

Some Problems with Genetic Horoscopes

Eva Jablonka¹

The Charms and Perils of Genetic Astrology

The view that dominated biological thought in general, and evolutionary biology in particular, from the 1940s until the beginning of the 21st century, was focused on the gene. Development was seen as the product of genes' action, and ecology as the context for the natural selection of genes. This view was well reflected in textbooks of evolutionary biology and in popular books like the *Selfish Gene*. The dominance of the gene and its material incarnation, DNA, led, as Nelkin and Lindee (1995) have noted, to the perception of DNA as the secular equivalent of the soul. But since, unlike the soul, DNA can be sequenced, this view was soon translated into future-telling. In 1989, for example, James Watson declared: "We used to think our fate was in the stars. Now we know in large measure, our fate is in our genes."

And soon enough, the time of genetic horoscopes arrived.

GenePlanet is one company that provides a DNA analysis through which, it promises, you will "get to know everything you ever wanted to know about yourself: to which diseases you are susceptible, what effect do medications have on you, what are your talents and special abilities, and who are your ancestors." Again and again the

¹ Based on a lecture delivered at the conference "Bioscience and Society: Biodiversity", held on October 1-2, 2009, in Ljubljana, Slovenia.

following slogan appears: **Discover your genes, know yourself!** And then, after a question (how hard is it for me to quit smoking? Which medications work for me? Will I gain weight if I eat fatty foods? Who are my ancestors?) a large announcement is seen:

“It’s all in your genes !”¹

Knowing oneself does not require too much meditation and introspection any more. We are modern people living in the 21st century, not Greeks in ancient Athens. Instead of discussing philosophy and studying geometry, you spit into a container and send it to the company. After a while, for a modest sum of 399 Euros (special offer), you get your genetic horoscope: its all there – what your health risks are, what your hidden and apparent talents are, and many other wonderful things. Wiser and more realistic, you return to your daily life, to live in more responsible and fulfilling way. Future possibilities are unraveled: **“It’s all in your genes !”**

But what exactly is in your genes? If we take two people with identical genes (e.g. identical twins) who live in more or less the same environment, will they have identical futures? If we clone a human being, will the cloned person develop identical propensities as those of the originator of its DNA? And if not, why not? If it’s not all in your genes, where else does “it” (you) come from?

I must start by stating very clearly that there are cases where genetic counseling on the basis of DNA testing is enormously useful. There are “monogenic” diseases— diseases for which the presence or absence of symptoms depends on which alleles of one particular gene are present. Knowing which allele one has can then be very informative and can lead to important existential decisions (such as whether or not to abort a fetus that will develop into a terribly sick and long-suffering baby, doomed to die). Tay Sachs,

a disease which is relatively common among Ashkenazi Jews, is such a monogenic disease, and thankfully, simple blood tests help in giving important counseling to pregnant mothers. However, such simple monogenic diseases are not very common: they make up less than 2 percent of all the diseases that are known to have a genetic component. For the remaining 98 percent of “genetic” disorders, the presence or absence of the disease and its severity are influenced by many genes and by the conditions in which a person develops and lives. And it is for the complex metabolic diseases, as well as for the interesting cognitive traits that are influenced by many genes, that the genetic testing companies promise to provide their horoscopes.

Last year, a British journalist (Sunday Times 8th of September 2008) decided to see how compatible the answers were from three different genetic testing companies (GeneticHealth, a British firm, deCODEme, based in Iceland, and American group 23andMe) about the health prospects written in his DNA. The results were a little baffling. This is part of what he discovered:

“deCODEme said my risk of developing exfoliation glaucoma, which causes loss of vision, was 91% below average. Yet according to 23andMe, I was 3.6 times more likely to get it than average. For age-related macular degeneration, deCODEme put my risk at 20% lower than average, while 23andMe said it was 62% higher.

According to deCODEme, my risk of developing Alzheimer’s was 74% above average, while GeneticHealth said my genes were associated with “a fourfold increased risk of developing Alzheimer’s disease by your late 80s”.

According to deCODEme, my risk of a heart attack, angina or sudden cardiac death is 54.8%, which is 6% above average. By contrast, 23andMe said my risk of a heart attack between the ages of 45 and 84 was 17.5%, below average. Yet another assessment came from Paul Jenkins, the clinical director of GeneticHealth, who said my risk of cardiovascular disease was “low to moderate.”²

Clearly the situation is not very good. There may be many reasons for that (for example, different databases used by the different companies). But there may be something about the assumptions made about correlating genes to health outcomes that is inherently problematic.

So let us make the assumptions of the genetic horoscope companies clearer. The main assumption is that genetic differences between individuals, differences between the base sequences of their DNA, are the most important reasons for the phenotypic, visible, differences between them. If you have allele A1 you will be prone to disease X, and if you have allele B2 you are likely to develop skill Y. A difference in a given gene, it is assumed, makes a difference at the level of the phenotype. Of course the gene's effects can be somewhat modified by the environment (diet, exercise). However the gene represents a program which the environment only *triggers*. The gene is assumed to be switched on or off by different conditions (e.g. “on” in environment 1, “off” in environment 2). Adaptation to unanticipated novelty which is not already embedded in the program is not possible. Nor is developmental plasticity very extensive, for if it is, a lot more than the DNA sequence would be necessary to anticipate the future.

How extensive is plasticity? Does a difference in a single gene usually make a large phenotypic difference? Or do differences in *many* genes make a phenotypic difference? If so, how many? Do genetic differences make the same phenotypic difference under different conditions? What conditions? How many types of conditions? Whose conditions? Are the conditions your own developmental conditions? Your parents'? Your more remote ancestors'? These are crucial questions, so let us look at them. Let us first start with what we have learnt about the complex relations between genes and traits.

Genetic Networks

We have learned quite a lot about gene interactions during the last two decades, and it is common for geneticists these days to talk not about single genes and their effect on traits, but about *gene networks*. The regulatory architecture of gene networks is a focus of study, and it is clear that the organization of the network is essential for understanding the relation between the genetic and phenotypic levels. The identity of a single network component may be unimportant in many cases. Even the absence of a network component may not make a difference at the phenotypic level. We know that many knockout mutations are selectively neutral, having no visible effect on the phenotype.

In *Evolution in Four Dimensions*, my collaborator Marion Lamb³ and I discussed a study of a genetically complex common metabolic disease, coronary artery disease, and the role played by some alleles at an important relevant locus, APOE (Jablonka and Lamb 2005 and references therein). The APOE locus codes for a protein (apoprotein E,

or apoE) that helps to carry fats around in the blood. This locus has three common alleles: allele 2, allele 3, and allele 4. The three alleles are associated with differences in the incidence of coronary artery disease. The three most common genotypes (two alleles per genotype) are 2/3, 3/3, and 4/3, and in this population, people with genotype 3/3 have a below-average chance of developing the disease, those with genotype 2/3 are average, and people with genotype 3/4 are twice as likely as the average person is to suffer from coronary heart disease.

Since the apoE molecule helps to transport cholesterol around in the blood, and the population surveys have shown that people with allele 4 do have, on average, high cholesterol, the conclusion that allele 4 leads to high cholesterol seemed logical. However, the combination of high cholesterol and allele 4 is not the worst possible combination. Having allele 2 and high cholesterol leads to the highest probability of coronary artery disease, while having high cholesterol and a 3/3 genotype leads to average disease risk (like that of people with normal level of cholesterol)! Moreover, APOE is just one locus out of more than a hundred that affect the development of this disease, and environmental conditions (e.g. diet, exercise) also have an effect. The complex interrelationships among genes and environment in this case, as in others, means that we cannot just add the average effects of genes together, and from this predict what a person's strengths and weaknesses will be.

This is not the only problem, however. Let us look at a variable trait like height. Height is a straightforwardly quantifiable trait (unlike intelligence), and does not seem too complex. It has a significant genetic component: its heritability (the part of the variance in the trait that is attributable to genetic differences) is 80 percent and 20

relevant SNPs (single nucleotide polymorphisms) are known. On the basis of the 20 DNA sequences that we have found to be associated with height, can we infer how tall an individual is likely to be? An extensive discussion in the *New England Journal of Medicine* (April 2009, volume 360(17)), was centered around the predictive value of genetic association studies, which included the study of height. In these studies, alleles or SNPs are drawn from the population, and the frequency with which certain alleles or SNPs are present in different groups (e.g. short and tall people) is tested for association with the trait (short or tall).

It turns out that in spite of the high heritability, the 20 SNPs explain only 3 percent of its heritability. If we extrapolate from this 3 percent, it was estimated that approximately 93,000 SNPs are required to explain 80% of the population variation in height. A similar result is found for Diabetes type II and many other traits. David Goldstein, the director of the Center for Human Genome Variation, Institute for Genome Sciences and Policy in Duke University, concludes: “If effect sizes were so small as to require a large chunk of the genome to explain the genetic component of a disorder, then no guidance would be provided: in pointing at everything, genetics would point at nothing.”⁴

So does the sum total of the heritability reside in the specific detailed configuration of the whole genome of the individual? If so, maybe one should try to discover *different types* (not too many types) *of network architectures*, rather than focus on the component units. Maybe when the number of identical genes is beyond a certain threshold, a particular network organization is generated, and this affects the trajectory of development. If so, the genetic architectures underlying distinct developmental

trajectories are of interest, rather than the individual genes involved. And perhaps the formation of particular developmental trajectories and their heritability are not just due to DNA sequence variation, but to another type of heritable variation as well: to *heritable epigenetic variation*, which may be the result of plastic, dynamic responses to critical environmental conditions in ancestors. If this is part of the answer, a few words about the possibility of inheriting environmentally induced (plastic) responses, are necessary.

Plasticity and Epigenetic Inheritance

Plasticity is defined as the ability of one genotype to generate different phenotypes depending on environmental cues that act as inputs into the organism's development (West-Eberhard 2003). Genetically identical animals can have different appearances (morphs) depending on the environment to which they or their ancestors were exposed. Identical twins can be rather different and have different fates; for example they can be discordant for various diseases (e.g. diabetes, schizophrenia). Why this is so is not always clear. Sometimes it is different childhood or adult life experiences, sometimes it is something that presumably happened in the womb, and often we cannot pinpoint the reason/s. But we are beginning to learn about the mechanisms involved in generating some of the persistent plastic responses (Jablonka and Lamb 2005, Gilbert and Epel 2009). They are called epigenetic mechanisms, and they can sometimes lead to epigenetic inheritance.

Epigenetic inheritance occurs when environmentally-induced and developmentally-regulated variations, or variations that are the result of developmental noise, are transmitted to subsequent generations of cells or organisms (Jablonka and

Lamb, 2005). The term “epigenetic inheritance” is used in two ways, a broad way and a narrow way. *Epigenetic inheritance in the broad sense* is the inheritance of any developmental variations that do not stem from differences in DNA sequence or persistent inducing signals in the present environment. This includes cellular inheritance through the germline, and soma-to-soma information-transfer that by-passes the germline. For example, soma-to-soma transmission occurs through developmental interactions between mother and embryo, or through observation-based social learning. *Epigenetic inheritance in the narrow sense* is cellular epigenetic inheritance, where the cell is the unit of transmission, and variations that are not the result of DNA differences are transmitted from mother cell to daughter cell. Cellular epigenetic inheritance occurs during cell division in prokaryotes, mitotic cell division in the soma of eukaryotes, and sometimes during the meiotic divisions in the germline that give rise to sperm or eggs. In this latter case offspring inherit epigenetic variations through the germline.

Marion Lamb and I distinguish between four types of cellular epigenetic mechanisms:

- (i) *Self-sustaining feedback loops*. Gene products act as regulators that directly or indirectly maintain their own transcriptional activity. Such positive feedback can lead to the transmission of these products during cell division, resulting in the same states of gene activity being reconstructed in daughter cells. Alternative and heritable cell phenotypes can be generated, and this is indeed seen in fungi, bacteria and other microorganisms (Smits *et al.*, 2006).
- (ii) *Structural inheritance*. Cellular structures act as templates for the production of similar structures in daughter cells, which then become components of daughter cells. This type of templating is seen in prion-based inheritance in fungi (Wickner *et al.*,

2004), the inheritance of cortical structures in ciliates (Grimes and Aufderheide, 1991), and the reconstruction of what Cavalier-Smith (2004) calls “genetic membranes.” For example, the mitochondrial membrane can reproduce only if there is a preexisting mitochondrial membrane that acts as a template. It cannot self assemble from its constituent elements without “guidance” from a “template membrane.”

(iii) *Chromatin marking*. Chromatin marks are proteins and small chemical groups that are attached to DNA and influence gene activity. Relics of these marks segregate with the DNA strands after replication, and nucleate the reconstruction of similar marks in daughter cells (Henikoff and Smith, 2007). Chromatin marks include modifiable histone and non-histone proteins that are non-covalently bound to DNA, methyl groups that are covalently bound to the DNA, and patterns of bound non-histone proteins. Methylation patterns are different in different cell types, and are different in imprinted genes. Identical twins are born having similar patterns of methylation, but as the twins age they grow more and more different (Fraga et al 2005). Methylation patterns and other chromatin marks can be also inherited between generations. The inheritance of a testis disease in male rats (for 6 generations so far, Crews, personal information) following a single injection of their ancestral great-grandmother with vinclozolin, is based on the inheritance of chromatin (methylation) marks. The heritable effects of maternal under or over nutrition during pregnancy also involve the transmission of chromatin (methylation) marks between generations (Anway et al 2005, 2006, Crews personal information).

(iv) *RNA-mediated inheritance*. Transcriptional states are induced and are actively maintained through interactions between small, transmissible, RNA molecules and the mRNAs or the DNA/chromatin regions with which they pair. Such interactions can be transmitted between cell and organism generations through an RNA-replication system, via the interaction of the small RNAs with chromatin, which leads to heritable modifications of chromatin marks and? Something else here? (Bernstein and Allis, 2005). An example is the inheritance of a heart disease induced by the injection of small RNAs, which are then transmitted for at least 2 generations (Wagner et al 2008).

Between-generations epigenetic inheritance is not rare. In a recent survey Jablonka and Raz (2009) found over 100 cases of epigenetic inheritance in 42 species:

- Twelve cases of epigenetic inheritance in bacteria. Most were of self-sustaining loops, but examples of chromatin marking and structural inheritance were also found.
- Seven cases in protists. Most were in ciliates, where structural inheritance (the transmission of cortical morphologies) is common, and all loci may be modified through the RNA-mediated EISs. The other two types of EISs were also found in protists. What does EIS stand for? Epigenetic Inheritance System?
- Nineteen cases in fungi, involving many phenotypes and loci. Examples of all four types of EISs were found.

- Thirty-eight cases in plants, involving many loci and many traits. Among the 36 cases, four were in plant hybrids, and in all of these many loci were heritably modified. Genomic stresses such as hybridization and polyploidization, especially allopolyploidization, seem to induce genome-wide epigenetic changes, some of which are transmitted between generations through the chromatin-marking and the RNA-mediated EISs. No evidence was found for gametic between-generation inheritance based on self-sustaining loops and structural templating.
- Twenty-seven cases in animals, some of which involved many loci. As with plants, stress seems to induce multiple epigenetic changes. Epigenetic variations were transmitted through the chromatin marking and the RNA-mediated EISs; and there was no evidence for between-organism gametic inheritance based on self-sustaining loops and structural templating.

And this is just the beginning. Since we conducted this survey there are additional cases, especially in plants. Studies of methylation inheritance in *Arabidopsis thaliana* show that a considerable part of the genome's methylation patterns can be inherited for many generations, while other regions show instability (Reinders et al 2009, Teixeira et al 2009).

In addition to epigenetic inheritance in cell lineages there are also soma-to-soma routes of transmission that bypass the germline. Soma-to-soma transmission includes transmitting substances that affect development through feces ingestion, through the placenta and milk of mammals, and through the soma-dependent deposition of specific chemicals (proteins or hormones) in the eggs of oviparous animals and plants (Avital and

Jablonka, 2000). Maternal morphological features, such as size, can constrain offspring development and lead to heritable and self-perpetuating developmental effects (Gilbert and Epel 2009). For vertebrates, socially-learned behaviors that do not require the transfer of materials are an important source of transmissible variations. Famous examples are bird and whale dialects, which are transmitted in ways similar to the transmission of dialects in human populations and form “linguistic” traditions. Another example is the traditions found among different groups of common chimpanzees – 39 different behavioral phenotypes in chimpanzee population in East Africa are attributed to cultural evolution, since the behaviors are transmitted from one generation to the next via social interactions between members of a group. (Avital and Jablonka 2000).

Developmental interactions among organisms that form coherent and persistent symbiotic communities (Zilber-Rosenberg and Rosenberg, 2008), also contribute to soma-to-soma inheritance.

The evidence suggests that organisms can respond to changing environments by changing their development and their phenotypes, and sometimes these plastic responses can be stabilized and inherited. Moreover, it seems that the responses are not “programmed” in any rigid way, and that the organism can respond to new conditions by launching an exploration processes at the intra-organismal level (for example at the intracellular level); when a response which relieves the stress of the changed environment is found, this response is stabilized and may even be inherited. This was shown by Braun and his colleagues (Stern et al 2007), who confronted yeast cells with a rewired regulatory circuit with a severe and unforeseen challenge, and studied how the population adapted to it. They placed the essential *HIS3* gene from the histidine biosynthesis

pathway under the exclusive regulation of the galactose utilization system. In a glucose containing medium the GAL genes are repressed and so is (in these rewired cells) the essential *HIS3* gene. The rewired cells were then placed in a medium containing glucose but lacking histidine and were therefore severely challenged (they could not synthesize histidine). The cells never had the *HIS3* gene regulated by carbon sources availability, so a substantial adaptive response was required for them to survive in a histidine-lacking medium. However, the cells did manage! Over 50 percent of the population overcame the challenge within only a few generations, and the adapted state was propagated stably for hundreds of generations. The study suggests that a biochemical exploration process was initiated under these stressful conditions, and among the many regulatory networks that emerged, those that relieved the stress were stabilized. This exploratory response to a new stress, followed by stabilization when a solution is found, is similar to neural learning by trial and error. It is possible that this type of response is a general strategy which is applied at many levels of biological organization, within cells and within systems. This type of entirely novel plastic response is not “programmed”, and is impossible to predict except in a very broad functional sense.

Back to the Future

If we return to the puzzle of the missing heritability in the case of height (and other similar cases) , then maybe part of the answer to this puzzle is that it resides not in the genotype, but rather in the epigenotype, in the cellular networks that underlie genetic and epigenetic memory and heredity. And if indeed the epigenomic developmental networks are the relevant level of analysis for studying the heritability of complex

diseases and traits, then it is not surprising that the genetic testing companies cannot produce very consistent results. For a reliable estimate, a very large number of tested individuals is necessary, and once this is done it is likely that although a statistically reliable association is indeed found, the contribution of the SNP or even the actual gene to the trait's heritability is rather modest. This does not mean that we cannot study the developmental pathways involved in common complex diseases which will help us design new medications. We can study them either by looking, as Goldstein (2009) has suggested, at very rare genetic variants which *do* have large effect and change the architecture of the landscape, or by looking at the effects of severe environmental stressors that alter the developmental trajectory. According to this latter strategy gene expression profiles in the stressful conditions would be the focus of study, and the epigenetic factors that contribute to the persistence of the developmental trajectories during ontogeny, or between generations, would be specifically targeted.

Very obviously it is not all in our genes. It is also in our experiences (beginning in the womb), in the experiences and life style of our parents, and in the genetic and epigenetic history of our ancestors.

References

- Anway MD, Cupp AS, Uzumcu M, Skinner MK (2005) Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* 308:1466-1469.
- Anway MD, Leathers C, Skinner MK (2006) Endocrine disruptor vinclozolin induced

- epigenetic transgenerational adult-onset disease. *Endocrinology* 147: 5515-5523.
- Avital E, Jablonka, E (2000) *Animal Traditions: Behavioural Inheritance in Evolution*. Cambridge: Cambridge University Press.
- Bernstein E, Allis CD (2005) RNA meets chromatin. *Genes Devel* 19: 1635-1655.
- Cavalier-Smith T (2004) The membranome and membrane heredity in development and evolution. In: *Organelles, Genomes and Eukaryote Phylogeny: An Evolutionary Synthesis in the Age of Genomics* (Hirt RP, Horner DS, eds), pp 335–351. Boca Raton: CRC Press.
- Fraga M.F, (and 20 others) (2005) Epigenetic differences arise during the lifetime of monozygotic twins *Proceedings of the National Academy of Science USA*, 102:10604–10609
- Gilbert S. and Epel D. (2009) *Ecological Developmental Biology and Epigenesis: An Integrated Approach to Embryology, Evolution, and Medicine*. Sinauer
- Goldstein B. D. (2009) Common genetic variation and human traits. *New England Journal of Medicine*, 360: 1696-1698.
- Grimes G, Aufderheide K. (1991) Cellular aspects of pattern formation: the problem of assembly. *Monographs in Developmental Biology* 22:1-94.
- Henikoff S, Smith M. (2007) Histone variants and epigenetics. Pages 249-264 in *Epigenetics*, edited by D C Allis, T Jenuwein, D Reinberg and M-L Caparros. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Jablonka E, Lamb MJ (2005) *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. Cambridge, MA: MIT Press.

- Jablonka E, Raz G (2009) Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q Rev Biol.* 82(4): 131-176.
- Nelkin, D and Lindee, MS. (1996) *The DNA Mystique: The Gene as a Cultural Icon.* WH Freeman, New York.
- Reinders J, Wulff BB, Mirouze M, Marí-Ordóñez A, Dapp M, Rozhon W, Bucher E, Theiler G, Paszkowski J. 2009. Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. *Genes and Development* 23: 939-950.
- Smits W K, Kuipers O P, Veening J W. (2006) Phenotypic variation in bacteria: the role of feedback regulation. *Nature Reviews Microbiology* 4:259-271.
- Stern S, Dror T, Stolovicki E, Brenner N, Braun E (2007) Transcriptional plasticity underlies cellular adaptation to novel challenge. *Molecular Systems Biology* 3 Article number: 106 doi:10.1038/msb4100147
- Teixeira, F.K., Heredia, F., Sarazin, A., Roudier, F., Boccara, M., Ciaudo, C., Cruaud, C., Poulain, J., Berdasco, M., Fraga, M.F., et al. 2009. A role for RNAi in the selective correction of DNA methylation defects. *Science* 323: 1600–1604.
- Wagner K. D., Wagner N., Ghanbarian H., Grandjean V., Gounon P., Cuzin F., Rassoulzadegan M. 2008. RNA induction and inheritance of epigenetic cardiac hypertrophy in the mouse. *Developmental Cell* 14(6):962–969.
- West-Eberhard MJ (2003) *Developmental Plasticity and Evolution.* New York: Oxford University Press.
- Wickner R B, Edskes H K, Ross E D, Pierce M M, Baxa U, Brachmann A,

Shewmaker F. (2004) Prion genetics: New rules for a new kind of gene. *Annual Review of Genetics* 38:681-707.

Zilber-Rosenberg I, Rosenberg E (2008) Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol Rev* 32: 723-735.

¹ (<http://www.geneplanet.com/>)

² <http://www.timesonline.co.uk/tol/news/uk/science/article4692891.ece>

³ Eva Jablonka and Marion Lamb. *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral and Symbolic Variation in the History of Life* Cambridge, MA: The MIT Press, 2005.

⁴ *New England Journal of Medicine* 2009 360:1696-1698