

Chapter 9

The Role of Vehicles in Natural Selection

The organism is just the gene's way of making another gene.

— Dawkins, *The Selfish Gene*

I coined the term 'vehicle' not to praise it but to bury it. . . . The question 'what is the vehicle in this situation?' may be no more justified than 'what is the purpose of Mount Everest?'.

— Dawkins, *Burying the Vehicle*

Lewontin argues that selection can act at many levels. But selection only results in adaptation if it produces a change in gene frequencies. Gene frequencies are the bottom line of evolution. Dawkins (1976), following Williams (1966), has used this fact to argue that natural selection should be understood as a process that *only* acts on genes, rather than other levels in the natural hierarchy. In this chapter I argue that the role of organisms (and other genetic 'vehicles') cannot be so easily eliminated.

The argument for genic selectionism starts from the fact that acquired characteristics are not inherited. In its modern interpretation this has become the Central Dogma of molecular biology which states that information cannot pass from protein to DNA, but only the reverse; i.e. changes to the phenotype incurred during development will have no effect on the DNA passed on to subsequent generations through the germ-line. Therefore, as Dawkins puts it, an organism is not the object of Darwinian evolution because

to regard an organism as a replicator . . . is tantamount to a violation of the 'central dogma' of the non-inheritance of acquired characteristics. A stick insect looks like a replicator, in that we may lay out a sequence consisting of daughter, granddaughter, great-granddaughter, etc, in which each appears to be a replica of the preceding one in the series. But suppose a flaw or blemish appears somewhere in the chain, say a stick insect is unfortunate enough to lose a leg. The blemish may last for the whole of her lifetime, but it is not passed on to the next link in the chain. Errors that affect stick insects but not their genes are not perpetuated. Now lay out a parallel series consisting of a daughter's genome, granddaughter's genome, great-granddaughter's genome, etc. If a blemish appears somewhere along *this* series it will be passed on to all subsequent links in the chain. It may also be reflected in the bodies of all subsequent links in the chain, because in each generation there are causal arrows leading from genes to body.

But there is no causal arrow leading from body to genes. No part of the stick insect's phenotype is a replicator. Nor is her body as a whole. It is wrong to say that 'just as genes can pass on their structure in gene lineages, organisms can pass on their structure in organism lineages' . . .

The special status of genetic factors rather than non-genetic factors is deserved for one reason only: genetic factors replicate themselves, blemishes and all, but non-genetic factors do not. (Dawkins, 1982, p99)

This picture of evolution is summed up in Weismann's famous diagram¹ in which arrows of causal influence run solely from the genome of one generation to that of the next, with the expressed phenotype being no more than an epiphenomenal offshoot from the main branch: the organism is just the gene's way of making another gene (figure 9).

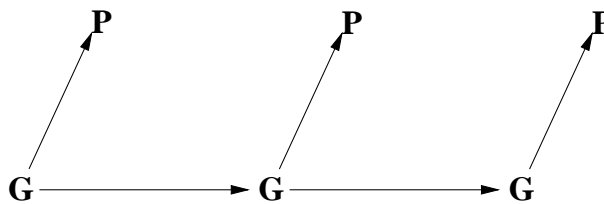


Figure 9.1:

Genes produce organisms, but genes are also *parts* of organisms. And the organism can bite back and affect the genes that they contain. This produces two kinds of problems with Weismann's diagram, depending on how those effects are understood. The first is that the replication of genes is *causally* dependent on organisms, and this dependence is the subject of the next section. The second is that the replication of genes is *conceptually* dependent on organisms, and this is the subject of sections 9.2 and 9.3. And in section 9.4 I show how we need to understand both kinds of dependency in order to understand evolution.

9.1 Burying Vehicles

Mayr (1963) describes two ways in which phenotypes causally effect the genes they contain. The first is that 'natural selection favours (or discriminates against) phenotypes, not genes or genotypes' [p184] (see also (Gould, 1978, p90) and (Brandon, 1982)). Gene frequencies change because of the effects they have on the ability of organisms to reproduce. The more successful the organism, then the more its genes will spread. The second argument is what Mayr calls 'the genetic theory of relativity', which is that 'no gene has a fixed selective value; the same gene may confer high fitness on one genetic background and be virtually lethal on another' [p296] (see also (Sober & Lewontin, 1982) and (Gould, 1978, p91)). The fitness of a gene — i.e. the rate at which its frequency changes — depends on its context (including the rest of the genome, the organism, and its environment) and is not a fixed property of the gene-in-itself.

Evolution, like all other biological processes, is a dense web of interacting causes operating at, and between, different levels of organisation. Mayr's argument is that we should pick out organisms as playing a privileged causal role in this process; but this argument stands or falls

¹This diagram has often been attributed to Weismann, though I can find no record of him actually using it.

according to how one chooses to define cause. Dawkins' defence of genic selectionism starts from the following point:

Philosophers, possibly with justification, make heavy weather of the concept of causation, but to a working biologist causation is a rather simple statistical concept. Operationally we can never demonstrate that a particular observed event *C* caused a particular result *R*, although it will often be judged highly likely. What biologists in practise usually do is to establish *statistically* that events of class *R* reliably follow events of class *C* . . . Statistical methods are designed to help us assess, to any specified level of probabilistic confidence, whether the results we obtain really indicate a causal relationship. (Dawkins, 1982, p12)

This Humean definition of cause dissolves the causal boundaries that pick out vehicles as ineliminable actors in evolution. According to this definition natural selection works *through* vehicles, rather than *on* them. To see why this is consider each of Mayr's objections in turn.

Mayr's first argument was that phenotypic properties, not genes, cause differential reproduction. Dawkins' response depends on causation being transitive. According to Dawkins' definition, if *C* causes *E1* and *E1* causes *E2* then there must be a correlation between each pair. This further implies that *E2* will be correlated with *C*, and hence that *C* causes *E2*. Now Mayr argues that phenotypic differences (*E1*) cause differential reproduction (*E2*), but those phenotypic differences (*E1*) are themselves caused by genetic differences (*C*). If causation is transitive then it is equally true that genes (*C*) cause differential reproduction (*E2*).

Mayr's second objection was that the effects of genes are context dependent. However *given* a particular context then allelic differences will be correlated with phenotypic differences (and hence with fitness). Therefore according to Dawkins' definition of cause it is still correct to say that a particular allele has a causal effect on fitness. For example, Lewontin and Sober (1982) discuss the case of heterozygote superiority in which a heterozygote (*Aa*) is fitter than either homozygote (*AA* and *aa*). In the case of human sickle-cell anaemia, for example, homozygotes for the normal allele have functional haemoglobin but are vulnerable to malaria, homozygotes for the mutant allele suffer anaemia, and heterozygotes are resistant to malaria *and* avoid anaemia. The effects of each individual allele (*A* or *a*) depends on the allele it is paired with, and so Sober and Lewontin conclude that they cannot be attributed with a unique causal role: 'if a gene raises the probability of a given phenotype in one context and lowers it another, there is no such things as the causal role that the gene has in general' (Sober, 1985, p313). However Sterelny and Kitcher (1988) argue that evolution does not require that a gene has a causal role 'in general', but only in specific contexts. A genic selectionist can perfectly well argue that an *A* allele is fitter than *a* when paired with an *a*, but less when paired with an *A*.

Both of these objections to Mayr's arguments prove only that the gene-centrist view is as valid as the vehicle-centrist and, in his first book, Dawkins compared the two points of view to the two possible views of the Necker Cube: neither is more correct than the other, they are just different views of the same process. However Dawkins later presented an additional argument that questions the validity of the vehicle-centrist view *per se*. This is the concept of the *extended phenotype* (1982, ch11–13). Consider the genes that contribute to the beaver's habit of building dams. What matters for the successful replication of these genes is not the particular actions of the beaver — the bites it takes out of trees, the driftwood it collects, and so on — but the size of the artificial

lake that it produces. The bigger the lake then the more protected the beaver will be against land-based predators and the easier it will find food. This turns the vehicle-centrist's arguments against themselves in a kind of *reductio* of Mayr's first objection: if what causally matters for evolution is the phenotypic effects of genes, not the genes themselves, then individuals do not have a privileged position in the ontology of evolution. The spread of beaver genes depends on the properties of beaver lakes, not the properties of beavers themselves. Causally speaking, individual organisms are no more, and no less, than one part of the extended phenotypic environment through which genes replicate themselves. The vehicle is buried, in Dawkins' phrase (1994).

If Dawkins (and Hume) are right about the nature of causation, then Mayr (and Lewontin) are wrong that natural selection acts on vehicles. There are two possible responses to this. The first is to argue that Dawkins is wrong about cause, as Sober does (1985). I believe that Sober's criticism is valid, but defining sufficient conditions for causation capable of supporting this claim would take us too far from the direct concerns of this thesis². Causation is a minefield that I would rather avoid if at all possible. However there is another way in which individual organisms play an ineliminable role in our understanding of natural selection. This role is *conceptual*, rather than causal, and is the subject of the rest of this chapter.

9.2 Counting Genes

What is a gene? The Mendelian answer is that a gene is a unit of heredity, but this describes what a gene *does*, not what it *is*. The biochemical answer is that a gene is a sequence of DNA, but not all possible DNA sequences are genes. So what makes a particular sequence a gene? The original answer to this question was given in the slogan 'one gene, one protein', i.e. a gene is a sequence of DNA that codes for a protein, but it is now clear that the situation is more complicated than that.

In prokaryotes, genes and the proteins they code for are co-linear; that is the sequence of amino acids in each protein is represented by a corresponding unbroken sequence of codons in the DNA. Therefore starting from the slogan 'one gene, one protein' we can clearly identify a gene as a contiguous sequence of codons on a DNA molecule. However since 1977 it has become apparent that the situation in eukaryotes is rarely that simple. Instead of forming a contiguous sequence the gene may be realised in a 'mosaic' of parts spread across the genome. The initial pre-mRNA transcript of the mosaic is then 'spliced': a process in which sections, *introns*, are edited out. The dihydrofolate reductase (DHFR) gene, for example, can vary in length from 25–31kb depending on the mammal it occurs in. Most of this variation lies in the sequence, position, and length of the introns which play no active role, however some of these introns themselves code for proteins that function independently of that coded for by the remaining exons. Moreover the exons and introns in the original DNA sequence may be spliced in many different ways depending on the actions of other regulatory genes, or on the presence of ancillary proteins. Thus two or more mosaics may intermingle across the same stretch of DNA (Breathnach & Chambon, 1981) (Wu, 1978).

Genes can also overlap on the DNA such that the same sequence of nucleotides code for more than one protein. One functional gene is often simply a truncated version of another, but in other cases the two expression processes may be read in different frames such that the divisions between

²In chapter 2 I only proposed a *necessary* condition on causal explanation (i.e. that the cause of *A* has an identity independent of *A*), and avoided the problem of sufficiency.

the codons for one gene do not coincide with those of the other. In such situations we cannot tell which amino acid — let alone which protein — a base pair codes for without observing the process of expression in action.

Genes do not even have to be tied to a particular place on a chromosome. Transposable sequences, or transposons, are capable of ‘jumping’ from one location to another either by detaching themselves from the main DNA sequence or by inducing the replication of copies that subsequently insert themselves at a new location. Most transposons are likely to be entirely ‘selfish’, with no further phenotypic effects, but it is also possible that they are a major source of effective mutations (Berg & Howe, 1989).

Genes may also be polymorphic; i.e. they may be encoded in a variety of different DNA sequences. We usually describe different sequences at a single locus as alternative alleles, but it is also possible to describe them as alternative forms of the same gene. Again the crucial factor is the effect on phenotype. There is a continuum of types of change of DNA sequences, including those that change DNA sequence but not protein sequence, those that change protein sequence without changing its secondary structure and/or function, those that create proteins with different activities, and those that create mutant proteins that are non-functional. Where on this continuum we draw the line between alternative forms of the same gene and alternative alleles depends on their role in the overall metabolism of the organism (Gusella, 1986).

In short, a gene is defined ‘semantically’ in terms of the role that it plays in the metabolism and development of an organism, not ‘syntactically’ in terms of a DNA sequence. As Lewin puts it

Genes can be isolated by working back from a protein . . . The concept of the gene itself, however, has recently evolved further. The question of what’s in a name is especially appropriate for the gene. We can no longer say that a gene is a sequence of DNA that continuously and uniquely codes for a particular protein. In situations in which a stretch of DNA is responsible for production of one particular protein, current usage regards the entire sequence of DNA, from the first point represented in the messenger RNA to the last point corresponding to its end, as comprising the “gene”, exons, introns, and all.

When the sequences representing proteins overlap or have alternative forms of expression, we may reverse the usual description of the gene. Instead of saying “one gene — one polypeptide”, we may describe the relationship as “one polypeptide — one gene”. Thus we regard the sequence actually responsible for production of the polypeptide (including introns as well as exons) as constituting the gene, while recognising that from the perspective of another protein, part of this same sequence also belongs to *its* gene. (Lewin, 1997, p146)

Suppose a copy of the DHFR gene is replicated, but in so doing mutates into one of its alternate forms. Should we describe this process as the successful replication of a single gene, or the death of an old gene and the creation of a new one? Has there been a change in gene frequencies? The answer depends on the effect of the new sequence on the organism. The concept of ‘gene frequency’, like that of ‘gene’ itself, is dependent on phenotypic properties.

Genes do not just determine the structure of proteins. Their other main function is to regulate the production of proteins, and this also raises problems in individuating genes. Sometimes a regulating gene will directly control the transcription of a neighbour. In the absence of this expressor

the regulator is functionally neutral, like a switch that is not connected to anything. In other cases the regulation is more subtle and long-range, involving the production of intermediate ancillary proteins. Gene regulation can also be affected by environmental factors that induce expression, or affect the initiation and termination of transcription (Reznikoff et al., 1985)(Platt, 1986). If the same DNA molecule is put in another organism, or *in vitro*, then different genes may be expressed.

Regulation can also occur at levels above that of a single gene. Eukaryotic genes that code for proteins whose functions are related are often organised into clusters, such as the three *lac* genes in *E. coli* which code for enzymes that decompose, transport, and aid the metabolism of, lactose. These three genes are grouped together on the genome and respond to a single regulator, forming a functionally unified *operon* that responds swiftly to the presence of lactose in the environment (Jacob & Monod, 1961). Therefore from the point of view of Mendelian analysis it is the operon, rather than its constituent genes, that is the unit of heredity.

The complexities of gene regulation mean that an apparently simple question, such as the number of genes on a genome, does not necessarily have a simple answer:

The major question about eukaryotic DNA concerns the number and types of genes in a genome. We may identify the coding potential of a genome directly, by identifying regions that have open reading frames. Large scale mapping of this nature is complicated by the fact that genes are interrupted in higher eukaryotic genomes, so that many separated open reading frames may be part of a single gene. . . . Since we do not necessarily have information about the functions of the protein products, or indeed proof that they are expressed at all, this approach is restricted to defining the *potential* of the genome. . . .

Another approach is to define the number of genes directly in terms of their expression in RNA or protein. This gives an assurance that we are dealing with *bona fide* genes that are expressed under known circumstances. It is of course the only approach that allows us to ask how many genes are expressed in a particular tissue or cell type, what variation exists in the relative levels of expression, and how many of the genes expressed in one particular cell are unique to that cell or are also expressed elsewhere. (Lewin, 1997, p645)

The first of these approaches tells us the potential capacity of the genome, but may not distinguish between functional genes, pseudogenes³, and ‘junk’ DNA. The second of these approaches can tell us more precisely how many functional genes are present, but then this number will be dependent on the cellular (and wider) environment: put the same genome in a different cell, or *in vitro*, and the number of genes changes. Gene frequencies are not determined solely by DNA sequence.

Genes, like mental states, are functional entities, defined by the relationship between the substrate in which they are realised (DNA), and its environment (the living organism). Genes are the role that DNA plays in the development of an organism, just as mental states are the role that brain states play in behaviour. If you take a DNA molecule out of its chromosomal, nuclear, cellular, and organismic environment then it does not contain any genes, just as a slice of brain tissue on a slide does not contain beliefs. Of course in order to be accorded a well-defined causal role — i.e. in order to understand how they do the job they do — it is necessary to understand how these functional genes are realised in DNA, just as we had to understand how mental representations

³Sequences of DNA that are similar to functional genes but play no active role.

were realised in brain states in order to understand how they played a causal role in our behaviour; but this should not tempt us into forgetting that it is the functional role that matters. Genes are emergent properties of the interaction between DNA and its environment. It is a truism that genes cannot replicate, or have any developmental effects, outside of the appropriate cellular and organismic environment. This is true, but is only half the story. Strictly speaking, genes do not even *exist* outside of the appropriate environment.

9.3 Counting Replicators

Natural selection necessarily involves a change in gene frequencies. But gene frequencies depend on how DNA expression is regulated by the organism. So perhaps we should do the ‘bookkeeping’⁴ for evolution in the hard currency of nucleotide sequences, rather than the fluid terms of genes? But what length of sequence? Is the unit of selection a single nucleotide, or is it the entire genome? Dawkins argues that any sequence can be a unit of selection — an ‘active replicator’ — so long as it has some effect on the phenotype, and so responds to selection (1982, p90–1).

The reason for measuring evolution in terms of arbitrary lengths of DNA rather than organism-dependent genes is that only changes in the universal currency of DNA matters for the evolutionary future. However not all DNA matters equally. Recall Dawkins’ argument why we cannot understand evolution solely in terms of the replication of organisms; namely that ‘errors that affect stick insects but not their genes are not perpetuated’. This is true, but does *not* imply that all errors that affect stick insect’s genes will be passed on either. Unless those affected genes are part of the stick insect’s germ-line cells, and unless those cells go on to form a viable zygote, then, like the stick insect’s lost leg, those genetic errors will not be passed on. If stick insects are not replicators then neither are their non-germ-line genes.

The gene-centrist view rests on a principled distinction between a replicator and the vehicle through which the replicator survives into the next generation. It is the fate of the replicator that matters, not that of the vehicle. Therefore we have to distinguish between the replication involved in vehicle-building and the replication involved in lineage-building — what Dawkins describes as germ-line, rather than dead-end replication.

The same distinction underlies the problem of ‘head-counting’ (Sober, 1985, p29). Suppose two land-dwelling organisms live on a rock surrounded by water. One has copies of gene *A* and gets fat, while the other has copies only of *B* and gets thin. The proportion of *A*-type *cells* in this two-organism population increases, but this does not constitute *adaptive* evolution: fatness does not constitute fitness. The same point also applies to the replication of sterile workers in groups of eusocial insects. The production of more workers increases the total number of the genes they carry, and may strengthen the nest, but they are as much as an evolutionary dead-end as stick-insect legs. The only germ-line replication in this case is that involved in the production of new swarming queens. In other cases there may be no prior separation between dead-end and germ-line cells. For example the apical cell of a strawberry plant runner may die after fruiting, or the runner may take root and form an offspring that lives on after the death of the original. The development of the plant thus turns a dead-end cell into a germ-line replicator.

⁴This term is due to Wimsatt (1980).

Consider another example. Many cancerous tumours are induced through the actions of proto-oncogenes that are present in the cells of the host organism and normally play a role in healthy development. The cellular proto-oncogene *c-myc*, for example, codes for a protein that helps initiate RNA transcription. These oncogenes can be implicated in the growth of cancerous tumours in two ways. The first involves the host being infected with a retrovirus, such as the mouse mammary tumour retrovirus (MMTV), which changes the way the cellular oncogene is regulated. The resulting tumour may then help the virus propagate. Or a tumour may form without viral infection due to a mutation in the regulatory genes of the mouse cell (Bishop, 1983) (Bishop, 1985) (Varmus, 1984) (Heldin & Westermark, 1984).

In both of these cases a change in the regulation of *c-myc* causes a tumour to grow. The lineage of the original cancerous cell will grow explosively, but will die with the host. But the evolutionary stories of each of these processes is very different. In the former case the replication of *myc* is a vital part of the reproductive cycle of the virus. In the latter case the cancerous replication is simply a by-product of having growth factors whose regulation is vulnerable to mutation. Thus a single act of DNA replication can be a dead-end with respect to one vehicle (the mouse) and constitute the replication of the germ-line of another (the virus).

Is cancer an adaptation and, if so, who for? In the case of mouse mammary tumours both the healthy and tumourous cells, *and* the virus, all contain the same *myc* gene. Therefore we cannot say that the cancer is an adaptation for *myc* genes *simpliciter*, but only with respect to the particular lineage of reproducing vehicles that carry it. It may be an adaptation for the *myc* of MMTV, or for the *myc* of a tumourous cell, but it is certainly not an adaptation for the *myc* of the healthy cells of the mouse. Unless we identify the vehicle through which the gene is transmitted then we have not yet answered Dawkins' question.

The point of doing our evolutionary bookkeeping in terms of DNA sequences is to determine whether evolution is occurring. But in order to do this we cannot just count the total number of copies of a sequence in the ecosystem, since it is only germ-line replication that matters for evolution. But what constitutes germ-line replication? Dawkins defines it as follows:

A germ-line replicator is a replicator that is potentially the ancestor of an indefinitely long line of descendant replicators. A gene in a gamete is a germ-line replicator. So is a gene in one of the germ-line cells of a body, a direct mitotic ancestor of a gamete. So is any gene in *Amoeba proteus*. So is an RNA molecule in one of Orgel's (1979) test-tubes. A *dead-end replicator* is a replicator which may be copied a finite number of times, giving rise to a short chain of descendents, but which is definitely not the potential ancestor of an indefinitely long line of descendents. Most of the DNA molecules in our bodies are dead-end replicators. They may be ancestors of a few dozen generators of mitotic replication, but they will definitely not be long-term ancestors. (Dawkins, 1982, p83)

The problem here lies in the words 'potential' and 'indefinite'. In one sense *all* DNA has the potential to replicate indefinitely, since we are able to extract the nucleic acid from the living organism and replicate it *in vitro*. In another sense *no* DNA can have an indefinite chain of descendents since the universe is bounded. Nonetheless the distinction between germ-line and dead-end replicators seems to have a clear biological basis; but what? Whether or not a DNA sequence

is capable of being passed on ‘indefinitely’ depends on the development and reproduction of the organism that it is part of.

Weismann (1904) believed that only germ-plasm contained the material of inheritance which was then irreversibly converted into differentiated somatic cells. According to this picture the distinction between germ-line and dead-end replication is explicit at the level of the individual cell: germ-line cells contained the material of heredity, but somatic cells did not. Weismann’s basic idea of a complete separation of germ-plasm from its phenotypic expression was absolutely correct but the mechanism he chose for making this distinction was not. We now know that *both* germ and somatic cells contain the material of heredity. Moreover some somatic cells are totipotent — i.e. they retain the capacity to produce both gametes and somatic cells. This is obvious in the case of plants which are able to grow from cuttings taken from almost anywhere on the adult. Indeed the development of cloning techniques means that, in a sense, *all* cells in *all* organisms are (potentially) totipotent. The division between germ-line and somatic cells is not as clear-cut as Weismann’s diagram implies. Whether or not a cell is destined to be a dead-end replicator, or has the potential to form part of the germ-line of a new generation, is not an intrinsic property of that cell.

Weismann pictured the relationship between development and evolution as involving a clear differentiation between somatic and germ-line cells (figure 9.2).

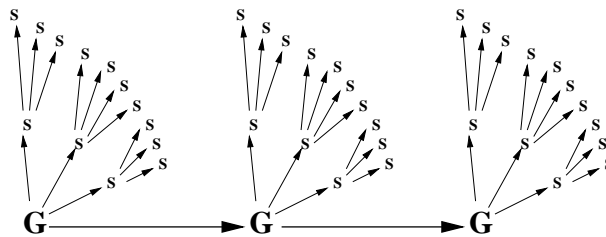


Figure 9.2:

But we now know that, in general, there is no such systematic differentiation. Whether or not a particular cell ends up as the germ-line for a new generation depends on the peculiarities of individual development (figure 9.3). Which cell ends up as the germ for a new generation is not fixed in advance⁵.

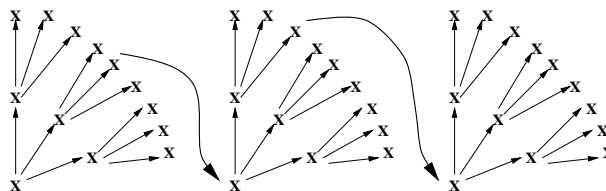


Figure 9.3:

In some situations — such as sexual reproduction — there is a more fixed differentiation between germ-line and somatic cells, and so the process of reproduction approximates more closely

⁵Dennett (1994) compares this to the ‘veil of ignorance’ in Rawls theory of justice.

to figure 9.2. But in many other cases — such as vegetative reproduction — it will not. Of course Dawkins is correct that all mutations must be instantiated in genes if they are to be inherited. But it is not determined which cells contain the genes that matter. As Buss puts it ‘knowledge of the molecular mechanics of heredity is *not* equivalent to knowledge of the units that prove heritable . . . Development controls heritability’ (1987, p14).

It is interesting to note that Weismann did not invoke his (mistaken) beliefs about germ-plasm to justify his assertion that acquired characteristics were not heritable⁶. Instead he based this argument on empirical facts about, for example, the pups of mice whose tails have been cut off, or the children of fathers who acquired duelling scars. Both of these traits were popularly believed to be inheritable — indeed I can remember being taught that Manx cats lost their tails in the same way — but Weismann demonstrated empirically that they were not. He also argued that it would be inconceivable that, for example, the well-developed fore-arms of a blacksmith could produce the appropriate changes in his sperm necessary to pass on the trait to his children. It would be akin to writing a telegram in English, sending it to China, and finding it arrived translated into Mandarin, as Weismann put it.

Weismann’s beliefs about germ-plasm stem from a consideration of development, not evolution. He argued that the development of large scale morphology in multicellular organisms could be controlled by genetic material through only two possible mechanisms. The first would be by having the entire genome copied into each new somatic cell. But then factors from outside the cell would have to ‘switch on’ the appropriate genes in the cells in all the various positions in the developing embryo: the ‘leg genes’ in the cells at the bottom, ‘liver genes’ in those in the middle, and so on. Weismann could not imagine how such a complex and delicate choreography of switches could be controlled so instead he plumped for the alternative, which was that each somatic cell would only inherit those genes necessary for their particular development. Liver cells would *only* contain liver genes, and so on. One look at a flower growing on a plant taken from a cutting should have persuaded Weismann that the former mechanism, however improbable, was responsible; but it is only relatively recently that we have started to scratch the surface of how this developmental choreography is possible.

To put it crudely, Weismann believed that given a phial of germ plasm and somatic cells then he would be able to distinguish between them. This is not true in general. Germ-line and dead-end cells contain the same DNA, and are only distinguished by the role they play in the development and reproduction of an organism. In this sense germ-line replicators are similar to straws that break camels’ backs: both are only distinguished by the role they play in the fate of a vehicle. When we look at a field of straw we cannot pick out the fatal ones. Similarly, we cannot look at a collection of cells and pick out the germ-line replicators. The object term ‘the straw that broke the camel’s back’ not only picks out a straw but also the camel for which it proved fatal; and the term ‘germ-line replicator’ not only picks out a sequence of DNA, but also the lineage of vehicles whose germ-line it is part of. In some cases, such as the *myc* gene in a tumour cell of a MMTV-infected mouse, a single DNA sequence instantiates two different germ-line replicators, each defined with respect to a different lineage of vehicles. The *myc* may be replicated either through propagation of the virus or through the spread of the tumour, and each lineage will define different adaptive

⁶Thanks to John Maynard Smith, personal communication.

pressures. In some cases the vehicle of a germ-line replicator may be what we naively think of as an individual organism. In other cases, such colonies of eusocial insects, it will be a higher group of organisms. In cases of ‘molecular drive’ (Crow, 1979) — in which genes multiply out of phase with the organisms they are in — the vehicle is lower in the hierarchy than that of the individual organism.

Dawkins *et al* argue that the bookkeeping of evolution should be done in units of germ-line replicators. But in order to pick out germ-line replicators from dead-end replicators we have to identify the vehicle that it is a germ-line replicator of⁷.

9.4 Fitness

In the last chapter I argued that the concept of ‘inheritance’, when clarified, reveals a genetic mechanism. Similarly the concept of ‘germ-line replication’, when clarified, reveals a reproducing vehicle. Therefore, although the bookkeeping for evolution may be done in the units of DNA sequences, the thing being measured is the reproduction of vehicles. Fatter is not fitter, even though it produces more copies of a gene.

This distinction between the property being measured and the units in which we measure it is explicit in the standard population geneticist’s definition of fitness:

The fitness of a particular type, *A*, is the expected number of offspring contributed by an *A* individual to the next generation. Fitness is estimated from one particular stage in the life cycle — usually the zygote — to the corresponding stage in the next. . . . Fitness is a property, not of an individual, but of a class of individuals — for example, of individuals homozygous for allele *A* at a particular locus. . . . Usually we ascribe fitness to a ‘genotype’, meaning a class of individuals with some genetic characteristic in common. . . . *Fitness is a property of a class of individuals, and not of genes.* (Maynard Smith, 1989b, p36-7, emphasis added)

Therefore when we say that a trait improves the fitness of (i.e. ‘is for the good of’) a gene, we are implicitly saying that it increases the ability of a vehicle to pass it on. Given the presence of a gene in a population the questions we must ask are not only how it aids fitness, but what vehicle it aids the fitness *of*. And the answer to this question is not always straightforward:

Genes are normally passed on by the reproduction of their vehicles, therefore fitness is usually operationally defined as the expected number of offspring of a individual member of a genotype. This definition not only covers the ‘usual’ cases of sexual or asexual reproduction of individual organisms, but also the case of transposons and other mobile DNA in which the vehicle of the gene is the DNA sequence itself: every act of replication is thus simultaneously a case of vehicle reproduction. But for the genes of infectious parasites it is not the rate of reproduction that matters *per se*, but the rate of transmission to new hosts. And parasites have evolved many ingenious mechanisms for achieving this, such as the diarrhoea induced by the cholera bacterium or the open sores produced by syphilis. Indeed in some cases too much reproduction can *reduce* the fitness of the genes of the parasite since they may kill the host before it is capable of infecting others. Therefore there is an evolutionary pressure to become *less* virulent. For example, when syphilis

⁷Indeed Hull argues that we should abandon the terms ‘organism’, ‘vehicle’, and ‘individual’ in favour of one that makes the concept of reproduction central. Thus he defines selection in terms of ‘interactors’: i.e. an entity that interacts as a cohesive whole with its environment *in such a way that reproduction is differential* (1980, p318).

first reached Europe at the end of the fifteenth century it could cover an entire body with pustules, cause flesh to fall off people's faces, and cause death within months. Now it has evolved into a much less virulent form that rarely kills the host, even if it is untreated. Therefore epidemiologists measure the fitness of such diseases in terms of the rate of transmission between host vehicles, rather than the replication of the vector or its genes (Williams & Nesse, 1991)(Ewald, 1994)⁸.

Recall that

the whole purpose of our search for a 'unit of selection' is to discover a suitable actor to play the leading role in our metaphors of purpose. We look at an adaptation and want to say, 'It is for the good of ...'. Our quest is for the right way to complete that sentence. It is widely admitted that serious error follows from the uncritical assumption that adaptations are for the good of the species. I hope to be able to show that yet other theoretical dangers, albeit lesser ones, attend the assumption that adaptations are for the good of the individual organism. I am suggesting here that, since we must speak of adaptations as being for the good of something, the correct something is the active germ-line replicator. (Dawkins, 1982, p91)

Dawkins is correct to warn against the assumption that an adaptation is good for the individual organism that carries it. However arguing instead that the adaptation is for the good of 'the germ-line replicator' does not yet answer the question, but just generalises it: germ-line replicators are always germ-line replicators *of* a lineage of vehicles, and these vehicles must be identified in order to answer the question.

In section 9.1 I argued that, although Mayr *et al* were correct to argue that individual vehicles play a causal role in the processes of selection that change the frequency of a gene, this was equally true of the rest of the extended phenotype. Therefore vehicles are not distinguished as an essential part of the characterisation of natural selection by their causal role *per se*. However in section 9.3 I argued that vehicles play an ineliminable *conceptual* role in picking out which DNA sequences count as germ-line replicators: i.e. the units in which the bookkeeping of evolution must be measured. When these two arguments — the causal and the conceptual — are put together the conclusion is that when we describe a trait as 'good for a germ-line replicator' we implicitly mean that it contributes to the ability of a *vehicle* to pass on more of its genes to the next generation. Beaver lakes can increase the frequency of beaver genes in two ways. The first is by producing more beavers. The second is by producing fatter beavers. Only the first constitutes evolution.

9.5 Conclusion

Lewontin *et al* argue that natural selection occurs whenever we have the differential reproduction of individuals on the basis of their inherited traits. This is true, but when we 'clarify' the notion of inheritance we find that correlations between parents and offspring must be mediated by a genetic mechanism. This implies that a change in gene frequencies is a necessary condition for natural selection. Dawkins argued that a change in gene frequencies is also a *sufficient* condition for natural selection. But the only changes in gene frequencies that matter are those that measure the

⁸This point will be important when we consider the analogy between infectious diseases and infectious ideas in section 10.2.

ability of a vehicle to pass on its genes to others. In the next chapter I will discuss how we can transfer this general characterisation of natural selection from natural history to social history.