SOCIOLOGY IN AN AGE OF GENOMIC INSTABILITY: COPY NUMBER VARIATION, SOMATIC MOSAICISM, AND THE FALLEN GENOME

Martine Lappé and Hannah Landecker

ABSTRACT

Purpose – *This study analyzes the rise of genome instability in the life sciences and traces the problematic of instability as it relates to the sociology of health. Genome instability is the study of how genomes change and become variable between generations and within organisms over the life span. Genome instability reflects a significant departure from the Platonic genome imagined during the Human Genome Project. The aim of this chapter is to explain and analyze research on copy number variation and somatic mosaicism to consider the implications of these sciences for sociologists interested in genomics.*

Methodology/approach – *This chapter draws on two multi-sited ethnographies of contemporary biomedical science and literature in the sociology of health, science, and biomedicine to document a shift in thinking*

Genetics, Health and Society
Advances in Medical Sociology, Volume 16, 157–186

ISSN: 1057-6290/doi:10.1108/S1057-629020150000016006

about the genome from fixed and universal to highly variable and influenced by time and context.

Findings – *Genomic instability has become a framework for addressing how genomes change and become variable between generations and within organisms over the life span. Instability is a useful framework for analyzing changes in the life sciences in the post-genomic era.*

Research implications – *Genome instability requires life scientists to address how differences both within and between individuals articulate with shifting disease categories and classifications. For sociologists, these findings have implications for studies of identity, sociality, and clinical experience.*

Originality/value – *This is the first sociological analysis of genomic instability. It identifies practical and conceptual implications of genomic instability for life scientists and helps sociologists delineate new approaches to the study of genomics in the post-genomic era.*

Keywords: Genomic instability; copy number variation; somatic mosaicism; health; difference; post-genomics

INTRODUCTION

Before the year 2000, the quality of instability was not frequently associated with the term genome. Yet these two words have bonded into a single phrase over the last decade as *genomic instability*. This is now a topic of intense research into how cellular genomes change and become variable between generations and within organisms over the life span. Technically, genomic instability refers to the tendency of the genome to acquire mutations when processes of genetic maintenance and replication are dysfunctional.[1] Conceptually, instability enables fundamental shifts in the idea of a singular human genome by suggesting that there are multiple genomes within individuals. Moreover, these multiple genomes age and change with time, producing molecular distinctions both within and between individuals that do not map neatly onto categories of health or disease. Rhetorically, the rise of instability has been accompanied by an anxious terminology of *fragility*, *vulnerability*, *mutational hotspots*, *breakpoints*, and *faulty repair*. This lexicon reflects a striking transformation in the language of

timelessness and implacability of DNA that was part of a universal (and universalizing) human genome dominant during the Human Genome Project.

For life scientists, genomic instability is a matter of both basic science and medical consequence. DNA sequence must be kept in order and chromosomes in correct configuration even as the genome is copied repeatedly during meiotic cell division that gives rise to germ cells, and mitotic cell division that forms the adult body. Otherwise, dysfunction can result, leading to conditions as diverse as cancer, diabetes, and neurodevelopment disorders. Yet understanding genomic instability – when it happens, why, and its effects – is proving to be very complex. Repeated sequences in genomes that are prone to deletion and duplication at replication, and to disrepair after damage, are present in everyone, not just those with disease. Further, genomic instability has been shown to increase over the life span and with environmental exposures, temporal and social dimensions of change that further complicate attempts to translate these findings into clinical settings and to understand their personal relevance.

The notion of genomic instability marks a dramatic shift from the genetics of the 1990s and early 2000s. The Human Genome Project (HGP) ran from 1990 through 2003 and was the largest biological research endeavor in history, promising an unparalleled understanding of human health and disease (Fortun, 2005). Extensive public and private investment in the HGP garnered popular and scholarly attention, and many questioned the effects that this deluge of genetic information would have on society. Sociologists spoke of "the dawn of a genetic age" (Conrad & Gabe, 1999), charting the iconicity of the double helix, and numerous gene metaphors (see Nerlich, Dingwall, & Clarke, 2002) alongside the social, political, and ethical consequences of the project. Efforts to identify a "gene-for" everything from cancer to political beliefs, sexuality to skin color, were critiqued as reductionist, essentialist, and deterministic, sparking concerns about the geneticization of social phenomenon and experiences (see Lippman, 1991 on geneticization; Shostak & Freese, 2010 for review). From the code to the blue print to the holy grail and book of life (Gilbert, 1992; Kay, 2000; Keller, 1995; Nelkin & Lindee, 1996), the "new genetics" initiated by the HGP simultaneously promised access to humanity's ancient history and a platform for the biomedical future.

In this period, DNA was largely represented as universal, stable, and timeless except to the extent that it was shaped by natural selection over the very long time span of evolution. The assumed stability of DNA across individuals over intergenerational time promised a guide to the genetic

basis of human health, disease, and difference. The mutation in the stable blueprint could be identified. The notion of genes as "fate" became both a dominant cultural representation of DNA and an organizing protocol for the HGP. Geneticist Jones (2011) described the Platonic view of DNA that underscored early promises of the HGP. He explains,

> Every human was built on the same plan, and to understand one was to understand all. The early DNA sequencers out-Platoed Plato in assuming not just that the essence of humankind could be expressed in a single DNA sequence, but that this sequence could be stitched together out of bits of double helix, taken from random donors across the globe, into a kind of genomic Everyman. (p. 6)

This vision proved both provocative and unsustainable. What has emerged in its wake is a profoundly more complex conceptualization of genomes and their manifold relationships to health and disease. Often characterized as the post-genomic era, in this period it is "widely accepted that most common traits and medical conditions will not yield straightforward genetic explanations" (Navon & Eyal, 2014, p. 330).

The post-genomic era is characterized both temporally, as the period of time following the sequencing of the "complete" human genome in 2003, and conceptually:

> Postgenomic and postgenomics are meant not only chronologically … but also epistemologically, as the recognition of those gaps in knowledge and unforeseen complexities surrounding the gene that have made our understanding of its function cautiously provisional and perennially contingent. (Meloni & Testa, 2014, p. 4)

Here we propose the rise of instability as a complementary analytical periodization to the "post-genomic." While the *post-* prefix signifies a shift from previous understandings and implementations of late twentieth century genome science and medicine, it operates as an umbrella term that loses its power to describe events of sociological interest except at the most general level. The *post-* prefix produces the question of what comes after it, as well as faintly tautological propositions such as the "post-genomic genome," in which change and uncertainty is produced in the genome in and through practices of genomics (Keller, 2014). It begins to mean everything that is not as it was: gene-environment interaction (Shim et al., 2014; Shostak, 2013), non-coding RNA (Keller, 2012), junk DNA (Bardini, 2011), epigenetic processes (Gilbert & Epel, 2008), the use of big data in molecular medicine (Cambrosio Bourret, Rabeharisoa, & Callon, 2014), or some combination of all of these (Griffiths & Stotz, 2013; Stevens & Richardson, 2015).

Instability offers a different cut. A new representation of the human genome as structurally variable and highly repetitive emerged around 2004, and genomic instability became an area of ferocious knowledge production. We track the rise of instability as a general axis of conceptual and technical change common to multiple subfields of biomedical science and multiple disease areas, that nonetheless is precise enough to specify new developments in old questions: the ones that drew sociologists to genetics and genomics in the first place, raised by genomics in prenatal testing, in the generation of biosociality through new notions of sameness, kinship and difference at the genetic level, in the remaking of health and disease, and in the designation of human historical and geographical distribution as genetic difference. In other words, genome instability is a specific empirical thread through the more general landscape of post-genomics, health, and society that we follow in order to make two arguments about key changes to ideas and practices of inter-individual and intra-individual genomic difference after 2004.

First, with instability, the notion of the *singular* human genome as a basis for distinctions between health and disease, and between sameness and difference has been unsettled. Mutations that were previously thought to be stable, heritable, and comparable between individuals are multiplying into a more complex landscape of molecular variants, many of which have unclear penetrance and are found in both "normal" and diseased populations. Below we illustrate this shift away from the singular genome through an analysis of the study of copy number variation: genomic sciences concerned with the deletions, duplications and translocations of large, often repetitive DNA sequence regions that generate differences in genome structure between individuals, even parents and offspring. Second, there is a growing sense that genomes are dynamic and heterogeneous not only between individuals, but within individual bodies as well. Somatic genome heterogeneity is affected by time, opening cellular genomes to life experiences and the effects of aging. We illustrate the temporal axis of genomic instability through the example of somatic mosaicism in aging and cancer. As genomic instability becomes an area of increasing focus for life scientists, it opens up a new landscape of genomic multiplicity and temporality in health and disease to sociological attention.

Methods

This chapter arose at the intersection of two research projects, focused on developments in autism science (Lappé, 2014) and chromatin biology

(Landecker, 2011, 2013). Both projects are multi-sited ethnographies (Marcus, 1995) that draw on in-depth interviews, peer-reviewed and popular literature, and participant observation at scientific meetings, community forums, talks, journal clubs, and pedagogical settings between 2007 and 2014. The social and historical dimensions of genome sciences figure centrally in both studies. While discussing these dimensions, we found common themes emerging from our respective projects. Utilizing the methods of grounded theory (Charmaz, 2013; Strauss, 1987), we identified *instability* as a conceptual and material instantiation of the practices we were seeing.

In order to systematically explore this theme and its practical and theoretical implications for sociologists, we conducted a key word search in PubMed using the MeSH term "genomic instability" to identify how and where the problematic of *instability* was emerging in relationship to genomics. Simultaneously, we mined the sociological literature on genetics and genomics to consider how these trends in the genome sciences aligned with theoretical and empirical scholarship in the field. Through processes of coding, memo-writing, discussing themes, and returning to the data and literature, we came to see copy number variation and somatic mosaicism as two examples that illustrate the importance of instability for emerging definitions of health and disease in the human lifespan from early development to aging. Below we review these two substantive areas of genomics research and their relationships to sociological studies of health, science, and biomedicine.

REPETITION WITH DIFFERENCE: COPY NUMBER VARIATION

We turn our attention first to the study of copy number variation in the human genome, a development that burst onto the genetics scene in 2004. Copy number variation (CNV) is just what it sounds like: variations – including deletions or duplications – in the number of copies of a genetic sequence. For example, Martine Lappé might have 10 repeated copies of a particular 100-base sequence of DNA all in a row, and Hannah Landecker might have 100 repeated copies at that same locus; our genomes would vary from one another in size by 9,000 base pairs. Then multiply this one example by a thousand or two: and you get two genomes that differ in structure and size, heterogenous with one another in their degree of

repetition or deletion at thousands of loci in the genome. Then extend this example to the human population. The study of this kind of heterogeneity – difference in copy number – is producing different accounts of heredity, as well as different accounts of health and disease, than those produced in a previous era focused on single nucleotide polymorphisms.

We are now, after several decades of DNA sequencing, used to the idea of sequence difference: where you have an A, I have a G. This allele derived from one parent might differ from that allele derived from the other by some nucleotide base substitutions. Copy number variations, in which people vary in number of copies of sequences, proposes a difference in *repetition*, a kind of mutation that is not at the scale of base substitution but of long sequence duplication or insertion. To understand this "different difference," we need to briefly revisit the large-scale sequence comparison that underwrote notions of genetic sameness and difference before 2004 (Fujimura & Rajagopalan, 2011). With the advent of the Human Genome Project and the technical capability to sequence large amounts of DNA, researchers began to compare genomes at scale. The focus was on the sequence of individual nucleotide bases of DNA: A, T, C, G. When single base substitution in one genome versus another is found in more than one percent of a sample population, it is referred to as a single nucleotide polymorphism, or SNP. The SNP became an important technical object in the rise of genomics, a tool for tracking difference between diseased and non-diseased individuals, or between different human populations (Hamilton, 2008). The principles of genomewide association studies, for example, are based on correlating SNPs with disease traits.[2]

The rise of this kind of genome sequencing in the name of health was premised on the idea that "genome sequencing will explain disease cases by revealing the causative genetic blemish – the mutation that stands out on a background of otherwise flawless molecular function" (Macosko & McCarroll, 2013, p. 564). It was a Platonic vision of the ideal genes (the state of health) from which copies with mutations were fallen versions (the state of disease). However, investigators looking for mutations in the genomes of cancer cells found, much to their surprise, that mutations were everywhere in their normal samples as well. A lead author of one of the first papers to describe the widespread abundance of copy number variations in humans underlines this paradoxical situation (Sebat et al., 2004). He explains in an interview how the search to identify disease mutations actually led to the identification of deletion, duplication, and rearrangement as a ground state of human difference:

> Back in the day, as a trainee in Mike Wigler's lab at Cold Spring Harbor, we were using some of his technology that he'd developed to look at the genome, specifically in cancer genomes to find mutations. At that time, what actually turned out to be the most interesting stuff were the mutations we were finding in healthy genomes. In fact there were large deletions, duplications, strange structural mutations in the genomes of everyone walking around on the planet. That was a shocker … we really didn't expect to see that much variation in chromosomal structure as part of our naturally occurring genetic variation, we just didn't expect it. That was really a big bomb that went off in the genetics field. That really highlighted this as an area of genetics that hadn't been explored.[3]

Thus "everyone walking around on the planet" was walking around with "strange structural variations" in their apparently healthy genomes. The "bomb" that was structural variations went off in 2004, and much work quickly ensued focused on a class of genomic changes between the tiny scale of the single nucleotide, and the massive scale of the mutations seen through visual inspection in microscopy: *DNA segments of hundreds of bases or more that repeat in varying numbers in one genome compared to another* (Redon et al., 2006).

Consequently, SNPs are now thought to be "only half of the story" – as attention has turned from sequence variation to structural variation (Malhotra & Sebat, 2012, p. 1223). With these findings, the 99.9% figure of human sequence similarity recedes, and a discussion of repetition with difference proceeds. It is estimated that the human genome is half to two-thirds made up of repeated sequences of various kinds; even the range in the estimates points to the technical difficulty of counting things present in multiple copies (LaSalle, 2011). Part of the shock at the 2004 findings arose exactly from the blind spots in the technologies developed for looking at how people differed one base at a time: these technologies tended to overlay repeated sequences with one another, collapsing them into a single consensus sequence. And even though sequencing technologies have rapidly become more sophisticated, repetitive sequences are still "largely refractory" to next generation sequencing (Lettre, 2014). In addition, this was a shock that belonged to the genre of finding out that humans have less protein encoding genes than rice do: somehow it doesn't seem very dignified to be riddled with repetition.

This change in perspective from single nucleotide base level variation to structural variation also required shifting from the linear sequence to more physical questions of genomic space and arrangement. Copy number variants are referred to as "structural variants," because they augment or diminish the size and alter the architecture of the genome. There are also inversions (reversal of orientation of a sequence relative to the rest of the

genome) and translocations (movement of a sequence to another location in the genome) that do not affect overall size. When scientists attempt to popularize this concept, they return to the book of life, but describe it as having been produced by an imaginary publisher that has duplicated pages, "dropped some pages, changed the order of some pages" so that the story is not continuous (Wigler, quoted in Hall, 2015, p. 40). While deletions, duplications, inversions, and translocations had been seen and studied for decades in cytogenetics, and formed the basis for the first prenatal testing, these changes have to be very large to be seen through the light microscope (Rapp, 1999). Structural variants, while involving far more nucleotide bases than the single changes detected through sequencing, are distinct from cytogenetics in part because they are still "submicroscopic" and cannot be seen with visual inspection of chromosomes through a microscope (Feuk, Carson, & Scherer, 2006, p. 85).

Now that they have been identified, it has become clear that copy number variations are not rare. In fact, they account for more of the variation between any two individuals than single nucleotide changes (with CNVs accounting for 1% of individual variation, where SNPs account for .1%). There can be thousands of CNVs between genomes, some of which are common to many individuals (polymorphic). The 1,000 genomes project estimates that there are more than 2,500 structural variants such as duplications, deletions, and insertions *between any two human individuals* (Conrad et al., 2010). These variants can be inherited across generations, or arise *de novo* in an individual (appearing in a child but not present in the parents). The rate of change in copy number mutations across generations is much higher than the change rate calculated for SNPs. Thus copy number variation and the repetitive nature of the human genome has come to the fore as a source of difference between individuals, both across and between generations.

There is no region of the human genome that does not contain copy number variants. A clear example of gene copy number variation comes from studies of the amylase gene that encodes the enzyme in spit that breaks down carbohydrate; 1−50 or more copies are found in one genome compared to another, with a corresponding physiological change in the amount of amylase made (Perry et al., 2007). This example shows that changes in gene dosage accomplished by having more or less copies of a gene are part of the normal landscape of human variation. However, structural variations can occur anywhere in the genome, and perhaps the majority of CNVs lies outside of the parts of the genome that code for proteins, where their functions and effects are much less clear. Even this very

simplified explanation of copy number variation should make clear its complexity: what we underline here is that it offers up a very different genome: one that is in equal proportions dynamic and unstable, where repeated sequences create variation with precarity during replication. It is to this complex and precarious character of instability that we now turn.

Copy Number Variation, Instability, and Human Disease

All of this structural variation might remain an issue of purely academic interest, were it not for the connection to human disease. Here we build on the work of a large number of sociologists of biomedicine interested in genomics and diagnosis: for example, how genomic technologies are designating phenotypically diverse conditions from the genome up instead of from clinical diagnosis down (Navon, 2011; Navon & Eyal, 2014). Other scholars have pointed out how genomic information is producing new categories of pre-symptomatic persons and pre-disease syndromes (Dumit, 2012; Konrad, 2003), blurring perceived boundaries between health and illness, normalcy and disease (Armstrong, 1995; Clarke, Mamo, Fosket, Fishman, & Shim, 2005; Lupton, 1999). Copy number variation will join a crowd of other examples of how the increasingly untidy relationships between genomic information and disease classifications are being negotiated in clinical settings (Bourret, Keating, & Cambrosio, 2011; Timmermans, 2014), and how such clinical vagaries have intense personal meanings for individuals and families (see Chilibeck, Lock, & Sehdeve, 2011; Lock, 2011). As shall become clear, research on CNVs provides a particularly poignant example of genomic designations of uncertain significance, with profound implications for prenatal testing, disease classification and diagnosis, and understandings of inheritance. They have a dual status as sources of genomic instability and sources of conceptual instability, as "major changes in our DNA lead to major changes in our thinking" (Sebat, 2007).

A core paradox of copy number variations is that they are simultaneously normal and pathological, their presence a source of dynamic instability in any one genome and in the inheritance from one generation to another. In other words, there is no "pure" genome and no "blemish" – there is a welter of difference in which some differences are strongly linked to disease, some are ambiguously linked to disease, and some are not linked to disease at all. The commentary on these developments from the scientific community is telling: instead of the mutation being the fallen version of the

healthy genome, it is all of us who carry "fallen genomes" (Macosko & McCarroll, 2013), genomes that are riddled with repetition (LaSalle, 2011), and populated with transposable elements and other remnants of parasitic mobile DNA (Kazazian, Haig, & Goodier, 2002).

However, copy number variation is not just another example of the post-genomic era's epistemological status as a period of recognition of uncertainties and gaps in knowledge. This research is a destabilization not just of what a sequence means. Rather, it highlights the literal physical instability of the sequence, an object previously assumed to be very stable and a-temporal. A particular material property of these repeated copies of DNA sequence is to destabilize the very process of copying, creating an unexpectedly high degree of instability in the biological processes of genetic inheritance. Here we turn to the connection between copy number variation and genome instability, explaining the different ways in which instability has come to be seen and measured as a property of human genomes. We then further connect CNV-linked genome instability to human disease, using the example of neurodevelopmental disorders to explain the clinical and social implications of the turn to unstable copy number variation as a form of difference between health and illness.

What is the link between varying numbers of copies of repetitive sequences in genomes and genome instability? Even if repetitive sequences, insertions, duplications, or deletions do not directly affect coding regions of a gene, they are understood to affect the *stability* of the genome in different ways. There is some slippage between different forms of genomic instability, but for clarity we will try to tease them apart. First, mistakes in copying occurring at cell division are more likely to occur because of DNA repeats. Repetitive sequences look a lot like one another, which can cause errors of alignment necessary to the complex process of copying the genome and distributing the copies evenly amongst daughter cells. Such mistakes are of profound significance in the meiotic cell division that gives rise to sperm and eggs.

Humans carry two copies of each chromosome, which every high school biology textbook will tell you come one from our fathers and one from our mothers. However, similarity in repeated sequences can cause a misalignment of the two matching copies, a literal wrinkle in the process of particular consequence during the crossover events that remix the chromosomes during meiosis. Repetitive sequences can cause these wrinkles within a chromatid, between chromatids, or even with a matching sequence on another chromosome. "CNVs suggest that an individual's genetic code may not simply be the sum of the genetic contributions of the individual's two

parents. Because the unequal crossover events responsible for CNVs occur during the production of sperm and eggs, children may have lost or gained additional copies of genetic information that were present in either of their parents' chromosomes" (Eichler, 2008, p. 3). Just in case the implications of this aren't clear: an offspring's genome is *not* the exact sum of bits received from mom and dad. Such change is referred to as *de novo* germline mutation, because the variant present in the child resulting from such gametes is *not* present in the parents. As areas of high repetition are more prone to copying errors in meiosis, they are seen as a source of instability in the fidelity of the transmission of genetic material between generations. Instability here is thus understood as a kind of disruption of historical stability: the given structure is not necessarily stable from generation to generation.

Second, a subset of copy number variants are *retrotransposons*, also known as "jumping genes." They encode the enzymes that mediate their own "copy-and-paste" mechanism of movement, by which their DNA sequence is copied into RNA, and the RNA is reverse-transcribed to DNA, and this DNA is inserted into a new location in the genome. The human genome is "littered with remnants" of retrotransposons that have been shortened, mutated, or inverted and thus become inactive; approximately half a million such remnants have been mapped (Kazazian, Haig, & Goodier, 2002). But there are also a small group of these mobile genetic elements that remain mobile. As with the copying errors described above, retrotransposon copying and new insertion during meiosis can cause *de novo* mutation, which is potentially very disruptive of gene function or regulation. This too is a source of instability in heredity.

At this point, a reader might ask, *so what?* If we all carry these variants, what's the big deal? A third usage of instability comes into play when describing the effects of CNVs once they have been inherited or arisen. Because these changes involve long base sequences, they can encompass whole genes and gene regulatory regions, and deletion or duplication can affect the *dosage* – the amount of a gene product made. Many CNVs affect so-called non-coding regions of the genome, but nonetheless change the transcription status of the part of the chromosome they are in, either affecting neighboring genes or changing the amount of non-coding RNA transcribed from that area. An insertion or deletion might affect the expression of the many enzymes and proteins that construct, maintain and open and close chromatin or repair DNA, thus interfering with the ability of the organism to repair itself after damage, for example from UV light. We will discuss this form of cellular genome instability further when we turn to

somatic mosaicism, but it is part and parcel of the same biology of instability found in the germ-line. Copy number variations are not necessarily deleterious, disease-causing, or compromising. But many of their effects are understood to destabilize or unbalance the genome, making it more prone to further mutation.

CNVs in Neurodevelopmental Disorders

The case of neurodevelopmental disorders demonstrates the broader effects of a framework of instability coming to genomic science, in which the genome has fallen and yet in falling has opened out to a new and productive research frontier. While copy number variations do not necessarily indicate disease, both inherited and *de novo* CNVs have been correlated with human diseases, particularly neurological and developmental disorders such as schizophrenia and autism (LaSalle, 2011; Vulto-van Silfhout et al., 2013). Indeed, chromosomal microarray analysis was recently recommended as a first-tier diagnostic test for individuals with developmental or congenital abnormalities, with guidelines for clinical use issued by the International Standard Cytogenomic Array Consortium and the American College of Medical Genetics (Miller et al., 2010). With something of a note of triumph, it has been noted that now, 10 years after genomic techniques came into routine clinical use, approximately 20–25% of children presenting with congenital malformations or intellectual disability and developmental delay can be offered at least a partial explanation for the condition based in genome analysis (Beaudet, 2014), what sociologist Daniel Navon has called "genomic designation" of disease (2011). Genomic designation, Navon argues, is creating new kinds of identity and new forms of illness identified primarily as genetic conditions. While some of those designations are based on SNP analysis, many of them are based on detecting CNVs.

A recent enormous study compared CNVs in 29,085 children with developmental delay to 19,584 controls, constructing a "CNV morbidity map" for intellectual disability by combining CNV detection with sequencing of the variants (Coe et al., 2014). This confirms earlier studies of CNVs in autism spectrum disorder (ASD), which found higher rates of de novo CNVs in affected individuals compared to unaffected siblings (Sebat et al., 2007). However, many of these findings are very different from earlier mutation studies that seemed to be finding the "gene for" certain diseases, the Mendelian or monogenic disorders, and indeed different from the identity formations and diffuse clinical categories arising from a mutation shared

by many individuals with different phenotypes, such as 22q11.2 Deletion Syndrome (Navon & Shwed, 2012). You might sequence a thousand autism patients and find a thousand different CNV mutations; what these individuals have in common is an increased rate of *de novo* CNVs, not necessarily a mutation *in the same place* in the genome repeated a thousand times. This is the concept of rare variants that taken together as a group account for some significant percentage of cases of a disease. It is an "aggregate burden" (Krumm, O'Roak, Shendure, & Eichler, 2014). The realization that these rare variants might explain more disease than the previously sought common variants for common disease has revived confidence that large-scale genomics will lead to genetic explanations for neurodevelopmental disorders – yet it must be noted that the percentage of cases explained in this way still accounts for a fraction of the whole.

While expanding the scope of genomic designation, CNV analysis has at the same time further complicated the genotype-phenotype link. Sometimes it is clear: this deletion or inversion leads to a disease or a cluster of symptoms. Most of the time, however, it is not. The same variants in different people can have a different phenotype, different variants in different people can have the same phenotype, different CNVs can modify the effect of other CNVs such that two CNVs in a person will have a different effect than either CNV alone. As human geneticist Beaudet (2014) puts it:

> Although there are databases and tracks on genome browsers with data for CNVs in healthy controls and in symptomatic individuals (Database of Genomic Variants, DGV; International Collaboration for Clinical Genomics, ICCG; Decipher; and dbVar, a database of genomic structural variation), *it is difficult even for those immersed in CNV biology and clinical practice to determine the likely pathogenicity of a particular variant*. (emphasis added, pp. 1046–1047)

In short, copy number variation has simultaneously reasserted and etiolated the explanatory power of genomics for neurodevelopmental disease. Beaudet (2014) writes that there is a spectrum of apparent effect of CNVs on human disease that ranges from all to none. He proposes a multiple category classification system that ranges from the most severe and penetrant CNVs, which if inherited also cause disease, to categories in which the parents of a patient presenting with deformity or intellectual disability might also turn out to carry the same CNV but are "superficially normal." Another category includes CNVs that appear to have subclinical penetrance, "a concept that has not been well documented but is thought by many clinicians to be a real phenomenon, whereby an individual has an IQ that is 10–15 points lower than would be the case without the CNV but

does not meet the criteria for any medical or disability diagnosis and has normal reproduction" (p. 1047).

To the ranks of the "pre-symptomatic person" identified by tests for Huntington's Disease (Konrad, 2003), the "patients-in-waiting" marked out by post-natal biochemical tests as potential metabolic disease patients (Timmermans & Buchbinder, 2012), and the recreational genetics consumer interested in ancestry or disease risk (Lee, 2013; Nelson, 2008), findings related to CNVs therefore add several more categories of genomic identity: the "superficially normal" bystander who would not normally have peered into his or her genome were it not for a relative undergoing diagnosis, the subclinical person, and the fetus with a "variant of uncertain clinical significance," or VOUS (Crolla, Wapner, & Van Lith, 2014). Beaudet (2010) estimates that such findings of "troublesome uncertain significance occur in about 1% of routine prenatal samples" (p. 42). The clinical detection of CNVs in patients has led to proposals to use these tests in routine prenatal testing. It is one thing to detect a copy number variant of unknown clinical significance in a patient; it is quite another to find it in a prenatal test of fetal DNA, where there is no "phenotype" or clinical presentation yet to consider in relation to the genetic finding. While we cannot discuss this fraught issue at any length here, the reality of prenatal testing for CNVs does underline the profound social consequences for these new modes of relating and confusing genotype-phenotype relations.

SOMATIC MOSAICISM: GENOMES IN TIME

What we would like to add to this story of genomic instability arising in the germ-line (as if it wasn't complicated enough), is a second, temporal axis of copy number variation: that which arises over time, *within* an individual, rendering the body a "somatic mosaic" of cells composed of structurally variant genomes. This idea that "we contain genetic multitudes" came on the heels of the concept of inter-individual structural variation described above, and was subject to much initial skepticism (Zimmer, 2013). Certainly, somatic mosaicism had been seen in various severe clinically identified conditions long before (Hall, 1988); what was new and controversial was the idea that it was a universal and common characteristic of everyone, not just people with a disease such as neurofibromatosis.

Somatic mosaicism is a temporal and intra-individual manifestation of copy number variation. It generates novel challenges to our understanding

of the biological individual, which has for many decades now rested on the assumption that all post-zygotic cells have identical genomes (Gage, 2013). In the example of neurodevelopmental disorders, we emphasized that copy number variation can generate instability in intergenerational inheritance, with copying or repair errors at meiosis that generate differences in gametes. However, the very same causes of instability occur with mitosis, the version of cell division that happens in the soma, where a copy of each of the chromosomes is made, and then the cell divides into two daughter cells, each containing its own copy of the genome. Copying errors, novel insertions due to retrotransposons, and DNA damage that is not correctly repaired also all occur during somatic cell life and *mitotic* cell division, particularly during early development in the embryo and the fetus, when cells are dividing at an extremely rapid pace to form tissues and organs.

While it is extremely likely that both germ-line and developmental copy number variation contribute to the neurodevelopmental disorders discussed above, somatic cells differ from each other genomically in every human being. Everyone is composed of somatic cells whose genomes are structurally variable. For example, any given person has multiple genomes in one brain, sometimes differing in thousands of base pairs or whole chromosomes from one another. Neurons taken from post-mortem adult human brain frontal cortex tissue, when investigated with single-cell sequencing, showed copy number variation compared to other cells from the same brain. Moreover, the amount of variation within one brain was variable between different brains (McConnell et al., 2013). Some of the neurons sequenced had "highly aberrant genomes marked by multiple alterations"; all this variability was seen in people who had in life been normal in clinical terms (p. 632). Thus there is variability even in degrees of variation. In other words, the germ-line can carry inherited and *de novo* CNVs, but in addition, the ensuing genomes of the many cells descended from the original zygote also variegate during cell division in development.

By early adulthood "somatic cells no longer possess identical genomes" (Vijg & Suh, 2013).[4] Thus the body is heterogenous with itself, and not just in brain tissue. Post-mortem comparison of DNA from diverse tissues such as liver, pancreas, and hippocampus taken from the same individuals showed widespread intra-individual tissue-specific copy number variation (O'Huallachain, Karczewski, Weissman, Urban, & Snyder, 2012). The individuals autopsied had no known hereditary disease or cancer, indicating that a variable somatic genome is a normal characteristic of humans (O'Huallachain, Weissman, & Snyder, 2013). The geneticist Carl Bruder, who has produced data showing monozygotic (identical) twins differ

genetically from one another due to copy number variation that increases with age, summed it up nicely by stating "I believe that the genome you're born with is not the genome you die with – at least not for all the cells of your body" (quoted in Casselman, 2008). Indeed, these findings of somatic mosaicism "question the long-standing notion that MZ twins are essentially genetically identical" by demonstrating that twins' somatic genomes can vary (Bruder et al., 2008, p. 765). Further, the degree of variance is a function of age; studies comparing the blood cell genomes of monozygotic twins showed more structural variation between older twins than younger ones; 55 being the age above which more variation was seen (Forsberg et al., 2012).

Perhaps paradoxically, somatic mosaicism also occurs in the body of research as it ages – in human cells cultured in the laboratory (Landecker, 2007). Sampled tissues are cultured to serve as proxies for the bodies they came from, and the assumption up until now has been that the original and sample are genetically identical. For example, lymphoblastoic cell lines (LCL) derived from virus-transformed B cells and cultured *in vitro* are supposed to create a long-term reference collections such that one can return to the collection and not the person, as was done in the Human Genome Diversity Project (Lock, 2001). Just as cells in the body can acquire genetic variation through multiple rounds of division during proliferation, cells *in vitro* can also "acquire a new genotype" (Forsberg, Absher, & Dumanski, 2013, p. 8). Induced pluripotent stem cells, around which many therapeutic hopes swirl, are also cultured *in vitro* from an original cell population; they also have been shown to acquire new copy number variations in culture (Abyzov et al., 2012). Indeed, researchers looking at iPS cells derived from human skin fibroblasts found new variations in the cultured cells, but also that about 30% of the original skin cells already were genetically variable, "suggesting widespread somatic mosaicism in the human body" (Abyzov et al., 2012, p. 438).

In short, assumptions about the self-sameness of cells through time, whether they are cells in a body descended from a zygote or cells in a dish descended from a tissue sample, were unduly influenced by the twentieth century machine replica, from the carbon copy to the photocopier. This has produced an assumption that cells in a body are clones premised on a notion of cloning as a form of precise copying of DNA (Friese, 2013). Yes, an amazingly complex set of copying and repair processes have evolved that maintain the fidelity of DNA replication. But the inviolability of the copy is further questioned when the template itself has a propensity for slippage, mismatch, and mix-up. The alternative – that there is a zygote,

from which cells descend to make a multitude of very alike but structurally variant genomes – generates philosophical challenges to Western ideas of individuality, as do other forms of biological mosaicism and chimerism.[5] Women who have been pregnant are home to cells from their fetuses for decades afterward, a phenomenon called feto-maternal microchimerism (Kelly, 2012; Martin, 2010), the donor stem cells received by transplant recipients go on to generate cells in tissues all over their bodies (Berger et al., 2013), and fetuses in utero together exchange cells that persist through their lives (Dupré, 2010). The recent explosion of interest in the human microbiome also presents an image of the body as a congeries of genomes functioning together, including human, bacterial, viral, and fungal. Somatic mosaicism is another example of this "polygenomic organism" characterized by philosopher of biology John Dupré – but again, it is an example that highlights instability as a key feature of biological existence in a way that other multifarious genomes might not.

Up to this point we have discussed the normal and surprisingly widespread nature of somatic mosaicism in the body, which seems to increase with time and age. Now we turn to the other side of the coin of instability, its connection with human diseases such as cancer. It is in the study of somatic mosaicism and human disease that environmental exposures come into the picture. In other words, genomes do not just age and become more heterogenous: they are perforated by the environments in which they age.

Time and Environment Come to the Genome

As with copy number variation more generally, the incidence of heterogenous genomes between cells in a single body is a normal condition found in everyone and yet is also increasingly investigated as a marker of disease risk. We have emphasized that somatic mosaicism generates intra-individual difference, but of course, the more variable an individual's genomes are within themselves, the more different they are from other individuals. Throw in a temporal axis that includes both development and aging, and you get a great deal of variation in variation, some of which is being linked to disease risk. This can occur early in life:

> Some individuals will experience their first mutation early in development and carry many somatically mutated cells. Other individuals will have their first mutation later in development and have fewer somatically mutated cells. The inevitable variation in the degree of somatic mosaicism arising early in life could determine much of the variation in the risk of disease later in life. (Frank, 2014, p. R577)

Or it could occur in adulthood, and affect only one organ or system. A growing number of studies looking at the impact of stress, obesity, or diabetes on the cellular genomes of various tissues in the body are finding that detectable changes in the epigenome, telomeres, and genomic stability are not homogenous throughout the body but manifest particularly in some tissues and not others (Horvath et al., 2014). Whether disease presages instability or instability presages disease, or both occur with increasing age, intra-individual genomic structural variation is a spectrum which shades toward disease risk at one end.

As discussed above, even the amount of copy number variation can vary between and within individuals. It is not at all clear why some individuals should have more copy number variation and genomic instability than others, in particular when one sibling is affected and another is not. This question opens the unstable genome up to the environment. Geneticist Janine LaSalle points out that CNVs can alter the size of the genome – literally how many bases it contains in its sequence, and this can in turn change the physical demand for the amount of epigenetic infrastructure that cells need to make to keep up chromatin conformation, methylate DNA, and maintain stability.[6] The molecules that constitute the epigenome come from food, and are affected in their metabolism by exposure to various pollutants (LaSalle, 2011). A developing organism exposed to more environmental toxicants will have a higher cellular demand for methyl donors to maintain genome methylation. This hypothesis thus proposes that genomic instability is always instability in context; inadequate environmental support for the stability of chromatin would then be part of the story of destabilization of genome inheritance between cellular generations (Waterland & Jirtle, 2003).

Similarly, extreme stress has been shown in experimental animals to produce changes in chromatin conformation in brain cells, and these changes are localized to regions encompassing active transposable elements. It is as if the animal has responded to stress by tightening control of mobile genetic elements, and the scientists reporting these results hypothesize that this is a "genomic stability mechanism" directed at protecting the neurons important to long-term memory formation (Hunter, McEwen, & Pfaff, 2013). Memory of life-threatening circumstances is presumably a good thing to maintain, so the persistence at the molecular level is theorized as subtending a behavioral adaptation beneficial to long-term survival. The experimentally induced chromatin changes persist long-term after the stressor is removed; the researchers see it as providing clues to the generation of human conditions such as PTSD, particularly in light of the finding that

repetitive elements are differentially methylated in combat veterans with PTSD when compared to combat veterans who did not develop the condition (Rusiecki et al., 2012). If the underlying variability in neural cells of the number, kind, and distribution of repetitive sequences is different between different individuals, perhaps this provides some insight into differential vulnerability or resiliency in relation to the same stressors; in other words, it might explain why not everyone gets PTSD.

Worries about faulty maintenance can also be seen in studies of the effects of stress and pollution on telomere length in humans, another form of copy number variation. Telomeres are the long repetitive sequences at the ends of chromosomes. With each cell division, the linear chromosome gets a little bit shorter, and telomere sequence buffers the rest of the chromosome's DNA from being eroded. Telomeres prevent different chromosomes from fusing to one another; these two kinds of chromosome protection are understood as essential to genome stability. Epidemiological studies have connected shorter telomeres (measured in blood leukocytes or in saliva) with a raft of human ills and conditions, such as cancer, diabetes, metabolic syndrome, long-term smoking, high blood pressure, and cardiovascular disease. A series of high-profile epidemiological studies also linked psychosocial stress to telomere length, comparing telomere length of women engaged in long-term caregiving of chronically ill children to control subjects with healthy children and finding that perceived stress was correlated with shorter telomeres (Epel et al., 2004). There is of course much more to say about telomeres as a new biomarker of human social life – as well as much debate about what telomere length means physiologically – but it serves as a further example of how genome stability under environmental pressure has come to the fore as a quantitative method connecting genomes to environments over time.

Finally, it has long been recognized that cancer cells have a high degree of genomic instability, and the cancerous cells that make up tumors are within themselves genomically heterogenous. Catastrophic "chromosome shattering" is sometimes seen in cancer cells, in which chromosomes break apart and get haphazardly patched back together by DNA repair mechanisms. The link between this kind of extreme genome stability and the increasing divergence of somatic cell genomes with aging is being actively pursued, although the cause-and-effect question of whether genome instability causes cancer or cancer causes genome instability is an open one (De, 2011). As there are many proliferative tissues, particularly blood, from which sub-populations of cells descend from progenitors, finding different copy number variation "clones" mixed in cell populations indicates that the

genomes of the blood progenitors have diverged with time. Several studies of blood from large numbers of individuals showed that "clonal mosaicism" is rarely seen in individuals under 50, but increases with age and with cigarette smoking; further, that detectable mosaicism in the genomes of blood cells was a risk for subsequent cancer (Jacobs et al., 2012; Laurie et al., 2012). Clonal mosaicism also increases with age in the cells of the intestinal tract and has been linked to colon cancer (Issa, 2014). Again, at issue is the maintenance of stability of the genome as the cells that make up the body continue to divide, particularly in tissues with high turnover such as blood and epithelium.

Neurodevelopmental disorders and cancer might seem to be conditions that are far afield from one another in kind and in the life course. What links the various hypotheses reviewed in the paragraphs above is not our certainty as observers that these scientific ideas are somehow more correct than previous ones. Rather, we wish to draw our readers' attention to how the environment is understood in each of these areas of study as a constant physical perforation of DNA. It is estimated that tens of thousands of DNA lesions arise *per day* in any given somatic cells, under the assault of reactive oxygen species, UV radiation, and environmental mutagens (Vijg & Suh, 2013). In every repair event, the processes of sequence matching and copying have to take place, and every repaired sequence will be itself copied again in cell division. The repetitive, inverted, duplicated, translocated sequences of copy number variation complicate these processes that depend so profoundly on sequence homology. Disease does not arise directly from environmental assault – we wouldn't live out a day if that were so – but from the failure of the processes of repair and replication fidelity. Thus, disease is refigured as a failure of maintenance.

Somatic mosaicism thus bears many implications for shifting our understanding of human disease within the context of social and physical environments, particularly in aging bodies. Findings related to somatic mosaicism add to a growing sense of genomes not as separate from or somehow beyond the environment, but constructed within and in response to it (Lappé & Landecker, 2015). This research is one part of a "much larger revolution in our thinking both about the relationship between genes, genomes, and organisms, and about the relation between all three entities and their environments" (Keller, 2014, p. 2424). This image of constant repair enhances the characterization of genomes being reconceptualized as reactive and responsive systems that can nevertheless remain stable in ways that matter for human evolution and health (Griffiths & Stotz, 2013; Stotz, 2006).

Instability is of course the inextricable counterpart to stability, and genomic instability is a necessary complement to concepts of reactive genomes in a world of signals. Intra-individual difference over time and in environmental context shows the very ground of response to be changeable and changing, in the fundamental sense of having to come into existence over and over again in any body, and in the course of time and repair and replication being shaped in sequence, structure and size. The example of somatic mosaicism reminds us that the environment is not all signals, and the genome as a singular cybernetic reaction apparatus receiving and responding to signals is a dominant contemporary representation with as many theoretical and cultural particularities as the timeless blueprint-codescript book of life it has displaced (Keller, 2012; Landecker, 2011; Meloni & Testa, 2014). DNA and chromatin are chemical matter existing in time in a material world, in which UV light and oxygen are pervasive elements, and humans exist in a social, dietary, and industrialized soup of matter and action that constitutes their biology just as it constitutes their sociality (Fausto-Sterling, 2008).

CONCLUSION

In sum, the venture to chart all inter-individual human difference, so as to generate the "normal" background for studying the genetic changes that cause disease, has resulted instead in a welter of differences. Genomes can vary in size and structure between people, within one person and even within one tissue, but they can also vary over time and the millions of genome replications and cell divisions that constitute a body in a lifetime. Structural variation and therefore genomic instability is a universal feature of human development, growth, aging, and disease. And, while it has produced a new frontier for genome science in analyzing human disease, genomic instability is at the same time a normal feature of life; genomic rearrangements have long been recognized as necessary to the generation of antibody diversity in immune cells. With the realization that structural variation is not a form of specialization, but is seen in germ cells, in embryos, in intestinal epithelium, in brain cells, and so on, the paradox is becoming clearer: repetition is a driver of difference, the site of plasticity and resilience and variation, but also of fragility and instability.

The implications of these changes in understanding the genome in time and context for patients and for society range from the extremely practical

to the existential. The manifest normality of all of this genomic variability and its change within an individual's lifetime raises the question of the adequacy of testing one tissue at one time in genomic diagnosis and study. If the genome you're born with is not the one you die with, then what does a genetic test at one point from one tissue tell you? The link between mutation and disease is simultaneously reopened for genomic medicine and further complicated, as the meaning of any given change ranges from the perfectly normal to the apparently normal to the pathological. The link between disease and heredity is likewise further complicated, as the very tenet that individuals are no more and no less than composite of a maternal and paternal genome shifts. These changes raise the important question of who will be genetically tested and how, given variants of unknown significance and their complex interactions. The genome and individuality are also opened up to the environment with these developments, changing the tenor of how we might think about chemical, industrial, and dietary exposure and aging. Finally and fundamentally, the genome is always actually genome*s*, repetitious, different, and changing.

The implications of genomic instability for sociology of health, disease, and genetics are therefore also manifold. We have followed the lead of the many social scientists observing and analyzing the post-genomic moment, but departed from this literature in focusing on genomic instability. In this way, we have demonstrated that it is not just the *meaning* of genomes and mutations that has shifted, but the very material object of genomics has come to be understood as an unstable entity in time and space. As the ground shifts uneasily under our feet, it seems a good time to ask how to do sociology of genomics in a time of instability. Instability is an empirical phenomenon that needs to be followed in order to grasp how notions of health, disease and human difference will change as medical and popular discourses accommodate changeable, fragile, and multiple genomes. At the same time, it is an analytic that has allowed us to cross between disparate diseases, different subfields of biomedical inquiry, and indeed separate research projects to note a common pattern of change across these various spaces. Thus we have been able to observe the rise of stability as a core matter of concern for a wide array of investigators and clinicians, the destabilization of DNA as a timeless object, the fallen genome as a source of renewed enthusiasm for genomics of disease and new classifications for copy number variant humans, the redescription of disease as failures of maintenance, and the resurgence of the time of the human life span as an important variable in understanding genomes. These are matters that need to be tracked going forward, as these new forms of difference and concepts

of stability, maintenance, and repetition disseminate through genomic science, medicine, and society.

NOTES

1. This definition comes from the PubMed MeSH term *Genomic Instability*, which was established in 2004.

2. Discussion of the implications of CNVs for the study of population structure, human evolution, the genomic analysis of race, and the sociological area of genetics and race, is beyond the scope of this chapter. Most of the literature, and most of the investigators who work on and with CNVs are driven by questions about disease, as "rare variants" that nonetheless carry big impacts for human disease. Because the attempt is not, as it was with single nucleotide polymorphisms, to find *common* variants for *common* diseases, the overlay between geography, population, polymorphism, and disease become less relevant for those interested in CNVs than it was for GWAS studies focused on SNPs (Montoya, 2011). We expect that scholars of race and genomics are just now observing the incorporation of structural variation into population genetics, and that the use of the same population assumptions and databases will extend the analyses made in this literature into new territory (see Sudmant et al., 2010), while also perhaps introducing new elements of instability into existing concepts of heredity and population (Bliss, 2012; Fujimura & Rajagopalan, 2011; Shim et al., 2014). The oft-quoted estimate that humans are 99.9% identical in DNA sequence (an underlying premise of talking about "the" human genome in singular terms, or about racial differences within the human genome) was based on looking at single nucleotide variation (Reardon, 2009).

3. Sebat (2012).

4. For simplicity, here we discuss the relationship between copy number variation and somatic mosaicism, but point mutations also contribute to this phenomenon and to some of the diseases traced to mosaicism.

5. Mosaicism is usually used to refer to mixtures of cells with variable genomes that arose from mutation, whereas chimerism is used to refer to mixtures of cells with variable genomes where some cells come from a different lineage, such as an organ donor.

6. For an in-depth analysis of studies of aging and genomic instability in epigenetics, see Lappé and Landecker (2015).

ACKNOWLEDGMENT

We would like to thank Stefan Timmermans, Daniel Navon, the members of the Institute for Society and Genetics, and the anonymous reviewers of this chapter for their insightful input. This research was supported by the Columbia University Center for Research on Ethical, Legal, and Social

Implications of Psychiatric, Neurologic and Behavioral Genetics (NIH grant number P50HG007257) (ML), and a grant (ID: 39147) to the Metaknowledge Network by the John Templeton Foundation (HL).

REFERENCES

Abyzov, A., Mariani, J., Palejev, D., Zhang, Y., Haney, M. S., Tomasini, L., … Vaccarino, F. (2012). Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature*, *492*(7429), 438–442. doi:10.1038/nature11629

Armstrong, D. (1995). The rise of surveillance medicine. *Sociology of Health and Illness*, *17*(3), 393–404. doi:10.1111/1467-9566.ep10933329

Bardini, T. (2011). *Junkware*. Minneapolis, MN: University of Minnesota Press.

Beaudet, A. L. (2010). Ethical issues raised by common copy number variants and single nucleotide polymorphisms of certain and uncertain significance in general medical practice. *Genome Medicine*, *2*(7), 42. doi:10.1186/gm163

Beaudet, A. L. (2014). Reaching a CNV milestone. *Nature Genetics*, *46*(10), 1046–1048. doi:10.1038/ng.3106

Berger, B., Parson, R., Clausen, J., Berger, C., Nachbaur, D., & Parson, W. (2013). Chimerism in DNA of buccal swabs from recipients after allogenic hematopoietic stem cell transplantations: Implications for forensic DNA testing. *International Journal of Legal Medicine*, *127*(1), 49–54. doi:10.1007/s00414-012-0687-5

Bliss, C. (2012). *Race decoded: The genomic fight for social justice*. Palo Alto, CA: Stanford University Press.

Bourret, P., Keating, P., & Cambrosio, A. (2011). Regulating diagnosis in post-genomic medicine: Re-aligning clinical judgment? *Social Science & Medicine*, *73*(6), 816–824. doi:10.1016/j.socscimed.2011.04.022

Bruder, C. E. G., Piotrowski, A., Gijsbers, A., Andersson, R., Erickson, S., Diaz de Ståhl, T., … Dumanski, J. P. (2008). Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *The American Journal of Human Genetics*, *82*(3), 763–771. doi:10.1016/j.ajhg.2007.12.011

Cambrosio, A., Bourret, P., Rabeharisoa, V., & Callon, M. (2014). Big data and the collective turn in biomedicine. How should we analyze post-genomic practices? *TECNOSCIENZA: Italian Journal of Science & Technology Studies*, *5*(1), 11–42.

Casselman, A. (2008). Identical twins' genes are not identical. Scientific American, April 3. Retrieved from http://www.scientificamerican.com/article.cfm?id=identical-twins-genes-are-not-identical

Charmaz, K. (2013). *Constructing grounded theory*. Los Angeles, CA: Sage.

Chilibeck, G., Lock, M., & Sehdeve, M. (2011). Postgenomics, uncertain futures, and the familiarization of susceptibility genes. *Social Science & Medicine*, *72*(11), 1768–1775. doi:10.1016/j.socscimed.2010.01.053

Clarke, A., Mamo, L., Fosket, J. R., Fishman, J. R., & Shim, J. K. (2005). *Biomedicalization: Technoscience, health, and illness in the U.S.* Durham, NC: Duke University Press.

Coe, B. P., Witherspoon, K., Rosenfeld, J. A., van Bon, B. W., Vulto-van Silfhout, A. T., Bosco, P., & Eichler, E. E. (2014). Refining analyses of copy number variation identifies

specific genes associated with developmental delay. *Nature Genetics*, *46*(10), 1063–1071. doi:10.1038/ng.3092

Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., … Hurles, M. E. (2010). Origins and functional impact of copy number variation in the human genome. *Nature*, *464*(7289), 704–712. doi:10.1038/nature08516

Conrad, P., & Gabe, J. (1999). Introduction: Sociological perspectives on the new genetics: An overview. *Sociology of Health and Illness*, *21*(5), 505–516. doi:10.1111/1467-9566.00170

Crolla, J. A., Wapner, R., & Van Lith, J. M. (2014). Controversies in prenatal diagnosis 3: Should everyone undergoing invasive testing have a microarray? *Prenatal Diagnosis*, *34*(1), 18–22. doi:10.1002/pd.4287

De, S. (2011). Somatic mosaicism in healthy human tissues. *Trends in Genetics*, *27*(6), 217–223. doi:10.1016/j.tig.2011.03.002

Dumit, J. (2012). *Drugs for life: How pharmaceutical companies define our health*. Durham, NC: Duke University Press.

Dupré, J. (2010). The polygenomic organism. *The Sociological Review*, *58*(s1), 19–31. doi:10.1111/j.1467-954X.2010.01909.x

Eichler, E. E. (2008). Copy number variation and human disease. *Nature Education*, *1*(3), 1.

Epel, E. S., Blackburn, E. H., Lin, J., Dhabhar, F. S., Adler, N. E., Morrow, J. D., & Cawthon, R. M. (2004). Accelerated telomere shortening in response to life stress. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(49), 17312–17315. doi:10.1073/pnas.0407162101

Fausto-Sterling, A. (2008). The bare bones of race. *Social Studies of Science*, *38*(5), 801–832. doi:10.1177/0306312708091925

Feuk, L., Carson, A. R., & Scherer, S. W. (2006). Structural variation in the human genome. *Nature Reviews Genetics*, *7*(2), 85–97. doi:10.1038/nrg1767

Forsberg, L. A., Absher, D., & Dumanski, J. P. (2013). Non-heritable genetics of human disease: Spotlight on post-zygotic genetic variation acquired during lifetime. *Postgraduate Medical Journal*, *89*(1053), 417–426. doi:10.1136/postgradmedj-2012-101322rep

Forsberg, L. A., Rasi, C., Razzaghian, H. R., Pakalapati, G., Waite, L., Thilbeault, A., … Dumanski, J. P. (2012). Age-related somatic structural changes in the nuclear genome of human blood cells. *The American Journal of Human Genetics*, *90*(2), 217–228. Retrieved from doi:10.1016/j.ajhg.2011.12.009

Fortun, M. (2005). For an ethics of promising, or: A few kind words about James Watson. *New Genetics and Society*, *24*(2), 157–174. doi:10.1080/14636770500184792

Frank, S. A. (2014). Somatic mosaicism and disease. *Current Biology*, *24*, R577–R581.

Friese, C. (2013). *Cloning wild life: Zoos, captivity, and the future of the endangered animal*. New York, NY: New York University Press.

Fujimura, J., & Rajagopalan, R. (2011). Different differences: The use of 'Genetic Ancestry' versus race in biomedical human genetic research. *Social Studies of Science*, *41*(1), 5–30. doi:10.1177/0306312710379170

Gage, F. (2013). Neuronal Plasticity and Neural Diversity. IMED Seminar, UCLA, March 20, 2013.

Gilbert, S., & Epel, D. (2008). *Ecological developmental biology: Integrating epigenetics, medicine, and evolution*. Sunderland, MA: Sinauer Associates.

Gilbert, W. (1992). A vision of the grail. In D. Kelves & L. Hood (Eds.), *The codes of codes: Scientific and social issues in the human genome project* (pp. 83–97). Cambridge, MA: Harvard University Press.

Griffiths, P., & Stotz, K. (2013). *Genetics and philosophy: An introduction*. Cambridge, UK: Cambridge University Press.

Hall, J. G. (1988). Review and hypothesis: Somatic mosaicism: Observations related to clinical genetics. *American Journal of Human Genetics*, *43*, 355–363.

Hall, S. S. (2015). Solving the autism puzzle. *MIT Technology Review*, 118(1), 36–43.

Hamilton, J. A. (2008). Revitalizing difference in the HapMap: Race and contemporary human genetic variation research. *The Journal of Law, Medicine & Ethics*, *36*(3), 471–477. doi:10.1111/j.1748-720X.2008.293.x

Horvath, S., Erhart, W., Brosch, M., Ammerpohl, O., von Schönfels, W., Ahrens, M., & Hampe, J. (2014). Obesity accelerates epigenetic aging of human liver. *Proceedings of the National Academy of Sciences*, *111*(43), 15538–15543. doi:10.1073/pnas.1412759111

Hunter, R. G., McEwen, B. S., & Pfaff, D. W. (2013). Environmental stress and transposon transcription in the mammalian brain. *Mobile Genetic Elements*, *3*(2), 17657–17662. doi:10.4161/mge.2455

Issa, J. P. (2014). Aging and epigenetic drift: A vicious cycle. *The Journal of Clinical Investigation*, *124*(1), 24–29. doi:10.1172/JCI69735

Jacobs, K. B., Yeager, M., Zhou, W., Wacholder, S., Wang, Z., Rodriguez-Santiago, B., … Charnock, S. J. (2012). Detectable clonal mosaicism and its relationship to aging and cancer. *Nature Genetics*, *44*(6), 651–658. doi:10.1038/ng.2270

Jones, S. (2011). *The Darwin Archipelago: The naturalist's career beyond origin of species*. New Haven, CT: Yale University Press.

Kay, L. (2000). *Who wrote the book of life? A history of the genetic code*. Palo Alto, CA: Stanford University Press.

Kazazian, H., Haig, H., & Goodier, J. L. (2002). LINE drive: Retrotransposition and genome instability. *Cell*, *110*(3), 277–280. doi:10.1016/S0092-8674(02)00868-1

Keller, E. F. (1995). *Refiguring life: Metaphors of twentieth century biology*. New York, NY: Columbia University Press.

Keller, E. F. (2012). Genes, genomes, and genomics. *Biological Theory*, *6*, 132–140. doi:10.1007/s13752-012-0014-x

Keller, E. F. (2014). From gene action to reactive genomes. *Journal of Physiology*, *592*(11), 2423–2429. doi:10.1113/jphysiol.2014.270991

Kelly, S. E. (2012). The maternal-foetal interface and gestational chimerism: The emerging importance of chimeric bodies. *Science as Culture*, *21*(2), 233–257. doi:10.1080/09505431.2011.628014

Konrad, M. (2003). Predictive genetic testing and the making of the pre-symptomatic person: Prognostic moralities amongst Huntington's-affected families. *Anthropology & Medicine*, *10*(1), 23–49. doi:10.1080/13648470301269

Krumm, N., O'Roak, B. J., Shendure, J., & Eichler, E. E. (2014). A *de novo* convergence of autism genetics and molecular neuroscience. *Trends in Neurosciences*, *37*(2), 93–105. doi:10.1016/j.tins.2013.11.005

Landecker, H. (2007). *Culturing life: How cells became technologies*. Cambridge, MA: Harvard University Press.

Landecker, H. (2011). Food as exposure: Nutritional epigenetics and the new metabolism. *BioSocieties*, *6*(2), 167–194. doi:10.1057/biosoc.2011.1

Landecker, H. (2013). Post-industrial metabolism: Fat knowledge. *Public Culture*, *25*(3), 495–522. doi:10.1215/08992363-2144625

Lappé, M. (2014). Taking care: Anticipation, extraction, and the politics of temporality in autism science. *BioSocieties*, *9*(3), 304–328. doi:10.1057/biosoc.2014.14

Lappé, M., & Landecker, H. (2015). How the genome got a life span. *New Genetics and Society*, *34*(2), 152–176. doi:10.1080/14636778.2015.1034851.

LaSalle, J. M. (2011). A genomic point-of-view on environmental factors influencing the human brain methylome. *Epigenetics*, *6*(7), 862–869. doi:10.4161/epi.6.7.16353

Laurie, C. C., Laurie, C. A., Rice, K., Doheny, K. F., Zelnick, L. R., McHugh, C. P., … Freedman, N. D. (2012). Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nature Genetics*, *44*(6), 642–650. doi:10.1038/ng.2271

Lee, S. S. J. (2013). American DNA: The politics of potentiality in a genomic age. *Current Anthropology*., *54*(S7), S77–S86. doi:10.1086/670970

Lettre, G. (2014). Rare and low-frequency variants in human common diseases and other complex traits. *Journal of Medical Genetics*, *22*(R1), R16–R21. doi:10.1136/jmedgenet-2014-102437

Lippman, A. (1991). Prenatal genetic testing and screening: Constructing needs and reinforcing inequities. *American Journal of Law and Medicine*, *17*, 15–50.

Lock, M. (2001). The alienation of body tissue and the biopolitics of immortalized cell lines. *Body & Society*, *7*(2–3), 63–91. doi:10.1177/1357034X0100700204

Lock, M. (2011). Dementia entanglements in a postgenomic era. *Science, Technology, and Human Values*, *36*(5), 685–703. doi:10.1177/0162243910388022

Lupton, D. (1999). *Risk*. New York, NY: Rutledge.

Macosko, E. Z., & McCarroll, S. A. (2013). Our fallen genomes. *Science*, *342*(6158), 564–565. doi:10.1126/science.1246942

Malhotra, D., & Sebat, J. (2012). CNVs: Harbingers of a rare variant revolution in psychiatric genetics. *Cell*, *148*(6), 1223–1241. doi:10.1016/j.cell.2012.02.039

Marcus, G. (1995). Ethnography in/of the world systems: The emergence of multi-sited ethnography. *Annual Review of Anthropology*, *24*, 95–116.

Martin, A. (2010). Microchimerism in the mother(land): Blurring the boundaries of body and nation. *Body & Society*, *16*(3), 23–50. doi:10.1177/1357034X10373404

McConnell, M. J., Lindberg, M. R., Brennand, K. J., Piper, J. C., Voet, T., Cowing-Zitron, C., … Gage, F. H. (2013). Mosaic copy number variation in human neurons. *Science*, *342*(6158), 632–637. doi:10.1126/science.1243472

Meloni, M., & Testa, G. (2014). Scrutinizing the epigenetics revolution. *BioSocieties*, *9*(4), 431–456.

Miller, D. T., Adam, M. P., Aradhya, S., Biesecker, L. G., Brothman, A. R., Carter, N. P., … Waggoner, D. J. (2010). Consensus Statement: Chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *The American Journal of Human Genetics*, *86*(5), 749–764. Retrieved from doi:10.1016/j.ajhg.2010.04.006

Montoya, M. (2011). *Making the Mexican diabetic: Race, science, and the genetics of inequality*. Oakland, CA: University of California Press.

Navon, D. (2011). Genomic designation: How genetics can delineate new, phenotypically diffuse medical categories. *Social Studies of Science*, *41*(2), 203–226. doi:10.1177/0306312710391923

Navon, D., & Eyal, G. (2014). The trading zone of autism genetics: Examining the intersection of genomic and psychiatric classification. *BioSocieties*, *9*(3), 329–352. doi:10.1057/biosoc.2014.18

Navon, D., & Shwed, U. (2012). The chromosome 22q11. 2 deletion: From the unification of biomedical fields to a new kind of genetic condition. *Social Science & Medicine*, *75*(9), 1633–1641. doi:10.1016/j.socscimed.2012.06.023

Nelkin, D., & Lindee, M. S. (1996). *The DNA mystique: The gene as a cultural icon*. New York, NY: Henry Holt.

Nelson, A. (2008). Bio Science: Genetic ancestry testing and the pursuit of African ancestry. *Social Studies of Science*, *38*(5), 759–783. doi:10.1177/0306312708091929

Nerlich, B., Dingwall, R., & Clarke, D. (2002). The book of life: How the completion of the human genome project was revealed to the public. *Health: An Interdisciplinary Journal for the Social Study of Health, Illness, and Medicine*, *6*, 445–469. doi:10.1177/136345930200600403

O'Huallachain, M., Karczewski, K. J., Weissman, S. M., Urban, A. E., & Snyder, M. P. (2012). Extensive genetic variation in somatic human tissues. *Proceedings of the National Academy of Sciences*, *109*(44), 18018–18023. doi:10.1073/pnas.1213736109

O'Huallachain, M., Weissman, S. M., & Snyder, M. P. (2013). The variable somatic genome. *Cell Cycle*, *12*(1), 5. doi:10.4161/cc.23069

Perry, G. H., Dominy, N. J., Claw, K. G., Lee, A. S., Fiegler, H., Redon, R., … Stone, A. C. (2007). Diet and the evolution of human amylase gene copy number variation. *Nature Genetics*, *39*(10), 1256–1260. doi:10.1038/ng2123

Rapp, R. (1999). *Testing women, testing the fetus: The social impact of amniocentesis in America*. New York, NY: Routledge.

Reardon, J. (2009). *Race to the finish: Identity and governance in an age of genomics*. Princeton, NJ: Princeton University Press.

Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D., & Hurles, M. E. (2006). Global variation in copy number in the human genome. *Nature*, *444*(7118), 444–454. doi:10.1038/nature05329

Rusiecki, J. A., Chen, L., Srikantan, V., Zhang, L., Yan, L., Polin, M. L., & Baccarelli, A. (2012). DNA methylation in repetitive elements and post-traumatic stress disorder: A case-control study of US military service members. *Epigenomics*, *4*(1), 29–40. doi:10.2217/epi.11.116

Sebat, J. (2007). Major changes in our DNA lead to major changes in our thinking. *Nature Genetics*, *39*, S3–S5. doi:10.1038/ng2095

Sebat, J. (2012, October 12). *We all carry genomic mutations*. Interview with the Center for Applied Genomics. https://www.youtube.com/watch?v=OfHJvt_BRm8&index=1&list=PL_qu3jTHDyy1utRIKnzTsQSyMnT6QZnVN. Accessed on December 12, 2014.

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., & Wigler, M. (2007). Strong association of de novo copy number mutations with autism. *Science*, *316*(5823), 445–449. doi:10.1126/science.1138659

Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., & Wigler, M. (2004). Large-scale copy number polymorphism in the human genome. *Science*, *305*(5683), 525–528. doi:10.1126/science.1098918

Shim, J., Darling, K. W., Lappé, M. D., Thomson, L. K., Less, S. S., Hiatt, R. A., & Ackerman, S. L. (2014). Homogeneity and heterogeneity as situational properties: Producing – and moving beyond? – race in post-genomic science. *Social Studies of Science*, *44*(4), 579–599. doi:10.1177/0306312714531522

Shostak, S. (2013). *Exposed science: Genes, the environment, and the politics of population health*. Berkeley, CA: University of California Press.

Shostak, S., & Freese, J. (2010). Gene-Environment interaction and medical sociology. In C. E. Bird, P. Conrad, A. M. Fremont, & S. Timmermans (Eds.), *Handbook of medical sociology*. Nashville, TN: Vanderbilt University Press.

Stevens, H., & Richardson, S. (2015). *Postgenomics*. Durham, NC: Duke University Press.

Stotz, K. (2006). With genes like that, who needs an environment? Postgenomics' argument for the ontogeny of information. *Philosophy of Science*, *73*(5), 905–917.

Strauss, A. L. (1987). *Qualitative analysis for social scientists*. New York, NY: Cambridge University Press.

Sudmant, P. H., Kitzman, J. O., Antonacci, F., Alkan, C., Malig, M., Tsalenko, A., & Eichler, E. E. (2010). Diversity of human copy number variation and multicopy genes. *Science*, *330* (6004), 641–646. doi:10.1126/science.1197005

Timmermans, S. (2014). Trust in standards: Transitioning clinical exome sequencing from bench to bedside. *Social Studies of Science*, *45*(1), 77–99. doi:10.1177/0306312714559323

Timmermans, S., & Buchbinder, M. (2012). *Saving babies? The consequences of newborn genetic screening*. Chicago, IL: University of Chicago Press.

Vijg, J., & Suh, Y. (2013). Genome instability and aging. *Annual Review of Physiology*, *75*, 645–668. doi:10.1146/annurev-physiol-030212-183715

Vulto-van Silfhout, A. T., Hehir-Kwa, J. Y., Bon, B. W. M., Schuurs-Hoeijmakers, J. H. M., Meader, S., Hellebrekers, C. J. M., & de Vries, B. B. A. (2013). Clinical significance of *de novo* and inherited copy-number variation. *Human Mutation*, *34*(12), 1679–1687. doi:10.1002/humu.22442

Waterland, R. A., & Jirtle, R. L. (2003). Transposable elements: Targets for early nutritional effects on epigenetic gene regulation. *Molecular and Cellular Biology*, *23*(15), 5293–5300. doi:10.1128/MCB.23.15.5293-5300

Zimmer, C. (2013). DNA Double-Take. *The New York Times*, September 16.