CHAPTER 7

Language in an Epigenetic Framework[*]

Donata Vercelli and Massimo Piattelli-Palmarini

I have to tell you a story and the story is that the reason I am here is that I can't say no to my friends. Juan Uriagereka was both very insistent and very eloquent in inviting me, so here I am, presenting something that Massimo and I have been thinking about. I have to tell you that the division of labor is such that Massimo takes all the credit and I take all the blame. So this, by way of disclaimer, that I think we acknowledge that there is a little element of absurdity in what we may be saying, but we hope that we also have something that may be relevant to you.

Today we would like you to think about a biological trait, and for reasons I hope will become clear to you, let us call it biological trait *L*. *L* has certain features. It is species-specific, and in particular is unique to humans. It has a common core that is very robust but allows for inter-individual and inter-group variation. It has both heritable and non-heritable components. It goes through critical developmental windows of opportunity: that is, its developmental patterns are time-dependent. It is very plastic, particularly in response to environmental cues. It has multiple and discrete final states, it is partially irreversible, and it is robust and stable over a lifetime.

The question we are trying to answer is, what kind of biology may underlie a trait such as *L*, or, how is a trait such as *L* implemented in our genome. Classical genetics (which I will define in a minute) can certainly account for some features of *L*: species specificity, uniqueness to humans, and a very robust common core that allows for variation. The problem is that classical genetics, we maintain, would not buy us the other features that *L* has. And this is where we think we need to go a little bit further. Let us qualify why.

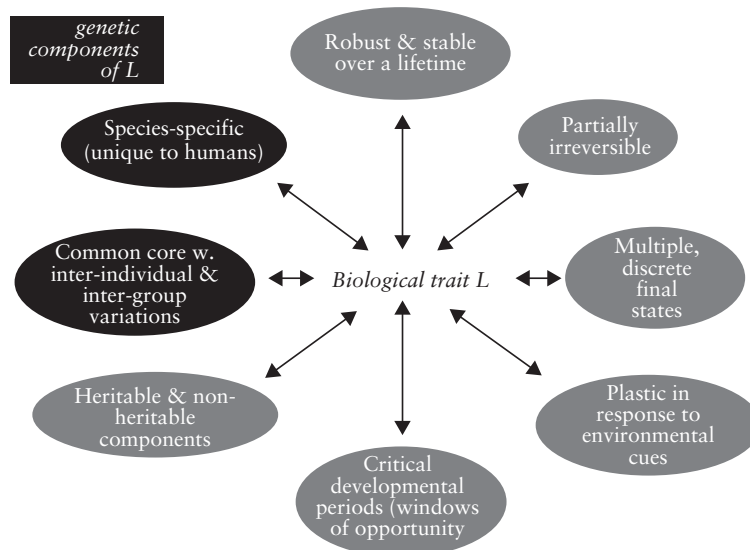[*] This paper was delivered at the conference by Donata Vercelli.

Fig. 7.1. Aspects of biological trait L

1953 is the year in which DNA, as we know it today, and classical molecular genetics were born. It is the year in which Watson and Crick published their rightly famous paper stating that the structure they proposed for DNA, the double helix, could be very effective to replicate, faithfully copy, and transmit information. The success of classical molecular genetics has been spectacular. In their labs, molecular biologists apply the paradigms of classical genetics every day. The notion that a DNA sequence is transcribed into an RNA sequence which is in turn translated into a protein is something we use, for instance, to make proteins *in vitro* starting with a sequence of DNA. This successful notion of genetics emphasizes the amount of information that is encoded and carried by the DNA sequence. What this genetics can give us is great fidelity and specificity in the transmission of information. What this genetics does *not* buy us is a fast, *plastic* response as well as environmental effects and memory of a functional state – nor does it buy us cell fate decisions. In essence, classical genetics is necessary, but not sufficient. This is where epigenetics comes in.

We are stressing the importance of plasticity, because we think plasticity is probably one of the defining features of our trait *L*. From a biological point of view, here is the puzzle. Let us consider the different stages our blood cells go through to become the mature cells circulating in our bloodstream. We have red cells and white cells, and they have quite different tasks. Red cells transport oxygen, some white cells fight infection from bacteria, some white cells fight infection from parasites. Therefore, all these cells do very different things, but they all derive from an initial common precursor cell – that is, they are

genetically identical, but they are structurally and functionally heterogeneous because they have different patterns of gene expression that arise during development. Such differences are epigenetically implemented.

To talk about epigenetics, we need to introduce a difficult but fascinating concept.[1] The DNA double helix is not linear in space. It is a very long structure, if you unfold it, but it is actually very tightly packaged, to the extent that in the cell it becomes 50,000 times shorter than it is in its extended length. Packaging is a stepwise process during which the double helix initially forms nucleosomes, that is, spools in which the DNA wraps around a core of proteins (the histones). In turn, each of these beads-on-a-string is packaged in a fiber that is even more complex, and the fiber is further packaged and condensed until it becomes a chromosome. All this packaging and unpackaging, winding and unwinding, provides a way to assemble a huge amount of information within a very small space, but also makes it possible to *regulate* what happens to the information encoded in the DNA.

This is the subject of epigenetics. Epigenetics is the study of changes in gene expression and function that are heritable, but occur without altering the sequence of DNA. What changes is the functional state of the complex aggregate formed by DNA and proteins. These changes – extremely dynamic, plastic, potentially reversible – occur in response to developmental and/or environmental cues that modify either the DNA itself (by appending chemical groups to the sequence, which remains unaltered) or by changing the proteins around which the DNA is wrapped (i.e., the histones). By modifying the core proteins around which the DNA is assembled, or the chemical tags appended to the DNA, the functional state of a gene is also modified (Vercelli 2004).

Deciphering these modifications is quite complex. For DNA to become active, to release the information it carries, the molecule needs to unwind, to become accessible to the machinery that will transcribe it and turn it into a protein. This cannot happen if the DNA is very compressed and condensed, if all the nucleosomes, all the beads-on-a-string, are so close to one another that nothing can gain access to a particular region. Such a state is *silenced chromatin*, as we call it – chromatin being the complex (which is more than the sum of the parts) of DNA and proteins. When nucleosomes are very close and condensed, chromatin is silenced. That happens when methyl groups are added to the DNA or the histones bear certain chemical tags. On the other hand, when other tags are added to the histones or the DNA is no longer methylated,

---

[1] Two recent classics are: Grewal and Moazed (2003) and Jaenisch and Bird (2003). For a recent exhaustive exposition, see Allis et al. (2006). For short accessible introductions see Gibbs (2003). (Editors' Note)

the nucleosomes are remodeled and open up, the distance between them becomes greater, and the machinery in charge of transcription can get in. Now, transcription can occur. Hence, active chromatin is marked by accessibility.

That epigenetics results in real changes in how genes function is a fact. A clear example of how this happens is provided by the case of the black mice. These mice are all genetically identical, in DNA sequence, but it does not take a geneticist to see that they are quite different phenotypically, in terms of the color of their coats. What has happened is that the mothers of these mice are given diets containing different amounts of substances that provide methyl groups. As we discussed, DNA methylation is a major epigenetic regulator of gene expression. After the mothers are fed different amounts of methyl donors and the pups are born, their coat color is checked. Depending on the amount of methyl donors the mothers received, and depending on the different colors of the coats, different levels of methylation are found in the DNA locus that regulates this trait, the color of the coat, with a nice linear relationship between methylation and coat color (Morgan et al. 1999).

This may be true not only of mice; there are interesting data in humans as well, for instance the famous case of the Dutch hunger winter, the famine in the Netherlands during World War II, when mothers who were pregnant at that time had very small children. The children of those children (the grandchildren of the mothers pregnant during the famine) remained small despite receiving a perfectly normal diet.[2] It is possible that this feature, this trait, was transmitted across generations.

What we propose is that this *kind* of mechanism may account for some of the features of *L* at least (those in red in Fig. 7.1). Here are some cases in support of our proposal.

Plasticity is certainly a paramount feature of biological trait *L*. A relevant well-known case is that of the Achillea, a plant. Plants are masters at using epigenetics because they are exposed to weather and heavy environmental insults and they need to react to light and temperature. This they do epigenetically. For Achilleas, the same plant at low altitude is very tall, at medium elevation is very short, and at high elevation it becomes again very tall. Nothing changes in the genome of this plant, but the phenotype changes heavily in response to environmental cues, in this case climate and altitude.[3] This is the concept of *norm of reaction* that Richard Lewontin, in the wake of the Russian geneticist and evolutionist Ivan Ivanovich Schmalhausen (1884–1963),[4] has so

---

[2] Described in Roemer et al. (1997).
[3] Studied ever since Hiesey et al. (1942).
[4] For an analysis of the history of this notion, see Levit et al. (2006).

clearly formulated: what the genotype specifies is not a unique outcome of development, it is a *norm of reaction*. A norm of reaction is constrained by genotype, but specifies a pattern of different developmental outcomes depending on the environment.

The concept of *windows of opportunity* is quite familiar to immunologists. In the stables of a Bavarian farm, the mothers work while their children sit in a cradle. As a result of that, we now know, these children are incredibly well-protected from allergic disease, but only if they sit in the stables up to the age of one year, or even better if the mother goes and works in the stables when she's pregnant. Prenatal exposure to stables and barns has the strongest effect. If exposure occurs when the child is 5 years old, it matters much less or not at all.

For *multiple discrete final states*, we already discussed how functionally and morphologically distinct cells (in our case, red and white blood cells) can derive from a single precursor. This process stresses two points. One is about plasticity, as we said, but the other is *partial irreversibility*. Once a cell becomes highly differentiated and its epigenetic differentiation program is fully implemented, this cell cannot go back. In fact, only stem cells retain plasticity all the time. For most other cells, the features acquired through epigenetic modifications are fixed and irreversibly preserved throughout life.

Now do we need to say the *L* we have been talking about is language? We think the genetic components of *L* are species-specificity and the common core (Universal Grammar) with room for large but highly constrained parametric variation (variation is going to become important to some extent, but it requires of course a robust common core). These components may correspond to FLN (the faculty of language in the narrow sense, in the terminology of Hauser et al. 2002). All the other plastic, dynamic components of *L*, we propose, are mechanistically implemented through epigenetic mechanisms – these could be the broader language faculty (FLB). We may have to go beyond this "division of labor" for another feature – the fact that *L* is or seems to be extremely robust, resistant to degradation, and also extremely stable, at least over a lifetime. From a strictly biological point of view, this feature suggests simplicity of design, because simplicity of design gives very high effectiveness. However, a simple design is also vulnerable to stress, unless it is balanced with some redundancy. The stability of a very small system is difficult to understand without postulating that somewhere, somehow, there is some compensatory repair pathway that allows a very compact core to repair. But this is even more speculative than our previous speculations.

Our last point, and this is entirely Massimo's doing, depicts two potential (alternative) scenarios: (1) *All parameters are innately specified*. This would put a very high burden on genetic encoding, something that we immunologists

are acutely aware of. And the problem of how you encode an enormous amount
of diversity in a limited genome would of course come back here. This possi-
bility would put very little or no burden on learnability. At the other end,
(2) *unconstrained variability*, would however put an excessive burden on learn-
ability. So I guess that what we are trying to say is that perhaps having
principles and parameters might represent an optimal compromise.

Discussion

DOVER: Epigenetics is a very active and important research field at the moment
and it is highly appropriate that you should attempt to link it to the supposed
difference between FLB and FLN as I understand it. But I need to add one
important caveat, which is that epigenetics is fast becoming a catch-all phe-
nomenon covering anything that moves in the workings of biology. The turning
on or off of any gene, whatever it's doing, requires the prior engagement of tens
upon tens of proteins which are the products of other genes of course. Now,
some of these other proteins are opening and closing the chromatin near
to our gene of interest in preparation for transcription; others are involved
with nearby DNA methylation; others with the initiation and termination
of transcription of the gene, and so on, so you can go on forever. If that is the
case, then everything is both epigenetic and genetic at one and the same time,
that is, no gene exists in a vacuum, its expression is carefully regulated and
depends on the state of its local chromatin, which in turn depends on the
comings and goings of many other gene-encoded proteins. In such a situation
we might well ask what is the real operational distinction between genetic and
epigenetic? Can this really be the basis to distinguish between *core* processes,
which are supposedly ancient and go way back, and the more recent *peripheral*
processes?

So just to get away from language, let me say something about legs, because it
is easier to make my point. We all have two legs, yet we all walk very differently.
Now it has long been thought that having two legs is one of those core, basic
things that universally characterizes our human species – any healthy fertilized
human egg will develop into an individual with two legs. But the shape and
manner of usage of legs, peculiar to each individual, is considered to be some-
thing peripheral, something that might be "epigenetically influenced" during
individual development. Now the whole point of Richard Lewontin's earlier
concept of "norms of reaction" (he might not have said this in precisely the same
way at the time, but it is certainly the way it's being interpreted now) is that the
developmental emergence of two legs, and not just the ways we use the two legs,
is as much "epigenetically" modifiable, and is as much a key part of that total

process of ongoing, ontogenetic nurturing that I spoke about earlier.[5] In other words, those complexes of genes that are involved in making two legs are no different in kind from the genes, or the very complex milieu of interactions of genes with genes, and genes with environment, that affect the individual shape and use of those legs. So it is very hard to distinguish between them, between "core" and "peripheral," given that this is happening from the moment a specific sperm enters a specific egg and on through each individual's highly personalized route of development.

Each individual's personal history of cell differentiation, tissue patterning, organogenesis, emergence of consciousness, language acquisition, and all the rest of it involves many complex and fluctuating networks of gene (protein) interactions, also subject to much environmental input. There is variation and constraint, simultaneously, at all times. The only thing we can be sure about is that, as a consequence of the sexual process of making sperm and eggs, we essentially get back to a genetic blank slate from which all human developmental processes, "core" and "peripheral," "genetic" and "epigenetic," "variable" and "constrained," need to re-emerge. Anything produced by evolution is bound to be a mess and even the original concepts of principles and parameters might be difficult to unravel when considering biological, ontogenetic processes and their inherently sensitive networks – but here I reach the edge of my understanding.

VERCELLI: I think we need to tread lightly because we are on tricky ground. That the development of an organism involves, as you put it, "many complex and fluctuating networks of gene (protein) interactions, also subject to much environmental input" I certainly will not deny. Nor will I argue against the continuous interplay between (and the likely co-evolution of) genetic and epigenetic mechanisms and processes, which at times may blur the distinction between them. But a distinction *does* exist and emerges when one thinks about the *kind of mechanisms* that may account for certain essential features of language as a biological trait. Some of these features (species specificity and uniqueness to humans, first and foremost) appear to be rooted so deeply and constrained so strongly that one would expect them to be inscribed in the genetic blueprint of our species – that is, to be genetically encoded. But most of the other defining features of language reveal a degree of *plasticity* in development and final states that best fits under the epigenetic paradigm. In other words, not everything in language is nurture – but not everything is nature either.

---

[5] See section 6.5 above.

104   DONATA VERCELLI AND MASSIMO PIATTELLI-PALMARINI

PIATTELLI-PALMARINI: Let me add to this the following: take the case you present of movement and the fact that we all have two legs and yet each walk differently. There is the famous two-thirds power law;[6] all biological movement obeys this two-thirds power law. All natural movement in humans and animals obeys the law that the two-thirds power of the ratio between linear speed and radius of gyration is always constant. It is universal and we immediately perceive it. Indeed, each one of us walks in a slightly different way. You can look at someone and say "Oh that's Jim, because see the way he walks." But it's very interesting to see that there is a universal law for biological movement. So, what are we interested in? The big effort that has been going on in language – we use different words, different accents, different tones of voice – but the big effort has been to go beneath these and see at what level there may be something universal, something that is common, that is deep. And it is no mean feat. You have seen these days what is in the lexicon, what is in the syntax, what is in the morpho-lexicon, what is in semantics – very, very difficult questions, all subdivided in order to deal with them one at a time. And so the FLB/FLN distinction is complicated to make, but it is a good way of distinguishing things, seeing which components are innate and which components are not. You are a geneticist but I have been a molecular biologist and continue to follow the field, so we both know that there are certain things you can do to genes with very specific effects. Of course, the effect of a gene on a phenotype usually depends on the effect of many other genes, that is called epistasis, and sometimes subtle or not so subtle effects come from apparently unrelated genes. But there are also clear examples of the effects of only one gene. For example, there is the outstanding phenomenon of Hsp90, with its chaperone protein which, if knocked out, gives rise to all sorts of mutations, all over the body of, say, a fruitfly.[7] That is, there are very specific things you can do to specific genes with very specific effects. Moreover, the distinction between genetic core processes and peripheral (also called exploratory) processes is unquestioned these days. I find it all over the current literature, often under the label of developmental and evolutionary modularity.[8] The biochemical pathways and their enzymes, for instance, just to name one clear case, are evolutionarily strictly conserved, often all the way down to bacteria.

DOVER: I don't think I've argued against genetics, otherwise I'd be out of a job; nor have I argued against universality, in terms of human-specific features which are shared by all humans. That's not my point. The point is that the

---

[6] Viviani and Stucchi (1992).
[7] Queltsch et al. (2002).
[8] For a vast panorama, see Schlosser and Wagner (2004).

ontogeny of a given individual is a highly personalized dynamic in which many factors are involved unavoidably nurturing each other. You cannot, with regard to the ontogeny of an individual, say that the "universal genes" and all their participatory networks for two legs are more of a "core" process than the genes and all their participatory networks for the manner in which we use those two legs. The two are ontogenetically unfolding together and there are many, many diverse and interactive influences at play in each unique individual – genes, proteins, environment, culture – the whole catastrophe!

Just one final thing: about the myth of the unique relationship between a specific gene and its very specific effect. First let us set aside the confounding property of rampant pleiotropy of most genes – that is, each and every gene having widely diverse effects at one and the same time – and let's just concentrate on one gene and one of its effects. Some of the best characterized of all molecular genetic diseases are the hemoglobin thalassemias. Now if you talk to David Weatherall and all those guys who have been working several decades on these genes,[9] they tell you the following. If you take a number of individuals, each of which has the identical mutation in say the beta-globin gene, which in turn is embedded in thirty kilobases of identical surrounding DNA (presumably with identical epigenetic patterns of chromatin condensation and methylation), you can then ask the question, what is the phenotype of all these individuals sharing the identical mutation in the same sequence neighborhood? Will they all have beta-thalassemia as part of their phenotype? And the surprising answer is "No." The disease phenotype is not just a specific effect of a specific mutation in a specific gene. They all have the specific mutant beta-globin allele but their phenotypes range from no clinical manifestations through to a requirement for life-long blood transfusions. This spectrum of effects arises because the rest of each individual's genetic background – all those other interactive genes (proteins) and metabolites, whether directly involved with blood metabolism or not, plus of course the internal and external environmental milieu – is absolutely crucial for the extent to which an individual goes down with beta-thalassemia. And the same story is emerging from the etiology of the majority of human diseases, once thought to be a specific consequence of single mutant genes. I think that in biology the pursuit of genetic subdivision, hierarchy, and specificity is not necessarily the appropriate approach to the seemingly indivisible, whether of legs or language. A recipe for despair or an exhilarating challenge?

FODOR: At the end of the presentation (I think this is perhaps especially Massimo's department), you had some speculations about the biological encoding of

---

[9] Craig et al. (1996); Weatherall (1999).

106   DONATA VERCELLI AND MASSIMO PIATTELLI-PALMARINI

parameters. I wondered if we could relate this somehow to some of the thinking we have been doing at CUNY about that huge grammar lattice of ours.[10] We worry about the biological status of this huge amount of information. I want to divide it into two aspects. One is that there is this huge amount of information, all those thousands of subsets of relationships; and then there is also the apparent specificity of the information. It codes for very particular relationships. This grammar is a subset of this one, but not this one of this other one, something like that. Now, wondering how that information got there, we should consider the possibility that it isn't really so specific at all, that in fact there are many, many other relationships equally coded but that they are invisible to us as linguists, as psychologists. We don't know about them because those languages aren't learnable, so imagine just for a moment you had two grammars in the lattice, so to speak the wrong way up, so that the superset came before the subset. Then we would never know of the existence of the subset language because nobody would ever learn it. It would be unlearnable. So you can imagine that behind the lattice that is visible to us as scientists there is a whole lot of other stuff just like it that we know nothing about because it is arranged the wrong way to be put to use by humans in learning. So: unlearnable languages. It may be that the specificity of the particular parameters that we know about is actually illusory.

PIATTELLI-PALMARINI: Well, this is really the core of the matter. I think that in the evo-devo approach to the evolution of language you have to take into account not just how we once got to the adult state; you have to take into account the whole process of getting there – how that evolved. And of course a very, very old puzzle is why we don't really have only one language. Since genetically we are predisposed to learn any language that there is, there is no specific inclination of a baby coming into this world in China to learn Chinese, nothing of the sort. So we have on the one hand the puzzle as to why we don't all literally speak the same language, and also on the other, why we don't have infinite variation beyond any limit, beyond any constraint. So the suggestion is that maybe what we have is a *minimax* solution, where you minimize the amount of genetic information and at the same time you optimize the amount of learning that there has to be in an acquisition somehow. Mark Baker (2001, 2003) has this hypothesis that the reason we don't all speak the same language is because we want to be understood by our immediate neighbors, but we don't want to be understood by people in the next tribe; which is a cute idea, but it really doesn't explain much, because you can only do that if you

---

[10] See sections 17.6–9 below.

already have an organ that is predisposed to have a large but finite set of possible languages. We could invent some codes that are different from having this parametric variation. So I think the consideration is in fact how complex the acquisition process is versus how much burden you have on the genetic or biological machinery. The guiding (and interesting) idea, in which Noam concurs, if I understand him correctly, is that you have a minimax, you have something close to the perfect compromise between loading the biology, loading the genetics, and having a reasonably complex acquisition process. You know, the things that you are doing and that Charles Yang is doing are closely related to this reflection.[11] We will have to learn from you how exactly these things developed, how much work has to be done there and then continue possibly with some data on other functions, on other species, to see if we can get a grasp on how much genetic information is needed for this or for that, and whether this hypothesis of a minimax solution can be tested.

FODOR: I guess I was trying to suggest that maybe there isn't as much biological design work to be done as we tend to think from our perspective, studying the particular cases, the particular languages that we observe, because in the case of language, if the design isn't optimal, we don't know about it, nobody is going to learn the language, nobody *has* to learn any particular language, so those languages just sort of disappear from view. So I am just wondering whether in fact there is so much specific biological design work going into what I still call universal grammar, and so the pattern of UG, as we tend to think.

VERCELLI: I can answer Janet's question only indirectly, using an intriguing analogy – that between the problem of encoding what there is in language, and the central problem my own field, immunology, faced for years. Our problem was to figure out how a large but finite genome could harbor a huge amount of information without clogging up. As you know, that problem was solved by an atomization of the encoding process, whereby the final molecular repertoire results from rearrangements of multiple, smaller units. That allows for a relatively limited core – then the information is rearranged and used, switched on and off. Systems of this level of complexity run into this kind of problem: how do you build information capacity effectively but not at the expense of everything else in a genome which is finite? The idea that you make space by erasing is a little hard for me to picture, because somehow you have to encode what you erase as well as what you don't. Thus, the encoding problem remains. I would argue a better way to solve it is, as Massimo was saying, by minimizing what you encode and then being very plastic in the way you use what you encode.

---

[11] See Yang (2002).