This chapter to be published as:

Burian, Richard (2009). "Selection Does Not Operate Primarily on Genes" in Ayala, Francisco and Arp, Robert, *Contemporary Debates in Philosophy of Biology*. Oxford: Wiley-Blackwell.

Selection Does Not Operate Primarily on Genes

Richard M. Burian

This chapter offers a review of standard views about the requirements for natural selection to shape evolution and for the sorts of 'units' on which selection might operate. It then summarizes traditional arguments for genic selectionism, i.e., the view that selection operates primarily on genes (e.g., those of G. C. Williams, Richard Dawkins, and David Hull) and traditional counterarguments (e.g., those of William Wimsatt, Richard Lewontin, and Elliott Sober, and a diffuse group based on life history strategies). It then offers a series of responses to the arguments, based on more contemporary considerations from molecular genetics, offered by Carmen Sapienza. A key issue raised by Sapienza concerns the degree to which a small number of genes might be able to control much of the variation relevant to selection operating on such selectively critical organs as hearts. The response to these arguments suggests that selection acts on many levels at once and that sporadic selection, acting with strong effects, can act successively on different key traits (and genes) while maintaining a balance among many potentially conflicting demands faced by organisms within an evolving lineage.

1. Introduction

This chapter addresses how selection "acts" on genes, organisms, and lineages of organisms and examines how to determine what sorts of "units" are selected.[1] I'll start by saying a few things about selection and about genes, and then review some traditional arguments about genes as units of selection. After that, we will

---

[1] The term *unit of selection* is due to Lewontin (1970), which is a fundamental paper for the theory of selection. A useful anthology of the literature to 1983 on this topic is Brandon and Burian (1984); an excellent brief review is Mayr (1997).

examine some of Sapienza's arguments in the companion to this chapter and develop some arguments against the view that selection acts primarily on genes.

2. Four Criteria for Natural Selection

This is not the place to examine closely what natural selection is or how it works. Ever since Darwin (1859) used artificial selection as a model for natural selection, it has been important to recognize that natural selection differs in important ways from other forms of selection. The conditions set forth here concern *natural* selection, not selection in general. It suffices to set out four criteria that, according to a wide consensus, must be met for selection to influence the course of evolution. The criteria include (1) variation that (2) affects fitness, and (3) is heritable. Furthermore, (4) environmental conditions must be sufficiently regular to permit selection to be cumulative.

The entities subject to selection must go through iterated generations and their properties (traits) must vary from individual to individual. When a biased sample of available variants survives over a series of generations, selection *may* be occurring. Effective trans-generational selection requires there to be heritable variation in fitness (Lewontin, 1970) and conditions permitting cumulative selection. *Variation* (criterion #1) is required for entities with different properties to be present in the next generation(s). Evolutionarily relevant variation must affect *fitness* (criterion #2), where fitness differences are, by definition,

differences in *the propensity to survive and reproduce.*[2] In addition to variation

and differences in fitness, *heritability* is required (criterion #3) – i.e., a positive

correlation, on average, between the trait of the parent(s) and the offspring. This

is necessary for differential survival to allow selection to favor or alter certain

traits of the selected entities within lineages. If an entity's traits had no

correlation with those of its offspring or with the distribution of those traits in the

next generation, cumulative change over long periods would be extremely

difficult to achieve and (if it occurred) would probably not be caused by selection.

Finally, if, the entities in question faced different regimes with the odds of

survival shifting arbitrarily in direction in every generation, the distribution of

traits or phenotypes would zigzag with no cumulative effect. Thus, the

environment (and environmental cycles) must be sufficiently regular for selection

to be effective (criterion #4). The rationale for combining all four conditions is

nicely presented in Sterelny and Griffiths (1999, pp. 29 ff.). Their initial example

of cumulatively effective selection in organisms is the cryptic coloration of a

ground-dwelling bird living in a reedy marsh habitat. If the habitat varied rapidly

enough, cumulative stabilization of the requisite pattern of mottling and barring

would be most unlikely.

---

[2] The addition of "and reproduce" causes difficulties because conditions increasing the
likelihood of survival may conflict with conditions favoring reproduction. Again, exactly
how should we calculate the number of offspring relevant to *reproductive survival*? Take
a pair of nesting birds: should we count the number of eggs, the number of hatchlings, the
number of fledglings, or the number of sexually mature offspring? If some offspring are
sterile, should that reduce the count?

Even without a definition of natural selection, we have set out fairly stringent necessary conditions for effective selection to occur: heritable variation of fitness, with sufficient regularity of (environmental) conditions to allow cumulative selection to maintain or alter the traits of the organisms (or other relevant entities) over a long series of generations. The four conditions do not, as such, exclude natural selection from acting primarily on genes, which reproduce over a series of (cellular or organismal) generations. Although *illustrations* of selection in terms of *organisms* may prejudice one against selection acting *primarily* on genes, the four conditions set forth above do not, by themselves, reinforce that prejudice.

3. Genes

The term *gene* is about a century old. A Danish biologist, Wilhelm Johannsen, coined it in 1909 to stand for something almost wholly unknown (Johannsen, 1909). He thought, following Mendel, that the only thing known about genes was that, when organisms were cross-bred, different variants of a gene caused organisms to inherit alternative versions of the affected trait(s) in a "Mendelian" pattern. Because male and female parents typically contributed equally to Mendelian traits, something contained in sperm and eggs (and therefore also contained in fertilized eggs) caused Mendelian differences between organisms. Mendel called the things that differed *factors*. When Mendel's work was rediscovered in 1900, many biologists tried to locate Mendel's factors, in specific

ways, in germ cells, and tried to understand them in terms of cytological theories (e.g., as parts of chromosomes) or theories of organismal development. Johannsen argued that such moves were a mistake: no one knew what genes are – they might, for instance, be properties or states of the relevant cells rather than separate entities. The term *gene*, therefore, was supposed to be free of theories or hypotheses about the material constitution of genes and about the means by which they brought about their effects. The claim that a particular allele (gene variant) is present simply summarizes the phenomena; it is another way of stating the Mendelian behavior of the trait(s) caused by that variant (Burian, 2000; Carlson, 1966).

For the rest of his life, Johannsen argued that we did not know the intrinsic properties of genes or how to localize them. It was not even safe to claim that there were two copies of genes in fertilized eggs: if genes were something like a state of the egg or sperm and the fertilized egg was composed of a sperm and an egg in the same state, the egg would be in a single state; what we now call homozygotes would have a single state of that gene, whereas heterozygotes would be fertilized eggs (and organisms) with a mixture of two states of the relevant gene – states that separated in the formation of sperm (in males) or eggs (in females). Little wonder that the material constitution of genes was contested well into the 1950s; even those who were strongly convinced that genes are material

parts of chromosomes did not know whether genes were made of proteins, nucleic acids, or some combination of the two.

Since 1909, there have always been multiple contenders for *the* definition of the gene and every fifteen or twenty years, or so, at least one new contender for *the* definition has been taken seriously by biologists (Burian, 2000, 2005; Carlson, 1966; Falk, 1986; Portin, 1993; Snyder & Gerstein, 2003; Stadler, 1954). This definitional problem sometimes resulted in sterile terminological debates, but the ruckus over defining genes raised major issues – many of which remain to be resolved – about the causes of biological inheritance and the extent to which specific factors caused inherited traits.

For the entire century since 1909, there have been intense struggles about how to define genes. A point that deserves immediate attention is that alternative definitions of genes always reflected major tension, never fully resolved, between attempts to provide *structural* (e.g., narrowly physico-chemical) and *functional* definitions of genes. A gene must, in the end, be some sort of material entity that performs immensely complex functions. It follows that structural and functional criteria for what should count as a gene are inextricably intertwined.

As it turns out, there is considerable advantage in combining structural and functional criteria for identifying genes. One can use criteria of each sort to help correct mistakes on the other front. This has been valuable ever since Johannsen's day, as geneticists and others worked to specify the factors that cause

inherited traits. An easy example comes from the early stages of the chromosomal theory of the gene. Work on this theory with the fruit fly *drosophila* began with T. H. Morgan's discovery in 1910 of a sex-linked mutation for white eyes in the fruit fly drosophila. He showed, in effect, that the mutant gene is inherited together with the X, i.e., a sex chromosome.

In drosophila, males have one X and females two Xs. Mutations carried on the X chromosome are therefore not recessive on males, so mutations are easily detected. X chromosomes could be followed from grandfather to grandson. By crossbreeding X-carrying males with the daughters of fathers with the same X, one can increase knowledge of the mutations it carries (by testing the effects on both homozygous and heterozygous females) and of how the mutation is transmitted. Cytological aberrations, such as an attached fragment of a foreign (previously tested) X chromosome allowed additional predictions and tests. Standard histories of the Morgan group's contributions to Mendelian genetics follow the development of this line of work.[3]

Morgan and others in his laboratory quickly developed techniques for locating genes on specific chromosomes. Once genes were closely associated with particular positions (loci) on chromosomes, two genes with identical-seeming effects located on different chromosomes were assumed to be different

---

[3] Darden (1991) provides a particularly useful philosophical treatment of the importance of cross-disciplinary interactions between cytological study of chromosomes, genetic study of mutations, and the co-association (linkage) between mutations when they are on the same chromosome.

structures that played a distinct role in the development or economy of the organism. This was verified fairly quickly by finding biochemical and other detailed differences in the effects of some genes yielding a common visible phenotype. A famous example concerns two genes that altered eye color in the same way: *vermilion* (located on the X chromosome) and *cinnabar* (located on chromosome 2). After about twenty years, it was shown that the *vermilion* mutation blocked a biochemical step in the formation of brown pigment in the drosophila eye and the *cinnabar* mutation blocked a later step in formation of the same pigment. The twisty path by which this was accomplished, combining biochemical, developmental, and genetic tools, is sketched in Burian, Gayon, and Zallen (1988, 389-400).

By the 1960s, when molecular biology was far enough developed to begin identifying genes with particular sequences of DNA, the traffic back and forth between structure and function became much more intimate. It turned out that genes could not be defined simply as sequences of nucleic acid, as lots of sequences are too short, too long, or seem to be junk, i.e., to have no known function or to perform a non-genetic function (e.g., to serve as spacers; see Sapienza's chapter). This complexity reflects and supports the fundamental commitment retained through most of the history of genetics, namely, that proper identification of genes must combine functional with structural criteria. By now, effectively by definition, genes are structures of DNA (or, occasionally, RNA, as

in RNA viruses) that are copied across organismal generations and that have some additional function(s) in virtue of being transcribed onto RNA that has specific potential to enter into, or to affect, additional biological processes. One consequence of this account is that there is no single answer as to what, exactly, should count as a gene, for that depends on exactly what parts of the genome count as having a function.

Interestingly, there are those who would retain the word *gene* only for protein-encoding genes. But this causes problems: for example, it would not count sequences of DNA encoding transfer RNAs and ribosomal RNAs as genes. Nor would it yield a univocal count of genes, since (by the standard criteria of current molecular biology) many plant and animal genes encode a few to hundreds of distinct proteins obtained by alternative splicing of their RNA transcripts, and some rare proteins are made from pieces assembled from distant transcripts of distinct genes.

4. Genic Selectionism

With this background, we are ready to turn to the question of whether or not selection operates primarily on genes. To begin, let's review some important early doctrines that influenced the debates about our topic. A book that influenced many philosophers is G.C. Williams' *Adaptation and Natural Selection* (Williams, 1966), in which Williams argued that group selection plays a minor role in evolution, in comparison with individual selection, and that

selection acts primarily on genes. The latter claim is what interests us. One of

Williams' key arguments is that genes can be beneficiaries of selection, but

organisms cannot because *organisms do not survive in evolutionary time, but*

genes do. Understood as sequences or structures of DNA, genes survive

recombination, death of the organism, and death of the species. Thanks to the

mechanisms by which DNA is copied and by which sequences are maintained,

genes have the potential to survive indefinitely. Thus, with an eye to meiosis, in

which chromosomes may be broken up by "crossing over," Williams (1966) writes:

> I use the term gene to mean "that which segregates and recombines with appreciable
>
> frequency." Such genes are potentially immortal, in the sense of there being no
>
> physiological limit to their survival, because of their potentially reproducing fast enough
>
> to compensate for their destruction by external agents. They also have a high degree of
>
> qualitative stability. Estimates of mutation rates range from about 10^{-4} to 10^{-10} per
>
> generation…
>
> Natural selection would produce or maintain adaptation [of genes] as a matter of
>
> definition. Whatever gene is favorably selected is better adapted than its unfavored
>
> alternatives. This is the reliable outcome of such selection, the prevalence of well-
>
> adapted genes. The selection of such genes of course is mediated by the phenotype, and
>
> to be favorably selected, a gene must augment phenotypic reproductive success as the
>
> arithmetic mean effect of its activity in the population in which it is selected. (pp. 24-25)

Mutation is inevitable, but evolved defenses against mutation protect

organisms (and genes) well enough that when a variant of a gene confers an

average selective advantage of one extra survivor/10^4 offspring that variant will

normally survive indefinitely. Although a number of qualifications are needed here, they don't matter for present purposes. On an evolutionary time scale, organisms, populations, groups of organisms, and even species exist for far shorter times than the genes (with their evolving variant forms) they contain. According to standard textbooks, species of mollusks and clams, which are longer-lived than most land species and are well preserved as fossils, have a half life of about 10 million years. A standard estimate suggests that at most 3-4% of the mollusk species extant 35 million years ago are still in existence. By standard criteria, most genes (that is such gene *kinds* as transfer RNA genes, alcohol dehydrogenase genes, etc.), preserving their function through a lineage of descent, are far longer lived than that.

If these are the relevant comparisons, it is the genes that are the beneficiaries of selection. However, I note that once we recognize that we should compare *lineages* of genes with *lineages* of organisms, it is not clear that these are the relevant comparisons; lineages of organisms, like lineages of genes, last until extinction.

In *The Selfish Gene*, Richard Dawkins (1976, 1982) popularized the above argument and expanded the scope of the controversy. Both Dawkins and David Hull (1980, 1981) sharpened the argument by distinguishing between replicators and interactors, arguing that replicators are the main unit of selection. For Hull, a replicator is an entity whose structure is passed on fairly directly via a process

(replication) that yields one or more reasonably accurate copies of the replicator. Evolutionarily relevant replicators must meet criteria of longevity, fecundity, and fidelity. The relevant form of longevity is *preservation of structure in the form of copies*. Hull's example makes the point clear: atoms of gold, structurally, are very similar to each other and quite long-lived; but because atoms of gold are not made by a copying process from previous atoms of gold, they do not count as replicators and are not subject to evolutionary processes like those affecting organisms and other biological entities. Dawkins' version of replication is more stringent; he requires nearly exact copying and, for relevance to selection, focuses on "active germ-line replicators," i.e., those that play an active role in favoring their own replication across organismal generations. Thus, on Dawkins' account, since organisms reproduce not by a process of copying, but by (re)constructing most of the materials (e.g., proteins) and structures out of which they are constituted, organisms are not replicators, not even single-celled asexual organisms. On Hull's less stringent account, single-celled asexual organisms are normally close copies of the parent from which they are produced, so they are perfectly good replicators. In spite of this difference, Dawkins and Hull agree that genes are exemplary replicators since their structure is preserved across enormous numbers of generations by means of the copying mechanisms that replicate DNA structure. Williams' (1966) definition of a gene as "that which segregates and recombines with appreciable frequency" is, effectively, a

definition of a replicator. It is non-committal, at least on the surface, about the structure or function of genes as units, and Williams explicitly points out that this definition treats as genes what, by more conventional criteria, might be counted differently, e.g., as several linked genes or as no genes at all.

Hull's (1980) term for the other principal sort of unit, *interactor*, is preferable to Dawkins' term, *vehicle*. An interactor "is an "entity that interacts as a cohesive whole with the environment in such a way that replication is differential" (p. 318). Dawkins' vehicles are explicitly restricted to entities within which genes (and perhaps other replicators) are contained. Organisms are the main entities that fit this description, but mitochondria and other DNA-containing organelles and groups and populations of organisms may also qualify. There are no exact boundaries for either replicators or interactors; they come in degrees of exactness of replication, and of cohesiveness plus relevance to differential replication. Still, organisms are built anew in each generation by an interactive process with the environment, not by a copying mechanism. This makes Dawkins's distinction between replicators and interactors quite rigid: replicators are made by copying, while interactors are made by some kind of construction. (Construction is my term: the classical embryological term is *epigenesis*, a process in which the entity is *built* rather than copied.) Hence, organisms cannot be replicators. Something is a vehicle for Dawkins only if it contains (and is built from, or by) replicators. A vehicle is evolutionarily important only insofar as its

properties bias the likelihood of its replicators making it into the next generation. For Hull, too, an integrated whole counts as an interactor because its properties or behaviors affect the likelihood of replicators making it into the next (organismal) generation, but only some interactors contain the replicators in question.

The take-home message of these arguments is that replication involves entities beyond the replicators themselves, and that evolution is the result of interactions that bring about transgenerational replication of sufficient exactness to enable construction or reconstruction of organisms and other evolving entities. All this is supposed to be required for evolution via natural selection. Orthodox evolutionary theory holds that genes are generally the primary replicators and that organisms interact as unitary beings with each other and with the rest of their environments in such a way as to affect their likelihood of reproducing and transferring copies of the genes within them into the next generation. As Hull insists, replication and interaction are distinct processes; both are required for evolutionary processes. But the boundaries of these processes may overlap. Thus, on Hull's view, the same entity – e.g., an asexual single-celled organism when it divides – can be correctly considered to have replicated or to have reproduced.

5. Three Traditional Arguments

We are ready to present arguments that show that genes are not the primary units (targets or beneficiaries) of selection. Sapienza argues that selection *can* and (at

least sometimes) *does* act on genes. He is right. But that's not our main question,

which is whether selection acts *primarily* on genes. Nor does the question

whether selection benefits individual genes or organisms solve our problem.

Selection does not benefit individual copies of genes (gene tokens) any more than

it benefits individual organisms. And successful *lineages* of organisms are just as

long-lived as successful lineages of genes, so a surviving lineage of organisms is,

as such, just as much or as little a beneficiary of selection as a surviving lineage

of genes is. Given this, I will leave the issue of beneficiaries behind and

concentrate on whether selection acts primarily on (lineages of) organisms or

genes.

The arguments considered here apply directly to sexual organisms,

especially multicellular eukaryotes, i.e., organisms whose cells have true nuclei

containing chromosomes that are complexes of histones and DNA. On this score,

the differences between prokaryotes and eukaryotes are important largely because

of the specialized regulatory processes affecting the development and evolution of

eukaryotes, and because the relationships between genes and their products, and

between genes and functional traits, are much more indirect in eukaryotes than in

prokaryotes. These features of eukaryotes make the argument easier and more

decisive, but similar issues can be raised for prokaryotes.

(1) The first argument focuses on the causal processes involved in

selection. Selection acts on marginal differences between organisms with

systematically different traits in the actual ecological and competitive circumstances faced by actual organisms. It "acts" when some heritable genotypic or phenotypic traits increase the probability of reproductive survival. The conditions for this are already familiar: heritable variation in fitness plus a fairly consistent selective regime. As Sapienza argues, genetic differences (almost) always underlie trait differences when these conditions are met, so favoring a trait typically favors some constellations of genes over others. However, this cannot be *universally* true since genetically identical organisms (e.g., identical twins) do not have strictly identical fitness-affecting traits and since, quite generally, all kinds of environmental differences and learned responses alter the course of development, the response repertoire, and even the anatomy of organisms. Not all heritability is genetic!

Thus, as a matter of bookkeeping, selection favoring a trait within a population favors not only that trait on average, but also some genes or gene combinations. Sometimes this may turn out to be one gene that is closely correlated with the selected trait. William Wimsatt (1980) offered a complex and influential argument that, in general, we should treat genetic changes that reflect phenotypic selection as a way of bookkeeping. The bookkeeping works because genetic changes track trait changes in evolutionary time – but not because changes of the individual genes, as such, yield a correct causal explanation of the evolutionary trajectory followed by the lineage of organisms.

One of Wimsatt's examples makes the point clearer. The case involves frequency-dependent selection involving simultaneous variation of two different genes, in which the selective effect of a given allele depends on the proportions of alleles present in the population at *both* loci. In such circumstances, whether selection favors or disfavors a given allele may depend on the frequencies of the alleles present at the other locus.[4] The result is a dynamic system driven by trait relations to ecological circumstances and by causal interrelations involving two genes and their products or effects; but it does not, in general, allow selection to favor any one gene variant across the full range of available environments and population distributions. Since seasonal and sporadic changes may dramatically affect the proportion of alleles at the other locus, no prediction of evolutionary outcome can be made from the selective advantage for a particular allele in a given situation in which the population has a particular composition. Although this only sketches the complexities involved, it makes the key points: without a clear and correct account of genotype-phenotype relationships and the relevant selective regime, one cannot determine, from the local fitnesses of alleles plus the specific trajectory of the genome in the lineage, whether the individual alleles that gained in frequency were the primary targets of selection. In general, then, even granting strong genetic causation of traits, it does not follow that gene-by-gene

---

[4] Concrete examples of such situations are well known; Wimsatt includes one in his paper.

analysis suffices to explain the evolutionary trajectory of the organisms in question. Genes (and their alternative alleles) have the effects they do as part of a system of tightly integrated interactions. This is the stumbling block at the heart of the issue we are debating.

(2) In a classic paper, Elliott Sober and Richard Lewontin (1982) present a second, related, argument using the familiar example of sickle cell anemia. The sickle cell allele of a gene for hemoglobin is prevalent in populations exposed to mosquitoes carrying the malaria parasite. Without medical intervention, which only recently became feasible, individuals homozygous for this allele nearly always suffer excruciating death before reaching puberty from sickle cell anemia. Nonetheless, the allele is maintained at fairly high frequencies in malarial regions because it provides human heterozygotes with considerable resistance to malaria. In most common ecological circumstances, one copy of the normal allele produces blood cells with enough normal hemoglobin to prevent them from sickling and enables them to carry enough oxygen for carriers of the allele to be unharmed, but the amount of oxygen carried is low enough to reduce the rate at which the malaria parasite can reproduce. In spite of the terrible cost of killing off homozygotic 'sickler' children before they reproduce (roughly ¼ of the offspring of matings between heterozygotes), the net effect of carrying the allele in severely malarial environments is more favorable than not having this protection against the disease. Thus, in favoring resistance to malaria, selection

has the most unfortunate side effect of carrying sickle cell anemia along with that resistance.

The sickle cell story is actually more complex than this exposition suggests, but the complications don't affect the point. We have here, approximately, a simple case of heterozygote advantage. What we need to determine is whether selection acts primarily on the sickle cell allele. According to Sober and Lewontin, the answer is *no*: in malarial environments, selection favors heterozygotes, penalizes homozygotes drastically, and does not act on the allele as such. The allele has an average effect, of course, in a given mixture of environments (this is Wimsatt's bookkeeping!), but selection acts on the effect of the distribution of different hemoglobins in blood cells in environments with varying degrees of malarial severity. The hemoglobins in blood cells of homozygous normals favor malaria; the hemoglobins in homozygous sicklers alter blood cells so that they block capillaries and cause lethal oxygen starvation, those in heterozygotes allow blood cells to function well enough and yet weaken the malaria parasite. These are the differences on which selection acts. Sober and Lewontin argue that if it acts on any genetic entity, it is the genomic states of being homozygous normal, heterozygous for sickling, or homozygous for sickling. That is, genetically speaking, the conditions causally relevant to the phenotypic states that influence survival are the combinations of alleles, not the

alleles themselves or their frequencies in the population. If so, at the genetic level, selection acts primarily on the genotypes, not the genes.

From this perspective, *selection acts on populations to increase the frequency of causally relevant genetic complexes* in the population over time. What this means in general depends on how genes or genotypes correlate with phenotypes. To get the causal story right, as the case of heterozygote advantage makes clear, one must examine the *genotypes* available in the populations and the dynamics of genotypic and phenotypic change over time.

Changes in the distribution of organismal genotypes will, of course, be tracked by changes in gene frequency, but the use of frequencies of individual genes, taken one at a time, as a surrogate for genotype frequencies ends up misrepresenting the actual causes of genetic change. This is made even clearer in a recent paper by Brandon and Frederik Nijhout (2006). They argue that genic selectionists must offer a false account of the causes that maintain a dynamic equilibrium when the sickle cell and normal alleles are maintained at a stable frequency. Baldly expressed, the correct causal account is based is on balancing selective pressures favoring the heterozygote and strongly disadvantaging the two homozygotes. However, since genic selectionists treat alleles (rather than combinations of alleles) as what is selected, they must view selection of alleles in this sort of case as frequency dependent – and when the alleles are at an equilibrium frequency, there is no selective pressure acting to change the

frequency of the alleles! This is obvious, because at equilibrium the average

fitnesses of the normal and mutant alleles are equal: *there is no net force acting

to increase or decrease the frequency of either allele*! In fact, however, *selection*

is acting. It acts on heterozygotes (who are moderately susceptible to malaria, but

otherwise undamaged) and, more harshly, on homozygotic normals (who are

highly susceptible to malaria) and homozygotic sicklers (who die from sickling of

blood cells). Thus, in cases of heterozygote advantage genic selectionists are

forced to offer a drastically false account of the effective selective forces in a

dynamic balanced equilibrium. More generally, *when the allele or the gene is

used as the unit of analysis and interaction effects alter heterozygote fitness non-

additively, calculating allele-by-allele yields a mistaken calculation of how

selection acts*.

(3) A more diffuse literature concerning complex life histories provides a

third argument. Take a relatively straightforward example: aphids. Here is a

simplified account of a typical annual aphid life cycle (Ragsdale, 2002). Most

species of these plant-sucking insects live in temperate climates and go through a

kind of alternation of generations. During early spring, the offspring of an

overwintering female, who mated at the end of the previous summer, are all

females. These females yield a series of all-female generations that typically live

3-4 weeks and reproduce parthenogenetically (without fertilization). The females

just described are viviparous, that is, they do not lay eggs, but produce free-living

daughters, already containing embryos. Even though these females and their daughters are genetically identical, what they eat can alter the size, morphology, and fertility of their daughters and granddaughters (Stadler, 1995). For example, it can determine whether those offspring are winged (and seek another host plant) or wingless (and stay on the same host). At the end of the summer, triggered by light, nutritional, and temperature signals, some of the females produce winged males and sexual females. Those females are oviparous, not viviparous, and lay overwintering eggs. The eggs that hatch in the spring produce stem females, who produce both winged and wingless parthenogenetic daughters and start the cycle anew.

Many organisms have yet more complex life cycles than those of aphids. Here are two major points to be taken from these sorts of cases. (A) Ecological (or nutritional, etc.) circumstances determine major features of offspring produced from organisms with a given set of genes. Indeed, in general, *the traits of the offspring of an organism are co-determined by environmental and genetic causes*. (B) Selection is episodic, but how it acts must be calculated over entire cycles of the right scale. This is part of the reason for the requirement that there must be suitable conditions permitting cumulative selection for natural selection to be effective. For example, the timing of environmental cycles must be short enough to be "visible" to selection, or selection cannot act to take them into account.

Annual cycles work, 10,000 year cycles probably do not (even for redwood trees) and million year cycles certainly do not.

Major genetic changes in aphids occur approximately once a year, during the sexual generation. That is when recombination and most other germ-line changes occur. The cumulative effect of many selective episodes on a lineage will be expressed, in part, by the robustness and the fecundity of the sexual female and the overwintering eggs she produces. And those properties of the female are affected by the circumstances encountered by the all-female lineage that preceded her for an entire season. There is, no doubt, a net average effect of the seasonal success and failures within her lineage on the female and a net average effect of the successes and failures of the clone within its local habitat, but there will be cases in which a key selective episode that affected one of the great, great,... grandmothers of a sexual female is a key determinant of some features of her offspring. Thus, certain seemingly genetic traits – e.g., the fecundity of the sexual female or what proportion of the daughters of a parthenogenetic female have wings – may be determined less by her genes (which determine her capacities in these respects), than by her nutritional status and that of her maternal lineage during the current season.

Generalizing over the life histories of organisms that succeed in reproducing and about the ways in which selection manages to take account of what happens on regular cycles over long periods of time (which is a major part

of what is at stake in evolution), *selection normally 'acts', not as a single force on an isolated gene or trait, but as a balancing device in response to situations affecting multiple traits in scattered episodes over extended periods of time. In consequence, a balance of selective pressures affecting multiple traits is the norm, not a single force affecting an isolated gene (or trait). Successful lineages must cross all of the distinct thresholds that recur regularly during the relevant long cycles that arise over multiple years and/or generations.* This result shows why it is necessary to analyze the conditions for *cumulative* selection with great attention to long cycles affecting life history patterns. If all this is right, systematic devices for handling diverse difficulties – at once, and in series – are typically involved in selective processes. For this reason, the genic selectionists' claim that selection acts primarily on single genes is prima facie implausible. The burden is on them to show how selection can work over such long periods in such complex ways by acting gene by gene in the manner described in Sapienza's contribution to this volume.

6. Response to Sapienza's Arguments: 1. What Does Selection Act On?

Let's start with hearts. Nearly everyone agrees that the function of a heart is to pump blood. (For an exception, see Cummins & Roth, this volume.) It must do so (in the absence of human intervention) for the animal's entire life. In every extant lineage animals have almost certainly faced recurrent contingencies affecting hearts. When they present significant challenges to the heart (as do

malaria and drought for example), the lineage is likely to have evolved some provisions to deal with, or avoid, those challenges. As an engineering matter, one cannot simultaneously optimize all of the traits of an organ to meet all expectable contingencies. If optimizing blood pressure causes increased risk of myocardial infarction, selection must act in a way that balances the resulting trade-off. And since ecological and seasonal cycles present various stresses and demands that hearts must meet, the heritable properties of hearts must permit or enable them to adjust to the full range of circumstances that recur sporadically or regularly over tens or, even, hundreds of generations. The penalty for failure to do so is simple, but drastic: extinction.

In game-theoretic terms, selection does not optimize traits or organs, it satisfices (does at least well enough to get by). For our example, heart designs must make (or at least allow) trade-offs adequate to meet the expectable tasks and challenges posed over a long series of generations. Organisms in a lineage must overcome (at different times) the effects of muscular and nutritional stress, low blood pressure, high blood pressure, blocked blood vessels, changes in oxygen supply, specific infections, drought, floods, temperature extremes, and so on, indefinitely. Lineages like ours, in which individuals grow hearts somewhat differently depending on ecological conditions and life histories (e.g., altitude, nutritional limitation, a sedentary life, or a life of long distance hunting), do not have rigidly fixed genetic programs that determine all the traits of their hearts, for

developing hearts respond differently to differing circumstances encountered in different ecologies, situations, and thanks to different life habits. Such phenotypic plasticity conflicts with rigid genetic control of traits.[5]

Two of Sapienza's claims about hearts are hard to reconcile unless they are carefully circumscribed. These are (1) that the variation in blood pressure between different individuals can be explained, in large part, by differences in genotype and (2) that the phenotypes under selection are controlled by a small number of genes. Take the first: variation in blood pressure can also be explained by variation in life styles (e.g., amount of exercise, daily hours of sleep), ecology (altitude and climate, availability of meat), and social and cultural factors (cf. correlations between heart health and various life styles). If Sapienza means (as is unlikely) that genotypic differences explain high blood pressure in general, independently of these other factors, the evidence is against him. Since we know that the blood pressure of most individuals can be altered considerably by controlling diet and exercise, and that early life styles and circumstances can have life-long effects on heart function, the claim I just ascribed to him is initially implausible and needs to be supported by specific evidence. If, instead, he means merely that the blood pressure has high (genetic) heritability, then he is right, but all that shows is that genetic variation has important effects on blood pressure,

---

[5] Many biologists have taken up phenotypic plasticity recently: Massimo Pigliucci (2001) and Mary Jane West-Eberhard (2003) published important books on the topic that is philosophically sophisticated, and Brandon (1985) put forward an early argument showing its importance for human genetics.

which is consistent with considerable environmental influence on it as well. Turning to the second claim, high heritability does not, as such, provide evidence that the phenotype in question is controlled by a small number of genes. Height is highly heritable, but it is not controlled by a small number of genes.

I believe that most organs and quantitative traits examined to date are like height in being controlled by a large number of genes. All in all, Sapienza's claim that blood pressure is controlled by a small number of genes is initially implausible. And since selection has been acting for many millennia to keep blood pressure within the bounds required by adequate health for reproductive survival, the fact that so much heritable variation is still available within human populations makes it likely that this phenotype, in particular, either is involved in trade-offs along the lines suggested above or is not controlled by a small number of genes.

This argument does not prove that Sapienza is wrong. Nor does it establish what claims in this neighborhood are right. Rather, it shows that specific and detailed knowledge of the relevant genotype-phenotype map is required to settle the matter. To argue more than this without going into serious molecular detail is, I think, hand-waving. Now that we can study such issues at the molecular level, their resolution requires successful analysis of the relevant genotype-phenotype relations.

Meiotic Drive. Professor Sapienza's remaining examples don't bring us any closer to a resolution of our debate. I shall deal with them briefly and then close by taking up a more general issue. Sapienza is entirely right that selection can act directly on genes. His example of meiotic drive in DDK mice illustrates the point nicely. Here, selection acts primarily on a particular allele because the relevant variation is between chromosomes containing the allele and chromosomes containing alternative alleles. Since the probability of the allele (and its chromosome) getting into the next organismal generation is altered by the direct effects of the allele itself, specifically its effect on the likelihood that its chromosome will make it into the egg, selection *acting primarily at the level of the gene* and not on other phenotypic traits of the organism secures the representation of this allele in the next organismal generation.

However, our main issue is not whether selection can act in this way, but whether it *typically* does so. And comparison with the heart example shows why this example of meiotic drive is not typical. The selectively relevant properties of hearts are not typically like the distribution of alleles or chromosomes into egg cells or polar bodies; they include such organ-level properties as the rate of pumping blood, the robustness of the organ, and its ability to operate continuously and withstand stress. In this connection, selection acts primarily on properties relevant to the demands of organismal survival imposed by the ecology and by competitors. In favoring organisms with more robust hearts, ability to maintain

blood pressures within a suitable range, etc., it alters genomic organization and the content of genes over generations, but it does so secondarily, for it acts, in the first instance, on heritable phenotypes. As Sapienza explicitly grants, meiotic drive and organismal selection often favor different outcomes; selection on multiple levels requires a difficult balancing act.[6] Since selection does not act purely at one level, one needs to know the balance of selective pressures at different levels to calculate the net selective forces acting on (and net fitness of) a gene, genotype, or trait. Such a calculation complicates (and I think will ultimately defeat) all attempts to treat selection as acting primarily on genes.

Robertsonian Translocations. This way of thinking, in terms of multiple levels and units of selection, carries over to the examples of Robertsonian translocations and G-proteins, both of which can be turned against the claim that selection acts primarily on genes. As Sapienza indicates, reproductive isolation of populations is a major step in speciation. Robertsonian translocations enforce reproductive isolation because members of a population that differ by a Robertsonian translocation cannot produce viable or fertile offspring. Where there is contact (or recontact) between populations separated by such a translocation, selection acts within each population to sharpen and tighten any other isolating mechanisms that reinforce the separation of the populations (e.g.,

---

[6] A classic paper by Lewontin and L. C. Dunn provides a beautiful demonstration of selection acting simultaneously in different ways at three different levels in house mice. The mice experience meiotic drive, organismal selection, and group selection between small demes in a relatively closed environment.

by preference for mates from their own population), or to reduce the competition between them. If contact is maintained in the absence of additional isolating mechanisms, it is likely that members of one or the other population will eventually be eliminated (at least locally) by competitive exclusion or by reproductive failure.

I argue that Sapienza has not yet shown that the chromosomal variation and allied phenomena are, as he claims they are, "examples of direct selection at the level of the gene." He assumes, but does not argue, that when a change in centromere affinities causes a chromosomal mismatch that leads to a Robertsonian translocation, *selection* is responsible for the chromosomal differences or acted to favor the allele(s) that favored the centrosomal change. But in principle and, perhaps, in fact, the chromosomal change may, rather, present the raw material on which selection may or may not act, acting (if it acts favorably) to form a separate population and, ultimately, a new species. If an isolated family, separated geographically from the main population, obtained the Robertsonian translocation by mutational accident(s), the new species might have been formed by mutation affecting chromosomes and drift alone! The fewer the mutations required and the smaller the initial populations, the harder it is to determine whether or not selection was involved. The role of selection in such a case is, thus, an open question. What is required, once again, is detailed knowledge of the relevant genotype-phenotype relations and enough information

about the relevant selective regimes. Knowledge of the relevant properties of

centrosomes and the distribution of genotypes is crucial, but it is not enough. *We*

also need to know the relevant grounds for the reproductive success of the

affected organisms of the different genotypes in relation to competitors within and

*between population*s before we can decide the issue.

G-Proteins. Sapienza's nicely nuanced account of the limitations

selection faces in acting gene-by-gene on genes for G-proteins is very helpful. I

agree with much of what he writes, but also with part of what he implies that his

colleague maintains (though not the strong gene selectionism the colleague

advocates). Sapienza is almost certainly correct that selection cannot act directly

on the thirty five genes encoding material used in various proteins to maintain,

simultaneously, all of the distinct contributions of each gene to the G-proteins

utilized by animals for recognition of scents, signal detection, and the many other

functions that those genes affect. As he argues, the combinatorics for gene-by-

gene selection are prohibitive. Yet, his colleague is surely correct that an

adequate theory or mechanism is needed to explain how selection maintains these

complex relationships. It is important to figure out the exact explanandum, that

is, exactly what it is about G-proteins that needs to be explained (and that is not

obvious!), but the coordination of G-proteins with distinct receptors is clearly the

result of selection acting in some manner or other on the substrates out of which

both the receptors and the proteins are manufactured. The puzzle about how

selection can do this has not been removed simply by showing, as Sapienza has shown in outline, that it does not do so by acting gene by gene.

Some important disciplinary differences and differences in background assumptions underlie the differences between Sapienza's and my positions. These differences affect which phenomena and explananda we think are most important, our divergent accounts of how selection integrates effects from many genetic and non-genetic sources, and our views about what sorts of evidence are required to decide the issue. These disagreements are partly on philosophical in character, but they are also scientifically productive. In the examples we have been discussing, they can lead us to experimentally concrete problems, potentially resolvable by experimental and populational studies.

7. Response to Sapienza's Arguments: 2. On the Importance of Epigenetics

Sapienza is right that biologists ought not (and, I am sure, will not) resort to arguments based on irreducible complexity; that would amount to the abandonment of science. But the first step in working toward new explanations in new domains is finding regularities that describe thus-far uncharacterized or unexplained phenomena. Recently, some philosophers of science have argued for the importance of exploratory experimentation in such contexts. It was suggested as a focal topic in philosophy of science a decade ago (Burian, 1997; Steinle, 1997) and its role in very recent molecular sciences explored in four recent papers (Burian, 2007; Elliott, 2007; Franklin, 2005; O'Malley, 2007). Given the

complex regularities being teased out in post-genomic molecular biology (those concerning G-proteins are the tiniest tip of an enormous iceberg), Sapienza's argument serves, I think, as a reductio of classical versions of the position he set out to defend – i.e., the position that selection acts primarily on genes. He showed that the combinatorics of gene-by-gene selection are incompatible with fine-grained selection of G-proteins or maintenance of matches between G-proteins and their receptors. What is sought are alternative mechanisms by which selection can act, effectively, to build and maintain complex traits (involving, in this case, coevolution of multiple independent complexes of genes). In this closing section, I suggest that new work on epigenetics (and epigenetic inheritance) offers the prospect of solving at least some of these problems. Sapienza rightly left this possibility open in the concluding section of his paper and has explored possibilities along these lines in some of his technical papers (e.g., Sandovici et al., 2006).

Sapienza maintains that when traits are too complex to be maintained by selection acting on genes, they must have enormous selective advantage or they will be lost. This would be true if selection acted primarily on genes, but that is the very question we set out to debate. Phenotypic studies establish, I maintain, that selection "sees" or acts on complex organismal traits – and that G-proteins are a plausible example of a case in which selection maintains traits that can't be explained by (classical) genic selection. Sapienza argues correctly that we do not

yet have a good theoretically-grounded understanding of how complex traits are altered and maintained. At least two lines of research, however, may provide serious mechanistic answers to such puzzles.

The lines of research I have in mind overlap somewhat, but address a variety of different mechanisms. One line concerns RNA regulatory networks, the other concerns epigenetics (and specifically the study of epigenetic inheritance). I close with a couple of paragraphs on each to indicate some of their promise and hint at how they overlap.

Recent genomic work has yielded some surprises concerning the amount of genetic material that is transcribed and the great variety of regulatory RNAs included in genomic transcripts. John Mattick has published several articles summarizing the evolutionary importance of regulatory RNAs in the evolution of eukaryotes (e.g., Mattick, 2004). In mammals, for example, it appears that at least 80% of the DNA is transcribed (in humans it is probably 97% or more), and that much of it forms regulatory RNAs that enter into complex networks. These play a major role in regulating development and determining what products will be made from which genes in which circumstances. In eukaryotes, the correspondence between genes and gene products is not 1:1: most protein-encoding genes make several (and some make up to several hundred) products by alternate splicing and/or by contributing to products by splicing of RNAs from

distant parts of DNA.[7] Many of the controls that regulate which products are

made in which cells and at which stage of development are located in RNA

regulatory networks. These controls deal with everything from nutritional

switching to responding to infectious agents, from switching cell type in

differentiation to coordinating responses to heat or cold or infectious agents.

There are at least a dozen major types of RNA involved in these networks, some

of them very short – e.g., short interfering RNAs and microRNAs, each about 21-

22 nucleotides long and capable of rapid response. Such controls alter the protein

content of a cell and its descendants by interfering with mRNAs or altering the

regulatory signals contained in mRNAs to regulate the developmental stage at

which they are translated. These controls also respond to nutritional state of the

organism, the presence of specific toxins or pathogens, the entry of signal

molecules into the cell, and so on.[8]

  Some mechanisms deployed to regulate development in RNA regulatory

networks coordinate gene expression for large numbers of genes. For example, a

developmental switch that coordinates the transition from maternal to zygotic

mRNAs in the zebrafish (a standard model organism for such work) is a

microRNA, labeled miR-430 (Schier & Giraldez, 2006). Expression of miR-430,

---

[7] There are many other regulated devices that alter the correlation between coding DNA and the products derived therefrom. Examples include RNA editing and post-translational modification and splicing of polypeptides.
[8] For a review of work on microRNA pertaining to the issues of this paper, see Burian (2007).

itself produced in large numbers at a specific moment in embryonic development, blocks a regulatory region in the untranslated regions of at least 700 maternal mRNAs in somatic cells or the embryo, thereby ensuring that they are degraded quickly rather than slowly over an extended period. It does so by base pairing with its target while attached to a protein complex that that then disrupts the target mRNAs. The disruption occurs coordinately in all somatic cells of the embryo. However, in the germ cells, the same microRNA protects the maternal mRNAs from degradation. Thus, the regulation is highly specific and coordinates the timing or loss of expression of immensely complex batteries of gene products (and, apparently, gain of expression of their zygotic replacements as well) in cell-specific and stage-specific ways (Schier & Giraldez, 2006). This example illustrates coordinate regulation of immensely complex traits and processes that can be regulated in many fine-grained ways by changing only a few nucleotides or a relevant environmental variable, e.g., by altering the timing of the expression of miR-430, by altering its sequence, or by altering the relevant regulatory sequences (and hence the mRNAs) of the genes of affected maternal proteins one at a time. We do not understand this or related mechanisms in detail, but it is immediately obvious that they provide strong controls of complex coordinated batteries of genes.

It has been experimentally demonstrated that such controls can be modulated in fine-grained ways to alter the timing of developmental changes, the

number of genes affected, and the size of the effect. The combinatorics of such regulation, unlike those of G-proteins proceeding allele by allele, appear to be tractable and their flexibility quite extraordinary. The connection with the second line of research, on epigenetics, is straightforward. Many changes in RNA regulatory networks are heritable, some of them genetically (e.g., by mutation of the DNA from which microRNAs are constructed), some of them epigenetically.

The term *epigenetic* goes all the way back to Aristotle (see the papers in Van Speybroeck, Van De Vijver, & De Waele, 2003). The historical use that fits most readily with present usage contrasts the theory of predetermination of embryos (which held that the fertilized egg already contains, in miniature, the "form" of the adult or all of the major organs of the embryo) and the theory of epigenetic development of embryos (which held that the fertilized egg does not contain the form of the adult, nor all or its organs, but that these are built by a series of developmental changes requiring internal and external resources in an extended developmental process). In the last few decades, the term has acquired a new meaning. An epigenetic change is a heritable change independent of any change of nucleotide sequence in the DNA (or genetic RNA) of the organism in question. A clear example of epigenetic change in development is provided by gene or chromosome silencing. For example, methylation of chromosomes (particularly at CpG sites – i.e., DNA sites at which the nucleotides cytosine and guanine are linked by a phosphate) can alter the conformation of the DNA in such

a way as to prevent transcription of that region. This is one way of achieving so-called genomic imprinting, in which chromosomal regions are silenced by epigenetic mechanisms. Genomic imprinting is a crucial control of gene expression in development, required, for example, for proper differentiation of secondary sexual characteristics. In mammals, it is now well established that many chromosomes are imprinted differentially according to whether they come from the male or female parent, and that imprinting is epigenetically inherited (see, e.g., Sandovici et al., 2006; Wood & Oakey, 2006). There is a large literature establishing that imprinting is removed from chromosomes very early in embryonic development, but various marks that serve as cues for reestablishing imprinting are left, that this process involves no change of DNA sequences, that one of the major mechanisms involved is methylation of CpG sites, and that 'correct' or nearly 'correct' imprinting is required for normal development.

I mention this example for a particular reason. At the close of his chapter, Sapienza cites the work of the Decode Consortium on differences in reproductive rates in Icelandic families and their finding that the women with the highest reproductive rates are those had the highest rates of meiotic recombination. Another of their findings suggests that there may be some connection with epigenetic inheritance involved in this result. The single genomic marker with highest correlation with a high rate of recombination was the fraction of the genome with a CpG motif, i.e., with the markers for methylation and silencing of

chromosomal regions. Like some of the work in Sapienza's lab (Sandovici et al.), this suggests, but does not provide definitive evidence, that epigenetic inheritance may play a role in the mechanisms affecting recombination rate. I do not understand these mechanisms and do not pretend to know whether this suggestion will pan out, but it is a nice illustration of the range of open questions still to be faced before we can reach a definitive answer about the mechanisms underlying selection of complex phenotypic traits. It also is a marker of the fact that an adequate case for selection acting primarily on genes has not yet been made.

One more example of an epigenetic process illustrates the power of epigenetic change to release hitherto hidden phenotypic potential of the organism and then to "lock in" the altered phenotypic properties of the survivor. This concerns the research on heat shock protein 90 (Rutherford, 2003; Rutherford & Lindquist, 1998; Wagner, Chiu, & Hansen 1999). In barest outline, the story goes like this: hsp90 is a 'chaperone' protein that marks cells to be killed if they have certain abnormal proteins. Experiments with mice and fruit flies have shown that if hsp90 is depleted by a heat shock early in embryonic development, it allows molecular variants with strong phenotypic effects to survive the molecular controls that normally would kill cells with those abnormal proteins. The result is unusual phenotypic variation in the surviving animals. Since hsp90 recognizes normal proteins by the conformation of the relevant proteins present in the earliest stages of development and the altered proteins are transmitted maternally, they are

present during early development of the offspring. Thus, the surviving conformal variants of the affected proteins are processed as normal in the offspring of the heat shocked animals, so the standards employed at the molecular level for recognizing normal proteins have been inherited epigenetically with significant phenotypic consequences. As I understand it, this outline is correct – but the details matter immensely and a great deal of important research must be examined with great care before it is clear whether an account like the one given here can be generalized to resolve significant puzzles of the sort raised by Professor Sapienza.

Postscript: Counterpoint

Like Professor Sapienza, I begin the postscript by pointing out some points of agreement. He is correct that selection 'acts' directly on genes that are able to spread themselves within genomes over organismal generations. Among such 'selfish' genes are those that produce meiotic drive or produce transposable copies that incorporate themselves in the genome. He is right that the importance of such phenomena has not been adequately appreciated and needs to be carefully assessed – the high prevalence of 'ghosts' of retrotransposition events in the human genome provides ample evidence of that! It will be particularly important to evaluate the impact of retrotransposition on genome organization, genome evolution, and the evolution of organisms.

Given this, I have no doubt that, as he argues, selection acting directly on genes has had important effects on evolution and evolutionary history. So far so

good. But our core issue concerns the centrality of such phenomena in evolution.

In this connection, his arguments do not establish that selection acts primarily on

genes (or on DNA). In spite of the important genetic findings and mechanisms he

uses to suggest difficulties for some of my arguments, the fundamental points I

made are not touched by those difficulties. The key issue is the balancing act that

is required when selection acts at several levels over periods of time sufficient for

evolutionary shaping of genomes and organisms. Genes and DNA are as

susceptible to this problem as any other traits that are affected by natural

selection. My counterarguments in this brief reply illustrate this point as well as

shortcomings in some of Sapienza's key arguments.

The fact that variants of some genes (currently) have disproportionally

larger selectively relevant effects than variants of other genes on phenotypes of

interest does not prove very much. Some of the recent literature on genetic

influences on human diseases shows that *which* genes have disproportionate

influence changes with the ecology. Our knowledge here is only now being

developed and is not very secure, but in a couple of examples it is both strong

enough and suggestive enough that I think it makes the point rather nicely. One

example concerns genetic resistance to AIDS. Recent work, reviewed by Dean et

al. (2002) has shown that some HIV viruses interact with chemokine receptors to

enter cells. The dominant HIV viruses in initial HIV infection are HIV-1 viruses

that enter solely through the CCR5 receptor. (These are called HIV-1 R5 strains.)

Several groups of humans (mainly with ancestors who lived in northern Europe)

have a 32-nucleotide deletion in the gene encoding the CCR5 receptor. (This

allele is known as *CCR5-Δ32*.) Homozygotes exhibit a near (but not quite total)

resistance to HIV-1 infection and heterozygotes exhibit significantly slower than

normal progression of AIDS. Subsequent population genetic studies, yielding

exceptionally high concordance with the data on the distribution of the *CCR5-Δ32*

allele in tested populations, explain the distribution of this allele as the result of

strong episodic selection favoring this allele because it also produced resistance to

the black plague (Duncan et al., 2005). Given very detailed knowledge of the

timing and incidence of plagues, the expected distribution of the allele matches

the available data closely; other suggested selection pressures and histories that

might account for the distribution of the allele do not come close to doing as

well.[9]

Such population genetic data and scenarios, though made plausible by our

understanding of the specific mechanism of action of both HIV-1 and our (less

thorough) understanding of its contribution to resistance to black plague, are not

---

[9] In the interest of brevity, I omit a second medical example discussed briefly by Dean et al. (2002). This concerns cystic fibrosis, a genetic disease (or condition) caused by a gene that appears to have been strongly selected because it produces to typhoid fever produced by *Salmonella typhi*, as was demonstrated by Pier et al. (1998). And because both the evolutionary issues about CF and AIDS resistance are not tractable to direct molecular resolution, I should also cite a paper that has demonstrated at the molecular level that severe selection such as that produced by typhoid fever and Black Death can produce stably modified populations with novel resistance genes present in high proportions. The paper (Navas et al., 2007) reports experimental evolution of *C. elegans* challenged to survive on food laden with a fatal pathogen.

by themselves wholly decisive. But in this instance they appear to provide the best currently-available account of a significant source of genetic resistance to HIV infection. And even though it does not decisively establish the selective grounds for the prevalence of *CCR5-Δ32* in various populations, it illustrates exactly the sort of scenario that I think will prove to be quite common: episodic, strong selection, acting sporadically, crucial in maintaining a balance in the arms races between hosts and disease vectors, predators and prey, invading competitors, and 'physical' environmental conditions such as drought, flood, climate, change, and the like. Such sporadic selection will heighten the impact of one or another genetic or genotypic variant in a population (especially at population bottlenecks) with a sometimes dramatic effect on the genetic and/or phenotypic composition of the population. Just like the 'ghosts' of episodes in which numerous transpositions occurred, so the 'ghosts' of these strong selection episodes will remain in affected populations, available for mobilization when ecological conditions prove favorable. Selection must balance the impact of such episodes over long intervals of evolutionary time, while maintaining sufficient genomic stability, organismic viability, and reproductive capacity for maintenance of the lineage. This is the image of evolution with which I operate. If it is sound, it is not plausible that selection operates primarily on genes.

I have just argued that to understand selection pressures on key genes we often need to understand the history of the relevant ecology/ies. Although I have

no direct evidence bearing on the genes that affect heart function, there is a plausible view in the semi-popular medical literature that illustrates the sort of issue that needs to be evaluated in connection with Sapienza's arguments about the high selective value of variation of only a few of the genes affecting heart function. It is often suggested that the ecological change that modern agriculture, together with increased wealth and food availability in many cultures, has radically altered the dietary regime of humans. Large populations now have access to sweets, fats, and an abundance of calories that was extremely rare in human history. (This is combined, of course, with radical changes in public health measures, medical care, protection from infection, and much more of the sort.) The increase in longevity, but also in obesity, type II diabetes, hypertension, and heart attack are all related to the changes in diet and life styles. Heart attacks occur predominantly at an age beyond that to which most humans lived until a relatively few centuries ago. It is not implausible that the selective value of the genes that Sapienza discusses, the ones with highly heritable effects on hypertension, etc., have altered significantly with the alteration of dietary regime, the altered longevity of humans, and the other changes just alluded to. Again, the genes that affect hypercholesterolemia surely have other effects than their effects on cholesterol levels. I do not know what these effects are, nor whether they are relevant to the long-term selective values of the alleles at issue. Before one jumps from the positive impact of the medications blocking excessive

cholesterol in medically relevant circumstances to the selective values of certain

alleles, a vast range of additional information is required. Such information may

be available (though I suspect it is not), but until it is presented and subjected to

careful evaluation, the case has not been made that these medical results suffice to

determine the selective value of the genes whose action has been blocked. As

best I can see, Sapienza's argument about the relative paucity of overdominance

in high-throughput tests does not address these points.

My argument is not decisive. It does not prove that selection acts

primarily via multi-level and multi-factorial causation, but neither does

Sapienza's argument make a decisive case that selection acts primarily on genes.

It is, of course, problematic for his view that selection can 'see' a relatively small

number of independent genes bearing on a selectively relevant trait at once,

caught up as they are in immensely complex networks. In general, it seems more

plausible that selection 'acts' on some of the specific traits that result from the

workings of genes and gene control networks and that the chips fall as they may

for individual genes. This would be one way in which selection could 'act' on

phenotypes (or intermediate steps toward phenotypes) such as pathogen and

drought resistance and leave the appearance that much mutation is neutral. But it

is also important to recognize (as the case of *CCR5-Δ32* suggests) can 'act'

episodically and when it does so, it need not 'see' all the genes on which it acts at

once. Selection can act sequentially on different genes, during relatively short

periods of intense selection when, due to the some salient circumstance (such as the flourishing of a major new pathogen) the impact of one allele rather than another given the stable configuration of the rest of the network, makes a large difference to organismal survival or reproductive success. It can act on the few salient genes in humans that affect hypercholesterolemia rather drastically when they lead sedentary lives, overeat drastically, and don't have the right medical remedies available for averting the consequences of doing so. Similarly, to support a point that Sapienza made in his postscript, when a transposable element breaks the restraints that have kept its transposition in check or enters an organism that has no such restraints, it can spread enormously rapidly until controls on its spread are established. But unless such controls are ultimately established, the genomes in which it is spreading will ultimately break down – and when adequate constraints to prevent genomic breakdown are established, they will involve a balance between organismal and genic factors. It is precisely the need for such a balance that I believe is omitted from Sapienza's account of the matter.

Let me return, finally, to Sapienza's position about the relatively few genes that, he suggests, are salient for the maintenance of hearts. It is a non-trivial question whether the situation here is more closely analogous to sporadic selection altering the genes affecting cystic fibrosis and resistance to AIDS along the lines I sketched or to genes affecting height in the manner Sapienza suggests. This question is far from closed by our arguments. My response to Sapienza's

argument is a bit like a just-so story – but with a difference. I don't pretend to know the truth in this matter. Rather, I am putting forward a plea for major extension of research to determine the balance of circumstances that influence the selective values of alleles *and other biological factors* that affect heart function. The research required should take full account of the wide range of relevant conditions at many different levels that must be considered before reaching a firm conclusion about the balancing act required of selection in cases like this. Arguments about the percentage of trait variance that is currently due to a given gene or allele don't carry much weight until they have been put through studies that cover relevant intervals of evolutionary time and relevant ranges of environments, and until the full range of (presumably pleiotropic) effects of the relevant genes have been taken into consideration. If my argument is correct, resolution of such questions requires an integrated examination of selectively relevant systems and contingencies at many levels, including the genetic, genomic, cellular, organismal, and ecological levels. Resolution of the debate about whether selection acts primarily on genes will remain unresolved unless it integrates considerations of life histories, diet, hormonal mechanisms, neurological mechanisms, and genetic and epigenetic mechanisms. But the very fact that all these considerations are relevant suggests that selection acts in multifarious ways. It also suggests that although selection 'acts' on genes it does not act primarily on genes.

References

Brandon, R. (1985). Phenotypic plasticity, cultural transmission, and human
sociobiology. In J. Fetzer (Ed.), *Sociobiology and epistemology* (pp. 57-
73). Dordrecht: D. Reidel.

Brandon, R., & Burian, R. (Eds.). (1984). *Genes, organisms, populations:
Controversies over the units of selection*. Cambridge, MA: MIT Press.

Brandon, R., & Nijhout, H. (2006). The empirical nonequivalence of genic and
genotypic models of selection: A (decisive) refutation of genic
selectionism and pluralistic genic selectionism. *Philosophy of Science, 73,*
277-297.

Burian, R. (1997). Exploratory experimentation and the role of histochemical
techniques in the work of Jean Brachet, 1938-1952. *History and
Philosophy of the Life Sciences, 19,* 27-45.

Burian, R. (2000). On the internal dynamics of Mendelian genetics. *Comptes
rendus de l'Académie des Sciences, Paris. Série III, Sciences de la Vie /
Life Sciences, 323,*1127-1137.

Burian, R. (2005). Too many kinds of genes? Some problems posed by
discontinuities in gene concepts and the continuity of the genetic material.
The Epistemology of Development, Genetics, and Evolution (pp. 166-178).
New York: Cambridge University Press.

Burian, R. (2007). On microRNA and the need for exploratory experimentation in post-genomic molecular biology. *History and Philosophy of the Life Sciences, 26 (3)*, in press.

Burian, R., Gayon, J., & Zallen, D. (1988). The singular fate of genetics in the history of French biology, 1900-1940. *Journal of the History of Biology, 21,* 357-402.

Carlson, E. (1966). *The gene: A critical history*. London: W. B. Saunders.

Darden, L. (1991). *Theory change in science: Strategies from Mendelian genetics*. New York: Oxford University Press.

Darwin, C. (1859). *On the origin of species by means of natural selection.* London: John Murray.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Dawkins, R. (1982). *The extended phenotype: The gene as the unit of selection.* Oxford: Oxford University Press.

Dean M., Carrington, M., & O'Brien, S.J. (2002). Balanced polymorphism selected by genetic versus infectious human disease. *Annual Review of Genomics and Human Genetics, 3*, 263-292.

Duncan S.R., Scott, S., & Duncan, C.J. (2005). Reappraisal of the historical selective pressures for the CCR5-D32 mutation. *Journal of Medical Genetics, 42*. 205-208.

Elliott, K. (2007). Varieties of exploratory experimentation in nanotoxicology. *History and Philosophy of the Life Sciences, 26 (3)*, in pres*s.*

Falk, R. (1986). What is a gene? *Studies in History and Philosophy of Science, 17,* 133-173.

Franklin, L. (2005). Exploratory experiments. *Philosophy of Science, 72,* 888-899.

Hull, D. (1980). Individuality and selection. *Annual Review of Ecology and Systematics, 11,* 311-332.

Hull, D. (1981). Units of evolution: A metaphysical essay. In U. Jensen & R. Harré (Eds.), *The philosophy of evolution* (pp. 23-44). Brighton: Harvester Press.

Johannsen, W. (1909). *Elemente der exakten erblichkeitslehre*. Jena: G. Fischer.

Lewontin, R. (1970). The units of selection. *Annual Review of Ecology and Systematics, 1,* 1-18.

Lewontin, R., & Dunn, L. (1960). The evolutionary dynamics of a polymorphism in the house mouse. *Genetics, 45,* 705-722.

Mattick, J. (2004). RNA regulation: A new genetics? *Nature Reviews Genetics, 5,* 316-323.

Mayr, E. (1997). The objects of selection. *Proceedings of the National Academy of Sciences, USA, 94,* 2091-2094.

Navas A., Cobas, G., Talavera, M., Ayala, J.A., Lopez, J.A., & Martínez, J.L. (2007). Experimental validation of Haldane's hypothesis on the role of

infection as an evolutionary force for Metazoans. *Proceedings of the National Academy of Sciences, USA, 104(34)*, 13728-13731.

O'Malley, M. (2007). Exploratory experimentation and scientific practice: Metagenomics and the proteorhodopsin case. *History and Philosophy of the Life Sciences, 26 (3),* in press.

Pier G.B., Grout, M., Zaidi, T., Meluleni, G., Mueschenborn, S.S., Banting, G., Ratcliff, R., Evans, M.J., & Colledge, W.H. (1998). Salmonella typhi uses CFTR to enter intestinal epithelial cells. *Nature, 393*, 79-82.

Pigliucci, M. (2001). *Phenotypic plasticity: Beyond nature and nurture*. Baltimore, MD: The Johns Hopkins University Press.

Portin, P. (1993). The concept of the gene: Short history and present status. *The Quarterly Review of Biology, 68,* 173-223.

Ragsdale, D. (2002). Aphid biology and life history: Implications for management. Available at: http://agronomy.cfans.umn.edu/sites/1e65c3d - ce73-46af-a196-12323314ecc5/uploads/ragsdalesbabiology.pdf.

Rutherford, S. (2003). Between genotype and phenotype: Protein chaperones and evolvability. *Nature Reviews Genetics, 4,* 263-275.

Rutherford, S., & Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature, 396,* 336-342.

Sandovici I., Kassovska-Bratinova, S., Vaughan, J.E., Stewart, R., Leppert, M. & Sapienza, C. 2006. Human Imprinted Chromosomal Regions Are Historical Hot-Spots of Recombination. *PLoS Genetics*, *2(7)*: e101.

Schier, A., & Giraldez, A. (2006). MicroRNA function and mechanism: Insights from zebra fish. *Cold Spring Harbor Symposia on Quantitative Biology, 71,* 195-203.

Snyder, M., & Gerstein, M. (2003). Defining genes in the genomics era. *Science, 300,* 258-260.

Sober, E. (Ed.). (1984). *Conceptual issues in evolutionary biology: An anthology*. Cambridge, MA: MIT Press.

Sober, E., & Lewontin, R. (1982). Artifact, cause, and genic selectionism. *Philosophy of Science, 49,* 157-180.

Stadler, B. (1995). Adaptive allocation of resources and life-history trade-offs in aphids relative to plant quality. *Oecologia, 102,* 246-254.

Stadler, L. (1954). The gene. *Science, 120,* 811-819.

Steinle, F. (1997). Entering new fields: Exploratory uses of experimentation. *Philosophy of Science, 4,* S65-S74.

Sterelny, K., & Griffiths, P. (1999). *Sex and death: An introduction to philosophy of biology*. Chicago: University of Chicago Press.

Van Speybroeck, L., Van De Vijver, G., & De Waele, D. (Eds.). (2003). *From epigenesis to epigenetics: The genome in context*. New York: The New York Academy of the Sciences.

Wagner, G., Chiu, C., & Hansen, T. (1999). Is Hsp90 a regulator of evolvability? *Journal of Experimental Zoology, 285B,* 116-118.

Williams, G. (1966). *Adaptation and natural selection*. Princeton, NJ: Princeton University Press.

Wimsatt, W. (1980). Reductionistic research strategies and their biases in the units of selection controversy. In T. Nickles (Ed.), *Scientific discovery: Case studies* (pp. 213-259). Dordrecht, Holland: Reidel.

West-Eberhard M.J. (2003). *Developmental Plasticity and Evolution*. New York and Oxford: Oxford University Press.

Wood, A., & Oakey, R. (2006). Genomic imprinting in mammals: Emerging themes and established theories. *PLoS Genetics, 2,* e147.