



# Genes, behavior, and behavior genetics

Evan Charney\*

According to the ‘first law’ of behavior genetics, ‘All human behavioral traits are heritable.’ Accepting the validity of this first law and employing statistical methods, researchers within psychology, sociology, political science, economics, and business claim to have demonstrated that *all* the behaviors studied by their disciplines are heritable—no matter how culturally specific these behaviors appear to be. Further, in many cases they claim to have identified specific genes that play a role in those behaviors. The validity of behavior genetics as a discipline depends upon the validity of the research methods used to justify such claims. It also depends, foundationally, upon certain key assumptions concerning the relationship between genotype (one’s specific DNA sequences) and phenotype (any and all observable traits or characteristics). In this article, I examine—and find serious flaws with—both the methodologies of behavior genetics and the underlying assumptions concerning the genotype–phenotype relationship. © 2016 Wiley Periodicals, Inc.

## How to cite this article:

WIREs Cogn Sci 2016. doi: 10.1002/wcs.1405

## INTRODUCTION

As a method, behavior genetics is now used in a variety of disciplines including psychology, sociology, political science, economics, and business. What unites them is the employment of statistical methods in an attempt to find genetic influences in complex social behaviors. Inasmuch as training in statistics is a common component of advanced education in all of these disciplines, many behavioral and social scientists have adopted the methodologies characteristic of behavior genetics with little difficulty.

All such studies stem from the so-called ‘first law’ of behavior genetics, according to which ‘All human behavioral traits are heritable.’<sup>1</sup> Although the expression ‘human behavioral trait’ is not defined, behavior geneticists claim to have demonstrated the heritability of ‘behaviors’ as diverse as, e.g., texting on a cellphone (52% heritable<sup>2</sup>), being a Born Again

Christian (65% heritable<sup>3</sup>), and being supportive of property taxes (41% heritable<sup>4</sup>). Furthermore, they often claim to have identified, via ‘gene association’ studies, specific genes that underlie this heritability and can predict, e.g., voting in a presidential election,<sup>5</sup> creative dance performance,<sup>6</sup> and utilitarian moral judgments.<sup>7</sup>

Most such gene association studies, however, have suffered from failures of replication.<sup>8</sup> Despite the supposed heritability of everything that could conceivably be called ‘human behavior,’ researchers have been unable to identify the genes underlying this heritability.<sup>1</sup> An explanation recently offered for this failure is this: Since hundreds or thousands of genes influence behavior, most genes are of such tiny effect size that they cannot be identified according to the standards of statistical significance currently employed in gene association studies.<sup>9</sup> There is another available explanation, however: heritability estimates are wrong in the first place. Going even further, they are based upon a faulty paradigm.

## CLARIFYING HERITABILITY

It is important to understand that heritability estimates address *similarities/differences* in traits rather

\*Correspondence to: echar@duke.edu

Sanford School of Public Policy, Duke Center for Brain Sciences, Duke University, Durham, NC, USA

Conflict of interest: The author has declared no conflicts of interest for this article.

than the *causes* of traits themselves, and that the difference here is enormously consequential. The claim, e.g., that lifetime income is 50% heritable<sup>10</sup> is not intended to mean that 50% of a person's income is due to her 'genes' and 50% to her 'environment.' Heritability estimates only apply to populations, not to individuals, and the word 'heritability,' in this context, must not be confused with 'genetic inheritance.' Unfortunately, although 'correlation does not entail causation,'<sup>11</sup> heritability estimates are frequently misinterpreted as showing that genetic similarities *cause* trait similarities in the study population (see Moore and Shenk, The heritability fallacy, *WIREs Cogn Sci*, also in the collection *How We Develop*).

## TWIN STUDIES

Most estimates of heritability in behavior genetics are derived using the 'classical' twin study methodology, based on the assumption that so-called 'identical' or monozygotic (MZ) twins, derived from the division of a single egg and sperm, have identical inherited DNA, whereas 'fraternal' or dizygotic (DZ) twins, derived from two separate egg and sperm, are, on average, 50% genetically identical, the same as non-twin siblings. If, in a study population composed of MZ and DZ twin pairs, MZ twin pairs exhibit a certain behavior to a greater extent than DZ twin pairs, this greater behavioral similarity is ascribed to the greater genetic similarity of MZ twin pairs and is used to derive a heritability estimate.<sup>12</sup> Widely available behavioral datasets, such as the National Longitudinal Study of Adolescent Health (AddHealth) have greatly facilitated the ease of deriving heritability estimates using the twin methodology. AddHealth is a large longitudinal study that consists of responses of thousands of adolescents and young adults to thousands of questions on every aspect of their behavior, and includes responses from ~1000 pairs of MZ and DZ twins as well as zygosity information (information as to whether the twins are MZ or DZ).<sup>13</sup> In a representative study using this dataset, several political scientists used the differences in the responses of MZ and DZ twin pairs to the question 'Did you vote in the most recent presidential election?' to conclude that voting was 72% heritable.<sup>14</sup>

Many, though not all researchers in BG who conduct twin studies, acknowledge that the 'environments' of MZ twins (e.g., how they are raised, how they are treated by others) are more alike than those of DZ twins. This could be a significant flaw in heritability estimates, since the greater similarities of MZ twins, ascribed to their greater genetic similarity, could be

due to greater environmental similarity, resulting in inflated or spurious heritability estimates.<sup>15</sup> However, according to the so-called equal environment assumption (EEA), the greater environmental similarities of MZ twins *do not make them more behaviorally alike than DZ twins*. Defenders of the EEA point to a series of influential studies involving small numbers of MZ and DZ twins ostensibly raised apart, the idea being that since they were raised in different environments, any possible confounding effects due to their shared rearing environments would be eliminated.<sup>16</sup> The authors of these studies reported that heritability estimates for several different behavioral traits of twins raised apart and twins raised together were basically the same.

Taking this as sufficient proof of the general validity of the EEA, most behavior geneticists rely upon populations of twins raised together. None of the studies of twins raised apart, however, involved only twins separated at birth. In some cases, twins as old as age *nine* were treated as raised apart for the purposes of the study, meaning that they spent the nine most formative years of their lives in the same environment.<sup>17</sup> In none of the studies of twins raised apart was it the case that the twin pairs were unaware of each other prior to being studied. In some instances, they had ongoing relationships prior to being studied lasting over 20 years, and since MZ twins are known to have closer relationships than DZ twins, this amounts to 20 years during which MZ twins could influence each other's behaviors to a greater extent than DZ twins.<sup>18</sup>

A separate serious concern is the shared womb environment. No study of twins raised apart, even if all of the twins were separated at birth and had no contact prior to being studied, could correct for the confounds due to a shared prenatal environment—that environment being universally recognized as being of profound importance. The prenatal environment is known to have an enormous effect on a wide range of developmental outcomes.<sup>19</sup> For example, fetal exposure to elevated gestational stress can affect infant and childhood neurological development. The prenatal environment of MZ twins is significantly more stressful than that of DZ twins, particularly in the case of the two thirds of MZ twins that share a single placenta and outermost fetal membrane (so-called monochorionic [MC] MZ twins<sup>20</sup>). MC-MZ twins, as compared to DZ twins, exhibit a significantly higher rate of early neonatal mortality, fetal growth restriction, preterm delivery, congenital abnormalities, and low birth rate. Low birth rate is one of the strongest predictors of negative developmental outcomes.<sup>21</sup> Hence, MZ twins, to a greater extent than DZ twins,

share certain developmental outcomes associated with their unique shared prenatal environment. Prenatally, the EEA clearly does not hold.

## GENE ASSOCIATION STUDIES

Despite these serious flaws, faith in the validity of the twin study methodology and the apparent confirmation of the first law of behavior genetics has prompted the search for the genes underlying all of this ostensible heritability.

Persons typically inherit 23 pairs of chromosomes, one copy of each of the 23 pairs of chromosomes is inherited from the mother and one copy from the father. While matched pairs of chromosomes contain the same genes (the exception being the X and Y chromosomes in males), they can exhibit various differences in their DNA sequences. For example, the DNA in a chromosome inherited from one parent might differ in various ways from the DNA in the matching chromosome inherited from the other parent. For example, it might have a different base pair in a certain location or repeated segments of varying lengths in another location. When a specific variation in DNA sequence in a particular location on a chromosome occurs in >1% of the population, it is referred to as a 'polymorphism.' Polymorphisms can be found at any position on the DNA molecule, including regions not transcribed to produce proteins, i.e., polymorphisms can be found in regions that lie outside of genes. These regions often have important regulatory functions for gene transcription.

The search for genes underlying the presumed heritability of human behaviors has essentially been the search for polymorphisms associated with behaviors. The finding of an association may be deemed 'causal,' e.g., it is assumed that the polymorphism results in a difference in the rate of transcription of a given gene, and the gene (or protein transcribed from the gene) itself is believed to play a role in the behavior. A polymorphism may also be deemed not to play a direct role in the behavior, but to lie on a region of the chromosome close to the 'real' causal gene.

One method for attempting to identify these polymorphisms is the candidate gene association (CGA) study. In a typical CGA study, a researcher proposes that those individuals with a given polymorphism are more likely to exhibit a given behavior. CGA studies have typically focused upon a small number of polymorphisms on the same regions of a small number of genes (e.g., MAOA, 5-HTT, DRD2, and DRD4) for two reasons: (1) These genes are

transcribed to produce proteins involved in the regulation of neurotransmitters and thus are believed to be important for human behavior; and (2) researchers believed, on the basis of mouse and *in vitro* experiments, that they could associate polymorphic differences in regulatory regions of these genes with differences in the level of certain neurotransmitters and thereby with differences in behavior.<sup>22</sup> In fact, attempts to associate these polymorphisms with actual differences in neurotransmitters in the human brain under physiological conditions have been largely unsuccessful.<sup>23,24</sup> Nonetheless, this did not stop researchers from proposing, and claiming to have demonstrated, associations between the same polymorphisms of these neurotransmitter-related genes and a bewildering array of different behaviors.

As with twin studies, CGA studies have been greatly facilitated by the use of large behavioral datasets. For example, the aforementioned AddHealth study, in addition to containing information on the zygosity status and questionnaire responses of ~1000 twin pairs, also contained limited genotype data for four neurotransmitter-related genes for a subset of the study population. As noted, several political scientists used the Add Health dataset to estimate the heritability of voting on the basis of MZ and DZ twin pair responses to the question, 'Did you vote in the last presidential election?' In a representative CGA study, these same political scientists used responses to the same question to identify two polymorphisms they claimed were associated with voting.<sup>5</sup> They reported that those who had a certain polymorphism of one gene (MAOA) answered 'yes' to this question more frequently than those who did not have this polymorphism. They also reported that those who were religious (on the basis of their answers to the question how often they attended religious services in the past 12 months) and possessed a polymorphism of another gene (5-HTT) were more likely to say that they had voted. On this basis, they concluded that 'two genes predict voter turnout.' This finding failed to be replicated.<sup>22</sup>

Because behavior geneticists the world over were using the same handful of genes to conduct CGA studies on hundreds of different behaviors, a situation arose in which each of these genes was reported to 'predict,' or be associated with, hundreds of different behaviors.<sup>22</sup> For example, polymorphisms on the same regulatory region of the 5-HTT gene that were reported to predict voting have also been associated with over one hundred different behaviors including agreeableness, alcoholism, anorexia nervosa, behavior in romantic relationships, bipolar disorder, creative dance performance,

depression, gambling, gang membership, intelligence, job satisfaction, obesity, optimism, persistence, premature ejaculation, schizophrenia, seasonal affective disorder, self-esteem, sexual novelty, shyness, and utilitarian moral judgments.<sup>25</sup>

Searching for correlations between a single polymorphism and hundreds of different behaviors is very likely, statistically, to generate numerous false positives.<sup>22</sup> This likelihood increases when a polymorphism is reported to be correlated with a behavior only when combined with a certain 'environment' or another behavior. As noted, a polymorphism of the 5-HTT gene was reported to be correlated with voting behavior, *but only for those who indicated they were religious*. Being religious (as measured by self-reported church attendance), however, is known to be correlated with voting. A different polymorphism of the same region of the same gene has also been reported to be correlated with depression, *but only for those who suffer stressful life events*. Stressful life events, however, 'have a potent relationship with the risk of depression, an association that has been one of the most widely studied environmental factors for a range of mental disorders.' Hence, it is not surprising that CGA studies have, without exception, been plagued by failures of replication. In 2010, the editor of the journal *Behavior Genetics* wrote: 'The literature on candidate gene associations is full of reports that have not stood up to rigorous replication ... it now seems likely that many of the published findings of the last decade are wrong or misleading and have not contributed to real advances in knowledge.'<sup>8</sup> Despite this, CGA studies, including ones involving the same polymorphic regions of the same four genes mentioned above, continue to be published.

If CGA studies have been plagued by *too many* associations (that have failed to replicate), genome wide association (GWA) studies have been plagued by the opposite: a pervasive failure to find any associations. Unlike CGA studies, GWA studies do not start with a given polymorphism and then look for a correlation between that polymorphism and a given behavior. Instead, large segments of the genome (a million or more base pairs) of typically thousands of subjects are investigated in an attempt to find a polymorphism that *cases* (those who exhibit the behavior of interest) have to a greater extent than *controls* (those who do not exhibit the behavior). Most GWA studies involve the search for single nucleotide polymorphisms (SNPs), a polymorphism involving the substitution of a single base pair in a given position on the DNA molecule. GWA studies have been largely unsuccessful, leading to the so-

called 'problem of missing heritability': Despite the supposed heritability of all human behavior, behavior geneticists have been unable to find any substantial associations between polymorphisms and behaviors.<sup>9</sup>

One recent explanation for the problem of missing heritability depends upon the resurrection of the 'infinitesimal model,' proposed by Fisher<sup>26</sup> nearly 100 years ago: Hundreds or thousands or even millions of polymorphisms underlie trait heritability. Each polymorphism, it is now claimed, contributes such a tiny amount to the heritability of a trait that it cannot be identified by a standard GWA study.<sup>27</sup> Standard GWA studies involve the search for a single polymorphism (typically a SNP) that those who exhibit a certain behavior (cases) possess to a greater degree than those who do not exhibit this behavior (controls), and require a strict standard of statistical significance to assure that any association is not a false positive. On the basis of the aforementioned assumption, behavior geneticists have significantly relaxed this standard. Rather than searching for a single SNP, all of the SNPs that (according to this relaxed statistical standard) cases possess to a greater extent than controls are added together. What is supposed to be statistically significant is the 'aggregate effect' of all of these individually insignificant polymorphisms.<sup>28</sup>

These assumptions underlie two methodologies in behavior genetics: (1) genome wide complex trait analysis (GCTA) is supposed to show, on the basis of whole genome scans of thousands of unrelated cases and controls, what percentage of the heritability of a trait can be accounted for by SNPs; and (2) all of the SNPs that cases have to a greater degree than controls (according to the aforementioned relaxed statistical standard) are used to derive so-called 'genome wide polygenic risk scores.' Neither method is meant to actually identify the specific SNPs involved (although some variations of these methodologies do attempt to identify specific SNPs). Working with these assumptions, behavior geneticists have gone from claiming that a single polymorphism could predict, e.g., whether an individual will make utilitarian moral judgments<sup>7</sup> to the claim that over *nine million* SNPs together account for only .43% (less than half) of the estimated heritability of 'educational attainment.'<sup>29</sup>

The claim that over nine million SNPs account for under half the heritability of 'educational attainment' lacks any independent empirical justification beyond the assumptions that the infinitesimal model is true and that genome wide heritability estimates and risk scores are valid. Adding together an untold

number of SNPs that individually fail to achieve statistical significance by the acknowledged standards of GWA studies is highly problematic. The expression ‘add together’ here is intended literally, since most GWA studies of heritability rely, in part or entirely, on an assumption of ‘additivity.’ For example, the claim that 8000 SNPs account for half the heritability of schizophrenia<sup>30</sup> depends upon the assumption that each SNP contributes 1/8000 to half the heritability. The assumption of additivity is used despite this candid acknowledgement by two researchers:<sup>31</sup>

We know this assumption of additive contribution is probably wrong. Epistasis—interaction between genetic loci—is important and pervasive. Unfortunately, we know very little about epistatic interactions. Given [that] the specifics of the interactions are unknown, additivity provides a workable default.

It is not clear, however, what a ‘workable default’ means in this context, or how or why one should maintain an assumption that is at once fundamental and wrong.

One known cause of false positives in GWA studies is population stratification,<sup>32</sup> the phenomenon that frequencies of particular polymorphisms can differ in different ethnic populations due to unique ancestral patterns of migration and mating practices.<sup>33</sup> Even populations that are considered relatively ethnically homogeneous, such as Swedes or Icelanders, are characterized by hidden or ‘cryptic’ population stratification.<sup>34,35</sup> In the aforementioned study in which researchers claimed that over eight million SNPs account for half the heritability of educational attainment, they also noted that 74 out of these 8 million SNPs showed a ‘stronger’ association. However, these 74 SNPs could, e.g., be more prevalent in a subsection of the population, e.g., Swedes from southern as opposed to northern Sweden, that holds a privileged socio-economic status in society, or culturally places a premium on higher education, and hence are more likely to exhibit higher ‘educational attainment.’ The genetics community has struggled to develop statistical techniques to guard against such spurious associations in GWA studies due to population stratification.

These techniques were developed for GWA studies that involve the search for a *single* polymorphism. Heritability estimates and risk scores derived from GWA studies add together the presumed effects of hundreds to millions of polymorphisms. Not surprisingly, these methods have been shown to be particularly vulnerable to population stratification.<sup>36–38</sup> Researchers have recently argued that, even assuming

the complete absence of population stratification (something that exists in no real population), GWA heritability estimates fail to produce valid results.<sup>39</sup>

## THE DIVERSE AND CHANGING GENOME

The focus of research in behavioral genetics is the nuclear genome, the DNA contained within the nucleus of the cell. The nuclear genome excludes the DNA located in mitochondria (mitochondrial or mtDNA). Both mitochondria and their DNA are inherited solely from the mother; thousands of copies of mtDNA exist per cell and the number of copies varies significantly by cell type; and mtDNA differs in different cells not only in term of copies, but also in terms of DNA sequence.<sup>40</sup> In addition to nuclear DNA and mtDNA, the nuclei of cells also contain tens of thousands of short extrachromosomal circular DNA molecules of varying sizes and sequences, but their manner of formation and function are unknown.<sup>41</sup>

Because of the manner in which mtDNA is inherited and because of the variation in amount and sequence of mtDNA in cells throughout the human body, MZ twins are not ‘genetically identical’ as far as their mtDNA is concerned, and individuals do not possess identical mtDNA in all the cells of their bodies.<sup>42</sup> It has been assumed, however, that MZ twins have identical nuclear DNA. Furthermore, all the cells of our bodies were assumed to have identical nuclear DNA (with the exception of egg and sperm cells, red blood cells, and certain cells in the immune system), an assumption that has allowed researchers to assume that the nuclear genome derived from a cheek swab or blood plasma would be identical in all of our cells. We now know, however, that the nuclear genomes within all of our cells are not identical.

One way in which the nuclear genomes in our cells differ is due to copy number variations (CNVs), variations in DNA sequence ranging in size from thousands to millions of DNA bases, involving deletions, insertions, duplications, and inversions.<sup>43</sup> CNVs can be inherited, but they can also occur post-conception. CNVs can be large enough to include whole genes, resulting in multiple copies of a gene or its complete deletion. In other words, there may be multiple copies, or no copies, of a gene in the nuclear genome of some of our neurons, but not others, or in our neurons, but not in blood plasma or cheek skin cells, and genomic differences are believed to be widespread in healthy individuals, particularly in the

brain.<sup>44,45</sup> Since the nuclear genomes in the cells of the same individual differ, the nuclear genomes of MZ twin pairs (who start out as the same individual) differ as well.<sup>46</sup> MZ twins are not 100% genetically identical.

## BEYOND THE DNA SEQUENCE

The exclusive focus on differences in DNA sequences (specifically, nuclear DNA sequences) as a way to understand differences in complex human social behaviors overlooks a key problem: Humans contain significantly fewer genes than wheat, ~20,000 as compared to ~100,000, and about the same number of genes as nematodes and sponges.<sup>47</sup> As far as overall genome length is concerned (which includes regions not transcribed to produce proteins but which play a role in gene regulation) the wheat genome contains over five times as many base pairs as the human genome. The genome of the flowering plant *Paris japonica* is fifty times larger than the human genome, the largest known genome in nature.<sup>48</sup> Many of the ~20,000 human genes are shared with other species. A homolog is a gene similar in structure and origin to a gene in another species and is transcribed to produce a similar (or the same) protein.<sup>49</sup> The MAOA gene, e.g., polymorphisms of which have been identified with voting behavior,<sup>5</sup> has a homolog in nematodes. 5HTT, which has been associated with utilitarian moral reasoning,<sup>7</sup> has homologs in zebrafish and frogs. Surely, if differences in these DNA sequences are not key to distinguishing humans from other species, they are not key to distinguishing, e.g., voters from non-voters. It might be objected that the human MAOA gene, e.g., acts very differently in a human than does its homolog in a frog. And indeed it does, but this is precisely the point. The differences in how homologous genes 'act' (how they are transcribed and translated) in any two species is not a matter of differences in their DNA sequence.

Although the human genome in its entirety is, of course, unique to our species, it is becoming increasingly clear that what makes the human genome a *human* genome, as opposed to a nematode genome, is something more than the DNA sequence. Of critical importance is also the manner in which DNA is transcribed and the manner in which the products of this transcription are translated into RNA molecules and proteins. To be sure, sometimes, differences in DNA sequence alone can tell us something about how a given sequence of DNA will be transcribed and/or translated. If an individual inherits a mutated copy of the HTT gene, an abnormal

huntingtin protein will be transcribed from this gene, resulting in the degenerative neurological disorder Huntington's disease. And sometimes, polymorphisms can be bona fide risk factors, although the physiological pathways are not well understood: a polymorphism of the ApoE gene is associated with an increased risk of Alzheimer's disease.<sup>50</sup> But examples such as this seem truly the exception.

Consider a key example of what differences in DNA sequence alone cannot tell us about phenotypic differences. A keratinocyte (a type of skin cell) is dramatically different from a retinal ganglion cell (a type of neuron) in both morphology and function. What differentiates these cells are not differences in their DNA sequence but differences in the manner in which these DNA sequences are transcribed and the products of transcription synthesized into proteins and RNA molecules.<sup>51</sup> Some genes transcribed in a retinal ganglion cell are 'turned off' in a keratinocyte, and vice versa. Gene transcription is regulated, in part, by the larger molecular structure of which the DNA molecule is only a part, known as chromatin. We inherit chromatin, of which the DNA molecule is a part, in the form of chromosomes.

Changes in chromatin structure can change the extent to which any given gene can be transcribed. Different types of cells and tissues are characterized by differences in their chromatin, which allows certain genes to be transcribed in some cells and not others. They are also characterized by differences in specialized proteins that are essential for transcription known as transcription factors; differences in the location and amount of special methyl compounds that are attached at various sites on the DNA and which, like chromatin, regulate transcription; differences in their non-coding RNAs, which also regulate transcription, and much more.

The key point is that on the basis of their DNA sequences alone, we would not be able to tell the difference between a keratinocyte and a retinal ganglion cell, or between any of the hundreds of other types of cells in the human body.

Furthermore, via a process known as alternative splicing, the transcription of a given gene in a retinal ganglion cell may result in a different protein, with very different physiological effects, than transcription of that same gene in a keratinocyte, and different proteins may be transcribed from these same genes in both cell types at different times or in response to different stimuli.<sup>52</sup> Alternative splicing entails that, normally, DNA sequence alone cannot tell us the specific form of a protein that will be transcribed from a given DNA sequence at any given time. Alternative splicing occurs in ~95% of human

genes and can result in numerous proteins being synthesized from the same gene, greatly expanding protein diversity. For example, three genes are transcribed to synthesize the neurexin proteins which help to connect neurons at their synapses, but there are well over 1000 alternatively spliced forms of this protein.<sup>53</sup> Proteins that are alternatively spliced from the same gene can have radically different and even opposing physiological effects. The spatial and temporal control of alternative splicing is a major mechanism used to generate protein diversity in the brain and is critical for normal development.<sup>54</sup>

Although our DNA sequences may differ, what differentiates a human from a nematode or a sponge is clearly not just differences in these sequences. Part of the difference lies in the different ways in which these sequences are transcribed and translated. To understand this, we must look beyond DNA sequence to the larger molecular structures that play a key role in this process—and we cannot truly understand DNA's role apart from the larger nuclear and cellular environment in which these structures are located. It is an anachronistic form of DNA sequence reductionism that leads to the belief that differences in the DNA sequence *must* be the ultimate reason why humans conceive humans and not sponges. We do not inherit simply DNA molecules. We inherit DNA as part of the larger molecular protein–nucleic acid structure of chromatin in the form of chromosomes. Nor do we inherit simply chromatin molecules. We begin life as a fully differentiated human cell—a fertilized egg cell—a cell that contains mitochondria, mitochondrial DNA, proteins, messenger RNAs, non-coding RNAs (involved in regulating transcription and translation), and all manner of other macromolecular structures. The zygote does not develop in a vacuum but in the unique environment of the human womb, which plays an intimate role in

every aspect of prenatal development, including changing the chromatin structure and gene transcription and translation of the developing embryo.<sup>51</sup>

## CONCLUSION

In 1694, Nicolaas Hartsoeker produced an image of a tiny human, a homunculus, curled up inside a sperm cell. This famous image embodied the theory of 'preformationism': A human develops from a tiny, preformed human that already contains, in miniature form, all the characteristics of an adult human. Historically, 'epigenesis' represented the opposing view: Humans do not develop from little preformed humans; rather, development is a process that gives rise to new and emergent structures and processes (see Lickliter, Developmental evolution, *WIREs Cogn Sci*, also in the collection *How We Develop*). We may smile at the naiveté of the homunculus but for the fact that, today, this homunculus has been resurrected in the form of the genome—not the genome of science but a supernatural 'homunculus-genome.'

We have entered a 'post-genomics' era, an expression widely used by geneticists themselves.<sup>55</sup> While these developments might matter less for 'monogenic' inherited disorders such as Huntington's disease, they are clearly important for more complex disorders such as Type I diabetes.<sup>56</sup> What, then, of the genetic connections to voting in a presidential election, creative dance performance, and utilitarian moral judgments?

It is a continuing goal of behavior geneticists to establish a firm connection between genetics and much of human behavior. But their current methods and underlying assumptions are so deeply flawed as to raise more questions than answers.

## FURTHER READING

Jablonka E, Lamb MJ. *Evolution in Four Dimensions*. revised ed. Cambridge, MA: MIT Press; 2014.

Krimsky S, Gruber J. *Genetic Explanations*. Cambridge, MA: Harvard University Press; 2013.

Gilbert SF, Epel D. *Ecological Developmental Biology*. Sinauer Associates: Sunderland, MA; 2015.

## REFERENCES

1. Turkheimer E. Three laws of behavior genetics and what they mean. *Curr Dir Psychol Sci* 2000, 9:160–164.
2. Miller G, Zhu G, Wright MJ, Hansell NK, Martin NG. The heritability and genetic correlates of mobile phone use: a twin study of consumer behavior. *Twin Res Hum Genet* 2012, 15:97–106.
3. D'Onofrio BM, Eaves LJ, Murrelle L, Maes HH, Spilka B. Understanding biological and social

- influences on religious affiliation, attitudes, and behaviors: a behavior genetic perspective. *J Pers* 1999, 67:953–984.
4. Alford JR, Funck CL, Hibbing JR. Are political orientations genetically transmitted? *Am Polit Sci Rev* 2005, 99:153–167.
  5. Fowler JS, Dawes CT. Two genes predict voter turnout. *J Polit* 2008, 70:579–594.
  6. Bachner-Melman R, Dina C, Zohar AH, Constantini N, Lerer E, Hoch S, Sella S, Nemanov L, Gritsenko I, Lichtenberg P, et al. AVPR1a and SLC6A4 gene polymorphisms are associated with creative dance performance. *PLoS Genet* 2005, 1:e42.
  7. Marsh AA, Crowe SL, Yu HH, Gorodetsky EK, Goldman D, Blair RJ. Serotonin transporter genotype (5-HTTLPR) predicts utilitarian moral judgments. *PLoS One* 2011, 6:e25148.
  8. Hewitt JK. Editorial policy on candidate gene association and candidate gene-by-environment interaction studies of complex traits. *Behav Genet* 2012, 42:1–2.
  9. Turkheimer E. Still missing. *Res Hum Dev* 2011, 8:227–241.
  10. Benjamin DJ, Cesarini D, Chabris CF, Glaeser EL, Laibson DI, Guethnason V, Harris TB, Launer LJ, Purcell S, Smith AV, et al. The promises and pitfalls of geneoconomics\*. *Annu Rev Econom* 2012, 4:627–662.
  11. Aldrich J. Correlations genuine and spurious in Pearson and Yule. *Stat Sci* 1995, 10:364–376.
  12. Plomin R, DeFries JC, McClearn GM, McGuffin P. *Behavioral Genetics*. New York: Worth; 2005.
  13. Harris KM, Halpern CT, Smolen A, Haberstick BC. The national longitudinal study of adolescent health (add health) twin data. *Twin Res Hum Genet* 2006, 9:988–997.
  14. Fowler JH, Baker LA, Dawes CT. Genetic variation in political participation. *Am Polit Sci Rev* 2008, 102:233–248.
  15. Kendler KS. Twin studies of psychiatric illness. Current status and future directions. *Arch Gen Psychiatry* 1993, 50:905–915.
  16. Bouchard TJ Jr, Lykken DT, McGue M, Segal NL, Tellegen A. Sources of human psychological differences: the Minnesota Study of Twins Reared Apart. *Science* 1990, 250:223–228.
  17. Kamin LJ, Goldberger AS. Twin studies in behavioral research: a skeptical view. *Theor Popul Biol* 2002, 61:83–95.
  18. Joseph J. *The Gene Illusion*. New York: Algora; 2004.
  19. Sandman CA, Davis EP, Buss C, Glynn LM. Prenatal programming of human neurological function. *Int J Pept* 2011, 2011:837596.
  20. Corsello G, Piro E. The world of twins: an update. *J Matern Fetal Neonatal Med* 2010, 23(Suppl 3):59–62.
  21. Poulsen P, Wojtaszewski JF, Richter EA, Beck-Nielsen H, Vaag A. Low birth weight and zygosity status is associated with defective muscle glycogen and glycogen synthase regulation in elderly twins. *Diabetes* 2007, 56:2710–2714.
  22. Charney E, English W. Candidate genes and political behavior. *Am Polit Sci Rev* 2012, 106:1–34.
  23. Alia-Klein N, Kriplani A, Pradhan K, Ma JY, Logan J, Williams B, Craig IW, Telang F, Tomasi D, Goldstein RZ, et al. The MAO-A genotype does not modulate resting brain metabolism in adults. *Psychiatry Res* 2008, 164:73–76.
  24. Nordquist N, Orelund L. Serotonin, genetic variability, behaviour, and psychiatric disorders – a review. *Ups J Med Sci* 2010, 115:2–10.
  25. Charney E. Four genes predict everything. Available at: <http://sites.duke.edu/evancharney/files/2014/12/4-GENE-PREDICT-EVERYTHING-12-1.pdf>.
  26. Fisher R. The correlation between relatives on the supposition of Mendelian inheritance. *Trans R Soc Edin* 1918, 52:399–433.
  27. Chabris CF, Lee JJ, Cesarini D, Benjamin DJ, Laibson DI. The fourth law of behavior genetics. *Curr Dir Psychol Sci* 2015, 24:304–312.
  28. Benyamin B, Pourcain B, Davis OS, Davies G, Hansell NK, Brion MJ, Kirkpatrick RM, Cents RA, Franic S, Miller MB, et al. Childhood intelligence is heritable, highly polygenic and associated with FBNP1L. *Mol Psychiatry* 2014, 19:253–258.
  29. Oakbay A, Beuchamp JP, Fontana MA, Lee JJ, Peters TH, Reitvald CA, Turley P, Chen GB, Emilsson V, Meddens SF, et al. Genome wide association study identifies 74 loci associated with educational attainment. *Nature* 2016, 533:539–542.
  30. Lee SH, DeCandia TR, Ripke S, Yang J, Schizophrenia Psychiatric Genome-Wide Association Study C, International Schizophrenia C, Molecular Genetics of Schizophrenia C, Sullivan PF, Goddard ME, Keller MC, et al. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat Genet* 2012, 44:247–250.
  31. Belsky DW, Israel S. Integrating genetics and social science: genetic risk scores. *Biodemography Soc Biol* 2014, 60:137–155.
  32. Liu L, Zhang D, Liu H, Arendt C. Robust methods for population stratification in genome wide association studies. *BMC Bioinformatics* 2013, 14:132.
  33. Hunley KL, Healy ME, Long JC. The global pattern of gene identity variation reveals a history of long-range migrations, bottlenecks, and local mate exchange: implications for biological race. *Am J Phys Anthropol* 2009, 139:35–46.
  34. Humphreys K, Grankvist A, Leu M, Hall P, Liu J, Ripatti S, Rehnstrom K, Groop L, Klareskog L, Ding



- B, et al. The genetic structure of the Swedish population. *PLoS One* 2011, 6:e22547.
35. Salmela E, Lappalainen T, Liu J, Sistonen P, Andersen PM, Schreiber S, Savontaus ML, Czene K, Lahermo P, Hall P, et al. Swedish population substructure revealed by genome-wide single nucleotide polymorphism data. *PLoS One* 2011, 6:e16747.
  36. Browning SR, Browning BL. Population structure can inflate SNP-based heritability estimates. *Am J Hum Genet* 2011, 89:191–193; author reply 193–195.
  37. Janss L, de Los CG, Sheehan N, Sorensen D. Inferences from genomic models in stratified populations. *Genetics* 2012, 192:693–704.
  38. Kumar SK, Feldman MW, Rehkopf DH, Tuljapurkar S. Response to commentary on ‘limitations of GCTA as a solution to the missing heritability problem. *BioRxiv* 2016. Preprint, Posted February 17, 2016. Available at: <http://biorxiv.org/content/early/2016/02/17/039594>.
  39. Krishna Kumar S, Feldman MW, Rehkopf DH, Tuljapurkar S. Limitations of GCTA as a solution to the missing heritability problem. *Proc Natl Acad Sci USA* 2016, 113:E61–70.
  40. Clay Montier LL, Deng JJ, Bai Y. Number matters: control of mammalian mitochondrial DNA copy number. *J Genet Genomics* 2009, 36:125–131.
  41. Shibata Y, Kumar P, Layer R, Willcox S, Gagan JR, Griffith JD, Dutta A. Extrachromosomal microDNAs and chromosomal microdeletions in normal tissues. *Science* 2012, 336:82–86.
  42. Mattias A, Silva S, Martins Y, Blickstein I. Monozygotic twins: ten reasons to be different. *Diagn Prenat* 2014, 25:53–57.
  43. Almal SH, Padh H. Implications of gene copy-number variation in health and diseases. *J Hum Genet* 2012, 57:6–13.
  44. Bushman DM, Chun J. The genomically mosaic brain: aneuploidy and more in neural diversity and disease. *Semin Cell Dev Biol* 2013, 24:357–369.
  45. Cai X, Evrony GD, Lehmann HS, Elhosary PC, Mehta BK, Poduri A, Walsh CA. Single-cell, genome-wide sequencing identifies clonal somatic copy-number variation in the human brain. *Cell Rep* 2014, 8:1280–1289.
  46. Bruder CE, Piotrowski A, Gijsbers AA, Andersson R, Erickson S, Diaz de Stahl T, Menzel U, Sandgren J, von Tell D, Poplawski A, et al. Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *Am J Hum Genet* 2008, 82:763–771.
  47. Eddy SR. The C-value paradox, junk DNA and ENCODE. *Curr Biol* 2012, 22:R898–R899.
  48. Pellicer J, Fay MF, Leitch IJ. The largest eukaryotic genome of them all? *Bot J Linn Soc* 2010, 164:10–15.
  49. Grotewiel M, Bettinger JC. *Drosophila* and *Caenorhabditis elegans* as discovery platforms for genes involved in human alcohol use disorder. *Alcohol Clin Exp Res* 2015, 39:1292–1311.
  50. Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms, and therapy. *Nat Rev Neurol* 2013, 9:106–118.
  51. Rosenfeld CS. *The Epigenetic Origins of Health and Disease*. London: Elsevier; 2016.
  52. Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. *Nature* 2010, 463:457–463.
  53. Treutlein B, Gokce O, Quake SR, Sudhof TC. Cartography of neurexin alternative splicing mapped by single-molecule long-read mRNA sequencing. *Proc Natl Acad Sci USA* 2014, 111:E1291–1299.
  54. Grabowski P. Alternative splicing takes shape during neuronal development. *Curr Opin Genet Dev* 2011, 21:388–394.
  55. Simons RL, Klopach ET. Invited address: “the times they are a-changin’” gene expression, neuroplasticity, and developmental research. *J Youth Adolesc* 2015, 44:573–580.
  56. Stankov K, Benc D, Draskovic D. Genetic and epigenetic factors in etiology of diabetes mellitus type 1. *Pediatrics* 2013, 132:1112–1122.