over the past several decades. These findings establish beyond doubt the potential for de novo mutations of both parental germ cell precursors and mutations of progeny somatic cells, especially tissue-specific stem cells, to exert substantial influence on medically important phenotypes.

In what follows, relevant mathematical concepts and some equations will be discussed. For other instances, we will give reasons why we believe a definitive mathematical treatment may be challenging.

Somatic Mutations that Influence Organismal Phenotypes

Physicians and biomedical scientists have recognized for decades the potential of mutations in somatic cells to influence the phenotypes of whole organisms. The best-known case is, of course, cancer. As long ago as 1960, Nowell and Hungerford identified a chromosomal abnormality associated with chronic myeloid leukemia cells, which was called the Philadelphia chromosome [2][3]. In 1973, Rowley demonstrated[4] this chromosomal anomaly was due to the breakage and exchange of portions of chromosomes 9 and 22, a translocation, with the modified chromosome 22 including a portion of chromosome 9 constituting the Philadelphia chromosome. In 1976, with evidence for a potential role of mutations of such large scale in one form of leukemia, Nowell proposed that tumors arise though a process of somatic cell genetic changes and selection resulting in clonal evolution [5]. Evidence strongly supporting his hypothesis has accumulated since [6].

There are many other examples of conditions of clinical relevance that involve somatic mutations. For example, Walsh and colleagues have delineated instances of post-zygotic mutations that lead to mosaicism in the brain and neural dysfunction such as seen in some forms of epilepsy or autism [7]. Goodell and Rando describe evidence that strongly supports the role of somatic mutation in affecting tissue-specific stem cells in aging [8]. Takeda *et al.* [9] summarize data demonstrating that the condition paroxysmal nocturnal hemoglobinuria results from somatic mutation in a bone marrow precursor that interferes with the biosynthesis of complement regulatory proteins CD55 and CD59 that are normally active on the plasma membranes of erythrocytes (i.e., red blood cells) and limit the probability of spontaneous complement activation causing the destruction of erythrocytes.

Below, we summarize information pertaining to two additional and compelling examples of the power of somatic mutations to affect overall phenotype. Both scenarios involve cell types that arise in the bone marrow, but as suggested above, somatic mutations in multiple tissues have the potential to alter host phenotypes.

WHIM Syndrome

The potential of somatic mutation and selection to transform not just cellular phenotypes but organismal phenotypes is dramatically illustrated by a case of a rare immunodeficiency condition, known as WHIM [warts, hypogammaglobulinemia (reduced serum immunoglobulin concentration), infection, and myelokathexis (retention of neutrophils in the bone marrow.)] Syndrome. In the 1960s, a young woman was evaluated for this condition [10][11]. This patient, who will be

referred to as WHIM-09, exhibited all of the hallmark clinical manifestations cited above at the time of this initial evaluation.

WHIM-09 eventually married and had a family. Two of her daughters were also afflicted with WHIM syndrome. A third child was unaffected. Her husband was also free of any manifestations of the condition.

Decades after the initial medical evaluation in the 1960's, WHIM-09 went to the National Institutes of Health (NIH) with her two affected daughters to be evaluated [12]. While the two daughters exhibited the same symptoms and signs as WHIM-09 did when she was first studied, WHIM-09 was mostly free of the clinical manifestations with which she was previously afflicted and had been so for approximately 20 years. The NIH investigators proceeded to study all three patients and in particular sought to determine the mechanisms responsible the seeming resolution of most aspects of WHIM-09's condition.

Important context is that by the time of the evaluation of WHIM-09 and her two adult daughters at the NIH, causative mutations for WHIM syndrome had been identified [13]. These disease-causing mutations occurred in a gene encoding the chemokine receptor CXCR4 and were associated with constitutive signaling. Instead of signaling (i.e., initiating a series of biochemical processes that ultimately alter which genes are transcribed and translated into proteins) only after binding the ligand CXCL4, the mutated version of CXCR4 would signal even without encountering its physiologic ligand.

So, what could have happened to WHIM-09 to cause the nearly complete disappearance of the clinical manifestations of WHIM syndrome? Since this special issue of *Axioms* is intended to commemorate the famous experiments of Luria and Delbrück on the nature of spontaneous genetic mutation, you will likely be anticipating that the answer to this question involves mutation. Indeed, mutation was involved but was not the only factor necessary for the astonishing clinical course of WHIM-09.

Since the WHIM-associated mutations affect cellular phenotype in single copy (only one of two alleles needs to mutate i.e., a dominant mutation), in the absence of any evidence that her parents or siblings exhibited symptoms or signs associated with WHIM syndrome, it is plausible that her condition was caused by *de novo* mutation in a parental germ cell precursor. Mutation in a bone marrow stem cell early in WHIM-09's ontogeny is rendered unlikely because two of her daughters had WHIM syndrome and presumably inherited a mutation in the gene encoding CXCR4 from WHIM-09. Why then did WHIM-09's overall clinical phenotype suddenly change?

The NIH investigators performed an analysis of the chromosomes in cells from WHIM-09's bone marrow using multiple methods. They found that one copy of chromosome 2 in some of her bone marrow-derived cells was shortened. Furthermore, in these altered copies of chromosome 2, the CXCR4-encoding locus was deleted along with about another 163 additional genes. Furthermore, the abnormal chromosomes 2 had many genes oriented differently than in the normal chromosome 2.

The best explanation for these unusual results was that there was a spontaneous instance of a process that has been associated primarily but not exclusively with cancer cells called chromothripsis or "chromosome shattering" [14].

Presumably this event occurred in one of WHIM-09's bone marrow stem cells or other progenitors [15] between the initial evaluation decades earlier and the spontaneous resolution of most of the WHIM syndrome-related symptoms. The progeny of this one cell must then have had a selective advantage and proliferated at the expense of the cellular progeny of the "normal" stem or progenitor cells.

Typically, while two copies of a chromosome may have different alleles at any given locus, the same loci will tend to exist in the same physical locations and in the same precise order on a chromosome, although such similarity is not always the case. Such exceptions, however, are usually limited in number and result from particular insertions or deletions, as opposed to the sort of global re-ordering and re-orienting of genes seen in chromothripsis.

The story is still more complicated and interesting. While the altered chromosome 2 was advantageous, that selective edge applied only to myeloid lineage cells and not to lymphoid lineage cells. The lymphocytes continued to have two full copies of chromosome 2, so the implication is that the gene deletions associated with the chromothriptic chromosome 2 must have put lymphocyte precursors at a selective disadvantage even as they increased the reproductive fitness, and coincidentally the functionality, of myeloid cells. It is notable that a presumably spontaneous and putatively unpredictable event on one chromosome in one cell in one tissue can massively transform the physiological state of an entire human being (Figure 1).

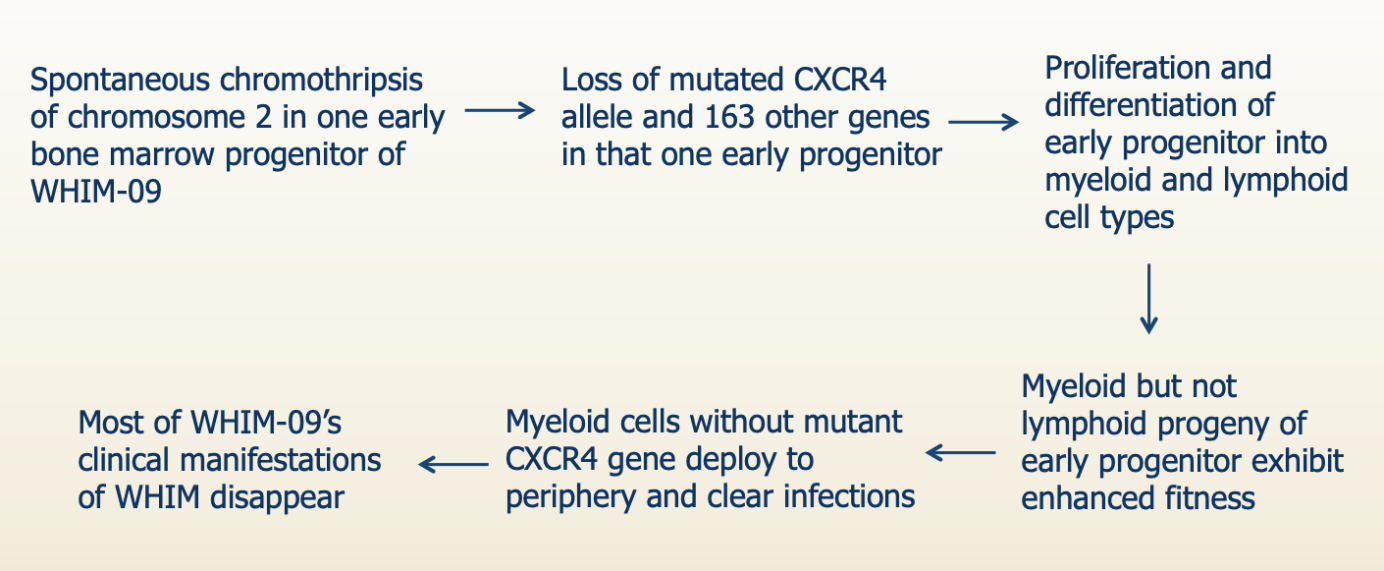

Figure 1. Sequence of key events in the spontaneous cure of WHIM syndrome for patient WHIM-09.

At least two mathematical challenges can be identified. One would be to create a model able to predict the probability of this sort of astonishing phenotypic transformation based on what can be regarded as a massive genetic alteration, i.e., a mutational event at a large scale both genetically and with respect to phenotypic consequence. This is a challenge for which we hope an inspired mathematician will see a path forward.

A second perhaps more tractable type of challenge is to estimate the magnitude of the selective advantage conferred by the abnormal chromosome 2 for myeloid (but not lymphoid) cells or to determine what magnitude of selective advantage

is required to account for the phenotypic transition in terms of clinical manifestations. There are many other situations that are clinically relevant in which reproductive advantage of a clone of hematopoietic cells (or stem cells in other tissues, such as intestinal mucosa or skin) leads to clonal dominance and phenotypic changes that have clinical significance. Hematopoietic examples include the disease, referenced above, known as paroxysmal nocturnal hemoglobinuria, the condition known as clonal hematopoiesis, which can precede actual leukemia, and instances of spontaneous development of resistance to therapy, pharmacologic or biologic, among populations of leukemic cells.

Perhaps an experimental system could be devised *in vivo* or *in vitro* to study the cellular dynamics in bone marrow derived or other cells and as a step towards developing the means to determine the magnitudes of selection coefficients and/or other quantitative aspects of these clinically consequential phenomena of somatic cell evolution. At least with respect to this challenge to mathematical biologists, one can imagine conceivable pathways towards increased understanding.

Humoral Immunity to *Plasmodium falciparum*

Another impressive example of the potential for genetic events to exploit mechanisms that would be difficult to predict is described in experimental studies from 2015 [16] and 2017 [17] from the laboratory of Antonio Lanzavecchia. These investigations offer extraordinary insights into the origins of at least some broadly reactive antibodies that contribute to immunity against *Plasmodium falciparum*, the cause of the most dangerous form of human malaria.

This research was motivated by the desire to find human antibodies that are broadly reactive with variable surface antigens of malaria parasites, specifically the stage, known as the merozoite, that infects erythrocytes, i.e., red blood cells. Such antibodies could be useful both as therapeutic agents and as guides to vaccine design by virtue of the antigens to which they bind and for which they might be useful for identification and isolation. With this goal in mind, Lanzavecchia and his associates used a clever assay to screen the plasma of individuals who had been infected with falciparum malaria for antibodies that possessed the ability to react with antigens displayed by multiple strains of the erythrocytic stage of the parasite.

Each of three culture-adapted blood stage parasites from Kenya were labeled with different fluorophores able to emit light of different colors. These tagged parasites were then used to infect red blood cells in culture. Plasma samples from infected individuals were tested to see if they could generate multi-color red cell agglutinates.

Once they found three such plasma samples, the memory B cells of the corresponding individuals were immortalized and grown in culture. Screening of supernatants for ability to bind red blood cells separately infected with one of eight different *P. falciparum* isolates identified B cell clones making the sorts of broadly specific antibodies they were seeking. The authors then used established methods to clone the relevant immunoglobulin genes from these B lymphocytes and sequenced them thereby revealing what turned out to be their unprecedented structures.

Unexpectedly, the genes encoding the heavy chain variable domains of some of the broadly reactive antibodies from two unrelated individuals included overlapping but non-identical insertions of unprecedented length (nucleotides encoding more than 100 amino acids) that did not correspond to any standard antibody-encoding gene segment. Through further

analysis these researchers determined that the insert sequences were derived from a gene located on a different chromosome.

This gene, Leukocyte Associated Immunoglobulin Like Receptor 1 or*LAIR1*, is located on the long arm of human chromosome 19, whereas the human immunoglobulin heavy chain locus is encoded on the long arm of chromosome 14. The *LAIR1* gene encodes a protein that binds collagen. It is therefore of great interest to learn that Lanzavecchia and colleagues showed that the inserts in the variable domains of antibodies that were broadly reactive with surface proteins of the malaria merozoites also had undergone numerous somatic mutations. Collectively, these mutations abrogated binding to collagen and conferred binding to merozoite surface proteins encoded by repetitive interspersed family *(rif) genes*. The proteins produced by transcription and translation of *rif* genes are called RIFINS.

So, in addition to a totally novel trans-chromosomal insertion of never-before-seen size, these potentially protective antibodies also had to undergo extensive somatic mutation. Once again, we believe it would be a major achievement to devise a mathematical model of this process that can predict when and for which antigen responses such phenomena will occur.

Mutations and the Concept of Heritability

The two heritability equations that stem from the ideas of two of the three founders of population genetics, Ronald Fisher and Sewall Wright, capture the range of factors of genetic and non-genetic origin that can influence a phenotype [18][19]. These ideas of Fisher, Wright, and perhaps a small number of others first began to take shape about a century ago.

In the narrow version of the equation (see*1* below), the focus is on the additive genetic variance that influences the extent of phenotypic variance and is intended to delineate the extent to which a phenotype can be modified in response to selection. The broad version of the heritability equation (*2*) is intended to offer a comprehensive accounting for the categories of factors that vary in ways that can contribute to phenotypic variation for a defined trait.

Narrow heritability equation

VP=VA+VE (*1*)

Where

- VP = phenotypic variance
- VA = additive genotypic variance
- VE = phenotypic variance due to variation based variation in environmental factors.

Broad heritability equation

VP=VA+VD+VI+VE+VG×E+COV(G, E) (*2*)

Where,

- VP = phenotypic variance
- VA = additive genotypic variance
- VD = phenotypic variance due to variation based variation on effects arising from genetic dominance
- VI = phenotypic variance due to variation based variation in interactions between genes at different loci, known as epistasis
- VE = phenotypic variance due to variation based variation in environmental factors
- VGXE = phenotypic variance due to variation based on variation in gene-environment interactions
- COV(G, E) = phenotypic variance due to variation based variation due to correlations between genetic and environmental factors.

The broad heritability equation has been largely accepted for decades as accounting for all of the factors that influence the phenotype of a human or other organism. Nevertheless, over several decades investigators have revealed factors able to influence phenotypes not covered or not covered explicitly in this equation. These factors include: 1) *de novo* mutations in parental germ cell precursors, 2) somatic mutations that occur post-fertilization during ontogeny, and 3) genes from other organisms that can influence a "target" organism (a phenomenon known as extended phenotype). Below, we provide brief examples for how each of these factors can play a role in substantially affecting organismal phenotype.

De Novo Mutations in Parental Germ Cells

New mutations occurring in the germ cell precursors in female or male gonads will ultimately generate, respectively, eggs or sperm that carry the mutation in question. Sometimes, an egg or sperm cell carrying the mutation engages in fertilization leading to an embryo with the mutation in the newly created genome. If the mutation alters the function of the resulting protein or RNA products or even the interaction of a portion of the DNA with a protein involved in some manner in gene transcription, an alteration in phenotype relative to what it would have been in the absence of the mutation may result.

Such a mutation will be in the somatic cells of the offspring but not the somatic cells of the parent in which the mutation arose. In some cases, a new clinical phenotype may appear in the child during the course of development from neonate to toddler, child, adolescent, and adult. Thus, parental germ cell phylogeny can exert substantial influence on progeny ontogeny.

Examples of the clinical importance of such*de novo* mutations include numerous syndromes that typically first appear in the early years of life. There is strong evidence for *de novo* genetic variants associated with and contributory to such conditions as autism [20][21][22][23] and intellectual disability [24].

Somatic Cell Mutations, Especially in Tissue-Specific Stem Cells

Although cells can influence rates of somatic mutation through standard biochemical mechanisms, as far as we know

mutation cannot be completely suppressed. Since many tissues, such as bone marrow and the intestine, need a constant supply of new cells to perform the necessary tissue-specific functions, these tissues require proliferating stem cells that are necessary to generate new cells that will differentiate so as to express the various phenotypes required to maintain the relevant physiologic functions. Examples of non-malignant conditions based on somatic cells mutations in the hematopoietic and central nervous systems were referenced above. We believe it is likely that more instances of clinically relevant phenotypes caused or influenced by genetic variants in somatic cells will be uncovered in the coming years.

Influences on Phenotype from Allogeneic or Xenogeneic Genes

As noted above, extended phenotypes, in the sense used by Dawkins in his book on the phenomenon[25] is defined as: "All effects of a gene upon the world," where "effect" refers to differential consequences for a particular gene and the allelic alternatives at a given locus.

Effects of a gene upon the world include but are not limited to influences on organisms other than the one in which the gene resides. Such trans-organismal effects can occur for two organisms of the same or different species. The most obvious examples are provided by the effects of microbes, whether pathogenic or not, on host organisms, such as humans [26]. Dawkins suggests limiting the concept to where the extended phenotypic effects have an impact on gene survival, but that additional constraint may not be as relevant in a medical context.

There is a massive literature on the influences of particular pathogen genes on host physiology, especially for bacteria and viruses but also for fungi and parasites. Similarly, although not dating back as far and presumably not as large, there is large and growing repository of published studies documenting the ways that elements of the human microbiome influence human physiology and phenotypes [27]. Knowledge at the level of specific genes may not be as voluminous as for pathogens, but this sort of information is likely to expand rapidly in the coming years.

The implications of the facts recounted in these immediately preceding sub-sections is that an equation presuming to take account of all factors affecting human phenotypes must include these potential sources of genetic variation.

We propose calling this heritability equation the "broader" equation to indicate that it is based on expanding the purview of the broad equation but leaving open the possibility that more is yet to be discovered about the influences of genetic variation and mutations on human phenotypes.

Broader heritability equation

$VP=VA+VD+VI+VENG+VEG+VDNM+VSM+VG\times E+COV(G, E)$ (2)

Where,

- VENG = phenotypic variance due to variation in environmental factors not directly based on genes
- VEG = phenotypic variance due to variation in environmental factors directly based on xenogeneic or allogeneic genes
- VDNM = phenotypic variance due to variation based on mutations in parental germ cell precursors

- VSM = phenotypic variance due to variation based on mutations in somatic cells

Figure 2 offers a diagrammatic summary of the factors that are currently known to influence phenotypes.

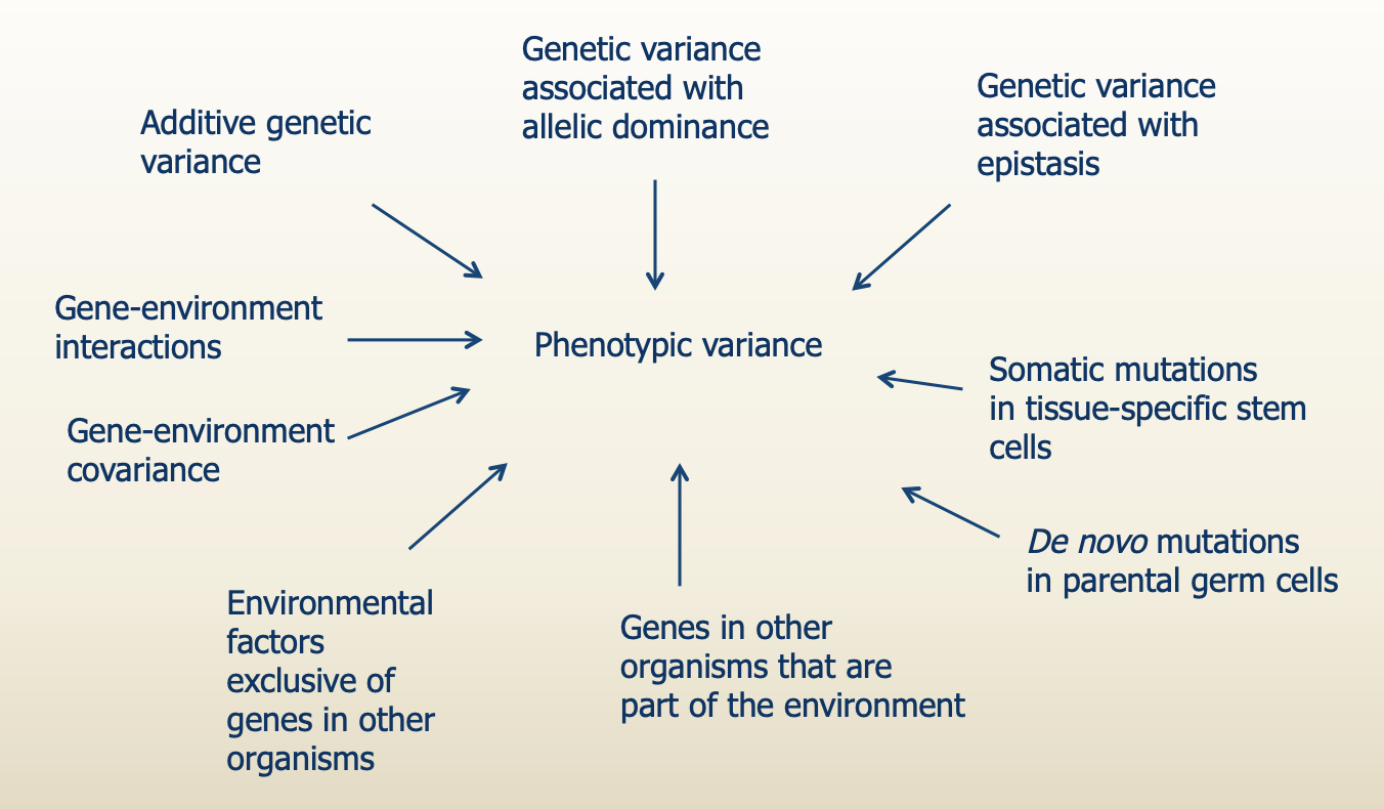

Figure 2. Schematic depiction of genetic and other factors that can vary in ways that influence phenotypic variation.

Is There a Useful Analogy between Genomes and Axiom Systems?

This section of the present essay is perhaps more relevant to meta-mathematics than mathematics proper. Many years ago in his magnum opus, **Gödel, Escher, Bach: An Eternal Golden Braid**[28], the computer scientist Douglas Hofstadter noted that in axiom systems, changing the meaning of any key term can carry consequences for the meanings of other, perhaps even every other, key term.

Hofstadter explicitly discussed comparisons of the meanings of words in Euclidean versus non-Euclidean geometries. He also directly addressed the levels of meaning associated with genomes and their units of information.

In the context of axiom systems, a particularly accessible example of the interesting effects associated with a single changed definition for a key term comes from the book “Taxicab Geometry: An Adventure in Non-Euclidean Geometry” [29], which describes a mathematic axiom system where the only “mutation” of the canonical Euclidean version is the definition of distance. In Taxicab geometry, distance is defined as a taxi driver confined to traveling on streets would define it, not as “the crow flies.”

One of us (NSG) has explored this circumscribed topic previously[30]. The main point of the comparison of Euclidean and

Taxicab geometries is that changing the definition of this one key term has consequences, for example, for the properties of circles, which can now have corners as we normally define that term. It also prompts the question, “What other definitions of distance might be useful in certain circumstances? The existence of still other definitions of distance suggests the possibility of a host of other non-Euclidean geometries that may not have been previously explored.

For our current purposes, the critical point is that in language the semantic context can influence the meanings associated with any individual word and in genomes both the genetic and environmental context can influence the biological “meanings” of genes and gene products. The ability for words to take on an indefinite number of meanings, what I have considered calling the “semantic multiverse,” roughly corresponds to the ability of genes to have multiple functions and to the ability of mutations in genes to have multiple context-dependent phenotypic consequences. Geneticists call the multiple functions of genes or the multiple effects of mutations “pleiotropy.” Examples of these effects are explored in some detail below.

The Variable Effects of Mutations

Sickle cell trait and sickle cell disease offer concrete examples of some of the above points. In sickle cell trait, changing one nucleotide in one out of 147 codons (nucleotide triplets) in one copy of the beta-globin gene leads to synthesis of a protein, which is one component of adult hemoglobin, with one changed amino acid (glutamic acid is substituted by valine at position six numbering from the amino terminus of the protein) [31] That a change in less than 1% of the primary structure of one out of thousands of polypeptides chains encoded in the human genome confers significant protection from the most severe forms of falciparum malaria, a function not associated with the wild-type beta-globin protein, is remarkable. The magnitude of this protective effect for severe malaria-associated anemia or cerebral malaria was found, by Hill and colleagues [32], to be greater than that associated with variants in HLA genes, which are known to play critical roles in immunity to many pathogens.

If both copies of the beta-globin gene have the same mutation as in sickle trait, the result in most individuals of sub-Saharan African origin is the condition known as sickle cell disease. In other words, changing two nucleotides out of the roughly six billion in a human genome (3.3 x 10e-8%) can result in a catastrophic cascade of pathophysiological effects that eventuate in anemia, impaired blood circulation in the smallest blood vessels (known as capillaries), and dysfunction and failure of multiple organs. Affected organs include the brain, which can be affected by stroke, and the spleen, which can be rendered dysfunctional resulting in increased susceptibility to infection. Sickle cell disease is also associated with extremely painful episodes termed sickle cell crises.

Summarizing the above two paragraphs, a single glutamic acid to valine substitution at position six in the beta-globin polypeptide chain confers a potential benefit of enormous significance where falciparum malaria is endemic but two such mutations can result in overwhelming morbidity and ultimately early mortality in the absence of, and perhaps even with, high quality modern medical care. This stark contrast illustrates the potentially dramatic context dependence of even the smallest possible genetic alteration.

This point is further exemplified by the finding that some individuals of Saudi Arab ancestry have two of the glutamic acid to valine substitution at codon six in the beta-globin gene, a component of adult hemoglobin, but exhibit only mild disease [33]. The attenuated phenotype appears to arise because in the erythrocytes of these individuals' fetal hemoglobin levels are elevated relative to the levels of fetal hemoglobin in the erythrocytes of African individuals with sickle cell disease. Increased intracellular fetal hemoglobin concentration, though a cell-intrinsic regulatory mechanism, leads to lower intracellular adult hemoglobin concentration. Since the pathophysiology of sickle cell disease depends with extraordinary sensitivity on the concentration of adult hemoglobin, even seemingly tiny decrements in the intracellular concentration of adult hemoglobin can substantially attenuate disease-related symptoms. Single nucleotide variants upstream of three loci, two of which are not globin loci, appear to associate with the increased production of fetal hemoglobin and substantially milder symptoms of sickle cell disease.

Completeness and Coherence in Axiom Systems and Genomes

Above, we call attention to the analogy between the ability for words to take on an indefinite number of meanings, what could be called the "semantic multiverse," and the ability of genes to have multiple functions and mutations in genes to have multiple context-dependent phenotypic consequences, which geneticists call "pleiotropy." This parallel prompts the question of whether there might be other instructive or useful analogies between axiom systems and genomes.

In mathematical logic, two features of great interest in an axiom system are: 1) whether any well-formed formula (i.e., statement within the system) can be determined to be decidable, i.e. determined to be or not to be a theorem (a formula derived by valid deductive inference from the axioms) and 2) whether all true statements can be derived from the axioms without also deriving false statements. Of course, it is precisely these attributes that in 1931 Kurt Gödel demonstrated were not attainable for an axiomatization of the arithmetic of whole numbers [34].

So, are there any additional analogies than can be applied to genomes as consisting of strings of nucleotides and axioms systems consisting of strings of symbols? Below, we discuss additional similarities.

We can ask, are there analogous attributes associated with a genome? If so, what are these genomic properties?

Here we offer some preliminary questions. For example, can we analogize individual genes to individual axioms? Can deductive consistency be related to genes contributing to reproductive fitness for the organism? If yes, then would genes that seem to decrease or undermine fitness be like axioms that are logically incompatible with other axioms? Could logical completeness, wherein the axioms permit deduction of all true theorems, be associated with the genome sequences providing all information necessary for survival and reproduction (at least in a given well-defined set of circumstances)?

In this context it will likely be necessary to grapple with some complexities inherent in DNA sequences: the ability of a nucleotide or a stretch of nucleotides of varying lengths to embody more than one type of information.

For example, Savisaar and Hurst [35] have demonstrated that nucleotides that encode amino acids in protein products can also influence the likelihood of splicing of the messenger RNA resulting from transcription at the relevant loci.

At present we are not prepared to offer definitive answers to the questions we pose in this section of the text. Nevertheless, we have reason to hope that if these questions can be posed in appropriate ways, there are individuals with the right skills and talents to determine whether pursuing them leads in productive directions.

Conclusions

In their seminal study of 1943[36], Salvador Luria and Max Delbrück focused on mutations in Escherichia coli bacteria that conferred resistance to infection by alpha bacteriophages, i.e. a type of virus that infects bacteria. These investigators noted that the alternative hypotheses for the mechanisms underlying this phenotypic transition had different deductive consequences for the distributions of the numbers of resistant cells in various experimental circumstances. Implicit in their analyses were assumptions about the mutation process that are most consistent with point mutations, changes in the nucleotide sequence of a single gene involving only one or a few nucleotides.

Even though Luria and Delbrück are known for demonstrating that the likely mechanism for the acquisition of resistance to infection for the particular pair of bacteria and bacteriophage that they studied involved relatively stable changes that appeared to occur prior to exposure to virus, they considered other classes of mechanisms for changes in other phenotypes of interest. The subsequent 80 years of research have vindicated their carefully and cautiously constructed thought processes. We now know that phenotypic changes in bacteria and eukaryotes can involve mechanisms other than point mutations, including much more extensive changes in nucleotide sequences (as noted above) and including (but not limited to) changes in transcription, translation, messenger RNA half-life, or gene product half-life for one or more genetic loci. Specifically for bacterial resistance to infection by bacterial viruses, we now know that still other more biochemically complex pathways, such as those involving clustered regularly interspaced short palindromic repeats (CRISPR) systems and the nucleases with which they associate, can be involved [37].

To make progress in quantitatively understanding some of these processes, the fundamental investigational approach of Luria and Delbrück, formulating alternative hypotheses, deducing hypothesis-specific quantitative predictions, and then designing experiments capable of decisively distinguishing among those predictions, is one plausible way forward. We would be gratified if this perspective stimulated even one investigator to follow up in this way to further our understanding of the diverse mechanisms underlying phenotypic change.

References

1. ^*Service, R. F. Could a popular antiviral supercharge the pandemic? Science 2023, 379 (6632), 526–526.*
2. ^*Nowell, P. C. The minute chromosome (Ph1) in chronic granulocytic leukemia. Blut Zeitschrift für die Gesamte Blutforschung 1962, 8 (2), 65–66.*
3. ^*Nowell, P. C.; Hungerford, D. A. Chromosome studies on normal and leukemic human leukocytes. JNCI: Journal of the National Cancer Institute 1960, 25 (1), 85–109.*
4. ^*Rowley, J. D. A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine*

fluorescence and giemsa staining. Nature 1973, 243 (5405), 290–293.

5. ^*Nowell, P. C. The clonal evolution of tumor cell populations. Science 1976, 194 (4260), 23–28.*
6. ^*Martincorena, I.; Campbell, P. J. Somatic mutation in cancer and normal cells. Science 2015, 349 (6255), 1483–1489.*
7. ^*Bizzotto, S.; Walsh, C. A. Genetic mosaicism in the human brain: from lineage tracing to neuropsychiatric disorders. Nature Reviews Neuroscience 2022, 23 (5), 275–286.*
8. ^*Goodell, M. A.; Rando, T. A. Stem cells and healthy aging. Science 2015, 350 (6265), 1199–1204.*
9. ^*Takeda, J.; Miyata, T.; Kawagoe, K.; Iida, Y.; Endo, Y.; Fujita, T.; Takahashi, M.; Kitani, T.; Kinoshita, T. Deficiency of the GPI anchor caused by a somatic mutation of the PIG-A gene in paroxysmal nocturnal hemoglobinuria. Cell 1993, 73 (4), 703–711.*
10. ^*Krill, C. E., Jr.; Smith, H. D.; Mauer, A. M. Chronic Idiopathic granulocytopenia. New England Journal of Medicine 1964, 270 (19), 973–979.*
11. ^*Zuelzer, W. W.; Evans, R. K.; Goodman, J. Myelokathexis — A new form of chronic granulocytopenia. New England Journal of Medicine 1964, 270 (14), 699–704.*
12. ^*McDermott, D. H.; Gao, J.-L.; Liu, Q.; Siwicki, M.; Martens, C.; Jacobs, P.; Velez, D.; Yim, E.; Bryke, C. R.; Hsu, N.; Dai, Z.; Marquesen, M. M.; Stregevsky, E.; Kwatemaa, N.; Theobald, N.; Long Priel, D. A.; Pittaluga, S.; Raffeld, M. A.; Calvo, K. R.; Maric, I.; Desmond, R.; Holmes, K. L.; Kuhns, D. B.; Balabanian, K.; Bachelerie, F.; Porcella, S. F.; Malech, H. L.; Murphy, P. M. Chromothriptic cure of WHIM syndrome. Cell 2015, 160 (4), 686–699.*
13. ^*Hernandez, P. A.; Gorlin, R. J.; Lukens, J. N.; Taniuchi, S.; Bohinjec, J.; Francois, F.; Klotman, M. E.; Diaz, G. A. Mutations in the chemokine receptor gene cxcr4 are associated with WHIM syndrome, a combined immunodeficiency disease. Nature Genetics 2003, 34 (1), 70–74.*
14. ^*Cortés-Ciriano, I.; Lee, J. J.-K.; Xi, R.; Jain, D.; Jung, Y. L.; Yang, L.; Gordenin, D.; Klimczak, L. J.; Zhang, C.-Z.; Pellman, D. S.; PCAWG Structural Variation Working Group; Park, P. J.; PCAWG Consortium. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. Nature Genetics 2020, 52 (3), 331–341.*
15. ^*Stephens, P. J.; Greenman, C. D.; Fu, B.; Yang, F.; Bignell, G. R.; Mudie, L. J.; Pleasance, E. D.; Lau, K. W.; Beare, D.; Stebbings, L. A.; McLaren, S.; Lin, M.-L.; McBride, D. J.; Varela, I.; Nik-Zainal, S.; Leroy, C.; Jia, M.; Menzies, A.; Butler, A. P.; Teague, J. W.; Quail, M. A.; Burton, J.; Swerdlow, H.; Carter, N. P.; Morsberger, L. A.; Iacobuzio-Donahue, C.; Follows, G. A.; Green, A. R.; Flanagan, A. M.; Stratton, M. R.; Futreal, P. A.; Campbell, P. J. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 2011, 144 (1), 27–40.*
16. ^*Tan, J.; Pieper, K.; Piccoli, L.; Abdi, A.; Foglierini, M.; Geiger, R.; Maria Tully, C.; Jarrossay, D.; Maina Ndungu, F.; Wambua, J.; Bejon, P.; Silacci Fregni, C.; Fernandez-Rodriguez, B.; Barbieri, S.; Bianchi, S.; Marsh, K.; Thathy, V.; Corti, D.; Sallusto, F.; Bull, P.; Lanzavecchia, A. A LAIR1 insertion generates broadly reactive antibodies against malaria variant antigens. Nature 2015, 529 (7584), 105–109.*
17. ^*Pieper, K.; Tan, J.; Piccoli, L.; Foglierini, M.; Barbieri, S.; Chen, Y.; Silacci-Fregni, C.; Wolf, T.; Jarrossay, D.; Anderle, M.; Abdi, A.; Ndungu, F. M.; Doumbo, O. K.; Traore, B.; Tran, T. M.; Jongo, S.; Zenklusen, I.; Crompton, P. D.; Daubenberger, C.; Bull, P. C.; Sallusto, F.; Lanzavecchia, A. Public antibodies to malaria antigens generated by two LAIR1 insertion modalities. Nature 2017, 548 (7669), 597–601.*

18. ^*Downes, S. M.; Matthews, L. J. Heritability. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/heredity/ (accessed 2023-07-24).*
19. ^*Visscher, P. M.; Hill, W. G.; Wray, N. R. Heritability in the genomics era — concepts and misconceptions. Nature Reviews Genetics 2008, 9 (4), 255–266.*
20. ^*Gaugler, T.; Klei, L.; Sanders, S. J.; Bodea, C. A.; Goldberg, A. P.; Lee, A. B.; Mahajan, M.; Manaa, D.; Pawitan, Y.; Reichert, J.; Ripke, S.; Sandin, S.; Sklar, P.; Svantesson, O.; Reichenberg, A.; Hultman, C. M.; Devlin, B.; Roeder, K.; Buxbaum, J. D. Most genetic risk for autism resides with common variation. Nature Genetics 2014, 46 (8), 881–885.*
21. ^*Iossifov, I.; O'Roak, B. J.; Sanders, S. J.; Ronemus, M.; Krumm, N.; Levy, D.; Stessman, H. A.; Witherspoon, K. T.; Vives, L.; Patterson, K. E.; Smith, J. D.; Paeper, B.; Nickerson, D. A.; Dea, J.; Dong, S.; Gonzalez, L. E.; Mandell, J. D.; Mane, S. M.; Murtha, M. T.; Sullivan, C. A.; Walker, M. F.; Waqar, Z.; Wei, L.; Willsey, A. J.; Yamrom, B.; Lee, Y.; Grabowska, E.; Dalkic, E.; Wang, Z.; Marks, S.; Andrews, P.; Leotta, A.; Kendall, J.; Hakker, I.; Rosenbaum, J.; Ma, B.; Rodgers, L.; Troge, J.; Narzisi, G.; Yoon, S.; Schatz, M. C.; Ye, K.; McCombie, W. R.; Shendure, J.; Eichler, E. E.; State, M. W.; Wigler, M. The contribution of de novo coding mutations to autism spectrum disorder. Nature 2014, 515 (7526), 216–221.*
22. ^*Krumm, N.; Turner, T. N.; Baker, C.; Vives, L.; Mohajeri, K.; Witherspoon, K.; Raja, A.; Coe, B. P.; Stessman, H. A.; He, Z.-X.; Leal, S. M.; Bernier, R.; Eichler, E. E. Excess of rare, inherited truncating mutations in autism. Nature Genetics 2015, 47 (6), 582–588.*
23. ^*Turner, T. N.; Coe, B. P.; Dickel, D. E.; Hoekzema, K.; Nelson, B. J.; Zody, M. C.; Kronenberg, Z. N.; Hormozdiari, F.; Raja, A.; Pennacchio, L. A.; Darnell, R. B.; Eichler, E. E. Genomic patterns of de novo mutation in simplex autism. Cell 2017, 171 (3), 710-722.e12.*
24. ^*Sánchez-Luquez, K. Y.; Carpena, M. X.; Karam, S. M.; Tovo-Rodrigues, L. The contribution of whole-exome sequencing to intellectual disability diagnosis and knowledge of underlying molecular mechanisms: a systematic review and meta-analysis. Mutation Research/Reviews in Mutation Research 2022, 790, 108428.*
25. ^*Dawkins, R. The Extended Phenotype: The Long Reach of the Gene; Oxford University Press, 2016.*
26. ^*Valentino, S. I.; Greenspan, N. S. Extended phenotype in evolutionary medicine. Evolution, Medicine, and Public Health 2019, 2019 (1), 48–49.*
27. ^*Bäckhed, F.; Ley, R. E.; Sonnenburg, J. L.; Peterson, D. A.; Gordon, J. I. Host-bacterial mutualism in the human intestine. Science 2005, 307 (5717), 1915–1920.*
28. ^*Hofstadter, D. R. Godel, Escher, Bach; 1979.*
29. ^*Krause, E. F. Taxicab Geometry; Addison Wesley Publishing Company, 1975.*
30. ^*Greenspan, N. Taxicab geometry as a vehicle for the journey toward enlightenment. Humanistic Mathematics Network Journal 2004, 1 (27).*
31. ^*Ingram, V. M. Gene mutations in human hæmoglobin: the chemical difference between normal and sickle cell hæmoglobin. Nature 1957, 180 (4581), 326–328.*
32. ^*Hill, A. V. S.; Bennett, S.; Allsopp, C. E. M.; Kwiatkowski, D.; Anstey, N. M.; Twumasi, P.; Rowe, P. A.; Brewster, D.; McMichael, A. J.; Greenwood, B. M. HLA, Malaria and dominant protective associations. Parasitology Today 1992, 8 (2), 57.*

33. ^*Perrine, R. P.; Brown, M. J.; Clegg, J. B.; Weatherall, D. J.; May, A. Benign sickle-cell anæmia. The Lancet 1972, 300 (7788), 1163–1167.*
34. ^*Gödel, K. Über Formal Unentscheidbare sätze der Principia Mathematica und verwandter systeme I. Monatshefte für Mathematik und Physik 1931, 38–38 (1), 173–198.*
35. ^*Savisaar, R.; Hurst, L. D. Exonic splice regulation imposes strong selection at synonymous sites. Genome Research 2018, 28 (10), 1442–1454.*
36. ^*Luria, S. E.; Delbrück, M. Mutations of bacteria from virus sensitivity to virus resistance. Genetics 1943, 28(6):491-511.*
37. ^*Richter, C.; Chang J. T.; Fineran, P. C. Function and regulation of clustered regularly interspaced short palindromic repeats (CRISPR) / CRISPR associated (Cas) systems. Viruses 2012 4(10):2291-2311.*