

It's the End of the Gene As We Know It

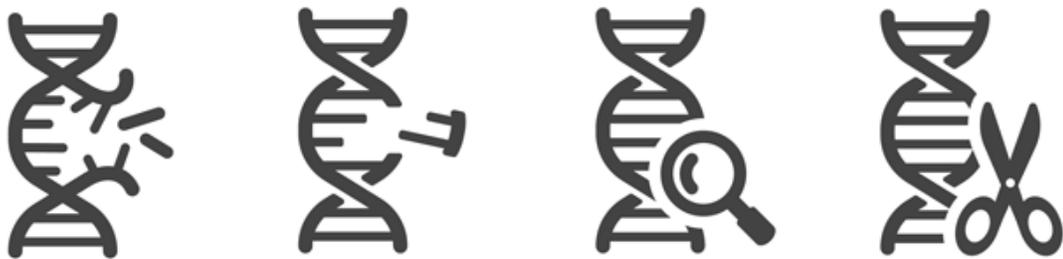
nautil.us/issue/68/context/its-the-end-of-the-gene-as-we-know-it

By Ken Richardson January 3, 2019

We've all seen the stark headlines: "Being Rich and Successful Is in Your DNA" (*Guardian*, July 12); "A New Genetic Test Could Help Determine Children's Success" (*Newsweek*, July 10); "Our Fortunetelling Genes" make us (*Wall Street Journal*, Nov. 16); and so on.

The problem is, many of these headlines are not discussing real genes at all, but a crude statistical model of them, involving dozens of unlikely assumptions. Now, slowly but surely, that whole conceptual model of the gene is being challenged.

We have reached peak gene, and passed it.



It is, of course, an impressive story. Today, most people know about Gregor Mendel's breeding experiments with pea plants in the 1850s. He concentrated on simple traits with well-defined, easy to count variations: purple or white flowers; long or short stems; smooth or wrinkled seeds; and so on. After cross-fertilization the patterns of variation in offspring suggested correlations with variation in single "heredity units."

Mendel's inherited factors—hitherto imputed, but unidentified—are what came to be called the genes. In the early 1900s, it was tempting to equate them with the information and instructions for the comprehensive development of the whole offspring, mental and physical.

In a famous paper in 1911, Wilhelm Johannsen warned against doing that. We do not know, he said, how those inferred, but invisible, factors can possibly carry such complex information. But Johannsen was ignored, for reasons, as it turned out, more to do with ideology than biology.

There is no correlation between the complexity of living things and the number of genes they have.

The preferred dogma started to appear in different versions in the 1920s. It was aptly summarized by renowned physicist Erwin Schrödinger in a famous lecture in Dublin in 1943. He told his audience that chromosomes “contain, in some kind of code–script, the entire pattern of the individual’s future development and of its functioning in the mature state.”

Around that image of the code a whole world order of rank and privilege soon became reinforced. These genes, we were told, come in different “strengths,” different permutations forming ranks that determine the worth of different “races” and of different classes in a class–structured society. A whole intelligence testing movement was built around that preconception, with the tests constructed accordingly.

The image fostered the eugenics and Nazi movements of the 1930s, with tragic consequences. Governments followed a famous 1938 United Kingdom education commission in decreeing that, “The facts of genetic inequality are something that we cannot escape,” and that, “different children ... require types of education varying in certain important respects.”

Post–war research sensibly focused more on the biochemistry, but with similar preconceptions. The existence of a powerful code–script seemed to be confirmed with the discovery of the structure of DNA by Watson and Crick in 1953. They revealed how the sequences of components (called nucleotides) in DNA could serve as a template—a code—for a protein, much as a typewriter sequences letters to form words. So the accepted “central dogma” could be conceived as the one–way flow of information from the code in the gene:

DNA template → proteins → developing characteristics;

as if production of the words alone is tantamount to writing the whole “book” of a complex being.

Then came the brilliant technology for sequencing genes (the components or “letters” in the DNA) in the whole genome. Its application, at enormous cost, in the *Human Genome Project* would, we were told, reveal “what it is to be human.” Extravagant promises were made that genes would soon be found that control human intelligence, social behavior, and complex diseases.

Now, in low–cost, highly mechanized procedures, the search has become even easier. The DNA components—the letters in the words—that can vary from person to person are called single nucleotide polymorphisms, or SNPs. The genetic search for our human definition boiled down to looking for statistical associations between such variations and differences in IQ, education, disease, or whatever.

For years, disappointment followed: Only a few extremely weak associations between SNPs and observable human characteristics could be found. Then another stroke of imagination. Why not just add the strongest weak associations together until a statistically significant association with individual differences is obtained? It is such “polygenic scores,” combining hundreds or thousands of SNPs, varying from person to person, and correlating (albeit weakly) with trait scores such as IQ or educational scores, that form the grounds for the vaulting claims we now witness.

Today, 1930s–style policy implications are being drawn once again. Proposals include gene–testing at birth for educational intervention, embryo selection for desired traits, identifying which classes or “races” are fitter than others, and so on. And clever marketizing now sees millions of people scampering to learn their genetic horoscopes in DNA self–testing kits.

So the hype now pouring out of the mass media is popularizing what has been lurking in the science all along: a gene–god as an entity with almost supernatural powers. Today it’s the gene that, in the words of the Anglican hymn, “makes us high and lowly and orders our estate.”

In her 1984 book, *The Ontogeny of Information*, the philosopher of science Susan Oyama warned, “Just as traditional thought placed biological forms in the mind of God, so modern thought finds ways of endowing the genes with ultimate formative power.”

In scientific, as well as popular descriptions today, genes “act,” “behave,” “direct,” “control,” “design,” “influence,” have “effects,” are “responsible for,” are “selfish,” and so on, as if minds of their own with designs and intentions.

But at the same time, a counter–narrative is building, not from the media but from inside science itself.

The long–suppressed logic of Johansenn that has stalked the gene–god for decades has come home to roost. Scientists now understand that the information in the DNA code can *only* serve as a template for a protein. It cannot possibly serve as instructions for the more complex task of putting the proteins together into a fully functioning being, no more than the characters on a typewriter can produce a story.

This can seem confusing to those of us indoctrinated in the idea that there must be a set of genetic instructions prior to development: If not in the DNA code, then where? By the 1980s, research findings started to turn that notion on its head.

First, laboratory experiments have shown how living forms probably flourished as “molecular soups” long before genes existed. They self–organized, synthesized polymers (like RNA and DNA), adapted, and reproduced through interactions among

hundreds of components. That means they followed “instructions” arising from relations between components, according to current conditions, with no overall controller: *compositional information*, as the geneticist Doron Lancet calls it.

In this perspective, the genes evolved later, as products of prior systems, not as the original designers and controllers of them. More likely as templates for components as and when needed: a kind of facility for “just in time” supply of parts needed on a recurring basis.



Then it was slowly appreciated that we inherit just such dynamical systems from our parents, not only our genes. Eggs and sperm contain a vast variety of factors: enzymes and other proteins; amino acids; vitamins, minerals; fats; RNAs (nucleic acids other than DNA); hundreds of cell signalling factors; and other products of the parents’ genes, other than genes themselves.

Molecular biologists have been describing how those factors form networks of complex interactions. Together, they self-organize according to changing conditions around them. Being sensitive to statistical patterns in the changes, they anticipate future states, often creating novel, emergent properties to meet them.

Accordingly, even single cells change their metabolic pathways, and the way they use their genes to suit those patterns. That is, they “learn,” and create instructions on the hoof. Genes are used as templates for making vital resources, of course. But directions and outcomes of the system are not controlled by genes. Like colonies of ants or bees, there are deeper dynamical laws at work in the development of forms and variations.

Some have likened the process to an orchestra without a conductor. Physiologist Denis Noble has described it as *Dancing to the Tune of Life* (the title of his recent book). It is most stunningly displayed in early development. Within hours, the fertilized egg becomes a ball of identical cells—all with the same genome, of course. But the cells are already talking to each other with storms of chemical signals. Through the statistical patterns within the storms, instructions are, again, created *de*

novo. The cells, all with the same genes, multiply into hundreds of starkly different types, moving in a glorious ballet to find just the right places at the right times. That could not have been specified in the fixed linear strings of DNA.

So it has been dawning on us is that there is no prior plan or blueprint for development: Instructions are created on the hoof, far more intelligently than is possible from dumb DNA. That is why today's molecular biologists are reporting "cognitive resources" in cells; "bio-information intelligence"; "cell intelligence"; "metabolic memory"; and "cell knowledge"—all terms appearing in recent literature.^{1,2} "Do cells think?" is the title of a 2007 paper in the journal *Cellular and Molecular Life Sciences*.³ On the other hand the assumed developmental "program" coded in a genotype has never been described.

Another wrench in the works has been the discovery that a gene product typically undergoes rearrangements before being put to use.

It is such discoveries that are turning our ideas of genetic causation inside out. We have traditionally thought of cell contents as servants to the DNA instructions. But, as the British biologist Denis Noble insists, "The modern synthesis has got causality in biology wrong ... DNA on its own does absolutely nothing until activated by the rest of the system ... DNA is not a cause in an active sense. I think it is better described as a passive data base which is used by the organism to enable it to make the proteins that it requires."

Of course, it's easy to see how the impression of direct genetic instructions arose. Parents "pass on" their physical characteristics up to a point: hair and eye color, height, facial features, and so on; things that "run in the family." And there are hundreds of diseases statistically associated with mutations to single genes. Known for decades, these surely reflect inherited codes pre-determining development and individual differences?

But it's not so simple. Consider Mendel's sweet peas. Some flowers were either purple or white, and patterns of inheritance seemed to reflect variation in a single "hereditary unit," as mentioned above. It is not dependent on a single gene, however. The statistical relation obscures several streams of chemical synthesis of the dye (anthocyanin), controlled and regulated by the cell as a whole, including the products of many genes. A tiny alteration in one component (a "transcription factor") disrupts this orchestration. In its absence the flower is white.

This is a good illustration of what Noble calls "passive causation." A similar perspective applies to many "genetic diseases," as well as what runs in families. But more evolved functions—and associated diseases—depend upon the vast regulatory networks mentioned above, and thousands of genes. Far from acting as single-

minded executives, genes are typically flanked, on the DNA sequence, by a dozen or more “regulatory” sequences used by wider cell signals and their dynamics to control genetic transcription.

This explains why humans seem to have only a few more genes than flies or mice (around 20,000), while a carrot has 45,000! There is no correlation between the complexity of living things and the number of genes they have. But there is a correlation with the evolving complexity of regulatory networks. Counting genes to understand the whole is like judging a body of literature by counting letters. It can't be done.

All of this provides a fraught background for modern gene association studies. What's more, the statistical analyses that power these studies are, themselves, full of pitfalls. First, the methods for computing polygenic scores, in which millions of variables are analyzed by statistical manipulation, provides huge opportunities for false positives. Very large databases—even randomly generated ones—can contain large numbers of meaningless correlations; and statistical significance values can be hugely inflated by invalid assumptions.

In polygenic score estimations, for example, it is assumed that SNP associations can be simply added together, as if beans in a bag, with no effects on each other, or from the environment. Then, as the National Institute of Health website reminds us, the majority of SNPs are functionally irrelevant anyway.⁴

More importantly, all modern societies have resulted from waves of migration by people whose genetic backgrounds are different in ways that are functionally irrelevant. Different waves have tended to enter the class structure at randomly different levels, creating what is called genetic population stratification. But different social classes also experience differences in learning opportunities, and much about the design of IQ tests, education, and so on, reflects those differences, irrespective of differences in learning ability as such. So some spurious correlations are, again, inevitable.

| The startling implication is that the gene as popularly conceived does not really exist.

As Jeremy J. Berg and colleagues warned this December in the online journal *Biorxiv*, polygenic scores “suffer dramatically from stratification bias, as even small differences in ancestry will be inadvertently translated into large differences in predicted phenotype.”⁵

Another wrench in the works has been the discovery that a gene product typically undergoes rearrangements before being put to use. It means that different proteins, with potentially widely different functions, can be produced from the same gene: not

one for one, as the central dogma has told us. Again, the instructions for such rearrangements are not in the genes themselves.

More startling has been the realization that less than 5 percent of the genome is used to make proteins at all. Most produce a vast range of different factors (RNAs) regulating, through the network, how the other genes are used.

Increasingly, we are finding that, in complex evolved traits—like human minds—there is little prediction from DNA variation through development to individual differences. The genes are crucial, of course, but nearly all genetic variations are dealt with in the way you can vary your journey from A to B: by constructing alternative routes. “Multiple alternative pathways ... are the rule rather than the exception,” reported a paper in the journal *BioSystems* in 2007.⁶

Conversely, it is now well known that a group of genetically identical individuals, reared in identical environments—as in pure-bred laboratory animals—do not become identical adults. Rather, they develop to exhibit the full range of bodily and functional variations found in normal, genetically-variable, groups. In a report in *Science* in 2013, Julia Freund and colleagues observed this effect in differences in developing brain structures.

In the same vein, we can now understand why the same genetic resources can be used in many different ways in different organs and tissues. Genes now utilized in the development of our arms and legs, first appeared in organisms that have neither. Genes used in fruit flies for gonad development are now used in the development of human brains. And most genes are used in several different tissues for different purposes at the same time.

In a paper in *Physics of Life Reviews* in 2013, James Shapiro describes how cells and organisms are capable of “natural genetic engineering.” That is, they frequently alter their own DNA sequences, rewriting their own genomes throughout life. The startling implication is that the gene as popularly conceived—a blueprint on a strand of DNA, determining development and its variations—*does not really exist*.

So it is, in a review in the journal *Genetics* in 2017, that the geneticists Petter Portin and Adam Wilkins question “the utility of the concept of a basic ‘unit of inheritance’ and the long implicit belief that genes are autonomous agents.” They show that “the classic molecular definition [is] obsolete.”

These radical revisions of the gene concept need to reach the general public soon—before past social policy mistakes are repeated.

Ken Richardson was formerly Senior Lecturer in Human Development at the Open University (U.K.). He is the author of Genes, Brains and Human Potential: The Science and Ideology of Intelligence.

References

1. Marijuán, P.C., Navarro, J., & del Moral, R. On prokaryotic intelligence: Strategies for sensing the environment. *Biosystems*, **99**, 94–103 (2010).
2. Lyon, P. The cognitive cell: Bacterial behaviour reconsidered. *Frontiers in Microbiology* **6**, 264 (2015).
3. Ramanathan, S. & Broach, J.R. Do cells think? *Cellular and Molecular Life Sciences* **64**, 1801–1804 (2007).
4. National Institute of Health. What are single nucleotide polymorphisms (SNPs)? <https://ghr.nlm.nih.gov>
5. Berg, J.J., *et al.* Reduced signal for polygenic adaptation of height in UK Biobank. *bioRxiv* (2018). Retrieved from doi.org/10.1101/354951
6. Wagner, A. & Wright, J. Alternative routes and mutational robustness in complex regulatory networks. *BioSystems* **88**, 163–172 (2007).