Genes as Difference Makers[1]

Evelyn Fox Keller, MIT, 2009.

What, in fact, is genetics, as a branch of biology, about? How one answers this question does of course depend on when, and where, one looks. But for classical genetics, and especially for the paradigmatic school of T. H. Morgan, genetics was about tracking the transmission patterns of units called "genes". What was a gene? No one knew, but notwithstanding this ignorance, a gene was assumed to be a unit that could be identified by the appearance of mutants in wild-type populations. That is, a phenotypic difference in some trait (a mutant) was taken to reflect a difference (a mutation) in some underlying gene associated with that trait. But to argue for such an identification between phenotypic difference and underlying gene in fact requires a two-step move. First, change (a "mutation") in some underlying entity (the hypothetical gene) is inferred from the appearance of differences in particular phenotypic traits (e.g., white eye, bent wing, narrow leaf), and second, the existence and identity of the gene itself is inferred from the inference of a mutation. In other words, the classical gene was on the one hand identified *by* the appearance of phenotypic differences (mutants), and on the other hand, it was simultaneously identified *with* the changes (mutations) that were assumed to be responsible for the mutants. Thus, the first map of the mutations thought to be responsible for the observed phenotypic differences in Drosophila was called not a map of mutations, but a "genetic map", a map of genes.

---

[1] This is an excerpt from my new book, The Mirage of a Space Between Nature and Nurture, Duke Univ. Press, forthcoming.

This is the sense in which the classical gene is often said to be a "difference maker" (see, e.g., Sterelny and Griffiths, 1999; Moss, 2003). But a gene was taken not only to be a difference maker; it was also assumed to be a trait maker. It was both the entity responsible for the difference observed, and (at least implicitly) the entity responsible for the trait which has undergone a change – i.e., the trait in which a difference has been observed.[i]

One might say, then, that a certain confounding of traits and trait differences was built into the science of genetics from the very beginning; moreover, one might argue, necessarily so. The occurrence (and frequency) of trait differences was what geneticists had observational access to: By examining phenotypes, they could detect phenotypic differences which, in turn, were taken as indicative of changes in some underlying, internal entity. Through breeding, the locus of such changes could be mapped. As Horace Freeland Judson has observed, "In 1913, Alfred Sturtevant, a member of Thomas Hunt Morgan's fly group at Columbia University, drew the first genetic map — "The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association." Ever since, the map of the genes has been, in fact, the map of gene defects." (2001, p. 769)[ii] Indeed, Wilhelm Johannsen, the man to whom we owe the word "gene", was himself clearly worried about this problem when he asked: "Is the whole of Mendelism perhaps nothing but an establishment of very many chromosomal irregularities, disturbances or diseases of enormously practical and theoretical importance but without deeper value for an understanding of the 'normal' constitution of natural biotypes?" (1923, p.140; quoted in Moss, 2003, p. 62).

It is hard to imagine that the slippage was entirely accidental. To think of genes

simply as difference makers would have been to detract from the very power of the gene concept. Mapping "difference makers" and tracking their assortment through reproduction may have been all that the techniques of classical genetics allowed for, but the aims of these scientists were larger. What made genes interesting in the first place was their presumed power to mold and to form – in a word, their presumed power to act. *Gene action* was the term invoked to refer to the process by which genes exerted their power in the development of the characters or traits themselves. But for illuminating the nature of this process (the developmental process), studies of trait differences would not by themselves suffice. In fact, neither successful mapping of the locus of the factors (difference makers) presumed to be responsible for such differences, nor analysis of their emergence and their intergenerational patterns of transmission, taught us anything about the causal dynamics of the developmental process by which the traits themselves came to be. As John Dupré puts it, "Classical genetics was about invisible features that could trigger different developmental outcomes, but not about the causal explanation of these outcomes."[iii] Furthermore, classical geneticists were for the most part well aware of this distinction. Nevertheless, the easy slide between genes as difference makers and genes as trait makers perpetuated the illusion (as widespread among geneticists as it was among their readers) that an increased understanding of the effects of gene differences would enhance our understanding of what it is that the entities called "genes" actually do.

5. What is a Disease?

A similar confounding of the etiology of traits with that of trait differences pervades virtually all of the current literature of medical genetics. Indeed, the very notion

of a disease as an individual trait   -- in the sense, i.e., that brown eyes is a trait -- already incorporates this confusion. We may commonly speak of an individual as "having a disease," much as he/she might have brown eyes, but in fact, and as has long been understood by many writers, disease is a state that only in relation to another state already established as normal.[iv] In his inquiry into the scientific rules for distinguishing the normal from the pathological, written more than a hundred years ago, Emile Durkheim stressed that "a trait can only be characterized as pathological in relation to a given species … " – in other words, in relation to a standard of normality or state of health which is, itself, inextricably confounded with the norm of a species. "One cannot even conceive, without contradiction, a species that could, by itself and in virtue of its fundamental constitution, be irremediably sick. [The species] is the norm par excellence and, accordingly, can harbor within itself nothing of the abnormal."[v]

It is true that the French philosopher of science Georges Canguilhem, following Kurt Goldstein, made valiant efforts to internalize the diagnostic criteria of pathology, locating them within the individual,[vi] but for all his efforts, by far the most common understanding of disease has continued to rely on comparison (or contrast) with a pre-established conception of 'normal'. English language dictionaries routinely define disease as a relational state: it is a dis-ease, "an abnormality of the body or mind" (Wikipedia); "a departure from the state of health" (OED); "a deviation from or interruption of the normal structure or function of a part, organ, or system of the body" (Dorland's Illustrated Medical Dictionary, W.B. Saunders); "An interruption, cessation, or disorder of a body, system, or organ structure or function" (Medilexicon). And indeed, French dictionaries do the same. In virtually every dictionary I have consulted, a *maladie*

is an "Altération de l'état de santé."   In French as in English, an animal can be said to

have brown eyes whether or not a comparison with other animals is at hand, but it cannot,

without such a comparison, be said to have a disease.

Like medicine, genetics too might be said to be a comparative science. Comparing

organisms with differing phenotypes, attempting to correlate these phenotypic differences

with corresponding genetic differences (mutations), has been the bread and butter of

geneticists from the earliest days of that science. But genetics aims beyond comparative

judgements, seeking an understanding of the developmental dynamics, and as I have tried

to show, its language invites us to lose sight of the complex moves – first, in attributing

the cause of a phenotypic difference to a genetic mutation; second, in the assumption that

the presence of a mutation automatically signals the presence of a gene; and third, in

attributing responsibility for the trait in question to the gene in which the mutation is

assumed to have occurred -- that are routinely made in effecting this shift from

comparative to individual. It seems therefore no accident that the adoption of a lexicon of

illness that refers to disease as an individual attribute comes with the emergence of a

medical science grounded in genetics.[vii] A disease, I am arguing, is not a trait but a trait

difference. The attempt to associate disease states with genetic mutations responsible for

"inborn errors of metabolism" can be seen simply as part of the more general effort to

associate mutations with particular phenotypic differences.   And subject to exactly the

same sorts of conflation.

Today, the genetic basis of a disease is more likely to be associated less with a

genetic difference -- with a mutation, a departure from a presumed normal genome – than

with a putative gene or genes; in short, medical geneticists seek the cause of a disease in a

genetic defect. Classical geneticists relied on "gene maps" to identify the gene presumed

to be involved; today, medical geneticists tend to rely more on analysis of nucleotide

sequences.   Although they continue to talk about *genes*, in actual practice, the

identification of one or more such differences may or may or may not point to a defect in

a particular gene (however that term is defined); these days, the direct object of interest is

the change (or defect) located somewhere in the nucleotide sequence that is correlated

with the expression of the disease. Despite widespread talk of "disease genes," or

"disease causing genes," it is the DNA itself that has become the focus of investigation.

I claim that the notion of a gene "causing" a disease (or even of a particular sequence

'causing' a disease) has exactly the same status as the notion of a gene "causing" a

mutation, But perhaps more important is the fact that, for diagnostic purposes (at least for

diagnoses based on genetic tests), the attempt to correlate a disease state with an

underlying gene is in many if not most cases largely irrelevant. Contemporary genetic

medical diagnostics rely on the identification of aberrant or anomalous sequences, and

not on the causal pathways such anomalies may disrupt. Such anomalies may be

anywhere in the genome – indeed, only rarely are they found in protein-coding sequences

(that is to say, in the segments of DNA usually associated with *genes*).

For most of us, the crucial question is, can the identification of such sequences

can be useful in the treatment or prevention of disease, and if so, how? Most

immediately (and perhaps most obviously even if generally left unstated), such

information can be used to promote selective abortion.   But if we are interested in

therapeutic medicine, we need more than simple correlation between aberrant sequence

and aberrant phenotype.   It is true that the early days of the Human Genome Project

brought the promise that in time we would be able simply to replace defective sequences with normal ones (gene therapy), but that hope has failed to materialize, and at least one of the reasons for this is that the relation between DNA sequence and phenotype has turned out to be far more complicated than originally expected. As to the possibility of other kinds of treatment (and sometimes of prevention) in a particular individual carrying the aberrant sequence, this depends on understanding something about the biological function that has been disrupted by the change in sequence that has been identified. Such a quest takes us beyond the analysis of phenotypic differences induced by mutant forms. Indeed, it requires an altogether different kind of analysis, and almost always, of a far more difficult nature.

There are, however, examples that have proven to be relatively simple. Phenylketonuria (PKU) is one, and it is probably the most celebrated case of therapeutic intervention in the history of medical genetics. Indeed, it is everyone's canonical example, mine as well. Phenylketonuria is a disorder (now recognized as genetic) associated with a range of disabling symptoms, including mental retardation, and caused by the inability of the body to properly metabolize the essential amino acid, *phenylalanine*. A major breakthrough in the treatment of this disease came with the recognition that its symptoms can be significantly alleviated if the affected individual adheres to a (carefully monitored) low phenylalanine diet for his or her entire life.[viii] However, the development of a strategy to treat PKU had nothing to do with either the identification or mapping of the gene(s) or genetic sequence(s) involved. Today we know that this disorder is caused by one or more mutations in the gene encoding the enzyme that breaks down phenylalanine (to date, as many as 400 such mutations have

been identified). In point of historical fact, however, neither our understanding of the (at least proximal) cause of this disease (Jarvis, 1937), nor the development of a therapeutic intervention (Bickel et al, 1953), depended on any sort of genetic analysis. It is now believed that the phenotypic expression of PKU (the disabling symptoms) are the direct consequence of the accumulation both of high levels of phenylalanine and of toxic intermediates resulting from its faulty metabolism, but we learned this from direct biochemical analysis; ordinary medical observation showed how dreadful the symptoms can be. Furthermore, we have now acquired the ability to precisely characterize many of the mutations responsible for the absence of the necessary enzyme, and doing so has certainly been instructive. The bottom line, however, is that, thus far, that ability has not significantly added to the possibilities of therapeutic intervention.

Of course, it needn't have happened that way. Identification, location, and characterization of the guilty mutation(s) *could have* provided the starting point for a research program that over the course of time led to an understanding of the disease and possibilities for treatment. But whatever the sequence of events of historical developments, there is no way in which the genetics of difference could have achieved that understanding by itself. What was required for a causal analysis of the disease – and hence, for the possibility of therapeutic intervention -- was a biochemical analysis of the metabolic pathway that had gone astray. That analysis might have begun with the identification of a particular gene (even though it in fact did not), but if it had, what would have been required is an understanding of the gene's down-stream effects, of the particular role that the gene in which the mutation had occurred actually played in development.

[i] Much the same can be said – and indeed, needs to be said – both of the gene in population genetics and of what Lenny Moss calls "Gene-P", or phenotypic gene. In both cases, the unit of interest is clearly a difference maker. Moss writes that "Gene-P" is a "phenotype predictor," and cannot be defined by its nucleic acid sequence. But in fact, it is a phenotype difference predictor. Indeed, as he himself acknowledges, the reason that "Gene-P" cannot be defined by a specific sequence is that "invariably there are many ways to lack or deviate from a norm" (2003: 60).

[ii] Judson proposes as a solution to this problem that we "revive and put into public use the term 'allele'. Thus, 'the gene for breast cancer' is rather the allele, the gene defect — one of several — that increases the odds that a woman will get breast cancer." (2001: 769) It seems to me, however, that the equation of allele with gene defect risks perpetuating precisely the same confusion: It is not the allele itself that is responsible for the phenotypic difference, but the difference between alleles.

[iii] Dupré (2006), p. 118.

[iv] Of course, the concept of normal is itself fraught with difficulty, subject to its own ambiguities that have primarily to do with persistent confusion between properties of individuals and those of populations. But most commonly, it too is understood as a relational property, pertaining not to comparison between individuals but to the statistical norm of a population. Durkheim, e.g., wrote, "The state of health, insofar as it can be defined, never conforms exactly to that of an individual subject, but can only be established in relation to the most common circumstances" (1894: 62). [L'état de santé, tel qu'elle le peut définir, ne saurait convenir exactement à aucun sujet individuel, puisqu'il ne peut être établi que par rapport aux circonstances les plus communes.] (See, also, Hacking, 1990, pp. 160-164.)

[v] The full quotation in the original reads as follows: "On voit qu'un fait ne peut être qualifié de pathologique que par rapport à une espèce donnée. Les conditions de la santé et de la maladie ne peuvent être définies in abstracto et d'une manière absolue. La règle n'est pas contestée en biologie ; il n'est jamais venu à l'esprit de personne que ce qui est normal pour un mollusque le soit aussi pour un vertébré. Chaque espèce a sa santé, parce qu'elle a son type moyen qui lui est propre, et la santé des espèces les plus basses n'est pas moindre que celle des plus élevées. … Le type de la santé se confond avec celui de l'espèce. On ne peut même pas, sans contradiction, concevoir une espèce qui, par elle-même et en vertu de sa constitution fondamentale, serait irrémédiablement malade. Elle est la norme par excellence et, par suite, ne saurait rien contenir d'anormal." Émile Durkheim (1894), *Les règles de la méthode sociologique.* Paris: Les Presses universitaires de France, 16e édition, 1967, p. 4. (http://classiques.uqac.ca/classiques/Durkheim_emile/regles_methode/regles_methode.html)

[vi] E.g., in *The Normal and the Pathological* (1943), Canguilhem writes, "We think with Goldstein that the norm concerning pathology is above all an individual norm" (p. 72). See also, his essay, "Le Concept et la Vie" (1966).

[vii] For an extensive discussion of the relation between the language of medical genetics and that of classical genetics, see Childs, 1999,

[viii] This is not to suggest that maintaining such a diet is an easy task, or that the almost inevitable relapses are not without dire risks of their own. Probably the best discussion of the history and politics of PKU, as well as of risks associated with its treatment, is to be found in the work of Diane Paul (see, e.g., Paul, 1998; 2000, and Paul and Edelson, 1997).