To the Graduate Council:

I am submitting herewith a dissertation written by Massimo Pigliucci entitled "Dangerous Habits: examining the philosophical baggage of biological research." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a Major in Philosophy.

Jonathan Kaplan

Major Professor

We have read this dissertation
and recommend its acceptance:

Kathy Bohstedt

John Nolt

Daniel Simberloff

Acceptance for the Council:

Anne Mayhew

Vice Provost and

Dean of Graduate Studies

(Original signatured are on file with official student records.)

Dangerous Habits:

Examining the Philosophical Baggage of Biological Research

A dissertation presented for the

Doctor of Philosophy degree

The University of Tennessee, Knoxville

Massimo Pigliucci

May 2003

Acknowledgments

None of this would have been possible without the help of many people, and I wish to single out a few for explicit thanks. First and foremost, my wife, Melissa Brenneman, to whom I solemnly promised that this was the last degree I would ever sought in my life. Then my advisor, colleague, and friend, Jonathan Kaplan. The colleagues who served on my committee, Kathy Bohstedt, John Nolt, and Dan Simberloff. My other philosophy teachers, Richard Aquila, Shari Clough, Sheldon Cohen, and Betsy Postow. And finally the members of my bio-lab, who have been very patient with me throughout this project: Josh Banta, Jeff Dole, and Norris Muth.

Abstract

Science is about conceptualizing the natural world in a way that can be understood by human beings while at the same time reflecting as much as possible what we can empirically infer about how the world actually is. Among the crucial tools that allow scientists to formulate hypotheses and to contribute to a progressive understanding of nature are the use of imagery and metaphors on the one hand, and the ability to assume certain starting points on which to build new avenues of inquiry on the other hand. The premise of this work is that, in the words of philosopher of science Daniel Dennett, "There is no such thing as philosophy-free science; there is only science whose philosophical baggage is taken on board without examination." The purpose here is precisely to examine some of this philosophical baggage, and to see where such analysis may lead us. Obviously, it is not possible for a single project to take on science as a whole, or even an entire discipline such as physics, or ecology. Therefore, the core of this work is constituted by a series of eight case studies drawn from the field of evolutionary biology, with which I am particularly familiar as a practicing biologist. The hope is that combining expertise in both philosophy and science in one person, the resulting insights might be useful for the practicing scientist as well as because of their value as philosophical inquires. The dissertation is a combination of conceptual analysis and science criticism applied to specific questions in organismal biology, largely of an evolutionary nature, with the eight case studies grouped into three broad categories: I-Unexamined Baggage, II-Bad Habits, and III-Good and bad metaphors.

Table of Contents

List of Tables:

List of Figures:

List of Text Boxes:

"There is no such thing as philosophy-free science; there is only science whose philosophical baggage is taken on board without examination."

—Daniel Dennett (*Darwin's Dangerous Idea*, 1995, p. 21)

Introduction

Science is about conceptualizing the natural world in a way that can be understood by human beings while at the same time reflecting as much as possible what we can empirically infer about how the world actually is. Conceptualization itself is a complex phenomenon that has been at the center of both philosophical (e.g., Margolis and Laurence 1999) and scientific (Fedlman 2000; Wolford et al. 2000) investigation. Be that as it may, among the crucial tools that allow scientists to formulate hypotheses and to contribute to a progressive understanding of nature (Kitcher 1995) are the use of imagery and metaphors on the one hand, and the ability to assume certain starting points on which to build new avenues of inquiry on the other hand.

Assumptions, as well as visual and verbal metaphors, are an inescapable part not only of doing science, but of functioning as cognitive animals, and there is no suggestion to be found in this dissertation that we should avoid using such tools in either science or everyday life. However, it is one of the tasks of philosophy to continuously examine the soundness of the assumptions and metaphors that are used by other disciplines (Berlin 1979; Prawitz 1997), in order to provide a level of critical analysis often neglected by practitioners within a given field, too busy solving more specific problems to engage in meta-analyses.

The premise of this work is that, in the words of philosopher of science Daniel

Dennett (1995), “There is no such thing as philosophy-free science; there is only science whose philosophical baggage is taken on board without examination.” The purpose here is precisely to examine some of this philosophical baggage, and to see where such analysis may lead us. Obviously, it is not possible for a single project to take on science as a whole, or even an entire discipline such as physics, or ecology. Therefore, the core of what follows is constituted by a series of eight case studies drawn from the field of evolutionary biology, with which I am particularly familiar as a practicing biologist. The hope is that combining expertise in both philosophy and science in one person, the resulting insights might be useful for the practicing scientist as well as because of their value as philosophical inquires.

Not that the relationship between science and philosophy hasn’t been controversial, on both sides of the divide. For example, physicist Steven Weinberg (1992a) -- whom I will refer to several times in what follows -- has argued in an essay appropriately entitled “Against philosophy” that philosophers are not only useless to science, but in fact at times positively harmful. Weinberg maintains that philosophy has never solved a scientific problem (and is thus useless to science), and that furthermore some philosophical schools of thought have temporarily hindered the progress of science. His favorite example is the alleged slow acceptance of quantum mechanical theory in the early part of the 20th century, which Weinberg attributes at least in part to the fact that some scientists were buying into the logical positivist’s contention that science should never make use of unobservable quantities (of which quantum mechanical theory is in fact full).

There are several lines of argument that can be used against Weinberg’s position, of course. Concerning his specific example, the evidence that there actually was a slow

acceptance of quantum mechanics by physicists is slim at best. Furthermore, the historical work to show that this alleged slowing effect was due to the philosophical positions of individual scientists (like Einstein), rather than to their legitimate skepticism about a completely novel way of conceptualizing the world, has simply not been done. More fundamentally, though, one could rightfully ask Weinberg why is it that philosophers are expected to solve scientific problems in order to be "useful." Why not ask when was the last time that scientists helped resolving philosophical questions instead? Or, better yet, why not accept that science and philosophy can move along parallel and intercommunicating tracks, without those tracks having to converge or replace each other?[1]

On the other hand, many philosophers have not been particularly comfortable themselves with too snugly a relationship between science and philosophy. Tom Sorell (1991) has published one of the most comprehensive (and rather balanced) presentation of the skeptical view internal to philosophy, with an eye toward the "colonization" of philosophy by science (an attitude referred to as "scientism," and of which both scientists and philosophers have been accused, in particular the Churchlands for their work on "eliminativism" in philosophy of mind [e.g., Churchland 1988] and E.O. Wilson for his book *Consilience: the Unity of Knowledge* [1998]). Having frequented both philosophy and science departments and academic meetings, I can personally testify to the strange mix of reciprocal curiosity, distrust, and occasionally contempt that the two camps feel for each other.

[1] This discussion, of course, could lead to the exploration of what philosophy itself is, something that is clearly beyond the scope of my work. However, another way to think of the relationship between some philosophy and science is that the former plays the role of "place holder" and conceptual clarificator of issues that cannot at the moment be addressed empirically. When this can be done, what was formerly a philosophical question has been turned into a scientific one. Even so, it would be unwise to dismiss the role that philosophy may have played in preparing the terrain for scientific investigation.

The purpose of my work here, of course, is certainly not as ambitious as to attempt a reconciliation of the "two cultures" (Snow 1959), but more simply to use the critical tools of philosophy of science to examine some of the often unexamined baggage that goes into biological research. Before proceeding any further, however, it may be useful to briefly consider what the scope of philosophy of science itself may be, in order to put in sharper focus the ambitions and limits of this dissertation. Broadly speaking, one can think of philosophy of science has having three aims: to inquire into the nature of science itself; to conduct a conceptual analysis of actual practices in which scientists engage; and to practice what we may term "science criticism," i.e., the critical meta-evaluation of claims and methodologies adopted within specific fields of science.

As far as **studying the nature of science** itself is concerned, philosophy can be either prescriptive or descriptive, with a great tradition of inquiry in both areas. Prescriptive philosophy of science attempts to produce workable suggestions for how science could best be practiced by scientists themselves. For example, Karl Popper (1968) clearly thought that science must work by falsifying hypotheses, since it is not possible to actually prove them. Popper even thought to have solved Hume's famous problem of induction (Hume 1748), a claim that was tempered by more sophisticated work on falsificationism, for example by Lakatos (1977). Despite the widespread perception that scientists don't read philosophy, it is nonetheless astounding that the overwhelming majority of scientists and science textbook authors do in fact at least formally accept the idea that good science proceeds by falsification (ironically, philosophers of science have instead largely abandoned "naive," Popper-style, falsificationism). Descriptive philosophy of science in this realm attempts not to tell scientists how to do their work, but rather to observe them as if they were laboratory

animals and then think about what their behavior might tell the philosopher. The classical example here is the work of Thomas Kuhn (1970) and his idea of the contrast between scientific "revolutions" and "normal," puzzle-solving science. The work presented below does not engage directly into the kind of philosophical analysis that goes into this sort of "big picture" work on how science does (or should) work. Instead, what I attempt to do is a mixture of the next two levels of analysis.

Conceptual analysis of how scientists think and talk about their subject matters also has a long tradition in philosophy, and has in fact probably produced the bulk of the recent literature in the field. Among the examples that come to mind are investigations into the nature of natural selection (e.g., Sober 1984; Waters 1991; Hull 2000; Matthen and Ariew 2002; Millstein 2002), the use of the concept of causality in science (e.g., Sober and Lewontin 1982; Rieppel 1990; Pigliucci in press-b), and the never ending discussions on what biological species are (see Chapter 4 and references therein). The point here is not -- like Weinberg would argue -- to see if philosophers are better than scientists at solving scientific problems. Rather, the idea is to put the essence of philosophical inquiry -- the ability to conduct rigorous analysis of the meaning of concepts and of how they are applied -- at the service of scientists to hopefully help them cut through the conceptual fog and get to questions that are actually answerable on empirical grounds.

The second component of what I try to do here, and the third overall area of inquiry in philosophy of science, deals with what could loosely be termed **science criticism**. Here again there are two distinct, albeit often interdependent, roles that philosophy plays. On the one hand, there is the idea of a critical analysis of the soundness of certain scientific claims, from the point of view of how good the science

itself is in any given case. An example would be the rigorous criticism of sociobiology and, more recently, of evolutionary psychology, that aims at showing how much of a disconnect there often is between the empirical evidence adduced to back up certain claims within those disciplines and the claims themselves (e.g., Lewontin 1998; Kaplan 2001, 2002). On the other hand, science criticism can have a decidedly "applied" component, in that one can take scientists to task for the social and ethical implications of their claims, especially when such claims appear to be only weakly justified. For example, talk of genes "for" certain traits (see Chapter 2 and references therein), can lead to heated debates affecting social policy and societal attitudes (see Chapter 6 and references therein). Science criticism can be irritating for the scientist, but an argument can be made that -- if properly carried out -- it actually benefits both science itself (by keeping scientists from making claims that find little justification in their science) and society at large (by providing a system of checks on the increasingly prominent and powerful role of science in people's lives).

What follows, then, is a combination of conceptual analysis and science criticism applied to specific questions in organismal biology, largely of an evolutionary nature. I present eight of these case studies, grouped into three broad categories. The categorization is not meant to be strict, and it is mostly for organizational purposes within this body of work, but it does reflect what I think are different aspects of the general question of the philosophical baggage that accompanies scientific research. The sections and chapters are as follows:

1. **Unexamined Baggage.** These four chapters are those that can be most properly conceptualized under the idea of unexamined philosophical baggage that goes into

biological work. They include a discussion of the following questions:

1. Why do we consider sciences that use statistical analyses and probabilities as "soft" when compared to the so-called "hard" sciences such as physics and chemistry?
2. What do biologists mean when they talk of genes being "for" a given phenotypic trait?
3. Why do both biologists and philosophers largely insist in denying the existence of human races, while the first ones then go on and apply the very concept of race to non-human organisms?
4. Why do we think in essentialist terms when we talk about the species "problem"?

2. **Bad Habits.** These chapters question some underlying philosophical positions that have been integrated into the way scientists think about their work, to the point that the paths by which they pursue such work is biased or otherwise affected by these bad habits of mind. In particular, I will focus on:

    1. Why do we insist in using the tools of "null hypotheses" and "p-values" even though there is plenty of reasons to believe that they are singularly uninformative, and furthermore given that there are better alternatives readily available?
    2. Why do we keep debating nature-nurture questions in humans when we know what the answer looks like in other organisms, and moreover we know that we simply cannot do the right experiments (for ethical as well as logistical reasons) in humans?

3. **Good and bad metaphors.** Scientists, like lay people, make ample use of metaphors and imagery to inspire their work and to think about it in more creative ways. Sometimes these metaphors are useful, at other times they are more likely to hinder scientific understanding and progress. The two chapters in this section discuss questions such as:
    1. Why do we keep talking about "selfish" genes even though excellent arguments have been made that the concept is unhelpful at best?
    2. To what extent do architectural metaphors help think about the relative role of natural selection and constraints in shaping adaptive phenotypic evolution?

Obviously, many more similar case studies could be examined, and even those presented here are certainly not treated in any way close to an exhaustive manner. However, I hope the reader will find plenty of food for thought while going through them. If such can be found regardless of the fact that the reader is a philosopher or a scientist, then a large part of my broader objectives will have been achieved.

Part I - Unexamined baggage

The premise of this work is in Dennett's (1995) suggestion that science simply cannot function without adopting certain (philosophical) assumptions as starting points. These assumptions are often not directly testable in an empirical fashion, and at any rate scientists usually do not even attempt any such test. One could think of the assumptions in question as reflecting the conceptual framework within which scientific research is being conducted. It then becomes interesting for both the scientist and the philosopher to ask what are the basis of this framework, and what would happen if one were to adopt a different one.

The first part of the present work deals with what I have loosely termed "unexamined baggage" in the sense of Dennett. This is a set of ideas and assumptions about specific objects of study in biology that guide the empirical research without surfacing to a visible level of debate within the biological community. Philosophers, on the other hand, will be more familiar with some of the themes discussed in these four chapters, as especially three of them (Chapters 2, 3 and 4) have generated a fairly large amount of literature in philosophy journals, literature that in the case of the subject matter of Chapter 3 is not even confined only to the field of philosophy of science.

The reader needs to bear in mind throughout the following that my aim is to explore the boundary between science and philosophy, not to engage in either science or philosophy *per se*. What this means is that the chapters in this dissertation may seem too "scientific" to philosophers and too "abstract" to scientists. That, however, is the very nature of the beast.

I begin with an analysis of the difference between "soft" and "hard" sciences, a rather sensitive matter for those (like social scientists and organismal biologists) who often found themselves on a defensive position, attempting to justify why is it exactly that their disciplines have not been characterized by the spectacular results of, say, physics or molecular biology. My analysis is couched in terms of the four Aristotelian causes which, when translated into the sorts of questions that biologists tend to ask, provide an excellent framework to understand the difference between so-called soft and hard sciences. Moreover, I maintain that this framework makes it possible both to understand why the soft sciences are no less rigorous than the hard ones, and why there are intrinsic limits in the currently popular "integrative biology" that is all the rage among biology departments and funding agencies.

The second chapter explores what we might possibly mean when we say that certain genes are "for" certain traits. This is a very commonly encountered locution, and most biologists don't seem to be aware of (or bothered by) the fact that they are using it in very different manners, depending on both the context of the discussion and the sub-discipline within which they happen to be working. Especially in this emerging era of "genomics," "proteomics," "metabolomics" (Hurst 1999; Jablonka et al. 2002) and several other fuzzily defined new disciplines, it seems to me that a rigorous analysis of when one is justified in actually using talk of genes being "for" something is very timely. Indeed, this is an area of philosophical criticism that can easily branch into very practical consequences at the societal level: just consider how often we come across media reports of a new gene "for" X (e.g., obesity, schizophrenia, sexual orientation, etc.) being discovered by biologists.

The third chapter in this section deals with the assumption that human biological

races don't exist because folk races are "obviously" not biologically meaningful. This position is challenged on the basis of an analysis of the actual usage of the word "race" by biologists, as well as on what we actually know of the recent evolution of human populations. I will suggest that there *are* indeed human biological races, and that to deny it smacks of political correctness on the part of practicing biologists that should know better. On the other hand, I will also propose that there is no simple mapping between biological and "folk" conceptions of what a human race is, which explains why biology is not very helpful when it comes to issues of social policy and attitudes toward race.

The last chapter of this group deals with what is perhaps one of the most discussed, controversial, and frustrating topics in the evolutionary biology literature: the concept of species. Rivers of ink have flown in endless attempts at settling the issue, and plenty of biologists and philosophers have weighed in over several decades of discussion. It seems to me that when a controversy lasts that long, with plenty of smart and largely rational people completely unable to come even close to an agreement, this is because of one (or both) of two reasons: either there is simply not enough empirical evidence to settle the matter, or the conceptual framework within which the question is posed has not been well thought out. Recently, at least one of the participants to this debate (Hey 2001a,b) has argued for the first possibility, which I however consider highly unlikely. The second alternative is more akin to Wittgenstein's famous suggestion that many philosophical (and, dare I say, some scientific) problems are such only because people have not clarified what they mean when they use certain terms or ask certain questions. In this spirit, I shall provide an entirely Wittgenstenian (dis-)solution to the species problem in biology.

Chapter 1—Are ecology and evolutionary biology "soft" sciences?

"Eliminate all other factors, and the one which remains must be the truth."

(Sherlock Holmes, in Arthur Conan Doyle's *The Sign of the Four*)

I begin this dissertation with the consideration of some of what I refer to as "unexamined baggage" in the practice of ecology and evolutionary biology. This chapter deals perhaps with one of the most important, and yet seldom questioned (by biologists), assumptions concerning the philosophy underlying biological research: the status of the physical sciences as "models" for how all "hard" science should proceed. This is obviously a topic on which much can be and has been written, but I will focus here on some aspects of the question that are of direct relevance to practicing organismal biologists. My conclusion is that organismal biology is as much of a science as any other, and that it is its very nature that dictates an approach partially distinct from those of physics and chemistry, without because of this implying that there is anything lacking or "soft" in ecological and evolutionary studies.

A recurring complaint

Are ecology and evolutionary biology making progress, and if so, in what sense? On the face of it, the question may seem nonsensical. Just consider the number of papers on all sorts of evolutionary and ecological questions that are been published at a steady rhythm, and one cannot have doubts about the fact that progress is being made. And yet,

one also cannot avoid the nagging feeling that what we see in most published papers is the accumulation of new information, not necessarily progress in the sense of a conceptual understanding of the objects of study. The latter, after all, doesn't necessarily follow from the former.

I contend that little conceptual (as opposed to special problem-solving) progress has been made in what I will refer to as the evo-eco disciplines when compared to physics (the queen of the so-called "hard sciences") or even (but to a lesser extent) other biological fields such as molecular biology. Furthermore, I will argue that this is nothing to be ashamed of, because the situation is at least in part the result of the intrinsic nature of ecology and evolutionary biology and of other "soft" sciences such as psychology and sociology. Of course, much depends on what one means by "progress" and how this is measured. However, in my other incarnation as a practicing evolutionary ecologist, I have often shared my frustration with too many colleagues about the fact that our fields don't seem to be going anywhere in particular for the problem to be only imaginary. As Michael Turelli, a leading mathematical biologist, put it at the 2001 evolution meetings, "questions in our field are never settled, they simply go in and out of fashion."

Before considering why progress in evo-eco seems slow, one should of course be prepared to answer the more general question: is there progress in science at all? Again, this is a much more thorny issue than most scientists realize, and it requires familiarity with the philosophical concepts of truth, discovery, induction, and the like. I will not attempt to expound on these matters here, instead giving the reader my answer and some references to dig deeper. While clearly science cannot arrive at "the truth," whatever that is, it is equally clear that it has made progress in understanding the natural world (Maxwell 1979; Kitcher 1995; Hull 2000). In fact, somewhat ironically, Kitcher (1995)

has used evolutionary theory as a paradigm of progress in science. Darwin's theory of evolution was "adopted on the basis of compelling reasons," and it was not simply the replacement of a theory with another, but a transition involving multifaceted consequences on the practice of biology. It indubitably led to cumulative theoretical and empirical knowledge in the biological sciences.

The question that I wish to address here is more limited in scope and more focused on sub-theories within evo-eco: does research in evo-eco resemble more the hard sciences such as physics and chemistry, or the soft ones like psychology and sociology? This now familiar distinction was introduced by Windelband (1894) in his *History and Natural Science*: what he called "nomothetic" knowledge is the one sought by most natural sciences and it consists in the discovery of general laws in order to understand and eventually master nature; what Windelband referred to as "idiographic" knowledge is typical of historical sciences and relies on descriptions of individual and unique aspects of reality, the main aim of which is reconstruction of events using a coherent narrative.

Some biologists, ironically from the ranks of paleontology – the most historically contingent of biological disciplines – have made a conscious but spotty effort to move the field as close as possible to the nomothetic "ideal" (Raup and Gould 1974; Gould 1980). Some further discussion of the potential and limits of the idea of evo-eco as hard science has been carried out since then, mostly by ecologists. The following brief discussion is not meant to be comprehensive, but only to assure the reader that I am not making this up. Plenty of serious researchers in our discipline have raised similar problems before, every time only to be largely ignored by the everyday torrent of puzzle-solving scientific papers that characterize most of what philosopher Thomas Kuhn (1970) called "normal science."

One of the loudest salvos was fired back in 1982 by none other than Ernst Mayr (1982), one of the fathers of the neo-Darwinian synthesis (Mayr and Provine 1980). In a section of his *The Growth of Biological Thought* entitled "Laws in the physical and biological sciences," Mayr (1982) comes to several of the conclusions I discuss in more detail below, and in particular to the idea that evolutionary biology, being an inherently historical science, has by nature to follow a very different modus operandi than physics and the other "hard" sciences. As we shall see, I depart from Mayr when he insists that the focus on concepts is the key to understanding evo-eco.

A comprehensive analysis of the historical conceptual developments in ecology has been presented by McIntosh (1985) in his *The Background of Ecology: Concept and Theory*, which is a good starting point to understand the current status of this discussion. Peters' (1991) *A Critique of Ecology* contains reflections on the epistemology of ecological theory that actually apply to most of the scientific enterprise at large. Peters makes clear that an in-depth analysis of ecological statements from the point of view of the information they carry reveals troubles with the whole discipline (how many of us have jokingly told our students in introductory classes that "ecology is the elucidation of the obvious"?). Peters concludes that ecology should refocus on simple questions of fact and observation, especially those of general relevance to science and society. This also happens to be my prescription, though reached by different means than Peters'.

From a different perspective, Lawton (1999) questions in what sense there are laws in ecology, concluding that while ecologists can produce generalized formulations based on observable tendencies, there is no place in ecology for laws in the sense meant in physics, i.e., statements about the natural world that are universally true. As a corollary, Lawton then suggests that ecologists should pay less attention to the "middle ground" of

community ecology (an inflammatory statement, to be sure) and rely less on reductionism and experimental manipulation. As we shall see, my starting point below is similar, but I reach quite different conclusions.

Additional soul searching has been attempted by Shrader-Frechette and McCoy (1994) with their *Method in Ecology: Strategies for Conservation*, which while focusing on the practical aspect of conservation biology, has much to offer to evo-eco scientists in general. These authors begin by discussing the shaky grounds of such key concepts as "community" and "stability" and conclude that general and predictive theories in ecology are simply impossible. Their recipe for success falls along the lines of Peters' suggestions: what they call the "case study method" amounts to a serious focus on what ecology is (arguably) all about: natural history, good autoecology, and the clearest possible definition of hypotheses. Their discussion of the scientific vs. ethical implications of an emphasis on type I errors is interesting, though it suffers from a lack of reference to the obvious alternative to classical hypothesis testing that will be mentioned below: Bayesian analysis. In a similar context, my discussion of the tyranny of null hypotheses (Chapter 5) is particularly germane to the long, and still ongoing, controversy in ecology about the usefulness and relevance of "null models" (e.g., Gotelli and Graves 1996; Hubbell 2001).

These recurring self-examinations notwithstanding, most biologists are much more preoccupied with carrying out business as usual, relegating questions of philosophy of science to coffee breaks and beer outings. Similar discussions, with similarly low impact on the everyday practice of science, have taken place for some time within the soft science par excellence, psychology, and the results are worth considering in some detail given what I think are very close parallels with evo-eco research.

Meehl (1978) has been perhaps one of the sharpest critics of the possibility of

general theoretical progress in the soft sciences. His remark that "most theories in psychology never die, they slowly fade away" closely resembles Turelli's comment mentioned above. He went on to predict that we will probably never have a substantive general theory in personality or social psychology. Gergen (1973) went further to suggest that if the events of interest to psychologists are capricious, the discipline should be replaced by the equivalent of natural history (similar to Shrader-Frechette and McCoy's take on ecology discussed above), because the continued attempt to build general laws of social behavior may be misdirected. Since the study of social psychology is primarily a historical undertaking, he suggested that it would be best to think in terms of a continuum of historical durability of our empirical and theoretical findings. Along such continuum, phenomena highly susceptible to historical influence represent one extreme (mostly the field of psychology, and some natural sciences), while more stable natural processes are found at the other end and are mostly studied by chemistry and physics. For similar reasons, which I will examine shortly, it might be more difficult than we thought to achieve a general theory of ecology that cuts across population, community and ecosystem levels. And as far as evolutionary theory is concerned, there has been no conceptual unification of its various subdisciplines following the synthesis of the 1940s, and despite recent calls for such unification to which I participated (Schlichting and Pigliucci 1998), evo-eco biology currently looks more like a badly tossed salad than a melting pot of conceptual understanding.

So, what's the problem?

If any of the above should be of concern to evolutionary biologists and ecologists, we need ask what the roots of the problem might possibly be. As we shall see, the answers identified by psychologists resonate with some of the well-known problems that plague evo-eco as well. **Box 1.1** lists a subset of the difficulties highlighted by Meehl in psychology that are relevant to ecologists and evolutionary biologists.

Box 1.1. A list of problems and conceptual issues common to psychological and evo-eco research, from a subset in Meehl (1978):

- Difficulties of slicing up the phenotype into meaningful intervals identified by causally relevant attributes.
- Difficulty in achieving an adequate classification and sampling of environments and situations.
- Difficulty in choosing appropriate scales of measurement and transformations.
- Individual differences (e.g., genotypic variation, reaction norms).
- Polygenic heredity of traits of interest.
- Divergent causality (non-linearity), where differences in the exact character of the initial conditions are amplified over the long run.
- Properties and relations (i.e., contingency) that make the study of living organisms rather more similar to such disciplines as history archeology, and geology.
- Unknown critical events; sometimes these are observable events that were not actually observed, such as a demographic crash.
- "Nuisance" variables: a non-negligible class of variables that are not random but systematic, exert a sizable influence, and are themselves also sizably influenced by other variables. Biologists refer to these as higher-order interactions.
- Feedback loops and autocatalytic processes, the complexities of which are refractory to quantitative decomposition.
- Sheer number of variables.
- Limited correspondence between the results of lab and field experiments.
- Inconsistency of results across labs.
- Importance of exceptions and outliers.

Some of the same points have been discussed by Cronbach (1975), who emphasized that the presence of higher-order interactions poses strict limitations to our understanding of complex phenomena. The usual rebuttal offered by classically quantitative-oriented biologists (and psychologists) is that interactions are not that important because they explain a small portion of the total phenomenological variance. This is misleading for a variety of reasons. First, as pointed out by Lewontin (1974) long ago the general linear model used for most statistical analyses is biased in favor of main effects and against interactions. Second, implicit in the rebuttal is the fallacy that small

variance implies lesser "importance." Yet, evolutionary biologists in particular should know that small effects can be of overwhelming importance, as in the case of small but sustained selection pressures, or of mutations of small effect that provide the long-term fuel for evolutionary change.

A particularly simple but striking example will make the point (Cronbach 1975). After the National Institutes of Health refurbished a lab for the study of how animals metabolize drugs, mice that used to sleep 35 minutes after injection of hexobarbital woke up after only 16 minutes. After a careful analysis, investigators found out that the red-cedar bedding of the new cages made the difference, stepping up the activity of several enzymes that metabolize hexobarbital. The solution of the mystery required an explicit search for apparently insignificant interactions, one that was successful only because of the confined setting and the very limited number of potential factors. As Cronbach put it, "once we attend to interactions, we enter a hall of mirrors that extends to infinity." Any ecologist or evolutionary biologist who has ever worked with complex systems will feel déjà vu all over again.

Cronbach also realized that to get at the bottom of higher level interactions would require sample sizes that are simply logistically impossible, thereby setting an upper limit to what we can do given current approaches. He also highlighted a problem that is very well known to evo-eco researchers, but that he thought should be approached positively as an incentive to further research, as opposed to being minimized or even worse completely ignored. It is hardly the norm that results from field and laboratory research match. Worse yet, results obtained in different laboratories may not correspond either. In recent years, a number of papers in evolutionary biology has confirmed and quantified this phenomenon (Matos et al. 2000; Sgro` and Partridge 2001; Ackermann et al. 2001;

Hoffmann et al. 2001). Should we then just throw our hands up and run for cover? No, because inconsistencies have causes as well. While some of these causes may not be of theoretical interest (such as experimental error, for example), others must lie in higher order interactions that can – with a subtle investigative work[2] – be identified and dissected. It is not by chance that a recent book devoted to a discussion of the most sophisticated methods in ecological analysis was entitled *The Ecological Detective* (Hilborn and Mangel 1997), and that some papers published in the main evolutionary literature explicitly use the detective metaphor to suggest the author's approach to the problem at hand (Wills 1995).

A further problem that deserves a brief mention is the neglect that a classical quantitative approach to biological research has implied for the exceptions and the "outliers." As it has been occasionally pointed out (Lewontin 1966; Levin 1995), an obsessive focus on means and other general descriptors of population behaviors often leads investigators to ignore exceptional behaviors or phenotypes as "anomalous" and "non-representative." But again, there is a real possibility that at least some of these anomalies actually fuel a significant amount of evolutionary change, or can provide ecological steppingstones that connect diverging populations. By concentrating exclusively on the average, do we risk ignoring crucial individual differences?

A possible solution? Strong inference

A problem similar to the one I am trying to bring into focus here was noted back in 1964 by Platt (1964). He asked himself why certain fields of research were making rapid progress and others were not, even within physics or biology themselves. His

[2] As opposed to brute force as manifested in large experiments.

suggestion was that differential training and historical-cultural effects have led practitioners of some disciplines away from the application of what he called "strong inference." This can be described by the following sequence:

1. Devise (several, not just two) alternative hypotheses.
2. Devise a crucial experiment(s), with alternative possible outcomes, each of which will, as nearly as possible, exclude one or more of the hypotheses.
3. Carry out experiment(s) so as to get as clean results as possible.
4. Recycle the procedure, make sub-hypotheses or sequential hypotheses to refine the remaining possibilities.

Platt described Francis Crick's lab at the time, in which a blackboard was always covered with "logical trees" of alternative hypotheses and students were excited at the prospect of discussing which experiments might cut some of the branches off the tree. This is of course also Karl Popper's (1968) idea that the best way to proceed in science is by eliminating as many wrong answers as possible. Science is a procedure based on falsification more than on confirmation.

If this sounds familiar, it should, since strong inference is an elaboration on the basic proposal for a method of scientific investigation laid out by Francis Bacon (1620) in his *New Organon*. More recently, Kitcher (1995) has also emphasized the central role of eliminative induction. This addresses the so-called problem of under-determination of theories by data (the fact that the same data can be explained by more than one theory: Okasha 2000; Shipley 2000). As Platt puts it: "When a group of hypotheses is at odds with some observational or experimental report, it is crucial to explore the explanatory

losses that would occur if some members of the inconsistent set were eliminated. Making up the deficiency may involve new contradictions or new losses, generating a treelike structure of possible adjustments to the corpus of beliefs." Furthermore, as Platt cunningly observed so many decades ago, getting into the habit of considering more than two possibilities (your favorite hypothesis and the "null" hypothesis, see Chapter 5) minimizes the always-present danger of falling so much in love with your theory that you become blinded to the actual verdict of the evidence (see also: Chamberlain 1897; Monod 1971; Hilborn and Mangel 1997). **Box 1.2** summarizes Platt's complaints about the sorry state of some scientific disciplines, which again I think closely resemble a rather common sentiment among ecologists and evolutionary biologists.

Platt's strong inference, however, is no panacea for the problems of evo-eco. While biologists can certainly make use of many of Platt's (and Bacon's) suggestions, it is not by chance that Platt's example of a successful application of the method was Francis Crick's lab: molecular biology is the least historical of biological disciplines, though it surely doesn't lack historical effects, e.g., the "freezing" of an arbitrary genetic code early on during the evolution of life on the planet. It therefore most closely resembles physics and chemistry, and is the one field within biology in which logical trees of sharply

Box 1.2. Platt's (1964) complaints about some bad practice in the natural sciences:

- There are two kinds of biologists: those who are looking to see if there is one thing that can be understood, and those who keep saying it is very complicated and nothing can be understood.
- Scientists become method- rather than problem-oriented. Stop doing experiments for a while and think.
- "How many of us write down our alternatives and crucial experiments every day, focusing on the exclusion of a hypothesis?"
- Small studies add another brick to the temple of science. But most such bricks just lie around the brickyard. They become substitutes for thinking, "a sad waste of intelligence in a research laboratory."
- "We substitute correlations for causal studies, and physical equations for organic reasoning. Measurements and equations are supposed to sharpen thinking, but, in my observation, they more often tend to make the thinking non causal and fuzzy. They tend to become the object of scientific manipulation instead of auxiliary tests of crucial inferences."
- THE crucial question to ask after a seminar: "Sir, what experiment(s) could disprove your hypothesis?" Or: "Sir, what hypothesis does your experiment disprove?"

differentiated hypotheses represent a productive approach to inquiry. For psychology as well as evolutionary biology and ecology, things are a bit more complex, and we shall see that a more nuanced approach mirroring the philosophy embedded in Bayesian analyses (without necessarily endorsing the specifics of that statistical approach) is likely to bring us further.

Hard vs. soft science: what's the difference?

Despite the fact that things would surely improve if we were to stick to the principle of strong inference, the situation is just not that simple. There are other fundamental differences between hard and soft sciences, which we should be cognizant of in order to avoid channeling our efforts into frustrating dead ends. In fact, within biology, the modern distinction between genetics/molecular biology on one hand (hard sciences in a sense close to that of chemistry and physics), and evolutionary biology and ecology on the other was evident from the beginning. Genetics was born when Mendel switched from natural history observations to rigorous quantification and controlled experiments employing the strong inference approach. But the milestone in evolutionary biology was Darwin's *Origin of Species*, which includes little or nothing in the way of calculations and refers to few experiments. Instead, the latter is a masterpiece of detective work, putting the pieces of a complicated puzzle together one by one.

One of the major differences between the two kinds of endeavor is that in evo-eco sciences (as Cronbach 1975 already noticed for psychology) the context in which research is done changes often, unlike in physics. An atom is an essentially a-historical object, so that it doesn't matter when and where you split it, you will get the same results.

Biological organisms (and higher-level groupings such as populations and communities), on the other hand, are inherently and inextricably the outcome of many historical events (evolution, community assembly). This does not mean that biological research is hopeless, but it does imply that generalizations in evo-eco have short half-lives, so to speak. To be sure, things aren't quite that easy in physics either. While most people refer to the first part of Newton's *Principia* as the quintessential example of how hard sciences proceed, they also tend to neglect the second part, which deals with the complications of real – as opposed to ideal – bodies. Still today, the theory of nonlinear dynamics has shown why the three-body problem (the calculation of the exact positions of three celestial bodies orbiting around each other) is insoluble: the system is highly sensitive to initial conditions, so that even if it is entirely deterministic, a precise solution valid for even the near future is not obtainable. And we are talking about a very simple situation with no historicity to deal with. Imagine how difficult it is to achieve a realistic mathematical treatment of objects as complex as populations or communities!

A more general point can be made concerning the complexity of evo-eco research that involves an understanding of basic philosophy of science. As we have seen, Popper (1968) suggested that the distinction between science and pseudoscience (the so-called "demarcation problem") can be drawn on the basis of the criterion of falsifiability. When a theory makes a risky (i.e., nontrivial) prediction, if it does not square with the empirical data the theory has been falsified. If a hypothesis is unfalsifiable (i.e., there is no way in principle to disprove it), then it is not scientific (which doesn't mean it's not true, of course).

Lakatos (1977) understood that this sort of falsificationism is too simplistic: often we don't reject a previously supported hypothesis outright simply because the empirical

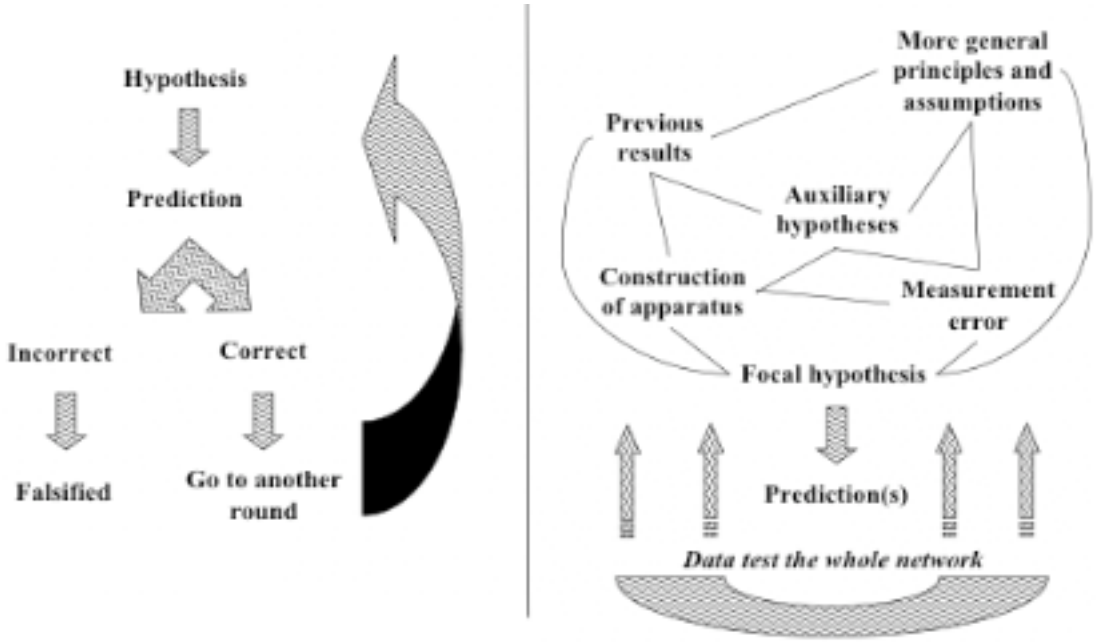

Figure 1.1. Two models of hypothesis testing: on the left, Popper's "naïve" falsificationism (on which the standard statistical approach is based), where a hypothesis makes falsifiable predictions which are directly tested by data; on the right, a complex web of knowledge in which assumptions, hypotheses, and reliability of measurements are simultaneously tested during an ongoing process of research.

results of one experiment don't fit it. Instead, we look first for possible experimental errors or systematic biases, and then we broaden our search to question the assumptions on which the hypothesis itself rested. So, empirical results cannot falsify a hypothesis, only the ensemble of hypothesis-assumptions-methods (the latter two components usually termed "auxiliaries"), which makes a strict application of the strong inference approach rather problematic (**Figure 1.1**). The difficulty is that in psychology (and evo-eco) the range of research circumstances in which auxiliary hypotheses are knotty is greater than in the exact sciences or in some but not all of the biological sciences (such as molecular biology). It is therefore hard to fulfill the Popperian requirement of stating beforehand what counts as a strong falsifier. This implies that some of evo-eco hypotheses may come close to pseudoscience, the just-so stories of some (but not all)

sociobiologists and evolutionary psychologists come to mind as good examples (Kaplan 2002).

There are however some myths concerning the difference between hard and soft sciences which need to be exploded because they too stand in the way of our understanding of science as a process. One of these myths is that physics, for example, is a paragon of consistency in its results, while soft sciences are so frustrating because things tend to be different between different experiments or situations. Meehl (1978) has tested this prediction by conducting an extensive meta-analysis of results from several sub-fields of physics and psychology. It turns out that there is no large difference between the consistency of results from the social and physical sciences. The idea that results in physics are strikingly consistent and in psychology strikingly inconsistent is simply not supported by the empirical evidence. Therefore, the results of social science research are reasonably empirically (though not necessarily conceptually) cumulative when compared with the results from physics. Might the same be true for evo-eco research? What Meehl also found was that the main difference between physical and social sciences' results lies in the fact that the first are much more *accurate* (where accuracy is the ratio between the measured value and its standard error). However, this measure might mean very little given some astounding data reported in his paper. Meehl shows a figure with two curves relating temperature to thermal conductivity of gadolinium. The accuracy of the first curve was stated as within 1% and that of the second one as 0.5%. Yet, the two curves differed from each other by up to 500%! Are our beloved standard errors really measures of how far our estimates are from the true mean of the population, or are we fooling ourselves in the false security of “hard” numbers?

The method of multiple hypotheses and Bayesian inference

An obvious alternative to the classical way of thinking about how science makes progress is provided by the so-called Bayesian framework, instantiated by a suite of specific statistical techniques that are becoming increasingly popular in evo-eco research (Sinsheimer et al. 1996; Hilborn and Mangel 1997; Rudge 1998; Shoemaker et al 1999; Huelsenbeck et al 2000; Wade 2000). I cannot provide here a review of this field, its underpinnings and its usefulness in biological research (but see for example: Howson and Urbach 1991; Jeffreys and Berger 1992; Malakoff 1999). However, what is of interest to our discussion is Bayesianism as a *philosophical* model for doing science, rather than the actual nitty-gritty (and quite complex) details of how to implement it in real research applications.

The fundamental theorem of Bayesian statistics (which was *the* way to do statistics until Fisher took over in the early 20th century) is:

$$P(H_i \mid D) = \frac{P(D \mid H_i) \cdot P(H_i)}{P(D \mid H_i) \cdot P(H_i) + \ldots + P(D \mid H_n) \cdot P(H_n)}$$

Where we are considering a series of hypotheses $i \ldots n$, D stands for the observed data at a given moment during the research project, H for a certain hypothesis, P for a probability, and | is the conditional probability operator. What the equation says is that the probability of a certain hypothesis given the available data (the so-called "posterior" probability) is the ratio of two quantities: the product of the probability of the data given the hypothesis in question (the so-called likelihood of the data), multiplied by the a priori

probability of that hypothesis (the so-called "prior") at the numerator, and the same quantity summed over all considered hypotheses at the denominator (sometimes referred to as the total probability of the data).

It should be obvious why a Bayesian approach directly addresses many of the problems we have discussed so far. For starter, it embodies the idea of multiple competing hypotheses in the definition of the denominator of Bayes' rule: there are no privileged ("null") hypotheses, but only a fair competition among stated alternatives. Second, the question is posed in sensible terms from the viewpoint of a scientist (as opposed to that of a statistician): Bayesian analysis asks what is the probability of a certain hypothesis given the data, not the other way around (which is typical of standard approaches to statistics: see Chapter 5). Third, the concept of probability in Bayesian analyses is different from the standard one and more appropriate for scientific research: a probability here is not the frequency with which a certain outcome would occur if we repeated the experiment *n* times, but rather an estimate of the degree of belief (as in likelihood, not blind faith) we are entitled to attach to a given hypothesis because of what we know of the problem (including the newly collected data as the result of our latest experiment). Fourth, Bayesian analysis takes into account what we knew of the problem before starting our experiment (in the form of the priors), which makes sense because we never start a project with our mind set to the state of a *tabula rasa*, and it is a good idea to quantify as much as possible our prior thinking about the problem at hand. Finally, a Bayesian framework makes it very difficult to think in terms of either naïve falsificationism or naïve confirmationism: we cannot get zero (complete falsification) or one (complete confirmation) as posteriors for any given hypothesis. All we can reasonably hope for is to see our belief in some hypotheses (ideally just one) go

significantly up after an experiment and our belief in as many other hypotheses as possible go down accordingly. Indeed, this change in likelihood can be measured independently for any given hypothesis, which gives the researcher an estimate of how informative the experiment was in respect to each hypothesis (the so-called "Bayes factor": Kass and Raftvery 1995).

While Bayesianism is itself no panacea for either statistical inference or scientific methodology (for example, it is tricky to set priors reasonably, and a long discussion has been going on about the difference between "objective" and "subjective" priors), the reason it strikes me as the best model on the market (itself a Bayesian statement, incidentally) is that it resonates very well with the actual practice of science as I have experienced it in decades of academic research. We think like Bayesians, even though we unwittingly constrain our papers within a Fisherian (classical) straight jacket.

The compromise: evolutionary biology and ecology as nomothetic *and* idiographic sciences

As Francis Bacon noted (1620), it does little good to engage only in the *pars destruens* (i.e., the negative criticism) of an argument, unless one has a *pars construens* (positive suggestions) to add to it. Despite the criticisms and warnings that I discussed above, I do of course think that evolutionary biology and ecology are (mostly) real sciences and that progress has indeed occurred and still does. The real questions are: where does evo-eco lie between the two extremes of nomothetic and idiographic sciences, and what does this mean for the everyday practice of researchers in these fields?

To begin with, it is undeniable that ecology and evolutionary biology are partly

historical and partly a-historical sciences. While this is a truism, the force of its impact can best be understood when considering particular examples. One among many is offered by a recent study of changes in variance-covariance (**G**) matrices in *Drosophila*. **G** is a crucial parameter entering evolutionary genetics equations to account for and predict the evolution of natural populations, and summarizes the relationships among many traits in terms of the amount of genetic variation available for their evolution and the amount of genetic co-dependence on each other that may limit response to selection. Phillips and collaborators (2001) have investigated the effect of genetic drift on the similarity of **G** matrices in 52 independently derived inbred lines of fruit flies when compared to outbred controls. They found that the *average* results were in perfect agreement with the theoretical prediction that drift (as an example non-discriminative process: Millstein 2002) alters only the size but not the shape of **G** matrices (because it doesn't act differently on different characters, the way selection, say, would). However, these authors also found that it is impossible to predict the actual shape of any given **G** due to a large amount of individual variation among the inbred lines. Since this variation would translate in significantly different evolutionary trajectories, we are in a paradigmatic example of success of average predictions (which tend to be a-historical) and abysmal failure of specific predictions (which are markedly influenced by history). This is exactly what John Dupré (1993) predicted based on his theory of non-reductionism in the sciences: in the case of complex systems, theoretical reduction can explain the boundaries of variation of the observed phenomena, but will fail to tell us what exactly is going to happen in any specific case (on the general problem of the basis of biological explanations see Rosenberg 2001). This may be a general property of the world (Kauffman 1993) with which we simply have to be able to live and even take advantage of.

As a consequence of the partial historicity of evo-eco research, these sciences are really somewhere between the nomothetic extreme of fundamental physics and the idiographic end of, say, paleontology, anthropology, and some of the social sciences. This, of course, does not mean that historical disciplines cannot make testable predictions, only that it is more difficult to do so the more history is a factor. This does mean, however, that we should finally drop our physics envy and work more like puzzle-solvers, adopting a healthy balance of observation, experimentation, and mathematical theorizing. As of now, the balance seems to be much too much at the expense of observation, branded as a somewhat second-grade activity reserved for natural historians (a term that has itself unofficially become a slur within certain academic circles). Evo-eco researchers should think of themselves as Sherlock Holmes, not as Isaac Newton. Far from being demeaning, it can be a liberating feeling (for example, see a paper on data analysis as detective work by a statistician of the caliber of Tukey 1969).

Another consequence of a new attitude toward the complexities of evo-eco research is that their *modus operandi* should embrace the concept of consilience. The word "consilience," made familiar to biologists through a book by E.O. Wilson (1998), was introduced by English philosopher William Whewell (1840) to explain the phenomenon that often pieces of evidence from disparate sources "jump together" toward a common explanation, what he termed a consilience of induction (in psychological terms this is a Gestaltian experience, and it may be profoundly linked to the way our brain works and interprets reality: Gazzaniga 2000). While Whewell's work on induction was overshadowed by the later contribution of John Stuart Mill, he was the first philosopher to resurrect the importance of induction in science. In his seminal paper, Whewell stated that: "accordingly the cases in which inductions from classes of facts altogether different

have thus jumped together, belong only to the best established theories which the history of science contains. And, as I shall have occasion to refer to this particular feature in their evidence, I will take the liberty of describing it by a particular phrase; and will term it the Consilience of Inductions." In other words, we can reach (provisional) conclusions in complex matters by a sort of triangulation in logical space, when different types of evidence point toward the same answer. And the more stringent the triangulation, the more likely (but never certainly) we can pinpoint the "culprit" (notice the underlying "Bayesianism" of this approach).

Perhaps the most delicate consequence of realizing what kind of science evo-eco really is concerns the impact that such realization should have on funding and publishing priorities. During recent times we have witnessed a marked movement toward more "hard" disciplines and topics of scientific research, such as genomics, as well as an equally clear preference for mathematical and statistical approaches to evo-eco questions. While this is all certainly very valuable and should be continued, it is time to pause and realize the limitations intrinsic in these choices. Editors, reviewers and officers of societies and funding agencies should ask themselves what kinds of questions are best pursued and how, in order to make choices that are influenced by a solid philosophy of science rather than by fashion or novelty.

Warnings from a variety of authors concerning over-reliance on statistical testing and simplistic formulation of hypotheses (Chapter 5), cavalier interpretations of results, and a tendency to substitute technique for thinking are all topics that are very familiar to philosophers of science. Alas, most practicing scientists just plug ahead with their research, confining the tough questions to evening discussions over a beer with their graduate students (several colleagues have actually put it to me in exactly these terms,

usually while providing their most disarming smile). And yet, it does pay to occasionally look at the forest instead of individual trees, with the intent of searching for the best road through the vegetation by adopting a bird's eye view. Contrary to what is maintained by some "hard" scientists (Weinberg 1992), philosophers of science could be extremely useful to the practical scientist, if only the latter would stop a moment to listen to what they are saying (Wilkins 2001).

Chapter 2: Genes 'for' phenotypes: a Modern History view[3]

"There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable. There is another theory which states that this has already happened." -Douglas Adams (1952-2002), *The Hitchhiker's Guide to the Galaxy.*

Perhaps one of the more conspicuous examples of rarely examined baggage (in the biological literature) that significantly affects how biologists think is the idea that genes are 'for' something. But what does it mean to say that a gene is 'for' a phenotypic trait? How one interprets such statements can have a profound practical influence on what research avenues will seem promising or necessary, since ways of talking obviously shape one's thoughts, and can push researchers in particular directions rather than others. It is for this reason that I argue that some of the more metaphorical uses of 'gene for' talk are dangerous and potentially misleading habits.

In this chapter, I wish to explore possible answers to the following questions: What does it mean when we say that a gene is 'for' some trait, and, what do we have to know about the relationship between an organism's genotype, its phenotype, its environment, and its developmental and evolutionary history, if we are to speak properly about some of its genes being 'for' certain traits? The distinction I wish to stress is between genes being 'for' particular traits and genes merely being 'associated with'

[3] A shorter version of this chapter appeared as: Kaplan, J., and M. Pigliucci (2001), "Genes 'for' Phenotypes: A Modern History View", Biology and Philosophy 16: 189-213.

particular traits. It seems obvious that we can know how to find genes associated with various phenotypic traits, and that we in fact do know of a lot of genes so associated. But what do we have to know to make the *further* claim that the genes are *for* such a trait rather than merely being *associated* with it? I will first attempt to formulate an answer to these questions for the general case, and then bring out some of the implications of the answer I provide for the human case.

I argue here that the best sense that can be made of the 'gene for *x*' locution is to treat it as a kind of functional talk, specifically the 'Modern History' version of functional talk that Godfrey-Smith developed for thinking about biological functions (Griffiths 1993; Godfrey-Smith 1994). Within this framework, we can properly speak of a gene being 'for' some trait only when the gene was maintained by natural selection in the recent evolutionary history of the organism by its causal association with the trait in question (a more precise definition will be developed below). Based on this interpretation, it will turn out that to say that a gene is *for* some trait *x* is to say something about not merely what the gene happens to do in this particular case, but also something about what biological meaning that association has.

On this 'Modern History' view of 'gene for' talk, it will turn out that we know of rather fewer genes that are *for* things (as opposed to merely being associated with them) than some might have thought or implied. Further, in many cases genes will turn out to be 'for' less exciting things than we might have wished (contra Dawkins 1982, I will argue that there are no genes 'for' such things as reading-ability in humans). The advantages to this view are many. Unlike views which conflate genes being for traits with genes being merely associated with phenotypic traits, and unlike views which reject the legitimacy of any talk of genes being for phenotypic traits, the sort of work encouraged by the Modern

History approach focuses attention onto biologically important questions. Namely, on how evolution changes the associations between genes and phenotypes, i.e., on how genes can evolve new functions (Ganfornina and Sanchez 1999). This way of thinking about 'genes for' is not merely cautionary in the way that a refusal to countenance any usage of 'gene for' talk would be, but rather would work to further the exploration of important avenues of research. Indeed, another reason to support the Modern History view is that while as yet there are few genes about which enough is known to justify thinking of them as 'for' particular phenotypic traits, the research encouraged by the Modern History view would be useful for a variety of goals. It would for example aid: (a) discovering more such genes; (b) beginning to get a sense of how many such genes there likely are; and (c) beginning to get a sense of how important such genes have been evolutionarily. It may be that on the Modern History view there will always be relatively few genes that are known to be 'for' particular traits; however, if these have been of great evolutionary significance, a language that encourages a program of looking for them is still to be recommended.

Unlike more permissive usage, this way of limiting 'genes for' talk would discourage the sorts of misconceptions about the relationship between genes and phenotypes so common in the contemporary culture and the popular media. For example, consider the naive sort of genetic determinism and genetic essentialism encouraged by the bombardment of reports of the discovery of 'the gene for *x*' type. While sharing this feature in common with those approaches that reject all 'gene for' talk, the approach proposed here recognizes that there are situations in which using such language seems natural and appropriate. While on the Modern History view the conditions on properly using the language of a gene being 'for' a phenotypic trait are restrictive, they are not

impossible to meet. When we do meet these conditions, and hence can reasonably conclude that a given gene is *for* some phenotypic trait, we end up knowing something quite a bit more important and interesting than when we merely know of associations. We know something of real biological import.

'Genes for' or Not 'Genes for'?

Historically, arguments about whether it makes sense to claim that some gene is 'for' some trait have centered on two extreme positions. Richard Dawkins is usually taken (fairly or unfairly) as the representative of the view that there are *lots* of 'genes for' phenotypic traits, and that theorizing about the adaptive significance of traits is at least a good first step towards discovering that there are genes for it (Dawkins 1976, 1982, especially Chapter 2). On the other side, it is sometimes suggested that the attacks on 'adaptationism' and genetic determinism most closely associated with Gould and Lewontin (1979; Pigliucci and Kaplan 2000) imply that the language of 'genes for' traits (at least if used at anything above the most straightforward biochemical level) is entirely misguided (see also Levins and Lewontin 1985; Lewontin 1992). Further, the attacks of Oyama and others on the supposed centrality of the gene, and their proposed alternative ("Developmental Systems Theory," henceforth DST), are sometimes thought to render the 'gene for' language somewhat obsolete (Griffiths and Gray 1997; Oyama et al. 2001). However, as stated, positions are little more than caricatures; before trying to defend a version of 'genes for' talk that doesn't succumb to the difficulties that DST proponents have found with permissive versions of 'gene for' talk, it is worth attempting to flesh out these positions and see just what they imply.

Dawkins, in *The Extended Phenotype* (1982), sometimes defends speaking of 'genes for' a trait in a way that makes it roughly equivalent to the gene in question being statistically associated with the trait under study. He writes for example that the proper translation of a geneticist who "speaks of a gene 'for' red eyes in *Drosophila*" is something like: "there is variation in eye color in the population; other things being equal, a fly with this gene is more likely to have red eyes than a fly without the gene" (Dawkins 1982, p. 21). Finding a statistical correlation of this sort, or taking a more sophisticated quantitative trait loci (QTL) approach (e.g., Phillips 1999; Asins 2002), does undeniably find genes *associated* with the traits in question, at least within the given environments (both genetic and external). But ought these be described as genes *for* the trait?

Sober (1993, pp. 186-187) attacks this use of the language of 'genes for' traits by noting that "a gene *for* phenotype *x* presumably is a gene that *causes* phenotype *x*," and that it is obvious that simple correlations between genes and traits do not imply causation at all. This problem is especially acute in human genetic research, where experimental manipulations cannot be performed with controls; however, it should serve as a reminder that such experimental manipulations *are* generally a necessary step in moving from associations to claims about causality. Sober's example involves finding genetic correlations to languages spoken, not at all a far-fetched scenario (Sober 1993, pp. 187-188). If one looks for associations between particular genes and language spoken between, say, native-born Australians and native-born Sub-Saharan Africans, one would no doubt find many strong correlations. For example, the absence of the FY-0 gene would be highly correlated with speaking English – the gene is all-but absent in Australia, but is present in some 87% of the population in Sub-Saharan Africa. English is of course spoken by more or less all native-born Australians, and by a relatively small percentage of the native-born

population in Sub-Saharan Africa (Cavalli-Sforza and Cavalli-Sforza 1995, p. 125). But, obviously, the gene is not *for* not-speaking English! In this case, we have a good idea of what the gene actually does; namely, it affords resistance to a particular strain of malaria (Cavalli-Sforza and Cavalli-Sforza 1995, p. 125).

But even Dawkins isn't oblivious to this problem; he seems, in fact, ambivalent about what kinds of correlations matter. He argues that, if we found genes associated, for example, with severe dyslexia, we would have found not only a gene for dyslexia, but could also conclude that the "wild-type gene at the same locus ... would properly be called a gene 'for reading'" (Dawkins 1982, p.23). He is less comfortable with other possible associations. He proposes, for example, that a gene associated with blindness would not be properly called a gene for not-reading, despite the fact that it *would* be statistically associated with not-reading. Consequently, we would be forced to conclude that the 'wild-type' gene at the same locus would not be a 'gene for reading', either. Dawkins's defense of the difference between these two examples is that while the hypothetical dyslexia gene's "most obvious or debilitating phenotypic effect" is not-reading, this is not the case with the hypothetical blindness gene (Dawkins 1982, p. 23). Needless to say, the "most obvious" phenotypic effect of the difference at the FY locus between Australians and Sub-Saharan Africans could not be 'language spoken' because, following Sober (1993), that is not an *effect* of the gene at all. The 'most obvious' effect of the FY-0 gene is "the absence of the FY substance," and this absence "grants ... a certain amount of protection against [*Plasmodium vivax*, a particular malaria parasite]" by making "it hard for the parasite to multiply" (Cavalli-Sforza and Cavalli-Sforza 1995, p. 125). What Dawkins fails to give is any way of making precise the notion of a "most obvious" phenotypic effect, nor indeed any indication of why 'obviousness' should even

be a relevant criterion to figuring out what a particular gene is 'for'.

Another difficulty with this approach is hinted at by Dawkins himself. If there is no genetic variation for a trait, then, on the basis of the statistical approach to find 'genes for', one would be forced to conclude that there isn't any gene for that particular trait. But that conclusion would often be patently false: surely there are genes "for" the number of eyes in humans, and yet there is little if any genetic variation for that trait. Nevertheless, Dawkins does seem to suggest that there could have been 'genes for' a trait at some time in the historical past, when there was variation, but no longer (Dawkins 1982, p. 24). On the other hand, Dawkins seems firmly convinced that many traits are adaptations, and that these adaptations have a genetic basis, and often a complex one (Dawkins 1982, pp. 24-25). Oddly enough, however, his contention that all geneticists are ever interested in are *differences* seems to force him to argue that the hypothetical complex genetic basis does not represent the 'genes for' the traits in question at all. Indeed, Dawkins makes it clear that genes associated with lacking a particular adaptive trait might well be such in a quite "boringly destructive" way (cutting "a vital link in the neural machinery," in Dawkins's example). In this case, the 'wild-type' gene would have to be called a "gene for" that trait, despite perhaps not being a locus "upon which natural selection worked during the evolution of the adaptation" (Dawkins 1982, pp. 24-25). Dawkins thinks that this "consideration seems ... to be a reason for caution" in thinking about the relationship between 'genes for' and 'causal connections' (Dawkins 1982, p. 25).

But surely that is much too weak. If what we want to know is something about the relationship between genotypes and phenotypes, then the kind of 'genes for' that Dawkins's approach identifies are not of the right sort at all. Dawkins admits that many

of the genes identified as 'genes for' using his approach are of little or no evolutionary significance to the trait in question. Further, the type of genes that Dawkins misses are *exactly* those we would like to know about, if our goal is to untangle the relationship between developmental and evolutionary biology. While even Dawkins admits that correlations are not the last word, his definition is not helped by the addition of "most obvious" phenotypic effects, even if it could be made precise.

Sterelny and Kitcher (1988, pp. 348-352), in defending the genetic approach to selection, attempt a reconstruction of Dawkins's notion of a gene being for a trait that they hope will be more precise and properly account for the complex interactions genes can have with their environments. They state that the "intuitive idea" behind their proposal is that "we can speak of genes for *x* if substitutions on a chromosome would lead, in the relevant environments, to a difference in the *x*-ness of the phenotype" (Sterelny and Kitcher 1988, p. 348). They go on to claim that a particular locus, L, affects a particular trait, P, if there are allelic substitutions at L that will result in differences in P in "standard environments" (Sterelny and Kitcher 1988, p. 349). This is supposed to add a causal element to 'gene for' talk, in that mere associations without any causal import will be excluded by the substitution demands. Much work is done fleshing out the idea behind "standard environments." Nonstandard genetic environments involve "rare" alleles at another locus or "rare combinations" of alleles at other loci (Sterelny and Kitcher 1988, p. 349). Standard external environments are either those that are encountered relatively frequently or, alternatively, those that "do not produce some gross interference with the organism's development" which, they claim, can be rendered more precise by thinking of standard environments as those which do not substantially reduce population mean fitness (Sterelny and Kitcher 1988, p. 350). Kitcher, responding to a challenge presented

by Griffiths and Gray in their articulation of the Developmental Systems Theory (DST) approach, developed another set of definitions of a "standard" environment (Kitcher 2001, pp. 404-407). One possibility is that when we think about 'standard environments' to make sense of the 'gene-for' locution we take a statistical approach and carefully partition our available environments (Kitcher 2001, pp. 405-406). The other possibility is that we think only of environments in which the organisms are *capable* of displaying the general property in question . In this way, Kitcher concludes, "talk of genes 'for' traits can be coherently reconstructed" (Kitcher 2001, p. 406).

While neither of these suggestions is without difficulties, the main problem with this version of talk of 'genes for' is that it suffers from the same basic defect as Dawkins's articulation. Finding a 'gene for' in Sterelny and Kitcher's sense gives us no guarantee that we have found a gene associated with the trait in any way that is meaningful to understanding developmental or evolutionary biology. Kitcher at times seems to suggest that this criticism is not far from the mark. Towards the end of "Battling the Undead" he notes that it may be time for biology to move beyond the view he has been defending, not because it is "false or incoherent" but rather "too primitive for the important tasks of fathoming" the ontogeny of complex organisms and the relationship between developmental and evolutionary biology (Kitcher 2001, pp. 407-408).

In the end, the difficulty with the views of Sterelny and Kitcher is that they fail to do justice to the notion of something being *for* something else. It seems clear that 'being-for-ness' is not just a matter of association or causal mechanisms. This is a more general difficulty with accounts of biological functions that fail to take the selective significance of the functional ascription into account; indeed, Griffiths cites this problem explicitly as the motivation for developing direct proper functions (and rejecting Cummins's

functions) in biology. He notes that as far back as 1963, Lorenz has written that "unless selection is at work, the question 'What for?' cannot receive an answer with any real [biological] meaning" (Lorenz 1963, p. 14). If we want to use the phrase 'a gene for *x*,' then we have to take seriously the implications of the language involved in that phrase. "What's it for?" is *not* a question we can answer by pointing to a statistical association, or even by making use of causal connections that are not necessarily biologically significant. And, as Lorenz reminds us, the only biologically significant functions, the only answers to the question 'what's it for' that are biologically meaningful, are those that involve natural selection, the "designer" of organismal forms and functions (Pigliucci 2001). There is nothing wrong with the suggestions made by Dawkins, Sterelny and Kitcher, or Kitcher, if what they want to discuss are *associations*, but they are not justified in talking about finding genes *for* traits based on those suggestions alone. To do this is to risk mistaking an analysis of variance for a causal analysis, which is precisely the mistake that Lewontin, in one of his seminal paper, warned us against (Lewontin 1974b).

It has sometimes been proposed that *no* good sense can be made of the locution 'a gene for *x*' where *x* is a reasonably high level phenotypic trait. Lewontin and Gould (1979; Pigliucci and Kaplan 2000) have been interpreted in this way on occasion, as have proponents of the Developmental Systems approach to understanding ontogeny, evolution, and the relationship between them (Oyama et al. 2001). These interpretations are not wholly unfair. Lewontin, for example, often does write things like "it is a fundamental principle of developmental genetics that every organism is the outcome of a unique interaction between genes and environmental sequences modulated by the random chances of cell growth and division" (Lewontin 1992, p. 27), which, while a truism, does

seem intended in part to cast cold water on the idea that a gene could be 'for' a trait. Oyama implicitly attacks the very legitimacy of the "what for?" question when applied to genes, first by arguing that it is an attempt at an "ultimate" causal explanation, and then arguing that "ultimate" causal explanations in this sense are not "best understood ... through the genes" (Oyama 1985, pp. 53, 126). However, most of the effort in DST is not directed towards arguing that it is *impossible* to properly speak of a gene being for some trait. The point is merely that this way of talking has, as a matter of fact, been very problematic, because it has been interpreted to make genes out to have a special place in the developmental process, which is exactly what DST wants to deny. Gray (1992, p. 199) for example writes, "I wish to dislodge the gene from the privileged site it has occupied in our accounts of development and evolution." This dislodging task can be done, and done well, without having to argue that there is no possibility at all of finding genes that actually are *for* traits, a position I think is untenable.

The position that it is impossible to find genes for traits is untenable *not* because the gene needs to hold a central, special place in biologists' thinking, and not because I wish to encourage any old-fashioned adaptationism which would want to make every trait out to have been the product of natural selection acting on genes. Rather, the problem is that there *are* cases where the locution seems to me entirely natural. While I want to reject *both* parts of Dawkins's hypothetical gene for reading, the FY-0 gene seems to be naturally described as a gene *for* resistance to a certain sort of malaria. If we think of the question "what's it for?" in the natural way, as a request for an explanation in terms of the *function* of the thing in question, then to say that a gene is *for* some phenotypic trait is to say something that can *only* be analyzed in functional terms. Specifically, such an analysis can be carried out in a biologically meaningful functional way of the sort Lorenz

suggests. It is for this reason that Dawkins's hypothetical gene *for* dyslexia could be no such thing. Whatever functions a gene might have acquired through selective regimes, causing dyslexia (or even the absence thereof) is not among them. But what is it to say that a gene is *for* some trait, then?

I believe that a good starting point for thinking about 'genes for' in a functional way is the "Modern History" approach to biological functions developed by Godfrey-Smith (Griffiths 1993; Godfrey-Smith 1994). Godfrey-Smith was attempting a definition of biological functions in general, and concluded that they are "disposition or effects a trait has which explain the recent maintenance of the trait under natural selection" (Godfrey-Smith 1994, p. 344). Griffith's use of the notion of a "proximal selective explanation" serves a similar purpose to that of explanation in terms of 'recent maintenance' in Godfrey-Smith's position (Griffiths 1993, p. 421). This leads to the conclusion that we can properly speak of a gene being *for* some phenotypic trait within a population only when its association with that trait explains the recent maintenance of the gene in that population.

Kitcher (2001, pp. 405-406) develops an interesting example in his "Battling the Undead: How (And How Not) to Resist Genetic Determinism," in which he argues that a gene is 'for' (in his sense) root proliferation if, in every standard environment, a tree with this gene will have more roots than a tree without it would in the exact same environment. In this case, it is interesting to note that given Kitcher's description, we would have no idea on the Modern History view whether it was really a gene 'for' root proliferation or not. Even if the gene fulfills the requirements of a Cummins's-style function (and even this is not guaranteed given Kitcher's description), this is insufficient for biologically meaningful 'for-ness.' The criticism is equally applicable when the term is used with

'genes' as when it is used with more ordinary traits. In order to know whether or not the gene was 'for' root proliferation on the Modern History view, we would have to know more about the recent evolutionary history and selective pressures faced by the trees in question. Furthermore, we would need information about what effects the gene had within those environments of recent evolutionary significance (which may well have been 'nonstandard' on some of the interpretations of 'standard environments' developed by Kitcher, and Kitcher and Sterelny). In other words, to know that the gene was 'for' root proliferation we would need to know that in those key environments in which the gene was actively maintained or spread, it did so because it provided a selective advantage (increased average fitness), and that it provided such an advantage *because* of its causal association with root proliferation.

Contrast Kitcher's hypothetical gene for root proliferation with the FY-0 case again, where we not only know the gene (its location, sequence, etc.) and the phenotypic trait it is statistically associated with in various environments, but we also know what biochemical changes to the surface of the hemoglobin molecule are associated with that particular allele. We further know why those biochemical differences have an important phenotypic outcome, and we know that *that* outcome was the subject of some fairly intense selective pressures in the recent evolutionary history of humans in Sub-Sahara Africa (but *not* those in Australia or in Europe: Cavalli-Sforza and Cavalli-Sforza 1995, pp. 124-125). From the perspective of the Modern History view, these represent something of a *minimum* requirement on claims that a gene is for a trait.

Finding 'Genes For' – The challenges of the Modern History view

Unless we accept the view that the entirety of talk of 'genes for' traits is misguided and should be given up, we need to find some way of distinguishing legitimate uses from uses that force us to take the gene as primary or central, or that force us into thinking in the adaptationist mold. Ideally, such a solution would permit us to speak sensibly of a gene being for a trait only where there was no risk of moving, as Sterelny and Kitcher put it, "from the 'genes for *P*' locution to the claim that selection can fashion *P* independently of other traits of the organism." A slippage that they admit is "perennially tempting" (Sterelny and Kitcher 1988, p. 361). In this section, I will argue that the Modern History interpretation of talk of 'genes for' traits is far more resistant to this slippage than other views that have been proposed. In part, this will turn out to be because the work necessary to claim that there is a 'gene for' some trait in this view forces one to consider more facets of the relationship between the organism's development and its recent evolutionary history than other versions of 'gene for' talk, or indeed even giving up 'gene for' talk altogether, would.

The first thing to notice is that no statistical correlation, no matter how strong, will be enough on the Modern History view to talk of there being a gene for some trait. Indeed, it will turn out that no single *kind* of evidence will ever be enough. Merely knowing something about the evolutionary history, or the population genetics, or the developmental pathways, or any other *single* sort of biological information, will not suffice. Some combinations of these will be required in a consilience of the evidence (Whewell 1840). Just what combination may well depend on the case, however.

So, what do we want to know in order to claim properly that a gene is *for* a trait?

Ideally, we would want to know all of the following:

a) That the gene is associated (not necessarily statistically) with the trait.

b) The most important aspects of the biochemical and developmental pathway from the gene to the trait.

c) That the trait in question really was the likely subject of natural selection in the species' recent evolutionary history.

d) That the maintenance of the gene in the species in its recent evolutionary history can actually be the result of (c).

The first requirement (a), that there actually be an association between the gene and the phenotype, is the reason why it is sometimes suggested that a statistical approach may be a good 'first step' in finding genes 'for' traits. While this is necessary, it should now be clear that it is nothing like sufficient. It is worth noting, however, that there does not have to be a *statistical* association in the sense Dawkins developed, as one could have 'genes for' on this view even where there is no variation in the population for the genes involved or for the trait in question (as long as one is able to show an association between the gene and the phenotype, e.g., by analysis of artificially induced mutants).

Obviously, (b) is difficult from an experimental standpoint. There are not many organisms, nor many traits, for which we know most of the relevant biochemical pathways involved in moving from any given gene to any given trait. Part of the reason for this is that the complex relationship between genes and phenotypes will often not permit us to talk sensibly of genes for traits. For example, for traits in which complex

epistatic effects play a role, the notion of them being 'polygenic' will prove inadequate and it will turn out best *not* to speak of their being specific 'genes for' the trait. Rather, it is best to speak of the trait as an 'emergent' property of the epigenetic process (Schlichting and Pigliucci 1998, p. 337). In these cases, we will not wish to speak of any particular genes 'for' the trait at all; some complexities at the level of ontogeny will be sufficient to reject the use of 'for' talk. Nevertheless, there are some genes and some traits that we have a rather good handle on, and so for some kinds of traits, we can expect at least limited success in meeting the challenge of (b). However, the relationships that one finds out about in (b), even when coupled with (a) are not sufficient, either. The reason can be seen quite clearly – even where we know of a gene associated with a trait, and know the biochemical (developmental) pathways by which it is so associated, we do not yet have good reason to use the functional language implied by 'genes for ' talk, because we still lack information about the role of natural selection in shaping the gene-trait relationship.

The condition listed as (c) provides perhaps one of the most difficult challenges. With (c) we need more than a good just-so story about how the trait might have been useful (Kaplan 2002); we need to actually understand something about how it was the object of natural selection. This has to rely on a phylogenetically informed comparative analysis of the organisms in question; a necessity that originates from the core fact that evolutionary biology is a historical, not just experimental, science.

Finally, by (d) we need to know how natural selection on the trait in question could account for the prevalence of the corresponding gene. It may in fact be impossible to be certain that this requirement has been met (the difficulties are compounded by such phenomena as pleiotropy and allometry), but theoretical biology can be used to

demonstrate at least the plausibility of this condition in each specific case. The recognition of these difficulties is a good reason to remain cautious about 'gene for' talk, even in those cases where the other requirements would seem to have been well met.

Now, it should be obvious that rarely (if ever) will it be possible to fully meet all these criteria. In some cases, though, we may wish to cautiously suggest that a gene may be for some trait where they have not all been adequately met. For example, in the cases of purported genes for phenotypic traits in humans, it would obviously be impossible (ethically and practically) to perform many of the experimental manipulations that are at the heart of figuring out (b) (Kaplan 2000). Gene knockout experiments, gene substitution experiments, attempts to develop reaction norms for the relevant genotypes, environments, and traits in question, laboratory evolution experiments, etc., while part of the geneticist's and ecologist's standard tool kit for most organisms (Burian 1997; Culp 1997; Schlichting and Pigliucci 1998, pp. 16-20), are out of the question in humans. This is why, in part, all claims to have found genes for phenotypic traits in humans should be treated with caution. In other species more easily experimentally manipulated, claiming to have found a gene for some trait without performing such experiments would be bold at best.

There is, of course, nothing in the Modern History view that suggests that we should expect to find genes for those traits that we happen to find of immediate interest – indeed, quite the opposite. Many traits of interest to us (especially in the human case) were probably never under direct selective pressure, and so on the basis of the Modern History account, there could not be genes for them (examples in the human case might include reading ability, advanced mathematical abilities, etc.). Also, in the case of many complex traits, we should not expect to find genes *for* the trait in question at all, but

rather expect that genes *associated* with the trait will prove to be part of suites of genes with complex interactions, such that the trait is best viewed as an emergent property of the epigenetic process. Further, while neither historical nor mechanistic (e.g., developmental) complexities alone will be enough to *rule out* the possibility of genes being 'for' a trait, these complications will tend to make 'genes for' harder to identify, and the tests necessary to gain evidence for the relationship harder to perform adequately.

This may seem like a lot of work, and it clearly is. Recall, though, the *point* of pushing for a Modern History understanding of what it means for genes to be for phenotypes. The work suggested by this method addresses some of the most fundamental questions in evolutionary biology, ecology, and developmental biology. Avoiding this work by refusing to talk about genes being 'for' phenotypes may be just as misguided as avoiding this work by pretending that mere statistical associations or mechanistic causal chains provide enough evidence to declare that a gene is for a phenotype. As yet, little of this work has been done; until more of it is, it will be hard to be sure about just how important those genes that are, in the Modern History sense, for phenotypes actually have been in an evolutionary sense. It is possible, though, that these genes, even if they turn out to be rare, could be some of the most interesting genes in an evolutionary sense.

Complications: pleiotropy, epistasis, plasticity, QTL, and emergent properties

If the idea of a gene being for some phenotypic trait runs into considerable problems when we think of simple, relatively straightforward cases of genotype-phenotype mapping functions, things become much more muddled (and interesting) when

we turn our attention to some of the most complex (and relevant) situations.

Perhaps one of the most universal and clearly understood properties of genes is the fact that they have pleiotropic effects. **Pleiotropy** is the phenomenon by which the effects of a gene can be measured not just on one trait (however we define "trait": Wagner 2001), but on several. For example, the Hb^S mutation of the _-globin gene coding for human hemoglobin causes a distortion of the three-dimensional structure of the molecule itself (the mutation changes a single base pair, from A to T, substituting a valine for glucine in the protein). This diminishes the ability of hemoglobin to exchange oxygen and carbon dioxide with the blood; however, in some instances the same structural defect alters the shape of the red cells, thereby preventing the agent of malaria from reproducing. What are these alleles at the Hb locus 'for'? Strictly speaking, they code for a certain molecule with a given metabolic function, and they have presumably been selected in the past in order to perform this function. The alterations in the shape of the red cells and the consequent resistance to malaria are therefore not what the Hb^S allele was originally 'for'. But – depending on the environmental conditions – the mutation actually does confer a selective advantage in malaria-infested areas. Is it therefore now to be considered a gene 'for' malaria resistance? Or rather for both resistance and gas exchange? Actually, the gas exchange function is impaired to some extent (so much so that homozygosity for the condition is lethal). A side effect of this mutation in heterozygote individuals is a mild form of anemia (due to the fact that only one of the two genes produces a functional allele for gas exchange). Surely, we would not speak of a gene for anemia, however, even though it is being maintained by selection favoring one of its other effects (this might best be thought of as the selection of a gene associated with anemia, see Sober 1993, p. 84 on

'selection for' versus 'selection of'). So, one could think of this as an example of a gene for gas exchange, which later partially evolved into a gene for malaria resistance. In this way, the 'genes for' language being developed here can account for constraints and tradeoffs; a gene can be *for* a particular trait, and associated with another unfavorable trait, where the selective advantage of the trait it is for outweighs the disadvantage of the trait it is associated with.

The situation becomes even more complex when we take into account gene-gene interactions, i.e. **epistasis.** I am not referring here to epistasis in a statistical sense, a property studied at the population level in quantitative genetics. Instead, I am discussing what is sometimes termed "physiological" epistasis, i.e. actual physical interactions among gene products (the two are conceptually distinct, which has brought quite a bit of confusion in evolutionary genetics: Cheverud and Routman 1995). If gene products would only interact in an additive manner, i.e. the metabolic or phenotypic results would be directly proportional to the action of each gene, we would still be able to talk about 'genes for' in some meaningful sense. For example, if ten genes affect flowering time in a plant, but each gene adds a fixed (not necessarily equal) number of days to the final phenotypic outcome, each of those genes would legitimately be considered a gene for delaying flowering. However, modern physiological and molecular biology teaches us otherwise. In the model plant *Arabidopsis thaliana*, for example, several genes interact in complex ways to determine the timing of flowering (Simpson and Dean 2002). Some of these genes have antagonistic, or inhibitory, effects on each other. Therefore, some genes involved in flowering are actually 'for' inhibiting other genes involved in flowering, depending on the circumstances (usually, but not only, the conditions of the external environment, such as photoperiod, or temperature). Of course, each of these genes may also have (and in some

cases demonstrably has: van Tienderen et al. 1996) pleiotropic effects, thereby compounding the problems of pleiotropy and epistasis.

Closely related to issues surrounding the epistatic interactions of genes are those that concern the possibility that **coadapted gene complexes** might be the result of natural selection on particular kinds of traits. Given that many traits are associated not with single genes, but rather with complex suites of genes, it is worth considering whether such gene complexes could be 'for' particular traits in the sense we develop. It is certainly possible in principle that coadapted gene complexes might be the result of natural selection for a particular trait they are associated with (statistically and mechanistically), and hence candidates for 'genes for' talk. However, it should be noted that, at least in the case of eukaryotes, almost nothing is known about such complexes or the role they play in the development of phenotypes (Clarke 1993). Until we have far more information, which may well require rather more technical know-how than is currently available (for example through the further development of microarrays: Schenk et al. 2000), claims about the role that such complexes play will be conjectural at best. A partial exception is provided by the study of gene complexes in bacteria, often termed "operons," where the number of genes is reduced, their interactions fairly clear, and the phenotypes they induce are of a biochemical nature, and therefore more easily dissected (Prokudina et al. 1991; Silva and Dykhuizen 1993).

Pleiotropy and epistasis are also further complicated by **phenotypic plasticity** (Bradshaw 1965; Pigliucci 2001b). The environments in which the organisms live mediate the effects of genes as well as the interactions among genes. In a mathematical model of dominance and pleiotropy based on the details of the metabolic structure of organisms, Keightley and Kacser (1987) predicted that "dominance and any possible differences are

thus a function of the environment in which organisms are operating as well as of their genes." This is because enzymes (the direct product of many genes) are characterized by reaction curves that alter their efficiency in relation to environmental parameters such as temperature and pH, among others. That dominance depends on the environment was well known already during the first part of the 20th century, as is attested by the discussion of the *pennant/vestigial* system in *Drosophila melanogaster* to be found in Schmalhausen (1949, see **Fig. 2.1**). The *pennant* and *vestigial* alleles (at the same locus)

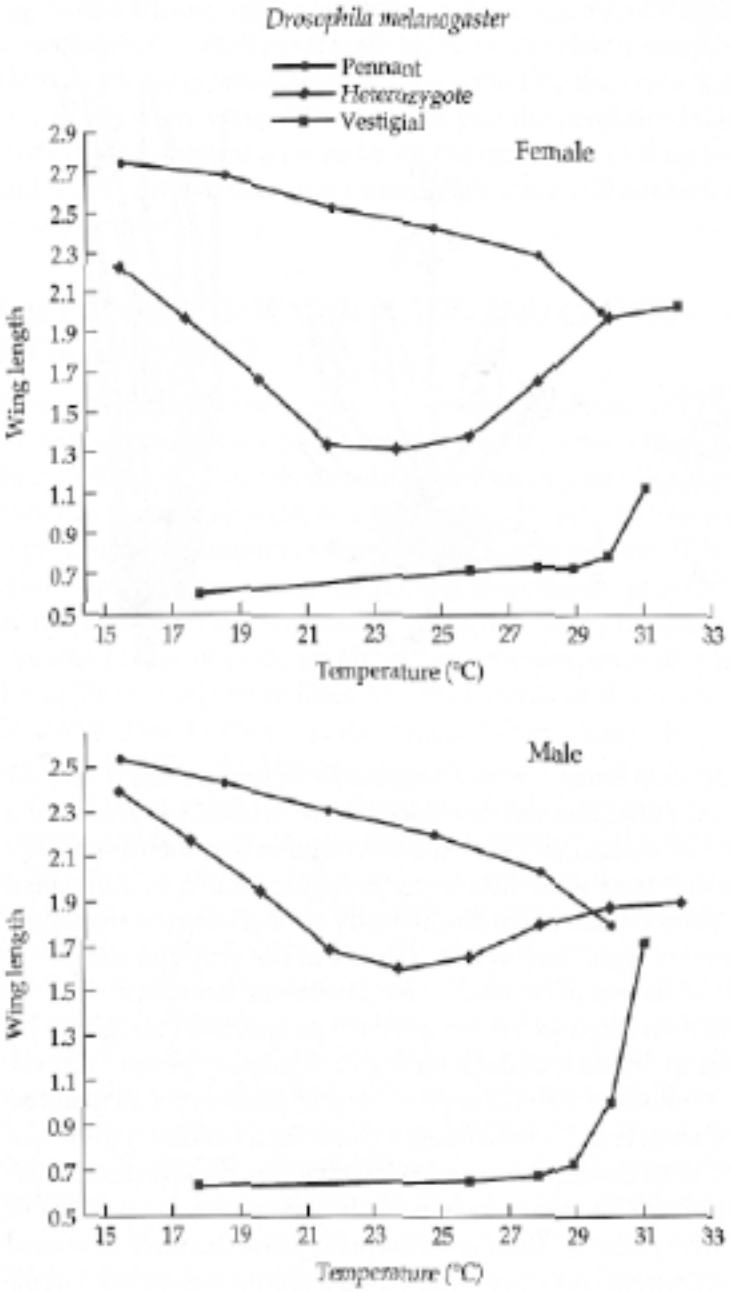

Figure 2.1. The *pennant / vestigial* (and their heterozygote) reaction norms in *Drosophila melanogaster*. The upper graph shows the reaction of the three genotypes to temperature and the corresponding wing length phenotypes for females. Lower graph for males.

change their degree of dominance in respect to the wild type allele along a gradient of temperature. The resultant phenotype (a more or less reduced wing) therefore also is affected by the combination of the genetic background and the environment in which this is expressed. The question then arises: are these alleles 'for' wing shape, or for the modulation of wing shape in response to temperature (i.e. for the plasticity of wing shape)? Without more information on the historical evolutionary context, the effects on fitness, and the interactions between the *pennant/vestigial* locus and other loci, the question cannot even be asked.

Two other points need to be discussed about plasticity: first, the fact that there are at least two types of plasticity, for which the current discussion on 'genes for' has very different implications; second, the special case of plasticity of fitness components. Smith-Gill (1983) discusses the difference between plasticity induced by phenotypic modulation and plasticity determined by developmental conversion. These two categories were actually defined (albeit using a more obscure terminology) by Schmalhausen (1949) in his seminal book on stabilizing selection. Schlichting and Pigliucci (1995) proposed an interpretation of the two categories in terms of modern molecular genetics, and this is the one I will follow here. Essentially, phenotypic modulation is the situation in which the phenotypic response to the environment is a continuous function proportional to the environmental input. For example, an animal's litter size may be directly proportional to the availability of food in the previous year. Schlichting and Pigliucci suggested that phenotypic modulation may be caused by allelic sensitivity, i.e. by a direct proportionality between the effects of gene products and the environment in which those products are expressed. Allelic sensitivity is very well known in biochemistry, where the reaction curves of enzymes are studied in response to a variety of environmental

conditions such as temperature or pH. Smith-Gill's developmental conversion, on the other hand, occurs when the phenotypic response is stable over a range of environments, but is "converted" to a different phenotype if the environmental signal surpasses a given threshold. For example, some plants will flower only if the photoperiod (day length) is longer than a certain number of hours, and will otherwise keep into the vegetative state indefinitely. Schlichting and Pigliucci suggest that developmental conversion may (but does not have to) be generated by specific plasticity genes, i.e., by regulatory elements that essentially control the position of the threshold and trigger one of a series of alternative developmental cascades. While one can make a reasonable argument that genes controlling developmental conversion may be genes 'for' the plastic response, the argument is much more difficult to substantiate in the case of phenotypic modulation. This is because while environmental receptors do not make sense if they are not triggering a specific (and presumably adaptive) response, gene products may be environmentally sensitive as a matter of unavoidable physicochemical properties, not because they were designed by natural selection. Of course, allelic sensitivity may be co-opted by selection to produce an adaptive plastic response. Indeed, Schlichting and Pigliucci (1995) proposed that this may be the intermediate step toward the evolution of developmental conversion. One would then regard this as an instance of exaptation at the molecular level. Of course, genes 'for' a continuous plastic response can determine phenotypic modulation, *if* the usual conditions of causal connection and historical continuity discussed above hold.

The last point on plasticity touches on the concept of fitness reaction norms (Weis and Gorman 1990). This is the idea that fitness (or fitness components) itself may be plastic, i.e. it may change as a function of the environment. Indeed, this is certainly the

case, as no organism is uniformly fit or unfit under all environmental conditions (the former case would be a "Darwinian monster" capable of driving every other species to extinction; the latter would simply be immediately selected out and go extinct). Furthermore, the idea of plasticity for fitness is important in evolutionary biology because it is directly related to the niche breadth of a given species. One could even define a niche in an evolutionary sense as the set of environments in which the fitness of a population remains relatively high. However, fitness plasticity has been used to model the evolution of plasticity per se (Via and Lande 1985; Via 1987). This latter use confuses the effects of a plastic response (of another character) on fitness with the environmental-dependence of fitness itself. One cannot reasonably argue that there are genes 'for' fitness plasticity, since by definition selection attempts to maintain as high a fitness level as possible under as wide an environmental range as actually encountered by the population. However, this fitness *homeostasis* can be achieved by the very presence of adaptive plasticity for other traits. For example, many aquatic plants produce two different types of leaves, one best adapted to gas exchange below water, the other to the conditions existing above water (Wells and Pigliucci 2000). As a result of this heterophylly, the plant is able to maintain high fitness under both flooded and terrestrial conditions. Plasticity for leaf shape leads to reduced plasticity for fitness. The first one has a genetic basis in the sense of there being genes 'for' it, the latter is a consequence of the former.

Perhaps paradigmatic of the 'gene for' problem in modern biology is the current fashion of studies on QTL, or **Quantitative Trait Loci** (Otto and Jones 2000; Loudet et al. 2002). The basic idea is in fact very simple and very old, in that it has been used in some form or another throughout the history of transmission genetics. It boils down to the possibility of mapping the position of a gene statistically associated with a given

character by examining the recombinant progeny of a cross between parents that are variable for that trait and a series of unrelated markers. If the phenotype of interest is significantly correlated with some markers rather than others, one can provisionally conclude that the genomic regions thus identified include one or more genes whose effects alter the phenotype of interest. Originally, this technique was used with phenotypic and then chromosomal markers. Today, it is much more effective (and almost as cheap) to use it with molecular markers (Basten et al. 2000).

However, there are some fundamental limitations inherent in the QTL approach. For one thing, it can be carried out only on one cross at a time, thereby limiting the extent of natural variation that can be sampled. Secondly, it biases the outcome toward the discovery of a few genes with large effects, because the power of the analysis is directly proportional to the sample size used in assaying the progeny. Nevertheless, QTL analysis allows the researcher to identify genes associated to trait variation in nature (as opposed to the functionally superior, but more artificial, method of mutational screening following chemical or radiation treatment: Koornneef et al. 1982). But is a given QTL a gene 'for' the phenotype under study? Not in any biologically meaningful sense of the term, and certainly not as it is developed here. For one thing, QTL analysis only identifies chromosomal regions that contain the relevant gene(s), it can rarely pinpoint *a* gene due to the limited number of markers and progenies usually tested. This, however, is only a technical limitation related to the power of the analysis and the sample sizes used. A more serious problem arises when one considers that QTL studies, as much as they are based on the apparently mechanistic approach of molecular mapping, are in fact the application of a statistical technique. As such, QTL mapping does not tell us anything at all about the *function* of the candidate genes, although it can provide a preliminary screen

toward further, more mechanistically oriented, studies. Therefore, claims such as those in Wu (1998) or Kliebenstein et al. (2002) that plasticity genes with regulatory functions can be identified and mapped through QTL studies are, to say the least, premature. This claim stems from the confusion between statistical and physiological epistasis referred to above, where the first has been detected and improperly translated into the second (the latter being the one required by the current theory of plasticity genes: Schlichting and Pigliucci 1995; Pigliucci 1996).

All of the above problems are compounded by the fact that phenotypes are actually the result of the emergent properties of developmental systems, not just of gene actions. This dreaded locution has been recently increasingly accepted by evolutionary biologists under the guise of "epigenetics" or "epigenetic rules" (Holliday 1998; Pal and Miklos 1999; Newman and Muller 2001). In fact, the new incarnation of sociobiology (sometimes referred to as Darwinian or evolutionary psychology) is recast around the idea that selection does not favor one gene or another, but one epigenetic rule or another (Wilson 1998). What is an epigenetic rule? While it is still frustratingly difficult to accurately pinpoint it, the idea is that phenotypes and behaviors are the result of a complex and continuous (but not necessarily unintelligible) interaction between genes and environments. Perhaps one of the best conceptualizations of epigenetics comes from the work of Atchley and Hall (1991). They pointed out that tissue-tissue communication (as in the induction of the formation of eyes in vertebrates by the adjacent neural crest) is not a genetic effect in the strict sense, but a response to the internal environment of the embryo. Moreover, this epigenetic effect catalyzes further gene action, necessary for the production of proteins specific to the new structure being developed (e.g., the crystalline in the eye). Local (intracellular) genetic effects, internal environments (cell-cell or tissue-

tissue interactions), and external environments therefore all combine to produce the phenotype and behavior of the organism. In cases of this sort, it will be difficult to make good sense of the idea of genes being 'for' specific attributes of those phenotypes and behaviors.

A particularly well-understood example of a gene that can have many, possibly evolutionarily profound, effects, while it cannot be thought of as a gene 'for' those effects, is the *hsp90* gene of *Drosophila* and *Arabidopsis* (Rutherford and Lindquist 1998; Queitsch et al. 2002; Sollars et al 2003). This gene codes for a protein that is part of the complex heat-shock response present in virtually every organism for which it has been searched in. Rutherford and Lindquist (1998) described *Drosophila* flies in which the function of the HSP90 protein has been impaired, either by mutation or by chemicals. These animals display an array of morphological changes, spanning virtually every aspect of *Drosophila*'s adult phenotype. The specific effects of mutations at the *hsp90* locus depend on the genetic background being considered. Most interestingly, selection can stabilize these effects so that after a few generations they are present even if HSP90 is functional. Rutherford and Lindquist conclude that HSP90 buffers *Drosophila*'s developmental machinery against this variation, thereby allowing other loci to behave neutrally and to accumulate mutations. A mutation at the *hsp90* locus or certain environmental conditions can "release" this hidden variation, promoting evolutionary change in spite of a conservative developmental system (a similar result was found later in the model system *Arabidopsis*: Queitsch et al. 2002).

From the perspective of our discussion so far, one can certainly argue that *hsp90* is 'for' buffering the fly's developmental machinery from heat shock. However, it would be hard to claim that *hsp90* is also 'for' accumulating a reservoir of genetic variation, or

'for' the occurrence of alternative phenotypes under certain environmental conditions. Yet, these additional roles may make *hsp90* and similar genes crucial for long-term phenotypic evolution, as Waddington suggested many decades ago (Waddington 1942, 1961; Pigliucci and Murren in press).

Genes for, exaptations, and genetic piracy

The Modern History view of genes being for phenotypes is a powerful tool for thinking through the way in which evolution can work to change the biological significance of various genes. On this view, genes that were for a particular phenotype can, in time, become for other phenotypes entirely, or indeed, for nothing at all (e.g., they may become pseudogenes). Similarly, genes that were not in fact 'for' anything in particular (at least at the gross phenotypic level) can, over time, gain that sort of biological significance (Wilkins 2002).

An example of gene that may once have been for some phenotypic trait but no longer is may be that of the alleles associated with cystic fibrosis (CF). It has been suggested that the prevalence of genes associated with cystic fibrosis in some populations is due to the heterozygote superiority such alleles afford. Heterozygotes for alleles associated with cystic fibrosis may well be highly resistant to typhoid fever, but homozygotes for genes associated with CF generally acquire the disease and die young (Pier et al. 1998). If true, the situation would be analogous to that of alleles associated with sickle cell anemia, discussed above with respect to pleiotropy. However, in contemporary western populations, the continued existence of the alleles associated with CF cannot, in all likelihood, be explained by the selective advantage heterozygosity

provides. Resistance to typhoid fever has not been a major selective pressure in the very recent history of most western populations; the existence of such genes in modern populations is probably due to simple inertia. Therefore, genes that may once have been *for* something, e.g. resistance to typhoid, are now not *for* anything at all. If at some future time, typhoid once again will become a major threat and will exert a significant selective pressure, and the maintenance of the alleles associated with resistance to it in the population will be once again actually due to the resistance they afford, the genes might de novo become *for* something. Otherwise, the genes in question are likely soon to accumulate mutations and become pseudogenes, literally molecular "fossils" of their former selves.

A similar example is provided by one of the current conjectures attempting to account for the prevalence of genes associated with HIV resistance (the so-called CCR5-Delta32 'AIDS-resistance' allele) in certain populations (Stephens et al. 1998). If the prevalence of this gene in some populations were the result of selective pressure for resistance to the bubonic plague (as some researchers have conjectured), then it seems that at some time in the recent past (about 600 years ago) the gene may well have been *for* resistance to the bubonic plague. More recently, however, its maintenance in the population has presumably not been due to any active selection for it at all, but rather to evolutionary inertia. If this scenario is correct, the gene's usefulness in providing resistance to HIV might make that feature out to be an *exaptation* in the sense described by Gould and Vrba (1982) and developed more fully by Griffiths (1992). If the prevalence of the gene in extant populations is not being actively maintained by natural selection, we would have what Griffiths calls an 'exaptation.' If at some later date, the gene's association with HIV resistance is responsible for its maintenance due to selective

pressures associated with that resistance, the gene would be in Griffiths' terms an 'exadaptation' (Griffiths 1992, p. 119). In the view being developed here, in the case of an 'exadaptation' the gene in question would have become a plausible candidate for being a 'gene for HIV resistance'. Obviously, this is very speculative, but the basic idea, that on the Modern History approach to 'gene for' talk, a gene that was for some trait can lose that (functional) association and later become 'for' some other trait is sensible.

Indeed, this aspect of the Modern History approach to thinking about genes 'for' traits is especially useful for considering the way in which new gene functions can (and do) evolve. Roth (1988) presents an interesting scenario in her discussion of genetic piracy. According to Roth's definition, genetic piracy is the process by which:

> "New genes, previously unassociated with the development of a particular structure, can be 'deputized' by evolution; that is, brought in to control a previously unrelated developmental process, so that entirely different suites of genes may be responsible for the appearance of the structure in different contexts." (Roth 1988, p. 7).

Roth cites de Beer on the eyeless mutants of *Drosophila*, in which it is possible to select for "other genes" which result in "a fly with restored eyes that still has the original mutation." From this, Roth concludes that in these cases "new genes are involved in the formation of eyes that previously had not been" (Roth 1988, p. 7). In this way, a different set of genes may become associated with the formation of a trait, and in some of these cases, different genes may end up being *for* the 'same' (a homologous) phenotypic trait. Interestingly, one of Davies' (1994) criticisms of functional ascriptions in biology is

based on the difficulty in specifying identity conditions for traits throughout recent evolutionary history. By forcing the researcher who wishes to claim to have found genes 'for' traits into focusing on the recent evolutionary history of the relationship between the gene and trait in question before making any such claims, the Modern History approach focuses attention onto the way in which this relationship is somewhat fluid. Indeed, this very fluidity is the key to understanding the evolution of 'new' gene functions and associations.

Genes for: a case study of how far we can push the concept

Philosophical discussions of the type engaged in here used to be limited by the scant availability of actual hard data pertinent to the subject matter. Modern molecular and evolutionary biology, however, afford us a unique opportunity to probe the advantages and limitations of the concept of a 'gene for' by testing it against concrete examples about which we know a great deal. Of course, in none but a few simple cases we know all or even most that there is to know about the genotype-phenotype mapping function. The black box of epigenetics has just recently being tackled by molecular developmental biologists, and a dim light is been shed on its interior. In the following, I will discuss the example of a gene, the *phyB* locus in the weedy crucifer plant *Arabidopsis thaliana*, as a guide to a more concrete understanding of what do we really mean by genes 'for' something.

The *phyB* locus codes for one of five alternative forms of a class of molecules named phytochromes. Phytochromes are light-harvesting molecules that function as photoreceptors in *A. thaliana* and in many other organisms, including all flowering plants,

non-flowering plants, green algae, and cyanobacteria (Wu and Lagarias 1997). The specific functions and effects of these photoreceptors differ to some extent, but the PHYB gene product is sensitive to the ratio of Red to Far-Red (R:FR) wavelengths of light. The molecule can exist in the cell in two forms, one capable of absorbing red light and thereby switching conformation to the alternative form; the other form is sensitive to far-red light and shifts to the red-form. PHYB is present in *A. thaliana* in both forms, whose chemical equilibrium depends on the R:FR ratio (Quail 1997). Phytochrome B has been associated with a number of phenotypic effects, generally grouped under the label of "shade avoidance response." The idea is that under natural conditions the R:FR ratio is an indicator of impending competition: the surrounding vegetation absorbs the red component of the spectrum, which is photosynthetically active, but reflects the far red, too weak to be energetically useful (Givnish 1982; Smith 1982; Casal and Smith 1989). This alters the R:FR ratio from about 1:1 under normal sunlight to values below one, in proportion to the density of the vegetation. If a plant can sense the level of competition before it becomes too strong (i.e., before actual competition for light, water, or nutrients ensues), it may be able to adopt alternative strategies to minimize the damage. These strategies include suppression of branching to concentrate resources on vertical growth of the main stem and accelerated flowering to complete the life cycle before the quality of the environment deteriorates further (Schmitt 1997; Dorn et al. 2000; Weinig 2000).

Indeed, physiological studies have confirmed that phytochrome B affects a variety of plant phenotypes: seed germination (Shinomura et al. 1994), cell elongation (Reed et al. 1993), cotyledon expansion (Neff and Van Volkenburgh 1994), seedling appearance (Casal 1995), flowering time (Halliday et al. 1994), as well as leaf production and branching pattern (Pigliucci and Schmitt 1999). Furthermore, evolutionary biologists have

amassed convincing evidence that the shade avoidance response is indeed adaptive, not only in *A. thaliana* but in other plants as well (Dudley and Schmitt 1996, although by no means in all plants: species that permanently live under shade in the understory do not show any appreciable shade avoidance syndrome, Bradshaw and Hardwick 1989). Finally, molecular genetic investigations have demonstrated that the mechanics of action of the *phyB* locus are complex. Its gene product interacts with a light-labile phytochrome termed PHYA (Reed et al. 1994), with phytochrome C (Qin et al. 1997), and with a completely different class of photoreceptors sensitive to blue light and known as cryptochromes (Mas et al. 2000). Phytochromes A and B interact with blue receptors in a complex manner, both synergistically and antagonistically, depending on the environmental situation and the characters studied (Callahan et al. 1999).

Of course, photoreceptors by themselves do not actually exert any biological function, and they have to act together with other gene products to do so. Here the literature is much more vague and incomplete, although we know that there are several "transduction genes" whose products carry the information from the phytochromes to other molecules (Quail et al. 1995; Casal et al. 1998; Whitelam and Devlin 1998). Eventually, the bio-effectiveness of light perception and transduction is mediated by one or more plant hormones such as gibberellin (Lopez-Juez et al. 1995; Chory and Li 1997), although other hormones (e.g., cytokinin) affect some of the same traits altered by light signals independently of the phytochromes (Su and Howell 1995).

So, what is the *phyB* locus 'for'? The obvious answer is that its function is to gauge the R:FR ratio. But surely light perception per se cannot be the target of natural selection, unless the information so acquired is actually used in some biological function related to fitness. Plants, after all, are not just curious about their environment. And here

is where the trouble begins. If we use the studies associating mutations at the phytochrome B locus with their phenotypic effects, we are led to conclude that the gene is for the control of germination, cell elongation, leaf production, flowering time, apical dominance, branching pattern, and reproductive output – to say the least. One problem with this sweeping generalization is that in fact a lot of other genes also seem to be 'for' the same characters, with only partial (and sometimes contrasting) overlap with the effects of *phyB*. Furthermore, some of these traits (e.g., leaf production) are altered by mutations at the phytochrome B locus not because the phytochrome has much to do directly with leaf production, but more likely because the general growth rate of these mutants is slower than the wild type plants (this may also account for part of the differences in flowering time). That is, some of the alterations associated to changes at the *phyB* locus are bound to be accidental byproducts of the mutation, and not indicators of the wild type function of the gene.

Another level of complication in order to attribute a specific function to the phytochrome B gene arises when we consider that its biological activity is strictly environment-dependent: under a variety of environmental circumstances (i.e., unless the R:FR ratio is altered) the gene does not really do anything at all. That is why this locus has been proposed as one of the best candidates for the role of "plasticity genes," genes whose effect makes sense only under environmentally heterogeneous conditions (Pigliucci 1996).

The wealth of molecular information accumulated on the action of *phyB* does not seem to help clarify things much. One could say, for example, that the gene product of this locus is 'for' stimulating a transduction cascade that eventually triggers the action of the gibberellin hormone. But more precisely one would have to say that *phyB* is for

triggering gibberellin through the use of intermediary molecules (and specify this mediation effect). How far from the immediate effect of the phytochrome transcript can we go and still meaningfully speak of its function? On the other hand, surely it would be limiting to say that the gene is for photoreception and activation of a transducer (which is all the gene product of the *phyB* locus *really* does). Ignoring the fact that these actions have biological consequences affecting the plant's fitness would be as absurd as ascribing the entire plant phenotype to the action of this particular gene. Yet, the reality is that one can place the PHYB molecule on any point of the sliding scale defined by this continuum. The problem is that genes do not do anything by themselves, but only in relation to their environment, both the external one, and the one provided by the actions of other genes.

Evolutionary and phylogenetic information, also available for *phyB* (Mathews and Sharrock 1997), does not solve the problem either. We know that it belongs to an ancient family of genes, certainly predating the origin and evolution of flowering plants. We also know that these genes have diversified their biological effects, which are now only partially overlapping. If we go back far enough in time, the phytochromes can be thought of as very ancient photoreceptors, whose DNA sequence has been highly conserved for the past billion years at least (Quail 1997). Such a long phylogenetic history also strongly suggests that they must have been maintained by natural selection, or they would have been lost (or at least their sequences would have been much more dramatically altered). Therefore, one could safely conclude that the phytochrome genes are really 'for' photoreception. However, their other functions/effects have changed considerably throughout the evolution of cyanobacteria, algae, and plants. Therefore, they have always been more than just photoreceptors, but this additional component has been constantly shifting through evolutionary time. This is exactly what I mean by the Modern History

view that considers the idea of a gene for something more fluid and fuzzy than the classical rendition. In this specific example, the *phyB* gene controls photoreception and – indirectly – a series of environment-dependent changes in the phenotype and phenology of the plant. This flexibility does increase the fitness of the plant and the gene does qualify as a plasticity gene. However, which bits of the response are adaptive and which are allometric byproducts of altered growth rates remains to be established, and it is a very difficult empirical question.

Conclusions: the advantages of the Modern History approach to 'genes for'

I have argued that no approach to finding genes that relies on their merely being correlated with phenotypic traits can properly claim to have found genes *for* those traits. To conclude that a gene is not merely associated with a trait, but rather is actually *for* that trait, one needs significant amounts of evidence about not only the association, but also the biochemical pathways involved, and the recent evolutionary history of the organism in question (see **Table 2.1**). The Modern History approach demands that to speak properly of genes for a trait, the frequency of the genes in question was maintained in the population *because* of their association with that trait, and the resulting positive selective pressures for that trait. That is, in moving from the discovery of a candidate gene (a first step which, unfortunately, many researchers feel inclined to stop at) to a 'gene for' some trait, one needs at least information on the biochemical pathways the gene is involved in, on the epigenetic effects (including pleiotropy, epistasis, regulatory plasticity, etc.), about the functional ecology of the phenotype in question, about the phylogenetic history of the organism and of the gene (or gene family), etc. Finding these things out

Table 2.1. A checklist for "genes for".

Type of necessary evidence	Methods to gather the evidence
Statistical association	Quantitative genetics, QTL analysis
identification of candidate genes	QTL analysis, mutagenesis
Information on gene regulation (developmental and environmental)	Molecular and ecological genetics
Functional ecology of associated phenotypes	Field studies on fitness consequences, experimental manipulation. functional ecology and quantitative genetics.
Phylogenetic history (of both the organisms and of the gene or gene family)	Molecular systematics, phylogenetic comparative methods

demands painstaking work; one cannot substitute assumptions, no matter how attractive, as many 'sociobiologists' and people partaking in 'evolutionary psychology' have tried to do. While it is obviously impossible to know all of the above in any particular case, a consilience of evidence may move us to speak (albeit cautiously) of a gene being 'for' a trait where some of the above evidence is wanting. These are not easy things to find out even in organisms in which experimental manipulations are possible. In the case of humans, where experimental manipulations are out of the question, they are more difficult still, perhaps so much so that we can never be confident to have found a gene 'for' a reasonably complex phenotypic trait in humans. However, I do not think that this is a reason to despair, or that this places undue burdens on current research.

Indeed, I feel that the recognition that we can rarely find genes 'for' specific phenotypic traits is valuable for many reasons. This recognition would focus attention on the evolutionary relationship between phenotypes and genes, including the fascinating

ways in which genes can acquire new functions, and 'old' structures become associated with different genes. Such recognition would also concentrate our efforts on the relationship between the organism and its environment, and the way in which gene-environment interactions shape the ontogenetic pathways and evolutionary histories of organisms. In drawing attention on the relationship between genetic and phenotypic development in an evolutionary context, this view of 'genes for,' as challenging as it may be, keeps the attention where it ought to be. These are positive benefits of this way of thinking about 'genes for' traits. But the dangers that this way of thinking avoids are equally significant.

The number of purported discoveries of 'genes for' various complex traits, often in humans, that are reported in the scientific literature often gives a misleading picture of the current state of research into the relationship between genes and phenotypes. But the misunderstandings and confusions that reports in the popular media engender, largely because of the very language of 'genes for,' is if anything even more worrisome. These misunderstandings are problematic because they systematically mislead the public. They raise illegitimate expectations about the power of genetic research, and these misunderstandings end up actually influencing public policy decisions, and doing so in ways that a proper appreciation of the current state of knowledge about the relationship between genes and complex phenotypic traits in humans would not countenance. To claim, as Dawkins does, that when geneticists say 'gene for' what they mean is 'a statistical association with' a phenotype borders on intellectual dishonesty. Geneticists (and other researchers), especially those that talk to the media, do not get to play Humpty Dumpy – their words do not get to mean whatever they choose them to.

Chapter 3: On the concept of biological race and its applicability to humans[4]

"Class, race, sexuality, gender—and all other categories by which we categorize and dismiss each other—need to be excavated from the inside."
–Dorothy Allison, *Skin* (1994).

It may seem strange to say that the concept of race is associated with some unexamined philosophical and scientific baggage, but my claim here is that this is indeed the case. More precisely, as we shall see, I think that the potential for the actual existence of human races in a different sense from the folk concept of race has not been explored. This lack of investigation has led to all sorts of misguided or misinformed statements about race, both on the part of biologists and of philosophers. This chapter attempts to clean up some of the resulting mess.

It has become commonplace to claim that, insofar as 'race' is a biological concept, there are no human races and hence that, biologically speaking, human races do not exist (e.g., MacEachern 2000). This claim, while defended by a wide variety of arguments, is misguided for several reasons. I suggest that careful attention to the uses of 'race' in the context of non-human biology can reveal in what ways human races might exist in a biologically meaningful sense, and what the limitations of the concept might be. While I argue that there likely are a variety of identifiable and biologically meaningful human races, these will not correspond to folk racial categories, nor will the fact of their existence

[4] Written with Jonathan Kaplan, currently in press in *Philosophy of Science*.

offer any support to racist views. As Hull pointed out, "The nature of biological species is a moral issue only for those people who ground human rights in human nature" (Hull 1998).

In the following, then, I will discuss the concept of race as used by biologists outside of human studies; I will then consider the specific applicability of the biological concept of race to humans and the problems it poses. I will examine specific objections that have been raised to the recognition of races in humans, and finally discuss the relationship—or lack thereof—between biological and folk races. The study of races is an interesting object of inquiry for biologists, including in the case of humans, but it has little or no consequence for social studies and for our understanding of the phenomenon of racism. Indeed, because the folk conception of human races is well-entrenched, politically and socially loaded, and does not for the most part align with the biological uses, I suggest that avoiding the term 'race' with respect to the human case would be advisable in order to prevent confusion. However, this suggestion does not imply that there aren't human races, but only that distinguishing between the biologically meaningful ones and the socially meaningful ones might best be done via the use of different terminology.

What are biological races?

For a concept that is allegedly in disuse in biology (Andreasen 1998; Futuyma 1998), an awful lot of papers have been published in the last few years that include the term "race" in their title or abstract. While many of these works refer to human races, a large number emerge from the non-human biological literature. Exploring the way that 'race' is currently used within the context of non-human biology is therefore a crucial first

step to determine whether there are biologically meaningful human races. Indeed, many of the current arguments against the existence of biologically meaningful human races fail precisely because they rely on a use of the biological race concept that is not in fact in wide circulation.

King and Stansfield's (1990) dictionary of genetics defines race as: "A phenotypically and/or geographically distinctive subspecific group, composed of individuals inhabiting a defined geographical and/or ecological region, and possessing characteristic phenotypic and gene frequencies that distinguish it from other such groups. The number of racial groups that one wishes to recognize within a species is usually arbitrary, but suitable for the purposes under investigation." Darwin (1859) had considered races as subspecies, as made clear in the very title of *The Origin of Species*, where he referred to the preservation of favorite *races* in the struggle for life. Darwin saw species themselves as fairly fluid entities, in accordance to his theory of gradual transformation of biological populations over long periods of time. *A fortiori*, subspecies had therefore to be even more labile. Interestingly, King and Stanfield do refer the reader to the term "subspecies" in their dictionary. It, in turn, is defined as "1. A taxonomically recognized subdivision of a species. 2. Geographically and/or ecologically defined subdivisions of a species with distinctive characteristics." Notice that the second definition is essentially the same as the one given above for race by the same authors. And yet, the actual biological literature is much more ambiguous about the relationship between races and subspecies, which is for example endorsed by Andreasen (1998 but explicitly rejected by Templeton (1999).

Given such confusion, it is perhaps more instructive to briefly consider how practicing biologists actually use the term race and how they relate it to subspecies. The

following examples are far from a complete survey of the literature, but they are representative of recent papers on the biology of races in animal and plant systems. Rehfeldt and Gallo (2001) have studied what they refer to as races of Douglas-fir from North America and Argentina with the express purpose of comparing their growth potential at different elevations. They concluded that the ponderosa pine land race was descended from a California low or middle elevation population, thereby implicating both eco-geographical characteristics and patterns of descent (cladistics) in their usage of the term race. Jiggins et al. (2001) expressly linked races to the speciation process in two butterflies. *Heliconius melpomene* and *Heliconius cydno* are sister species that diverged recently by mimicking different model taxa, and the authors have shown that there is assortative mating within this system, so that individuals of one species tend to avoid mating with those of the other, an event that creates hybrids lacking the mimetic coloration and therefore characterized by lower fitness. The point is that Jiggins and colleagues think of the stages preceding the actual speciation event in terms of the formation of two races/subspecies, which eventually led to the almost complete reproductive isolation observable today.

Yet another concept of race emerges from studies of pathogens. Vicente et al. (2001), for example, studied 164 isolates of *Xanthomonas campestris*, a bacterial pathogen that infects plants of the genus *Brassica* (a member of the mustard family). 144 of these strains were grouped in six races on the basis of their host specificity and pathogenicity, while the remaining 20 isolates could not be assigned to putative races because of their weak or absent pathogenicity. In this case, therefore, ecological adaptation to a specific host is considered the hallmark of a race, regardless of the geographical location of origin of the bacterial strains.

An interesting twist on the whole idea of race emerges from an unusual study conducted by Palmateer et al. (2000) on the definition of race in the nematode worm *Heterodera glycines*, which attacks soybean plants. These authors tested the currently accepted scheme to identify races of *H. glycines* by performing their experiments at three different temperatures. Interestingly, they found that while all three indices used to identify races gave consistent results when the worms were grown at 27_C and 30_C, this was not the case if the animals were raised at 20_C. This was apparently because of environmental effects (i.e., phenotypic plasticity, Pigliucci 2001), which differentially affected the adult female, egg and juvenile stages of the life history of these animals.

A final example makes the connection between races and the concept of ecotype, which will play a major role in the rest of our discussion. Stone et al. (2001) studied the differential success in northward expansion of two ecotypes of marble gall wasps, one attacking the Turkey oak, the other the cork oak. The authors found evidence for at least four independent lineages surviving in glacial refugia and expanding after the end of the last glaciation, with the currently existing Northern populations divided into two sets: southwest France and Spain on the one hand and the remainder of northern Europe, Italy and the Balkans on the other. Pertinent to our discussion here, the transition from the French to the northern European race was found to be abrupt in both ecological and genetic terms: the French populations attack the cork but not the Turkey oak and show the presence of "private" alleles, i.e., forms of various genes that are found only in one or the other race. While hybridization of the two forms has been detected, Stone et al. consider these two entities a good example of races, a term which they treat as synonymous with ecotype.

It then appears that within the current biological literature the term "race" is used

to pick out subspecies/incipient species, geographically distinct populations, ecologically differentiated units, genetically dissimilar entities, or independent phylogenetic lineages. All or a mix of these concepts are used by practicing biologists dealing with a range of non-human organisms. It is obvious that the underlying idea here is that there is *something* that distinguishes races, but that something can be one or more of different ways in which populations become somehow distinct in nature. Furthermore, several of these distinguishing criteria will likely overlap because of the nature of biological organisms. For example, if a population is ecologically distinct (e.g., it lives at high elevations) it is also likely to be geographically isolated (by virtue of occurring in a location at high elevation), and to be somewhat genetically differentiated. The latter property can be due to two phenomena: some genes may have undergone selection for specific adaptations to the ecological conditions of the population, or the genome may simply have become distinct enough from other populations of the same species because of accumulated mutations and random genetic drift.

But while genetic and phenotypic differences between local populations will often be associated with phylogenetic distinctiveness, such differences do *not* imply phylogenetic distinctiveness, nor, *a fortiori*, do they imply incipient speciation. Let us see why. For a lineage to acquire phylogenetic distinctiveness, i.e., for it to become a separate clade, that lineage has to essentially cease gene flow with other closely related populations. This is because if gene flow is still significant, then the lineage is evolving according to a reticulate, not cladistic (branching) pattern. While it is still possible for such an entity to maintain ecological distinctiveness (see below), its historical roots are continuously reshuffled by migration events. If the lineage does in fact becomes a new stable clade, then two outcomes are possible: it will either go extinct (which is the most

likely result), or it will eventually accumulate enough genetic differences from the rest of the original species that it will no longer be able to inter-cross and produce viable offspring: a speciation event will have occurred. Thus, while eco-geographical-genetic differentiation tend to correlate with each other they do *not* imply cladogenesis and speciation, though the latter two are themselves associated (in part by definition).

That biologically meaningful races do not have to be phylogenetically distinct is obvious when we consider the case of ecotypes. The concept of ecotype was introduced by Turesson (1922) to describe genetically-based specific responses of plants to certain environmental conditions, although the idea has been applied to the animal literature as well (Kinley and Apps 2001; Lu et al. 2001). The King and Stansfield's dictionary defines an ecotype as a "Race (within a species) genetically adapted to a certain environment." It is important to understand three things about ecotypes: 1) As the definition clearly states, there has to be a connection between genetic differentiation and ecological adaptation; but 2) This does *not* mean that large chunks of the genome of an ecotype have to be different from the average of the species to which that population belongs. In extreme cases, as in heavy-metal tolerance in a variety of organisms (Lovley 1993; Martinez and Levington 1996; Schat et al. 1996; Vekemans and Lefebvre 1997; Boyd and Martens 1998; Shirley and Sibly 1999), the status of ecotype may be due to the presence of one or very few genes that confer the ability to survive in a particular habitat, though it can involve a fairly complex system of molecular shuttling, distribution and detoxification (Clemens 2001). Finally, 3) Ecotypes are not (necessarily) phylogenetic units; rather, they are functional-ecological entities.

These facts about ecotypes have several important implications for our discussion of race. Similar ecotypic characteristics can and do evolve independently in geographically

separated populations, as clearly demonstrated by a study of large- and small-bodied ecotypes of the freshwater amphipod *Hyalella azteca* conducted by McPeek and Wellborn (1998). These similar phenotypic characteristics may, or may not, be mediated by similar genetic differences from other populations of the species (because of genetic piracy: Roth 1988; or because of possible genetic redundancy: Clegg and Durbin 2000). Further, gene flow between different ecotypes is relatively common (DeSalle et al. 1987; Berry and Kreitman 1993; Wang et al. 1997). As long as there is sufficient selective pressure to maintain the genetic differences associated with the different adaptive phenotypes, other genes, not so associated, may flow freely between the populations. Further, because different ecologically important characteristics are not guaranteed to coincide, a single population can consist of multiple overlapping ecotypes. In such cases, whether two organisms belong to the same ecotype will depend on *which* ecotype one is referring to.

These points will become particularly important when we discuss why I think that insofar as there might be human races of biological interest, these will best be thought of as ecotypes. Many of the arguments that conclude that there are no, or could not be, human races depend upon definitions of race that reject the ecotype interpretation of the concept. By focusing on the ecotype concept of race, I show that biologically interesting and significant human races may well be discoverable.

Human races: definitions and problems

Given the variety of ways in which 'race' is used in the biological literature, it is hardly surprising that a significant element of the debates surroundings the existence of

biological *human* races is the particular definition of 'race' used. Indeed, some authors have argued against the existence of biologically significant human races by suggesting that there is no acceptable 'race' concept in biology more generally (e.g., Futuyma 1998, p. 450); however, as noted above, the vagueness of the biological race concept does not prevent its useful application in many areas of non-human biology (as is true for the concept of species itself: Pigliucci in press, see also Chapter 4). The question, then, is not whether biological 'races' exist, but rather whether biological race concepts can be usefully applied to human populations, and whether the careful application of these concepts to the human case will reveal the existence of any biological races in *Homo sapiens*.

Insofar as one considers appeals to biological races to be attempts to pick out incipient species, it seems perfectly clear that currently there are no human 'races.' There are no extant populations of our species that are plausible candidates for being incipient species. Further, the current distribution of genetic variation within *H. sapiens* implies that at no time in the past were any of the lineages that led to the currently extant portions of the human population evolving independently (a necessary precondition for being an incipient species, see Templeton 1999 and cites therein). While the genus *Homo* very likely generated incipient species during its history (and perhaps even full-fledged separate species), none of these currently survive (Tattersall 1995, 1998, 2000). The evolution of contemporary *Homo sapiens* was likely *not* marked by populations that at one time had independent evolutionary trajectories but exist today as part of the larger population (Waddle 1994; Rogers 1995; Templeton 1999).

Rather, human evolution seems to have been marked by extensive gene flow. This immediately implies that there are not now, nor ever were, biologically significant human

races that corresponded to populations that had been phylogenetically separate for some significant period of time (contra Andreasen 1998). However, extensive gene flow does not imply, as it has been argued, that there can be no significant biological races in the human case. As we saw above in the case of ecotypes, adaptive genetic differentiation can be maintained between populations by natural selection even where there is significant gene flow between the populations. Templeton (1999), for example, notes that gene flow sufficient to ensure that distinct populations evolve together as a single species is compatible with the populations having distinct, genetically mediated, phenotypic adaptations. For example, he notes that there are populations of *Drosophila mercatorum* in Hawaii that "show extreme differentiation and local adaptation," yet have significant gene flow between them. In these cases, to talk of biological races is not to motion towards populations with their own evolutionary trajectory, but rather towards populations that differ from each other in some biologically significant way, where these differences are 'determined' by genetic differences maintained by natural selection. As we saw above, these biological differences can be maintained despite extensive gene flow where there is a strong adaptive significance to the difference.

Lewontin and Gould have made much of the fact that there is relatively little genetic variation in *Homo sapiens* (at least compared to other mammals, see Templeton 1999), and that most of what genetic diversity is known to exist within *Homo sapiens* exists within (rather than between) local populations (see, for example, Lewontin et al. 1984; Gould 1996). These facts are cited repeatedly in arguments concluding that there are no biologically significant human races. But the idea that this data might imply something about the existence of biologically significant human races emerges from a focus on the wrong sort of biological races, as far as humans are concerned. If one is arguing

against human races forming (or ever having formed) incipient species or reproductively isolated populations, such evidence might be important. The relative lack of genetic variation between populations compared with within populations samples does indeed imply that such populations have not been reproductively isolated for any evolutionarily significant length of time. But of course, this fact is *irrelevant* to the consideration of races based on adaptive variation; in this case, if there is extensive gene flow, genetic variation can be mostly within groups, rather than between groups, as variations not related to adaptive phenotypic differences between populations will be spread by gene flow relatively easily. The question is not whether there are significant levels of between-population genetic variation overall, but whether there is variation in genes associated with biologically important *adaptive* differences between populations.

So, if we conceive of races similarly to the way ecotypes are conceived of, it is clear that much of the evidence used to suggest that there are no biologically significant human races is in fact irrelevant to the question as it has been rephrased. As long as differences between populations can be maintained by selection because of their adaptive significance, races can exist despite extensive gene flow between populations. The question, then, is whether such conditions exist in the case of human populations, and whether such conditions existed during the course of human evolution and influenced population differences to an extent that might still be detectable today (though perhaps no longer being actively maintained).

Before addressing that question, it is worth taking a short detour to consider *why* so many authors writing about the (non) existence of human races have made use of such a strong definition of race, i.e., they have assumed that biologically significant races must be populations separated from other populations by serious barriers to gene flow. Part of

the reason undoubtedly has to do with the history of the term 'race' as it is applied to humans. Insofar as one is asking a question not about the existence of biologically significant races (of the sort that exist in certain species of *Drosophila*, for example), but rather about the existence of a biological *justification* for the 'ordinary' language sense in which racial categories are used, the concept of race appealed to will have to be quite strong. As for example Appiah (1996) and Hull (1998) point out, the races colloquially appealed to are generally supposed to differ from each other not merely in one particular adaptive trait, but in *many* traits simultaneously. This is a kind of racial 'essentialism' and, as Hull notes, a throwback to typological thinking (Mayr 1964, 1982; Hirschfeld 1998). Knowing someone's (biological) race, on this view, would permit one to make accurate predictions about a wide range of traits they possess – as Keita and Kittles (1997, p. 534) put it, that "visible human variation connotes fundamental deep differences within the species, which can be packaged into units of near-uniform individuals." This, however, will likely be impossible if there is little between-population genetic variation compared to variation within populations, and is in any event biologically unrealistic. Very few, if any, species have sub-populations that form groups of that sort, and the search for such groups seems to be a holdover of pre-Darwinian typological thinking (Futuyma 1998). So while the amount and distribution of genetic variation is largely irrelevant to the question of whether a species is divided into biologically significant races generally, it *is* relevant to the question of whether 'ordinary' conceptions of folk racial categories in humans have any biological support, and there is a broad consensus that the answer to this question is 'no.' Biology, it has been rightly noted many times, cannot underwrite the sort of racial concepts that have usually been applied to humans.

This conclusion, however, is often mistakenly thought to imply that there are no biologically significant human races at all, or at least that folk races must be utterly unrelated to biologically interesting human populations. While it seems clear that biologically meaningful races will not correspond particularly well to folk racial categories, this does not imply that folk racial categories are completely orthogonal to biologically meaningful racial categories. On the other hand, insofar as there is evidence that biologically significant human races do exist, that evidence points towards *most* biologically meaningful human races being quite a bit smaller (and far more numerous) than are folk races. Therefore, the idea that those groups picked out by folk races and those populations that form biological races will not, in general, correspond, is likely correct. And of course, as has already been noted, insofar as folk races are supposed to pick out populations that systematically differ from each other over a wide range of genetic and phenotypic measures, biology provides no support for the existence of such populations (and indeed, provides evidence that no such populations exist).

Confusion about these points is rampant, and far too much of the literature surrounding the biological basis, or lack thereof, of human races misunderstands these points. To take a trivial example, consider the controversy that has surrounded Entine's (2000) book *Taboo: Why Black Athletes Dominate Sports and Why We're Afraid to Talk About It*. While I agree with the critics who stress the dearth of hard data to support some of Entine's specific claims (e.g., Hoberman 2000), my main concern with the debate is that, as Michael Shermer (2000) notes, Entine's evidence, even taken at face value, doesn't support the contention that *blacks* dominate sports at all. Rather, even if all that Entine claims is true, the only conclusion that can be drawn is that smallish, particular populations generate the athletes that dominate *particular* sports. In other words, as even

Entine admits, "blacks" are not better runners – rather, some West African black populations produce more world-class sprinters than the proportion expected from their population size and the assumption of random distribution of athletic talents among humans would generate; similarly, Kenya (especially the Nandi region) produces far more than its share of great marathon runners. It is certainly possible that these regional differences in the production of top athletes reflect regional differences in athletic ability (or, better put, differences in physiology more generally), and it is even possible that these differences are the result of local adaptations to particular environmental (including perhaps long-term cultural) pressures. If this is so (and again, there is currently insufficient evidence to support such a claim), on an ecotypic conception of race, there would in fact be 'races' – and indeed races associated with athletic ability.

However, what one must remember is that the races in question do not, in this scenario, have much to do with folk races. If instead of phrasing the issue in terms of 'race', Entine had put it in terms of local adaptations within smaller populations (ecotypes), his book would likely have been seen as far less controversial. Further, the sorts of evidence necessary to support his conclusions would have been far more obvious as well. Just as one can gather evidence that particular ecotypes of the mustard-like weed *Arabidopsis* have the particular features they do in virtue of the particular selective pressures they've been under (e.g., Pigliucci 1998; Callahan and Pigliucci 2002), so too could one gather evidence in the case of human ecotypes (albeit with all the usual problems of ethical and practical restrictions on human experimentation: Kaplan 2002).

None of this should come as a surprise. The issue is not, as Gould et al. have been fond of claiming, that skin color is only 'skin deep,' but rather that 'skin color' is an ecologically important – but not a phylogenetically significant – trait. If skin color had

evolved only once, such that populations with different skin colors formed at least partially monophyletic populations, we would expect to find many other phenotypic differences associated with differences in skin color; some would be the result of different selective regimes, but some would no doubt be the result of drift or other evolutionary forces. The reason that skin color is not well correlated with other phenotypically important features is, at least in part, that skin colors evolved independently several times, and likely often evolved in populations that were not genetically isolated from other populations (Diamond 1997). Similar skin color therefore represents not shared ancestry, but rather similar selective pressures. The only thing that fair-skinned peoples share is that, at one time or another, their ancestors lived in an area with low levels of sunlight and ate a diet poor in vitamin D. As there were many such areas and many such times, fair skin says little or nothing about phylogeny. Occasionally, however, a folk racial category *will* happen to correspond to a biologically significant one. In areas where a folk racial population picked out, say, by skin color, is comprised entirely of recent immigrants from a particular area, it is at least likely that those people represent a biologically significant race (skin color, after all, is an adaptive trait!). We suspect that those people who claim that folk races are 'obviously' biologically distinct are thinking primarily of cases like these (see for example Sauer 1992 on the usefulness of the race concept in forensic anthropology).

But while skin color is not well correlated with other phenotypic traits of interest in humans, there is, despite Gould's claims (1996) to the contrary, no guarantee that particular populations of humans will not, due to particular features of their environment, share particular distributions of adaptive behavioral (including intellectual) traits, as opposed to simple physical traits. To the best of our knowledge, there is no evidence that

such populations exist, nor are there reasons to suppose that such populations *must* exist. Given the difficulty with testing hypotheses regarding the adaptive significance of behavioral tendencies in humans *simpliciter* (Lewontin 1998), the lack of evidence for behavioral (and/or intellectual) ecotypes in humans is not surprising. But it is intellectually dishonest to move from the lack of evidence for such differences to claiming that there is evidence for an absence of such differences, a move all too often made (oddly enough, both by Gould and by some of his opponents in "evolutionary psychology." See for example Tooby and Cosmides 1989, 1990; Gould 1996). The conviction that there are no such populations emerges not from research or principled arguments, but rather, I suspect, from fear that to even suggest the existence of such populations is to fall into the worst sort of racist thinking.

This is unfortunate. The study of the relationship between adaptive traits and expressed behaviors in humans is difficult enough without ideological posturing. Indeed, if there is any systematic variation in adaptive behavioral traits between human populations, discovering and studying such variation might provide one of the best entries into the study of human behavioral traits as adaptations more generally. Many of the most obvious problems with discovering and testing adaptive behavioral traits in humans are at least much less severe with respect to traits that vary systematically between human populations. Obviously, this is very speculative: again, there is no evidence that such populations exist, and if they do, discovering them and properly testing adaptive hypotheses concerning them may yet prove impossible, given our limited ability to test adaptive hypotheses regarding humans more generally. But looking for such variation does not commit one to racist thinking, partly because the populations displaying such variation would very likely *not* correspond closely to folk races.

Overlapping adaptations, clinal variations, and human races

Some authors have argued from the existence of sizable populations with phenotypes intermediate between those associated with particular folk races to the conclusion that there are no biologically significant human races (e.g., Keita and Kittles 1997). But this is just what we would expect to find if these ecotypic races are sometimes clinal in nature. A cline is a pattern of gradual variation of one or more characters, usually – but not exclusively – along a latitudinal or altitudinal range. Gene flow can be extensive through clines, as long as selective pressures are sufficient to maintain the genetic differences associated with adaptations to the ecologically important conditions (e.g., Munjal et al. 1997; Jordan et al. 2001). Given the wide geographical distribution of human populations over evolutionarily significant periods of time (Templeton 1999), it would be surprising if human populations did not show any clinal variation in ecologically important characteristics. The key points made above regarding ecotypes – that they may or may not be phylogenetic units and may or may not limit gene flow – also hold true for clinal variations, as does the observation that an individual may simultaneously be a member of multiple different ecotypes (as in multi-clinal variation).

Of course, this implies that insofar as we focus on an ecotype conception of race, there will not necessarily be a unique 'race' to which any given member of a population belongs. The same individual may in fact belong to a number of different ecotypic races, and/or be a member of one (or more) intermediate population within a series of clinal distributions. (Here my understanding of the ecotype concept follows, for example, Karan and Parkash 1998, and Joron et al. 1999 on multiple environmental clines and their

effects on the genetic architecture of natural populations.). However, this is hardly an unexpected complication in a discipline like biology, characterized by a high level of complexity of both the object of study and the conditions that induce variation in that object.

The problem posed by clines, then, is no different from that posed by any other gradual transition, and provides no reason to reject the possibility of the existence of biologically significant human races. Similar problems, after all, face any definition and practical application of the concept of species itself (Pigliucci in press); nonetheless, biologists have not given up the use of that most controversial biological category just yet (Howard and Berlocher 1998).

Ecotypes & folk races

As we have seen, insofar as biologically meaningful races are conceptualized as populations more like ecotypes than like incipient species, many of the arguments purporting to show that there are no human races miss their mark. While in non-human biology the term 'race' has been and is being used in a variety of fashions, the best way of making sense of systematic variation within the human species is likely to rely on the ecotypic conception of biological races. In this sense, there are likely human races (ecotypes) of biological interest. But, again, biology provides no support for the very strong, essentialist-style conception of 'race' that has, both historically and at present, underwritten racism (of both the individual and institutional varieties), and indeed, biology reveals that the assumptions underlying such a conception of race are false (at least in the human case).

This does not, of course, imply that our folk conception of race is not significant: while it does not pick out populations of *biological* interest, it does pick out populations of deep social and political interest. These populations do not, in fact, have many of the features they were historically supposed to have, but that does not prevent the application of the folk concept of race. Nor, I believe, should it. As long as the folk racial category to which one happens to belong is systematically related to other important aspects of one's life, there is obviously still a need to pay attention to race in formulating, e.g., social policy. And, it need hardly be said, it is. In the United States as well as in many other contemporary societies, one's race is systematically related to one's chances of acquiring most (if not all) important goods – everything from education to money to self-respect.

While it is valuable for biologists to note that the essentialist conception of human races has no support in biology whenever particular claims are made that seem predicated on such a conception (e.g., Herrnstein and Murray 1994, on race and intelligence), they should not fall into the trap of claiming that there is no systematic variation within human populations of interest to biologists. Studying human ecotypes could yield insights into our recent evolution, and perhaps shed increased light onto the history of migrations and gene flow. To some extent, this is already happening (Cavalli-Sforza et al. 1994; MacEachern 2000). However, the ambiguity affecting definitions of 'race' and the politically charged atmosphere surrounding discussions of race in humans has hampered research into these areas, a situation from which neither biology nor social policy can benefit.

Chapter 4—Species as family resemblance concepts: the (dis-)solution of the species problem?

"If the only tool you have is a hammer, you tend to see every problem as a nail." -Abraham Maslow (1908-1970).

The problem that never goes away

The so-called "species problem" is one of those topics of discussion among evolutionary biologists that has been present since before Darwin's publication of the aptly titled *Origin of Species* (Darwin himself referred to it as an already old problem), and will probably never go away. Furthermore, biologists have a schizophrenic attitude toward the whole issue: on the one hand, they tend to turn away in disgust when species concepts are brought up by colleagues, are the subject of papers, or are discussed at conferences. On the other hand, they simply cannot resist the temptation to offer graduate seminars on the topic and avidly reading anything that is published on the subject. In recent years two of the major journals of evolutionary biology have devoted several papers in special issues to the ever-burning question of exactly what species are (*Journal of Evolutionary Biology*, volume 14, 2001, pp. 889 ff.; and *Trends in Ecology and Evolution* volume 16, 2002, pp. 326 ff.). The situation -- I suggest -- can be traced to a particularly insiduous kind of "unexamined baggage," namely the unacknowledged philosophical assumption that it is possible to find an "essentialistic" definition of what species are; i.e., that there is a necessary and sufficient condition (or small set of

conditions) that -- when met -- allows one to determine if a biological entity represents a separate species. In fact, in this chapter I will attempt to show that the species problem cannot be resolved scientifically: it can only be "dissolved" philosophically. This may be one of the few clear situations in which philosophers can indeed "solve" an (apparently) scientific problem, contra to the suggestion by some scientists (Weinberg 1992) that such a feat can never be accomplished.

I will proceed by arguing two points: First, the reason why the species problem has not gone away is because it is not as much an empirical problem (contrary to what argued, for example, by Hey 2001), but rather one that has strong philosophical overtones. In fact, the philosophical literature on the definition of species is as extensive as the biological one, with some biologists contributing to both (e.g., van Valen 1988; Ridley 1989; De Queiroz 1992, 1999; Mayr 1996). This does not mean that empirical information is not relevant here, but rather that the problem represents a paradigmatic example of a philosophical question that requires empirical information (provided by science) to be settled, not of a scientific problem with unwelcome philosophical characteristics.

Second, the problem does in fact have a satisfying philosophical solution based on Wittgenstein's idea of "family resemblance" or cluster concepts (Wittgenstein 1953), as was proposed early on by Hull (1965) in a different context (he was interested in taxonomists' apparent historical inability of letting go of essentialist concepts of species). I wish to cast this solution in modern terms while naively and optimistically suggesting that there is no reason why this should not be the end of the controversy.

The controversy: the biological side

It is not my intention here to provide the reader with a history or comprehensive review of either the biological or philosophical literature on the species question. That would be a fascinating endeavor, but it would require a book length treatise in its own right, something I hope a philosopher and a biologist will eventually get together to write. However, in order to substantiate my two points it is helpful to have an idea of what has occurred so far in the debate.

From the perspective of the biological literature, besides the two recent special issues of major journals mentioned above, several books have been published (e.g., Otte and Endler 1989; Claridge et al. 1997; Howard and Berlocher 1998; Hey 2001), and of course countless articles have appeared. A brief history of the species concept can be found in Grant (1994), a recent empirical example of how easily different concepts of species don't mesh even for the same group of organisms is provided by Gleason et al. (1998), and Mayden (1997) provides a (incomplete) list and brief discussion of a whopping 21 species concepts!

Before investing any more time on this matter, one could reasonably wonder why bother. Part of the answer is obvious to any biologist: despite all the controversy surrounding their definition, species are considered a fundamental level of organization of the biological world (though this at time disputed), and as such they are pivotal to several fields of investigation as well as to practical applications of evolutionary biology. To mention a few, researchers interested in the study of the process of speciation (obviously), evolutionary geneticists, evolutionary ecologists, systematists, and conservation biologists all deal directly with questions for which—it would

seem—understanding what constitutes a species is of paramount importance in order to make progress.

Or is it? One can actually argue that progress in all of the above areas (including, paradoxically, our understanding of the process of speciation) has actually been achieved *despite* all the discussion on what species are (or, more charitably, independently of it). To paraphrase a famous American judge, many biologists seem to agree that—like the case of pornography—it is impossible to define species, but it is certainly feasible to recognize them when you see them (notwithstanding some taxonomic wrangling here and there). Going further, some (Noor 2002) even suggest that it is because of entanglement with such "semantic" (a pejorative term in biology) issues that evolutionary biology has not achieved the recognition as a science that, say, physics has. I actually think that particular problem is caused by much more fundamental factors, and in some respects is not a problem at all (Pigliucci 2002 and Chapter 1).

Be that as it may, it is instructive to go over the list of major species concepts that have emerged so far, in search of common themes. A more thorough discussion of individual species concepts and of their relationships to each other can be found in Mayden (1997). Commonalities among species concepts are actually not difficult to find: **Table 4.1** lists what one might consider the "top 9" species concepts (in alphabetical order, I am not taking sides here), together with brief definitions conveying the focus they put on specific biological aspects that are considered essential to each concept. We can see that there are broadly speaking only three factors entering into the equation: phylogenetic relationships, genetic continuity (sometimes specifically concerned with reproductive traits, sometimes more broadly defined) or similarity, and ecological similarities, broadly construed. Furthermore, one could argue that even these three themes

Table 4.1: Some of the major species concepts proposed so far (the list is incomplete and arranged in simple alphabetical order), short definitions, and some relevant references.

Species concept	*Brief definition*	*References*
Biological (including Recognition concept)	Groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups.	Mayr 1940
Cohesion	The most inclusive population of individuals having the potential for genetic and/or demographic exchangeability.	Templeton 1989
Ecological	A lineage which occupies an adaptive zone minimally different from that of any other lineage in its range, and which evolves separately from all lineages outside its	Van Valen 1976
Evolutionary	A lineage evolving separately from others and with its own unitary evolutionary role and tendencies.	Simpson 1961
Genetic	Group of organisms so constituted that a hereditary character of any of these organisms may be transmitted to a descendant of any other.	Simpson 1943
Morphological	The smallest groups that are consistently and persistently distinct, and distinguishable by ordinary means.	Cronquist 1978 (though proposed as early as the 1920s)
Phenetic	The level at which distinct phenetic clusters can be observed.	Sneath 1976
Phylogenetic (several variants)	The common feature of these concepts is the attempt to identify the smallest biological entities that are diagnosable and/or monophyletic.	E.g., Nixon and Wheeler 1990; McKitrick and Zinl 1988
Population-level lineages	Lineages identified at the level of evolving populations, before they become clades.	de Queiroz 1998, 1999

are obviously not independent, since genetics, phylogeny and ecology are tightly intertwined when it comes to determining the fate of any population of organisms. There are, it seems, more commonalities among the various species concepts than one might at first suspect.

Indeed, one author (de Queiroz 1998, 1999) claims to have solved the species problem by pointing out that no matter what specific characteristics one uses to study species in particular instances, all definitions share the fact that species are population-level lineages. I cannot here provide a detailed assessment of this proposal, but a few points deserve attention. First, while the population-lineage concept is often confused with variants of the phylogenetic concept, it is distinct enough from it because de Queiroz makes a distinction between lineages and clades, where the first may become an instance

of the latter, but not necessarily. Second, this proposal still looks for an *essence* defining species, an approach that I consider at the core of the species problem (see below), and that needs to be abandoned. Third, the "essence" proposed as common to species (i.e., being population-level lineages) is too broad for being useful for at least two reasons: on the one hand, it is not clear what sets aside species from population-level lineages that do not diverge enough to become species (if the criterion invoked is cladogenesis, than we fall back into the phylogenetic concept); on the other hand, many other characteristics are also necessary (but not sufficient) to talk about something being a species: for example, being comprised of living organisms, being subjected to a variety of evolutionary forces, and so on. The problem when one wishes to identify the *essence* of a concept is not to pick some necessary condition (there are often many available), but to identify a set of conditions that are necessary and sufficient. Being a population-level lineage is not sufficient for being a species.

The controversy: the philosophical side

As in the case of the biological literature, the philosophical outpour on the species problem is vast and complex, and I cannot possibly attempt here to do it justice. However, it is interesting to note the main threads, particularly as they partly overlap with the concerns of biologists (some of whom, as noticed above, have in fact participated actively to this side of the debate as well), while maintaining a characteristically philosophical flavor (biologists often do wonder what exactly do philosopher's write about biology anyway).

Like its biological counterpart, the philosophical discussion can best be

understood as focusing on a small number of basic themes, which are summarized in **Table 4.2** (which, again, is certainly not to be taken as an exhaustive list). First, in overlap with the biological literature, there are discussions of species concepts proposed by biologists, mostly focusing on the biological species concept and, to a lesser extent, on the many flavors of the phylogenetic species concept. These are the discussions that, understandably, have seen the major contributions by biologists to the philosophical literature (van Valen 1988; Ridley 1989; de Queiroz 1992; Mayr 1996). Ruse (1969) suggested early on that biologists prefer one well-defined species concept (and are therefore very unlikely to go for any of the pluralistic proposals briefly outlined below) because they are looking for *the* answer to the problem and because they want this answer to be embedded within general laws and be derived logically from first principles of biology. Giray (1976) produced one of the early criticisms of morphological species concepts and proposed a synthesis of different variants of the biological species concept that had been put forth by biologists up to that

Table 4.2: Some of the major themes dominating philosophical discussions of the species problem. The list is certainly not comprehensive, and the references mentioned are but a small portion of those available in the literature. However, the table should provide biologists with a feeling for the threads underlying the debate among philosophers, as well as an entry in the relevant literature.

Theme	Representative references
Discussions of species concepts proposed by biologists, mostly the biological and phylogenetic one.	Ruse 1969; Giray 1976; Mishler & Brandon 1987; Splitter 1988; Sterelny 1994; Horvath 1997.
What kinds of "things" are species? Are they individuals or sets? Are they natural kinds?	Giray 1976; Kitcher 1984; Splitter 1988; de Queiroz 1999.
Pluralism: do we need more than one concept because of inherent heterogeneity of purposes within the biological sciences?	Mishler and Donoghue 1982; Kitcher 1984; Dupré 1993.

point. Mishler and Brandon (1987) compared the biological and phylogenetic concepts in terms of the idea that species are individuals (which is fundamental to theories such as species selection: Stanley 1975; Gould 2002). Horvath (1997) examined the consequences of the phylogenetic species concept, while Sterelny (1994) is unusual (for a philosopher) in defending the evolutionary species concept as opposed to the two dominant ones (biological and phylogenetic).

On the matter of what sort of "things" species are, Kitcher (1984), contrary to Mishler and Brandon, proposed that species are not individuals, but rather sets, from which idea one can derive the limitations of the biological species concept. Splitter (1988) suggests that species are not natural kinds, an idea that—with an interesting twist that includes arguments concerning human brain's physiology—has been revisited by biologist Hey (2001) in a recent book.

We then come to pluralism, the suggestion—put forth in different fashions by Mishler and Donoghue (1982), Kitcher (1984), and Dupré (1993)—that the reason there are many species concepts is because biologists are legitimately interested in a heterogeneous group of questions, each of which requires logically independent, and equally valid, concepts of species. The idea is that there are at least two such components to biological endeavors, which Kitcher terms the historical / evolutionary and the structural / functional inquiries. These two (epistemologically equivalent) views of biological evolution translate for Dupré into a genealogical and an ecological conception of species, neither one of which has logical priority. In other words, if one wishes to emphasize historical relationships, one must adopt a phylogenetic concept of species and speciation; but one could also equally validly be interested in the ecology and function of organisms, in which case something along the lines of the biological or ecological species

concepts would serve the purpose better.

Why the problem has not gone away

I come now to the first of the two points I wish to articulate in this chapter: the species problem has not gone away for the all-important reason that it is not the sort of empirical problem that can be solved by biologists alone. This, however, is not to say that it cannot be solved or, worse, that there is no real problem (i.e., it's "just" a matter of semantics: Noor 2002). On the contrary, it is a prime example of a philosophical question that requires input from empirical science and that can provide a useful return to the practice of that science. I will defend this claim in two steps: first by briefly discussing the relationship between philosophy of science and science and why some scientists have a misguided conception of it; second, by detailing my case in the specific instance of the species problem.

The relationship between philosophy and science is a complex one, and entire books have been written on its turbulent nature on both sides of the cultural divide (e.g., Sorell 1991; Levitt 1998). Perhaps one of the best examples of the misgivings that scientists often have concerning philosophy can be found in an essay written by Nobel prize physicist Steven Weinberg (1992), aptly entitled "Against philosophy". In it, Weinberg accuses philosophers of not having contributed to the advancement of science in a single instance, and in fact of having positively retarded it in at least one case: the negative influence that logical positivism and its abhorrence of unobservable quantities allegedly had on the acceptance of quantum mechanical theory. This is not the place to mount a comprehensive criticism of Weinberg's position, but it is instructive to ask a

simple question: given that the principal aim of philosophy of science is to understand how science works, not to solve scientific puzzles, why *would* we expect philosophy to have contributed significantly to answering specific scientific questions? Another way to put it would be to ask, equally legitimately one would think, the question the other way around: how many cases can we think of in which science has actually solved a problem being debated in philosophy of science? It is interesting to ponder why the latter question sounds immediately specious, while its reverse seems quite plausible.

Let us now come to the specific problem posed by the debate on species concepts. Hey(2001b) has suggested that the main reason the problem is still unsettled is because "we [biologists] are not acting like scientists. We are acting like some philosophers." Besides betraying the same sort of questionable attitude toward nonscientific disciplines that characterizes Weinberg, Hey is saying that the knot of the problem lies in the fact that we simply don't have enough information to make up our mind on what species really are. Yet, he himself is incredulous at this suggestion, since in the same article he asks: "How could our knowledge, upon which the species debates have been built, be missing something?" How indeed: after all, we have been studying species and speciation for many decades now. Hey's answer (further elaborated in his book, Hey 2001) is in fact interesting, and it's probably part of the more comprehensive picture that we should draw as far as the species concept wars are concerned. According to Hey, the problem is that we evolved as pattern-recognition animals, but that this ability is far too crude to match the sophistication of nature when it comes down to the make-up of species. The result is that the species that we perceive as natural categories (a philosophical term) have in fact little correspondence with the real thing (he is adopting another philosophically contested assumption, as we have seen: that species are indeed

natural entities of some sort). I think Hey has a good point here, but he perhaps pushes it a bit too much as the principal explanation of why we have a species problem. At any rate, the presence or lack thereof of correspondence between natural categories as they are and as they are perceived by the human brain is an empirical question that awaits further study.

What I maintain, however, is that more empirical data will not be sufficient to solve the species question simply because this is by its nature a problem with strong philosophical overtones, not just a scientific one. Scientists have been able to function in practice very well without apparently agreeing on what the *essence* of a species is, even when it comes to empirical studies of species and speciation. I think this independence of scientific progress in an area from the solution of a "semantic" problem related to that same area is a hallmark of questions that are more philosophical than scientific in nature. Analogously, research on the neurobiology of consciousness is proceeding at a fast pace (Gazzaniga 2000) despite strong disagreements about what consciousness *really* is (Armstrong 1997). This line of reasoning, of course, opens up again the question of why should biologists care about philosophical questions (which I think they should, in a limited manner—see below), but right now my concern is with showing that biologists have simply being barking up the wrong tree in considering the species problem.

The (dis-)solution of the problem: species as family resemblance concepts

My second point is, as I mentioned at the beginning, that species are actually best seen as what Ludwig Wittgenstein (1953) saw as "family resemblance" concepts, something that was already realized among others by Hull (who, incidentally, is *both* a

biologist and a philosopher) in a different context back in the mid-60s (Hull 1965). The idea was endorsed by numerical taxonomists (Sneath and Sokal 1973), but has been rarely cited since the demise of that particular school of thought.

Let us briefly look at Wittgenstein's concept (**Figure 4.1**) and how it applies to the species problem (**Figure 4.2**). Wittgenstein was interested in the nature of human language and proposed the idea that humans engage in what he referred to as "language games," an iterative social negotiation of the meaning of terms made possible by the continuous interaction among individual human beings. At one point in his *Philosophical Investigations* he finally gets around addressing the obvious question of what he means by "language games." (He had used the idea up to that point to great effect, just as biologists use "species" in practice, without need for a formal definition.) As a way to answer this, Wittgenstein considers an example: what do we

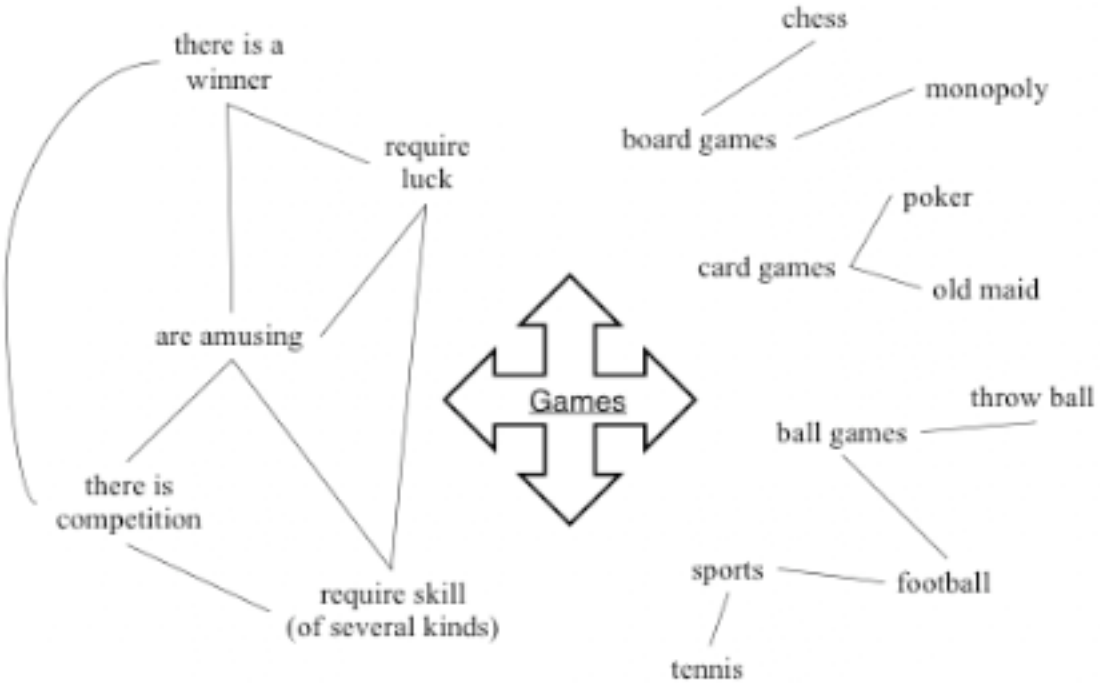

Figure 4.1: The idea of "game" as a family resemblance, or cluster, concept. On the left, som of the many characteristics we associate with games (the arrows signify that some of these characteristics are linked in the definition of particular games; not all possible relationships are visualized). On the right, types of games and some representative examples.

mean by a complex concept such as "game"? He immediately notices that, hard as one might try to, it is simply not possible to come up with a single, all-encompassing definition of what a game is. This is because things as disparate as board games, card games, ball games, and sports, to name a few instances of activities to which we attach the label of "game," do not share one essential quality. Rather, Wittgenstein suggests, there are many threads that crisscross the multidimensional linguistic landscape occupied by the concept of game. Some of these threads connect several types or instances of games, others connect additional instantiations of the word, and yet other threads run through some (but not all) examples of different classes of games. Game, in other words, is defined by a *cluster* of characteristics, or what Wittgenstein refers to as a "family resemblance" (in analogy to the very biological fact that members of a human family share some

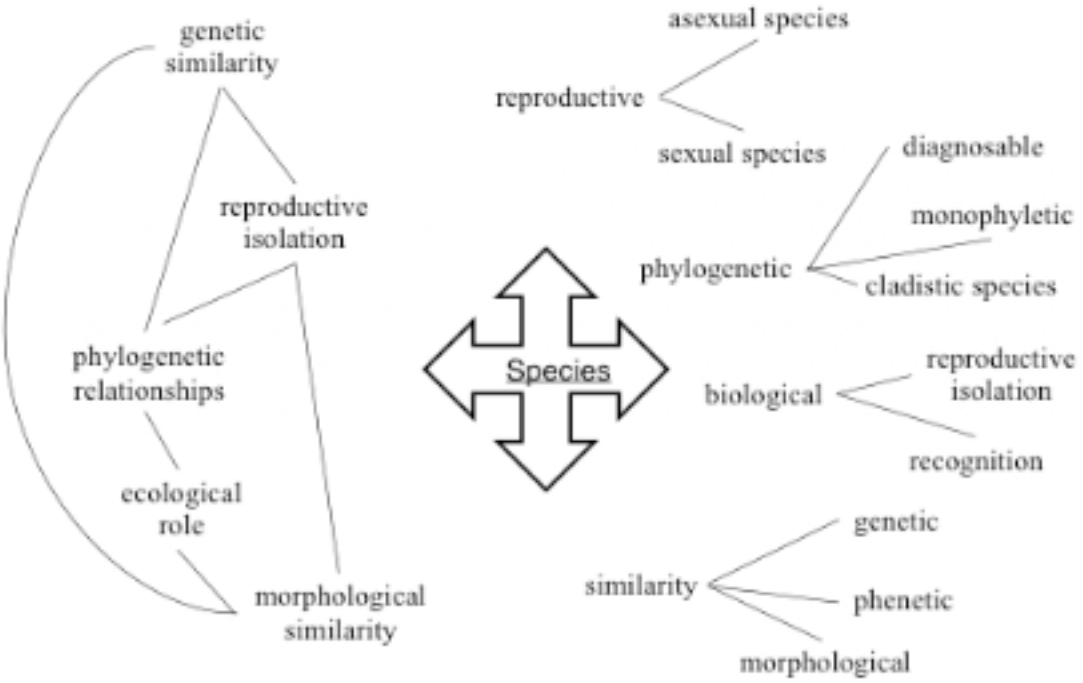

Figure 4.2: The idea of "species" as a family resemblance, or cluster, concept. On the left, some of the many characteristics we associate with species. The arrows represent the idea that more than one of these traits may enter into the appropriate characterization of any particular group of species (not all possible relationships are explicitly drawn). On the right, categories of species concepts based on different kinds of properties defining them.

characteristics or others, but that no single trait identifies, say, the Wittgenstein family as distinct from all others).

Interestingly, Wittgenstein directly addresses the practical problem posed by cluster concepts, a problem similar to the question of how we can possibly use the concept of species if we don't agree on what species are: "How should we explain to someone what a game is? I imagine that we should describe games to him, and we might add: 'This and similar things are called games.' And do we know any more about it ourselves? Is it only other people whom we cannot tell exactly what a game is?" (§69). Analogously, as biologists we *teach* our students what species are by example, exposing them to courses in systematic and natural history where they can see firsthand what the professionals consider the same or distinct species.

Furthermore, Wittgenstein goes on to explain why the above situation does not constitute a problem at all, since we can use the concept of game (or species) in practice very effectively: "But this is not ignorance. We do not know the boundaries because none have been drawn. … We can draw a boundary for a special purpose. Does it take that to make the concept usable? Not at all!" (§69). For the purposes of biological research, we can draw on one set of threads or another to work with particular species, depending on what taxonomic group we are considering. For example, in separating clonal or parthenogenetic taxa, the biological species concept's reliance on reproductive isolation is pretty much useless, while the same criterion is particularly appropriate for obligate outcrossers. In both cases there will be a phylogenetic component, but it will be easier to determine in the second than the first instance, and so on.

What, then, are species, and why do we care?

My suggestion here then is that species is a family resemblance concept whose underpinning is to be found in a series of characteristics such as phylogenetic relationships, genetic similarity, reproductive compatibility, and ecological characteristics. These traits take on more or less relevance depending on the specific group one is interested in, as a function of the particular biology of that group.

Adopting Wittgenstein's approach (dis-)solves several problems at once, both on the biological and philosophical side of the species problem. To wit:

- We can now move from our historical obsession with Platonic "essentialist" views of species to a cluster concept view that is more nuanced and realistic. This is a philosophical shift, although it is informed by the wealth of empirical information we have on species: our understanding of cluster concepts depends on our experience of them.
- Wittgenstein's family resemblance idea, when applied to species, is actually compatible with Hey's (2001, 2001a) suggestion, discussed above, that part of the problem may be in a mismatch between the categories recognized by the human brain (and hence language, Wittgenstein's main focus of interest) and whatever natural categories are really "out there."
- Wittgenstein's suggestion (and Simon's 1969 elaboration of it) that we may draw boundaries on subsets of family resemblance concepts for practical purposes at once erases the need for endless squabbles among biologists on what *the* best species concept is. The concept is fluid (but not arbitrary!) and gains enough

> flexibility to be applicable to the variety of real biological cases, which should be a welcome feature for whoever is accustomed to appreciate the extent of variation in the biological world. While scientists tend to be uncomfortable with fuzzy concepts, this is simply a philosophical prejudice: just because we cannot draw a precise line somewhere, it doesn't mean that there are no distinctions and that everything can be accommodated. Cluster concepts are not at all about abandoning the search for definitions, but they do force our mind to be less rigid about our solutions.

It is important to note that what I am suggesting here is conceptually very different from the idea of pluralism of species concepts advocated by Mishler, Donoghue, Kitcher, and Dupré (toward which, however, I am also sympathetic, I think without falling into a contradiction with my present proposal). The pluralist suggestion is that there are equally legitimate, *conceptually independent*, species concepts that can be used depending on the interest of the investigator. So, if a biologist's focus is on phylogenetic relationships, then a species concept that involves phylogeny is useful. If, however, the interest veers toward functional ecology, then a mixture of biological and ecological species concepts will be more appropriate. What I am saying here, on the other hand, is that species represent one large cluster of natural entities, quite independently of the interests of human observers. This cluster, however, is a loose one, with its members connected by a dense series of threads, possibly none of which go through every single instantiation of the concept. Among the species concepts listed in Table 4.1, Templeton's (1989; Templeton et al. 2000) "cohesion" concept comes closest to the idea of family resemblance, especially if a phylogenetic component is appropriately factored into it.

Biologists can benefit from the adoption of a cluster concept of species in a variety of ways: First, they can stop wasting their time by trying to empirically solve a problem that has philosophical components that cannot be settled by the accumulation of new data. Second, however, they should not therefore draw the conclusion that the solution of the problem is irrelevant to their aims and "just" a matter for philosophers to quibble about. On the contrary, they need to acknowledge and examine what amounts to their "bad philosophical habits" and be more cognizant of the implications of such habits.

Indeed, the latter two points represent in my opinion the best model of the relationship between science and philosophy: on the one hand some philosophical problems do require empirical input, and hence have to use information from science; on the other hand, science proceeds on the basis of philosophical assumptions, and at least occasionally it pays to be aware of such assumptions. In practice, biologists will go about their business of identifying and classifying species in the usual way, but thanks to philosophy they should now feel liberated of a (philosophical) burden that they are not trained to deal with.

Part II - Bad habits

The second part of this work deals with a slightly different form of philosophical baggage from what we have seen so far, although, as I mentioned above, this tripartition is meant to provide a convenient way to organize our thoughts on such matters, not as an indication of deep conceptual differences among the biases that accompany scientific practice. The following are two examples of what I think are misleading mental habits that biologists often display. I refer to them as "habits" because to me -- when I put on my biologist's hat -- they feel more like the sort of mental phenomenon associated with one's inability to do what is right, despite that individual's knowledge that what they are currently engaging in is not the best behavior. So, in the same way as I *know* that I should not slouch on the couch, channel surf and eat junk food, I think most scientists realize that certain ways of approaching problems are not quite the best that are available. The "health food-meaningful activity" equivalent does exist, and most people are aware of it. But, for a variety of reasons -- including convenience, access to resources, and last but not least the continuous pressure to "produce" publications and grants (and perhaps even occasional intellectual laziness) -- most of us scientists fall into these habits and sporadically even try to rationalize them to our advantage.

The first chapter in this group deals with the widespread contention that certain ways of conducting hypothesis testing and statistical analyses must be used in quantitative biology. As I will argue on the basis of work done primarily by social scientists and statisticians, it turns out that by now a very compelling argument has been articulated to the effect that common concepts in hypothesis testing, such as that of

"null hypothesis" and of "p-values" associated with the probability of specific experimental outcomes to be due to chance, should be abandoned as highly uninformative. Furthermore, there are obvious alternatives to be embraced that build on commonly available tools of statistical analysis. And yet, a large portion of the scientific community seems to be blissfully unaware of all of this and continues to do business as usual,[5] even though this may diminish the value of much that is currently published in the quantitative sciences.

The second chapter explores a topic that has been very close to my own research as a practicing scientist (Pigliucci 2001b): the issue of genotype-environment interactions or, as philosophers would put it, of nature vs. nurture. The problem here is the persistent inclination to underestimate the importance of interaction terms when studying genotype-environment questions. In particular, the crucial issue concerns the application of biological research in this area to the case of human beings, especially on the matter of cognitive traits such as intelligence (whatever that is) and IQ scores (whatever they measure). Again, I will argue that bad mental habits have settled in here: I will show that we should know better than to engage in endless discussions for which there is not an iota of pertinent or satisfactory empirical evidence, or to keep using concepts such as heritability as if modern research had not yet shown that they are quite different from what we thought they were. All of this is certainly not because the genotype-environment question cannot be settled. Indeed, great progress has been made in this field on plant and animal experimental systems. But alas, in the case of humans we are barred from conducting the crucial experiments for both ethical and logistical reasons. To paraphrase Lewontin (1998): it may be very interesting to find out how genotype-environment

[5] Worse yet, as I have experienced directly with several colleagues when confronted on this matter, the common reaction is to dismiss the whole issue as "well known," even though they demonstrably don't *act* on such allegedly settled matter.

interactions play in the case of human behavior; but (at least at the moment) we can't, though luck, and let us spend our energy and resources pursuing enterprises that are likely to be more successful.

Chapter 5-On hypothesis testing, a brief guide for the reasonably skeptical

"The great tragedy of science: the slaying of a beautiful hypothesis by an ugly fact." -Thomas Henry Huxley, 1825-1895.

One of the common insider's jokes in quantitative biology is that all biologists who have ever used statistical analyses think of themselves as experts in statistics. Furthermore, they become zealous about it, pounding on colleagues who don't stick to the established canons of the discipline. Conversely, statisticians with whom I have interacted seem by and large more casual about their own profession, thinking of it as providing nothing more than a convenient toolbox to explore data and to more or less rigorously aid in hypothesis testing. Indeed, a debate has been going on for decades about the usefulness of the two basic tools of statistical analyses in the quantitative sciences: null hypotheses and p-values. The suggestion of this chapter is that most practicing scientists have acquired bad habits of mind about both instruments. Many statisticians and some scientists who like to think outside the box claim that these habits are hindering scientific progress and propagating themselves into the mindset of new generations of researchers, and that we would be better off correcting them. This debate has the potential to change the way peer review and editorial decisions about publications and grant proposals are conducted and it represents a prime example of how habits of mind, once entrenched, are very difficult to correct or dislodge from the scientific culture.

Quantitative biologists simply take for granted the value of null hypotheses and

p-values as key concepts of their statistical practice, which they have been taught in graduate school and see used almost universally in technical journals. And yet, the debate on the usefulness of such concepts has intensified to the point of influencing the actual practice of researchers in fields such as statistics, psychology, sociology and some applied disciplines heavily relying on decision-making procedures (Cronbach 1975; Matloff 1991; Loftus 1993; Cohen 1994; Gregson 1997; Dixon and O'Reilly 1999; Anderson et al. 2000).

Scientific vs. statistical inference: the problem with null hypotheses

For the purposes of this chapter, I will define as the "standard" or classical approach to statistical analysis a mixture of Fisher's and of Pearson and Neyman's concepts of hypothesis testing, as it is in fact presented in most statistical textbooks for biologists (e.g., Sokal and Rohlf 1995). The main problem with the classical approach is rooted in the fact that it involves testing the probability of the observed data given an often unsatisfactory "null" hypothesis (Matloff 1991; Dixon and O'Reilly 1999; Anderson et al. 2000). That is, biologists (and social scientists) typically set up statistical tests to tell them if, for example, the correlation coefficient between flowering time and leaf production in a sample of plants is zero (null hypothesis) or different from zero (any other value). They then collect an appropriate set of data and ask what was the probability of getting that data if the null hypothesis were correct. If such probability is low, they "reject" the null hypothesis and "accept" the alternative hypothesis.

There are many problems with this procedure, which are summarized in the following list:

1. What scientists are actually interested in determining is the probability of a series of alternative hypotheses given the data, not vice-versa (i.e., scientific inference is about comparing hypotheses on the basis of data, not the other way around—See Chapter 1 and references therein on the Bayesian approach).
2. The null hypotheses tested by standard tests based on p-value are often not the best formulation of what is interesting to the researcher.
3. The standard approach does not take into account any prior knowledge one may have about the system (like the expectation that not only there will be a correlation between flowering time and number of leaves, but that it will be positive and linear, for example). If science is to be a progressive enterprise, then surely scientists don't start each project with a blank slate (again, see Chapter 1 on Bayesianism).
4. The so-called "alternative" hypothesis is in reality an infinite family of hypotheses, and concluding that the result is not zero is not very informative.
5. The best science proceeds by considering several alternative hypotheses (as discussed in Chapter 1), not just two. The more alternatives, and the better defined, the more scientists can design informative experiments and make progress.

A more formal way to think about the problem posed by the use of null hypotheses can be based on the concept of conditional probability: scientists are interested in P(H|D) (the probability of a hypothesis given the data), not in P(D|H) (the

converse), which is closer to what is actually tested in the standard approach. One of the most common fallacies in statistical reasoning is to assume that the two are related: if the probability of the data given the null hypothesis is low (i.e., we reject the null), then the probability of the null given the data must also be low. It can be shown both mathematically and by example that this is not the case, because one could pick a very unrealistic alternative hypothesis, or because the sample size may be inadequate (leading to what is known as a type I error, the incorrect rejection of the null).

It follows from all this that null hypotheses as usually intended are really a straw man: their rejection is not very informative, and a failure to reject them is even less so. Of course, most standard parametric statistics actually allow one to use any hypothesis in the test, not just H_0, though researchers rarely take advantage of this. For instance, one can take the equation for the p-value of a correlation coefficient (using the t-distribution), replace the zero with any number one has good reasons to believe should be used as a starting point, and get a p-value for that hypothesis. For example, a common problem in quantitative genetics is to determine if the correlation between two variance-covariance matrices is or is not one (not zero, e.g., Steppan et al. 2002). If we insist in using null hypotheses, at least we should think carefully about which ones are biologically informative.

The twin problem: p-values

P-values in the standard approach are also often misinterpreted in a variety of ways. The most important thing to understand about this is that the p-value is *neither* a measure of $P(H|D)$ (what's interesting, as discussed above) *nor* (exactly) of $P(D|H_0)$

(where H_0 here is the null hypothesis). In fact, it is the *cumulative* probability of the point null hypothesis and of all values more extreme than the one observed (Sokal and Rohlf 1995). This is related to $P(D|H_0)$, but clearly is not the same.

Furthermore, it is absolutely incorrect to interpret a p-value as an estimate of the strength of the effect under investigation (as in "given the small p-value the results are highly significant": significant in what sense, statistically or scientifically? See Dixon and O'Reilly 1999 for the difference between the two). This is because for a fixed effect size (large or small), the p-value is a function of the sample size. One can repeat the experiment with a much larger sample size, obtain a much smaller p-value associated with the null hypothesis, and yet the effect size (e.g., the strength of a correlation coefficient) can be unchanged (Loftus 1993; Gregson 1997).

Indeed, the null hypothesis is almost certain to be rejected given a large enough sample size. This is because the null hypothesis is stated very precisely (e.g., $r_{ft,ln} = 0$, the correlation between flowering time and number of leaves is exactly zero). In reality, what scientists mean to test is whether the correlation (or whatever other statistic they are interested in) is far enough from zero to be scientifically interesting (**Figure 5.1**). Another way to think about this problem is that the effect of any interesting variable on the system is highly unlikely to be exactly zero (because of small, often unaccounted for, systematic deviations), but that this doesn't imply an effect-size worth thinking about. Conversely, small and statistically not significant effect sizes can be biologically very important, as in the common encountered case of small selection coefficients (Kingsolver et al. 2001) that can nevertheless yield major evolutionary change if sustained over a long period of time. The problem in the latter case is that chance is lumpy (Abelson 1995, p. 21), and it is too easy to spot patterns that are not really there. Researchers have

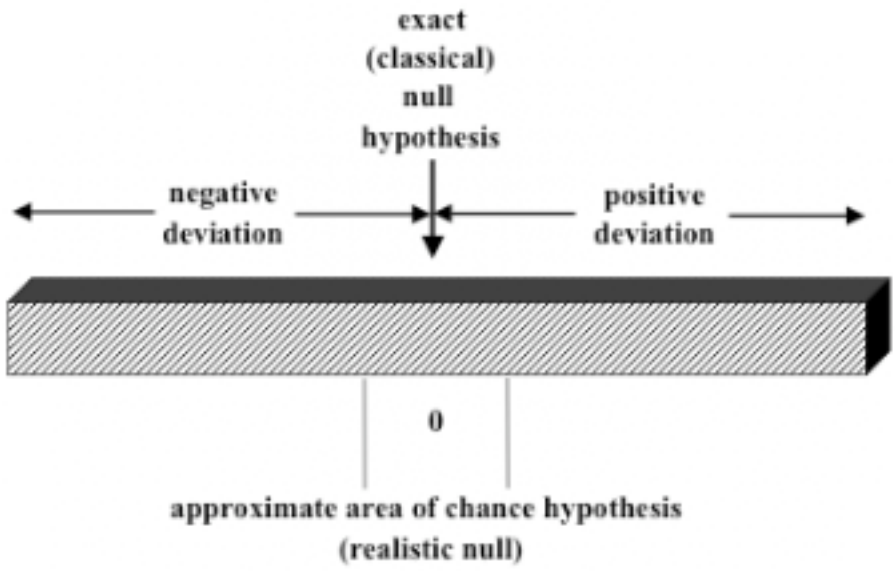

Figure 5.1. The structure of hypothesis testing in a hypothetical experiment. What one is really interested in is not the rejection of the exact null hypothesis (which is what standard statistical tests do), but in seeing if the outcome of the experiment is far enough from that point to seriously raise the possibility of a systematic effect of biological interest. The problem, of course, is that "far enough" is a subjective concept and requires a judgment call on the part of the scientist. As much as we would like it to be otherwise, this is rather unavoidable.

therefore to rely on their knowledge of the system, not on ready-made statistical cutoff points; in particular, if the small deviations can be confirmed by repeated experiments, and if they are in the direction predicted by suitably realistic biological hypotheses, *then* they become interesting regardless of the p-values that accompany them.

Of course, the difficult question in evaluating a biologically realistic null hypothesis is how far is "far enough" from the point null expectation? This question, unfortunately, does not have a ready made, "objective" answer and it calls for reasoned judgment on the scientist's part. Then again, simply calculating a p-value does not provide researchers with an objective answer either, because they still have to make an arbitrary judgment as to the "level of significance" of said p-value. If one takes a p-value of 0.05 as the dividing line between accepting and rejecting a hypothesis, or a likelihood

ratio (which will shortly be discussed in more details) of 10:1 between two alternative models, or a Bayes factor (which is related to the likelihood ratio: Dixon and O'Reilly 1999) of X or Y, all of these are arbitrary points along a continuum, and need constantly to be interpreted as such. Probabilities and likelihoods are continuous measures by their very nature, despite the persistent human tendency to wish to categorize things into true or false. In essence, the search for an objective decision-making algorithm (which inspired Fisher and has led most of practicing biologists to be more Fisherian than he was) is a chimera to be abandoned, the sooner the better (a moment of reflection will confirm that p=0.051 and p=0.049 are not that different from each other. Why then call the first one "borderline significant," as it is often done? If 0.051 is borderline significant, then so is 0.049). Decisions need to be made by humans while considering a broad range of factors. Statistics can only quantify some of these factors and aid in the decision-making procedure. They are no panacea, regardless of the fact that many practicing biologists have slipped into the bad habit of thinking of them as such.

The philosophical underpinnings of the standard approach

Fisher, and Neyman and Pearson developed what I call here the "standard" approach in statistics (though Bayesian analyses, which ameliorate several of the above mentioned problems, were actually popular in statistics for decades prior to Fisher's work). Fisher constructed a system of statistical inference that is echoed by the insights of philosopher of science Karl Popper. According to Popper (1968), hypotheses can never be proven; they can only be disproven (falsified). This is because no matter what the amount of evidence in favor of a hypothesis is, there is always the possibility that an

as yet unknown alternative hypothesis will explain the same evidence and then some. Popper thought that, on the other hand, once falsified, a theory could not possibly be revived. This is the idea behind the use of a null hypothesis: one is attempting to falsify a given possibility in order to move on and consider the others.

One problem with this should already be evident from my discussion above: even according to Popper the rejection of one hypothesis does not imply the acceptance of another (and Fisher did not in fact advocate the latter, though this interpretation has entered common usage). Furthermore, Popper suggested that the most interesting hypotheses to falsify are the most daring ones (because they are the most informative), while null hypotheses tend to be trivial and uninformative (in Popper's terminology, they don't "stick their neck out" far enough). Finally, philosophers of science have long moved past what is sometimes referred to as "naïve falsificationism" (Longino 1990; Schick 2000), and several of them have come to the conclusion that the scientific method is really a continuous competition among alternative models, with some gaining and some losing, depending on the available evidence (a very Bayesian view of science: Howson and Urbach 1991, see also Chapter 1). All of this notwithstanding, scientists begin their undergraduate and graduate students' training with a Popperian view of science, a habit that is then difficult to shed.

As philosophers of science have long realized, an infinite number of hypotheses can be constructed to explain (fit) a certain data set (the so-called problem of the under-determination of the theory by the data: Okasha 2000). In general, no matter what statistical approach one uses (standard, Bayesian, or the maximum likelihood discussed below), the resulting output can only be taken as a description of the data, not as a decision-making procedure. This is because there are always alternative explanations to be

considered and discussed, from the fact that the data may not be sufficient to be informative, or maybe are not reliable enough, to the possibility of problems with the experimental setup or with the instruments of measurements, to the likely action of factors not explicitly considered by the researcher up to that moment (see Figure 1.1).

Alternatives to the standard approach

Philosophers are often accused of not helping science solving its own problems (Weinberg 1992; but see Wilkins 2001), limiting themselves to sit in judgment of what scientists do (allegedly) wrong. This section is therefore meant as a recapitulation of the several alternatives to the "bad habits" of null hypotheses and p-values that have been identified in the course of this debate, and which are increasingly changing publications, and especially ways of thinking, mostly in the social sciences (Matloff 1991; Loftus 1993; Gregson 1997; Anderson et al. 2000).

The method of competitive explanations

One simple change scientists can enact immediately is to abandon the practice of pitting an often-uninformative null hypothesis against every other possible alternative and let the latter highly heterogeneous category easily win the day (especially if the sample size is large). As argued in Chapter 1, interesting and informative scientific papers are those that consider two (or more) plausible hypotheses and can actually provide empirical support for an increase or decrease in likelihood of one or more of said hypotheses. This was realized as early as the end of the 19^{th} century by Chamberlain

(1897) and forcefully reiterated in a recent book on ecological investigations by Hilborn and Mangel (1997).

To insist on publishing papers where the results are presented as if the authors were kicking a very weak opponent (the null hypothesis) makes for the appearance of simplistic thinking and more difficult advancement in evolutionary and ecological disciplines (Pigliucci 2002). In reality, most scientists do realize the distinction between statistical and scientific inference, and as a result write their papers in a rather schizophrenic fashion: the Introduction and Discussion are usually framed in terms of a series of alternative hypotheses whose relative merits are judged based on previous knowledge (other papers) and on the current results. However, these sections of the paper sandwich a rather different middle one that presents the results largely as endless tables of p-values purportedly rejecting a series of often uninteresting null hypotheses.

Thinking in terms of several alternative hypotheses is not only more natural and more effective, but easily avoids a typical occupational hazard of being a scientist, what Jacques Monod (1971) referred to as falling in love with one's own hypothesis. If one considers several, viable and interesting possibilities, one can avert more or less unconscious biases in favor of the "alternative" hypothesis. Furthermore, this practice might ameliorate the known problem of the reduced publication of "non significant" results which, even though may be highly informative, tend to be underrepresented in the literature (the so-called "file drawer effect": Moller and Jennions 2001), because editors do not want to waste precious journal space to publish a negative conclusion. Again, this is a result of the widespread Popperian attitude adopted by many practicing scientists, since according to Popper the *only* way to make progress in science is by rejecting hypotheses, never by confirming them; accordingly, failing to reject a null allegedly does

not advance science. It does, if the null is interesting.

An obvious alternative to the standard statistical approach to hypothesis testing has been mentioned already, and is discussed to some length in Chapter 1: Bayesian analysis. Essentially, Bayesian analysis asks the right question (what is the probability of the hypothesis given the data?), and attempts to answer it by taking into consideration not only the actual data obtained during a given experiment, but also whatever knowledge was available beforehand, summarized in the form of priors. I will not get into the ongoing discussion concerning the choice of objective or subjective priors: suffice it to say that Bayesian analyses can be carried out with both subjective and objective priors, and that in the first case the results—in the long run—tend to converge, while in the second case one gets good answers much faster than if no priors were available at all (Howson and Urbach 1991; Jefferys and Berger 1992; Malakoff 1999).

Sometimes, researchers prefer to compare hypotheses by means of what is known as Bayes' ratio:

$$B = \frac{P(D \mid H_1)}{P(D \mid H_2)}$$

That is to say, *B* is the ratio of the likelihood of the data under the first hypothesis and the likelihood of the data under the second hypothesis, and it does not involve the priors of either hypothesis. This ratio is also called the weighted likelihood ratio, and it is similar (indeed, in some cases identical) to the more familiar (and simpler to calculate) likelihood

ratio discussed below (Dixon and O'Reilli 1999).[6]

The Bayesian approach is becoming more popular within evolutionary biology, having been applied to ecological modeling (Hilborn and Mangel 1997), selection analyses (Rudge 1998), quantitative genetics (Shoemaker et al. 1999), phylogenetic studies (Sinsheimer et al. 1996; Huelsenbeck et al. 2000), and conservation biology (Wade 2000). The main problem with the Bayesian framework for most practicing scientists is that there is still little software available, and what is possible to find is rather difficult to use and usually tailored to specific applications. This situation obviously makes it more difficult to "kick the habit" of using the standard approach (while many scientists may concede the conceptual points made here, they will maintain that they need to deal with real problems of data analysis, and that they cannot afford to "go philosophical" when a grant proposal's deadline is due). However, things are likely to change in the near future, given that Bayesian methods are now being devised for many applications.

Simply graphing

The next alternative to be considered as helpful in "kicking the habit" is as simple and straightforward as it can be imagined: just plot the data, together with their confidence intervals. As Loftus (1993) put it, "a picture is worth a thousand p-values." There is no need to be skeptical about the power of this approach just because it is obvious, and—for the reasons mentioned above—it is simply not true that a graph is more "subjective" than

6 For the sake of completeness, I should add that when one considers an idealized population, the p-value from least squares model fitting is the same as that from the likelihood approach and the solutions are identical. Least squares is in fact a maximum likelihood approach, although it is limited in its applicability because it relies on the assumption of an idealized population and performs best with large sample sizes.

a p-value: note that graphs are quantitative tools and remember that the threshold levels for p-values are entirely arbitrary.

Loftus (1993) discusses a couple of examples from the primary literature from which it is clear that a graph of means and standard errors gives us all the information that a p-value provides, and then some. Moreover, graphs of means and their associated measures of dispersion provide scientists with a visual estimate of effect sizes, with an idea of the power of the analysis (which is inversely proportional to the standard errors), and with an immediate understanding of the patterns identified by the data (other measures of centrality and dispersion can be used if the data are not approximately normally distributed). Loftus points out that these quantities are completely invisible in a table of p-values, where all we are told is that certain null hypotheses are to be rejected.

The likelihood ratio

A more sophisticated alternative to the standard statistical habit combines the insights of the Bayesian approach (*sans* the difficulties related to the estimation of priors) and the simplicity of readily available statistical approaches (and computer packages): maximum likelihood ratios. A maximum likelihood ratio is a quantity that compares the likelihood of the data based on one model with the likelihood of the data based on a second, competing model (comparisons can also be made among multiple models, pairwise or cumulatively). The ratio provides a measure of the relative match of models and data, with most sources suggesting that a ratio of 10:1 is roughly equivalent to a p-value of 0.05, i.e., it should be the minimum ratio at which one should begin considering one model better than another one. However, remember the warning above about the fact

that any such threshold is arbitrary and that there is no shortcut that avoids using one's own judgment.[7]

The maximum likelihood ratio is conceptually(and mathematically: see Dixon and O'Reilly 1999) related to the Bayesian approach, but it is also conveniently close to the familiar standard statistics, which makes it a powerful practical alternative, a way to ease oneself out of the habit, so to speak. There are at least three ways to calculate likelihood ratios from standard statistics, as I summarize below. I will provide a simple example of these calculations for illustrative purposes, but a more detailed treatment (and additional worked examples) can be found in Dixon and O'Reilly (1999).

Likelihood ratios from standard deviations: One way to calculate the maximum likelihood ratio for comparing two models is to obtain the ratio of the standard deviations derived from the two models and raise it to the power of the sample size, n:

$$\Lambda = \left(\frac{s_1}{s_2} \right)^n$$

Where s_1 and s_2 are the maximum likelihood standard deviations based on the two models.[8] For example, if model 1 has an associated standard deviation (SD) of 0.65, model 2 has a SD of 0.58, and there are n=40 observations, then the ML ratio between model 1 and model 2 is:

[7] In fact, one could even calculate p-values to go with a likelihood ratio, since this measure tends to be distributed as a ?2, but that would defeat the purpose of this discussion.
[8] Note that the ML standard deviation is calculated by dividing the sum of squares associated with a given model by n-1 instead of n.

$$\Lambda = \left(\frac{0.65}{0.58}\right)^{40} = 95.4$$

This means that the data fit the second model (the one at the denominator) roughly 95 times better then they fit the first one. One should definitely prefer the second model.

Likelihood ratios from sums of squares (i.e., from ANOVA or similar tables): It is even easier to calculate the likelihood ratio from a standard table of sums of squares like those produced by any commercial software that carries out general linear model calculations. In this case the formula is:

$$\Lambda = \left(\frac{SSE_1}{SSE_2}\right)^{n/2}$$

Where SSE is the error sum of squares calculated after fitting either model, and *n* is again the sample size. The SSE's can be obtained either summing up the SS of various sources of variance within a single model, according to the corresponding sub-hypotheses, or—which may be simpler—by rerunning the analysis using each model and looking up each SSE value.

Likelihood ratios from coefficients of determination: An even simpler method is to run each model separately and calculate the R^2 (i.e., the amount of variance explained by the model), something that again is part of the standard output of any statistical package. The ratio between models then becomes:

$$\Lambda = \left(\frac{1 - R_1^2}{1 - R_2^2} \right)^{n/2}$$

All three of these methods give, of course, exactly the same numerical results. Dixon and O'Reilly (1999) also discuss the fact that this approach can be extended to repeated measures models, mixed and nested ANOVAs, and even to cases in which the data are not completely independent. In other words, likelihood ratios between competing hypotheses can be calculated for most problems that usually fall under the rubric of general linear modeling, which includes the majority of practical applications for quantitative biologists and social scientists. There really is no excuse, in terms of conceptual sensibility, ease of calculation, and flexibility, not to present results in terms of maximum likelihood ratios among competing models (together with the appropriate means and confidence intervals plots, of course).

A few important caveats apply to likelihood ratios. First, a likelihood ratio is not a measure of the posterior probabilities of the models (the Bayesian interpretation). This is because, as the attentive reader will have noticed, likelihood ratios calculate the probability of the data given a certain model, not vice versa. However, a quasi-Bayesian interpretation of the likelihood ratio is in fact defensible (Dixon and O'Reilly 1999), because the two are equivalent if the priors for the two models are taken to be equivalent (i.e., one has no a priori reasons to prefer one model over the other) and if the models are of comparable complexity (measured by the number of parameters they employ).

Second, unlike the situation in a Bayesian analysis, the models being considered

need not be exhaustive of all possible models: they are just the particular subset that the experimenter wishes to test (or can think of testing). This means that just because one model is much better than another, one cannot exclude the possibility that the winning model is not adequate for the task at hand: a much better explanation may exist but not have been thought of (yet?). This, however, is just part of life for a scientist and it's not a limitation of the likelihood ratio approach. It may seem at first as if model fitting is a hopeless enterprise given the potentially large number of models one should compare with each other. However, most standard statistical packages come equipped with model fitting algorithms using maximum likelihood or restricted maximum likelihood approaches to iteratively explore the parameter space and aid in finding the best model. More radically, Shipley (2000) has suggested an intriguing new approach to model selection based on fundamental properties of statistical analyses such as path analysis and structural equation modeling (see Magwene 2001 for an application to evolutionary questions).

Third, it is of course possible that the data will not strongly favor one model over another. This is important, because it means that either the data do not discriminate between the models or that both (or more) models are inadequate representations of the causal structure of the phenomena being studied. The advantage of the likelihood ratio is that in this situation one is not forced to accept a "null" hypothesis based on weak evidence.

One thing perhaps needs to be reiterated, although it should be clear after a moment's reflection: *any* statistical analysis is, in fact, an exercise in model fitting, followed by some approach to gain a measure of how well the model(s) actually did fit the available data. This is true if one uses or not null hypotheses, p-values, maximum

likelihood, or Bayesian approaches implemented on regression analyses, analyses of variance, or a host of other exploratory and confirmatory statistical techniques. The approaches advocated in this chapter simply make this more obvious and explicit, thereby helping the researcher to think more clearly through the problem at hand.

Taking into account the complexity of a hypothesis: the Akaike Information Criterion

Yet another approach to the solution of the same problems concerning model selection and hypothesis testing—and one that is favored by an increasing number of statisticians—is provided by the Akaike Information Criterion (AIC). This is a measure based on information theory and derived from the concept of entropy in physics (Anderson et al. 2000). Briefly, AIC measures the fit of the data to a given model, when the model is penalized in proportion to the number of parameters it employs. This is sensible because other things being equal a model with more parameters will always fit the data better than one with fewer parameters, but the improved fit will not always be scientifically relevant (again that all-important difference between statistical and scientific inference).

When one engages in model fitting using standard sums of squares (as is often the case in the biological literature, again encompassing all problems that can be treated within the framework of the general linear model), the AIC for a given model can be computed simply as:

$$AIC = n \cdot \ln(\frac{RSS}{n}) + 2K$$

Where *n* is the sample size, *ln* is the natural logarithm, *RSS* is the residual sum of squares for that model, and *K* is the number of parameters (the degrees of freedom taken up by the model).[9]

Once AIC has been calculated for each alternative model, one can rank models by computing:

$$\Delta_i = AIC_i - \min(AIC)$$

Where *i* indicates a particular model, and *min* is the value of the model with the smallest AIC.

Finally, one can compute the likelihood of a certain model given the data by using the transformation: $e^{(-1/2\,\Delta_i)}$. If this quantity is normalized to one (the so-called Akaike weight) it allows a convenient comparison among models:

$$w_i = \frac{e^{0.5\Delta_i}}{\sum_i e^{0.5\Delta_i}}$$

Where the summation is over all models.

The Akaike Information Criterion lends itself very easily to the comparison of several alternative models, while not being at all a "test" in the sense of requiring the

9 Notice that for small sample sizes, roughly when n/K < 40, an additional corrective term needs to be added to AIC; this term is: $\frac{2NK(K+1)}{n-K-1}$.

calculation of associated p-values. From a philosophical perspective, it is a particularly intriguing approach because it takes into account not just the fit of the model to the data, but also the complexity of the model itself, and therefore of its underlying hypothesis.

An additional bonus: the disappearance of the problem of multiple comparisons

One interesting byproduct of fighting the habit that has ingrained the standard approach to hypothesis testing is that one of the most troublesome and persistent problems in quantitative biology literally disappears into thin air. For a long time statisticians have recognized that when one carries out many "formal" (i.e., implying the calculation of p-values) simultaneous tests on a given data set—as it is often the case—one should correct for the fact that the nominal ?-value (the threshold used to decide about the statistical significance or lack thereof of a test) does not reflect the actual level, because the more tests one runs the more will come out significant just by chance. Several more or less conservative remedies have been proposed for this problem, two of the most popular being the Bonferroni or sequential Bonferroni corrections (Rice 1989). However, the problem with the corrections is that they tend to be too conservative (Turkheimer et al. 2000), thereby making it more difficult to reject the null hypothesis than it is reasonable. Furthermore, with complex data sets that can be analyzed in several ways, it is often not clear how many multiple comparisons one is actually conducting (Abelson 1995), an intriguing philosophical riddle in itself.

Because the approach advocated here does away with formal statistical testing altogether, one can conduct as many comparisons as one needs (graphically, via likelihood ratios, or using the Akaike criterion) without running into the problem of multiple tests.

This makes perfect sense within accepted statistical theory, as well as from a philosophical perspective on hypothesis testing, and it is not a "free lunch" of any sort. All that is required is the abandonment of the logic of p-values and its replacement by a habit of thinking in terms of competing hypotheses.

Should we throw away what we have published so far?

Fortunately, no. It turns out that in many cases inferences based on the classical p-value approach are similar to those arrived at using the alternative methods described here. Why bother with a new way of doing things, then? For two reasons, the second of which has decidedly philosophical implications. First, there are cases in which the two approaches will yield very different results. In particular, as we have seen, p-values are problematic when sample sizes are either too small or very large. In the first case, a failure to reject the null hypothesis may be premature, because the problem may be the low power inherent in small sample sizes. P-values do not carry any information about this problem, and one instead needs to plot the data and their confidence intervals.[10] In the second case, when sample sizes are very large, p-values may lead researchers to reject the null hypothesis when the difference between effects is scientifically irrelevant, though technically nonzero (as we have seen, this is because p-values refer to a point-hypothesis, which is never of interest to scientists). In this case again plotting the data will give a much better description of what is going on.

The second reason to opt for alternatives to standard hypothesis testing is

[10] If one insists in implementing p-value based analyses one should at least carry out a power analysis: to state that there is no relationship between X and Y because the accompanying p-value is large becomes meaningful only if a power test can demonstrate that there was sufficient power in the analysis to detect a moderate or small effect, thereby making the conclusion more robust.

philosophically more interesting. It is, again, a matter of habit of mind: often we are simply not interested in the hypotheses tested with the standard methods, and good science actually proceeds by comparing the relative fitness (using the data) of a series of reasonable alternative models (hypotheses). That is something that standard hypothesis testing simply cannot do but that scientists naturally tend to implicitly incorporate into the structures of their research proposals and papers. An explicit acknowledgment of this will help them and their students to be more clear about the questions they are asking and how to go about answering them.

A final note on the sociology of science

Finally, it is interesting to ask why these rather elementary considerations on p-values and null hypotheses, known to any statistician and to many researchers in other fields, have failed to bring about a change in the way organismal biologists (i.e., ecologists and evolutionary biologists) present their results in published papers. They may soon begin to, as a result of the fact that the discussion—in fields other than ecology and evolutionary biology—has been going on for some time now. In fact, statisticians and some scientists in other disciplines have been aware of the problem with the standard approach to hypothesis testing and with the concept of null hypothesis essentially since their inception (see references above). Yet, discontent has built up slowly, starting to surface first in the social sciences (Cronbach 1975; Meehl 1978).

There are several reasons why the p-value and null hypothesis testing approach is still so widespread: First, because—as I have argued above—quantitative biologists actually do think about their work and write the discussion sections of their papers

largely by disconnecting their conclusions from the formal results of their statistical tests. Most practicing biologists realize that what is important is the competition of different hypotheses and how much each is favored by the data, not the rejection of simple nulls. That is why the discussion section in a given paper focuses on the former, not the latter. Second, several of the alternative methods are not widely implemented in the available commercial software and—while actually conceptually simpler and more intuitive—they are therefore not familiar to most researchers. Third, once a system of textbooks and teachers built around the standard approach is in place, it is difficult to do without it because of simple intellectual inertia, which is what motivated the writing of this chapter to begin with. Once people learn to do things one way, and editors and reviewers are happy with the status quo, it is counterproductive to try to change things.

Yet, as I hope is clear from the above, the situation needs to change for very good reasons related to fostering better and clearer thinking at all stages of the scientific enterprise, from teaching undergraduate students to publishing intellectually more satisfying papers.

Chapter 6: Genotype-environment interactions and our understanding of the biological bases of human cognitive abilities

"Men have an extraordinarily erroneous opinion of their position in nature; and the error is ineradicable." – W. Somerset Maugham (1874–1966), *Writer's Notebook* (1949), 1896 entry.

Perhaps one of the most entrenched "bad habits" in the biological literature—both technical and especially popular—hinges on the false (but widely accepted) dichotomy between nature and nurture. Much has been written on this topic, and yet the problem simply doesn't seem to go away, as it will be clear from reading this chapter. I am under no illusion that my perspective will improve things much, but the practical consequences of this "debate" for human well being are so great, and it is such a quintessential example of what sound philosophical criticism of science practice can achieve, that I simply had to include it here.

The debate on the relative importance of nature (genetics) and nurture (environment) in determining human traits has been prolonged, acrimonious, and largely pointless. This claim may seem excessive, yet anyone familiar with the field of phenotypic plasticity (i.e., the way in which a genotype responds to the environment: Pigliucci 2001) in modern evolutionary biology will know that such a conclusion must follow from what we actually know of genotype-environment interactions in other organisms. My position is not that we cannot (in principle) know anything about reaction norms (the function describing genotype-environment interactions) in humans, or that we

do not know anything about genetic or environmental effects on human characteristics. But I submit that we currently do not know much that is pertinent to the discussion, and especially that we know very little that can sensibly inform our social policies. When eminent scientists in a field can draw such diametrically opposite conclusions about a given subject matter as Stephen Gould and E.O. Wilson do, it is a good bet that they are engaging in an intellectual exercise that is groundless enough from an empirical standpoint to allow for such latitude of positions.

Nature vs. Nurture?

We should start by asking if there really is a debate between opposite camps anymore, or rather if we are simply dealing with a sliding scale along a continuum, with people disagreeing on where human attributes fall onto that continuum. Let us make no mistake about it. While both sides pay lip service to the idea that the solution to the riddle falls somewhere in the middle, this is mostly rhetoric. Furthermore, what does it *mean* to be in the middle in this context? Are we in the middle if we say that X% of intelligence is determined by genes and Y% by the environment? Or is it more reasonable to propose that intelligence is the result of a vague and unspecified "emergent property" of nature and nurture?

The position of Gould, Richard Lewontin and others is that the environment is the major determinant of human nature. Now, these scientists clearly understand the fundamentals of genetics very well, so how can they claim that genes do not have anything to do with the human condition? Technically, they don't make such claim. The modern nurture school acknowledges that genes influence human behavior, but they haste

to add that whatever such influence may be, it can be overridden by the comparatively much larger effect of environmental conditions, chiefly education and socioeconomic status. How do they know this? They do not. For example, Gould (in the revised edition of *The Mismeasure of Man*, 1996, p. 355) says that "the *biological* basis of human uniqueness leads us to reject biological determinism" (his emphasis). Gould is here implying that what we know of the biology of humans negates a major role of genes in shaping our characteristics. Simply put, this is a sweeping statement for which there is no convincing empirical evidence whatsoever. A few pages later (p. 363), Gould continues: "most of the behavioral 'traits' that sociobiologists try to explain may never have been subject to direct natural selection at all–and may therefore exhibit a flexibility that features crucial to survival can never display." Notice the repeated use of the conditional. But if it is true that lack of relevant information is a valid objection to a particular theory (sociobiology), it is equally true that the same dearth of information is a problem for the opposite theory (nurturism). The problem is that the nurturists seem to implicitly advocate the notion that their theory is the default, which should be embraced if there is no evidence to the contrary. This is a peculiar stand, akin to the reasoning of creationists who maintain that any gap in the fossil record must be taken as prima facie evidence of divine creation (Pigliucci 2002b).

On the other side of the divide, Arthur Jensen, Herrnstein & Murray, Wilson (albeit in a category of his own) and many others are convinced that genetics and natural selection have shaped the physical as well as mental characteristics of all living beings, including humans. When Murray (1998) suggests that "IQ will put you in your place" he is assuming that IQ is written in stone in the DNA of each one of us. He, like the nurturists, also bows to a formal acknowledgment of the opposite camp by saying that

the environment may have "some" effect. But the role of the genes is so powerful that he goes on to suggest that governments should not waste too much time and money in improving education and economic conditions, because that will not alter one's place in society as determined by her IQ.

How do these researchers know that the balance between the two forces is so much in favor of genetics? Again, they don't. For example, Murray (1998) affirms that "the research bears out what parents of children with unequal abilities already know–that try as they might to make Johnny as bright as Sarah, it is difficult, and even impossible, to close the gap between them." This is a very specific conclusion about the shape of human reaction norms for IQ, a conclusion that simply is *not* supported by any published research. On his part, Rushton (1998), in an essay amusingly entitled "The Mismeasure of Gould," discusses the evidence favoring a link between characteristics of the human brain and violent human behavior. He cites: "Adrian Raine … tentatively concluded that frontal-lobe dysfunction was associated with violent behavior, including rape." Is this supposed to be compelling evidence? Of what, exactly? Of course damaged brains will cause altered behaviors, this is a finding that has been a persistent result of neurophysiology for over a century (Ramachandran and Blakeslee 1998; Damasio 1999; Adams 2000; de Oliveira-Souza 2000; Cardinal et al. 2001; Greene et al. 2001; Miller et al. 2001). But this does not tell us anything about the plasticity of such behaviors and how they respond to environmental change. Without this knowledge, any statement to the effect that genetics is more important than nurture is empty rhetoric. Evolutionary psychology, the more recent incarnation of human sociobiology, does not make things any better. Despite cautionary statements by some of its luminaries, such as Steven Pinker (1997, 2002), most of that discipline's literature is littered with the same kind of

wild guesses and unsubstantiated just-so stories that plagued sociobiology to begin with (Kaplan 2000; 2002).

As it is easy to imagine, the debate goes far beyond academia. In fact, most of the discussion has raged outside proper academic channels, with vitriolic attacks and counterattacks published in outlets such as the *New York Times Book Review* or the Internet. At times, the confrontation has gotten physical, as in the infamous instance of a scientific meeting at which an opponent of sociobiological ideas treated Wilson to a downpour of ice water on his head. Gould has used the unorthodox stratagem of reviewing the same book (*The Bell Curve*) twice, to make sure that his invectives will reach the largest possible audience. On the other hand, Rushton (1998) complains of Gould's unfair "character assassination" in *The Mismeasure of Man*, while beginning one of his own articles with the following sentence (quoting from one of Gould's arch rivals, John Maynard Smith):

> "Gould occupies a rather curious position, particularly on his side of the Atlantic. Because of the excellence of his essays, he has come to be seen by non-biologists as the preeminent evolutionary theorist. In contrast, the evolutionary biologists with whom I have discussed his work tend to see him as a man whose ideas are so confused as to be hardly worth bothering with, but as one who should not be publicly criticized because he is at least on our side against the creationists."

If that is not character assassination, it is hard to see what qualifies. So much for fair play and scientific integrity: this is obviously not a scientific debate.

One thing is clear: each side has done a marvelous job at poking large holes in the flanks of the other. Let me briefly discuss how some of the apparently sound arguments used by each camp can in fact be dismissed quite easily, and how the insistence upon such arguments on the part of either nurturists or naturists indicates poor understanding, less than objective analysis, or both.

One of Gould's major attacks against genetic determinism is based on what he calls the "reification" fallacy. Simply put, the fact that a statistical correlation exists among two or more variables does not imply that that correlation corresponds to a physical entity or demonstrates the existence of a real phenomenon underlying the observed correlations (Shipley 2000). So, for example, genetic determinists interpret the fact that the scores from different types of intelligence tests correlate in multivariate space as an indication of the existence of an underlying factor describing general intelligence, "g" (Duncan et al. 2000). g, Gould objects, is simply a statistical artifact due to the fact that the different tests are designed in a similar way, so that individuals who score high on one test will score high on another. This may be true, except that Gould then does not apply his preaching to his own scientific research. He published several papers (e.g., Gould 1984, 1989) describing the (statistical) covariation among morphological traits measured in land snails of the genus *Cerion*. In that body of work, he argues that these correlations show the *existence* of several "constraints," i.e., limitations on the future evolutionary trajectories of these populations (which, incidentally, turns out not to be true: Stone 1996). Furthermore, he goes on to *name* such constraints, a perfect example of "reification"! The practice is either acceptable or questionable to an equal measure in *both* cases, not just when it is not convenient for Gould's position.

But the opposite camp doesn't fare any better. Naturists have always argued that

their incontrovertible evidence is based on rigorous studies of twins reared together or apart. The statistical arguments are quite complex, but they boil down to the fact that identical twins have a high degree of genetic similarity, compared, for example, to fraternal siblings, or to unrelated individuals–such as adopted children. If twins show a higher correlation of their IQ scores than the control groups, this is taken as prima facie evidence for genetic determination of intelligence. But this approach ignores that the real problem is how to determine human reaction norms, not to simply estimate heritabilities. Twin studies are not very useful for the simple reason that they control for only one of the two factors, the genetics but not the environmental. Thus, while yielding evidence of *some* sort of genetic basis to human behavior, they provide us with no information on the all-important genotype-environment interactions.

A perfect example of how unreasonable and unscientific *both* parties are in this controversy is given by their arguing over the relationship between intelligence and brain size. The nurturist camp claims that there is no relevant connection between the two, while the genetic determinists insist that brain size has a lot to do with intelligence (both camps admit that the brain's fine structure is more important than simple size). The latter position is backed by the now incontrovertible correlation between brain size and IQ: about 0.44. This sounds impressive, but it is much less so once one realizes that a correlation of 0.44 between two variables means that one accounts for only 21% of the variance of the other. That is, while brain size is indeed associated with IQ scores, about 80% of the variation in the population does *not* depend on brain size. Surely there is ample space in that 80% for environmental effects, but of course Rushton (1998) trumpets this finding as a decisive blow to the nurturists. It does not come even close to that effect.

Back on the other side of the fence, Gould (1996) takes the preposterous position of flatly denying even the existence of such correlation between brain size and intelligence. He builds implausible scenarios of respected scientists of the 19^{th} and 20^{th} century "leaning" on their balances, or stuffing just the right amount of measuring pallets in skulls to come up with the "right" results. Gould goes even so far as to make the argument that since clearly intelligent individuals, such as the French anatomist Cuvier and Nobel Prize novelist Anatole France, had brains that span almost the whole gamut of human values (1830 and 1017 grams respectively), brain size doesn't have anything to do with intelligence. This is such a ridiculous claim that, made in a different context, would undermine a scientist's credibility. The fact that two variables are correlated, and even causally connected, does not imply that one cannot pick two arbitrary data points and use them to contradict the general trend. To pretend otherwise is questionable science to say the least.

What do we actually know about the biological basis of human characteristics?

This discussion does not have to be taken as a nihilistic statement on the state of our knowledge of genetic and environmental effects on human characteristics. On the genetic front, there is little doubt that genes affect the development of the brain. Since the brain and the peripheral nervous system determine human behavior, then in some sense it is undeniable that we behave because of our genes. A large and fascinating literature on brain damage provides direct and incontrovertible evidence that alterations in the physical structure of the brain (by accident or by mutations) directly affect all sorts of human behaviors, including subtle personality traits (e.g., Gazzaniga 1998; Ramachandran and

Blakeslee 1998; Damasio 1999).

Let us briefly consider a few examples that make exceedingly clear why the nature-nurture controversy is important, and at the same time why neither camp is "right." For instance, it is quite clear that homosexual behavior can be acquired or lost, so there must be environmental components to it. At the same time, the level of exposure to male or female hormones in the womb also significantly affects gender preferences and sexual behavior of the adult individual (Moir and Jessel 1992 and references therein). While hormonal levels can be considered an environmental factor, the way in which the brain responds to them is definitely genetically hardwired.

Another, equally emotionally charged (and actively avoided by most scientists) areas of inquiry concerns the biological basis of religious belief (Larson and Witham 1997; Pigliucci 2002c). It would be foolish to claim that the environment does not preponderantly affect religious beliefs, given all we know about the psychology of individual humans and the sociology of human societies. Nevertheless, Ramachandran and Blakeslee (1998) report that micro-seizures in the temporal lobes are associated with unusually vivid religious experiences, often complete with sounds, visual stimuli, and a sense of sudden cosmic "understanding." It is easy enough to imagine genetic mutations making some individuals more or less prone to such seizures, and therefore genetically affecting their propensity for religious beliefs (see also Newberg and D'Aquili 2001). In both the cases of homosexuality and of religious beliefs, it doesn't make any sense to take an ideological posture and defend it regardless of the evidence. However, the data do not allow us to disentangle the relative contributions of nature and nurture in a neat and simple way, they only lead us to conclude that both components must be there.

A clear and largely unemotional example of why both nurturists and genetic

determinists are partially right, but neither can address the most important questions, is given by the environmentally curable genetic disease known as phenylketonuria (Kaplan 2000). PKU is caused by a simple "inborn metabolic error," a mutation that does not allow the formation or proper functioning of an enzyme that metabolizes the amino acid phenylalanine. Often this results in accumulation of the amino acid in the brain during development, which in turn causes a host of phenotypic and behavioral effects, including severe mental retardation. The link between the genetic level and the phenotype is unusually straightforward, offering an ideal example of how genes can affect behaviors, although no genetic determinist would go so far as to say that this is a gene *for* normal intellectual development (Kaplan and Pigliucci 2001, see also Chapter 2). On the other hand, a very simple dietary change can completely neutralize the genetic effect: an individual can grow up normally by carefully avoiding phenylalanine, which is why many soda cans carry a warning label for PKU patients. This is therefore also an extreme case of plasticity in humans. So, one has to conclude that PKU is *both* genetically determined and environmentally plastic. The crucial question, however, remains: what are the reaction norms of PKU genotypes? There surely are several PKU genotypes, not only because the mutation can be caused by different alleles at the main locus, but also because the effects of the mutation depend on the other genes that interact with that locus (Kaplan 2000). Furthermore, there may be plenty of other environmental circumstances that affect the degree of occurrence and gravity of PKU symptoms. We simply do not know, and we will not know any time soon, given the obvious limitations on experimental research in human genetics.

Why we don't know what really matters

The fact that the quantitative genetics of human behavior is far less advanced than most "experts" would like the public to think is a logical consequence of what we know about phenotypic plasticity (Pigliucci 2001b). First, what we would really like to know is how genes and environments interact to produce a certain phenotype or behavior *in an individual*. Given that this has so far proven to be an elusive goal, human quantitative geneticists have shifted the target to a statistical understanding of *variation within and among human populations*. Enter the venerable concept of heritability, the ratio between the genetic and the phenotypic variance for a particular trait, measured by a variety of breeding designs (Lewontin 1974). The result of this shift is that, assuming a given trait turns out to have a heritability of 70%, all one can say is that 70% of the phenotypic variation in the population is associated to differences in the genetic constitution of its individuals, *when measured in a particular environment*. This is Lewontin's mantra: heritability is only a local measure; it estimates the amount of genetic variation for a quantitative trait *only* within a particular population (set of genes) and in a particular environment (Lewontin 1974, see also **Figure 6.1**. It cannot be reasonably extrapolated either to other populations or to other environments without further empirical research.

Using the concept of reaction norms, it is rather easy to understand where the problem lies. The extreme genetic determinist position can be represented by a series of reaction norms parallel to each other in a genotype-phenotype space. The norms are flat, indicating that there is no effect of the environment whatsoever. They are also widely spaced apart, suggesting that genetic differences do result in major phenotypic/behavioral differences. If this were indeed the case for human cognitive abilities, no social program of

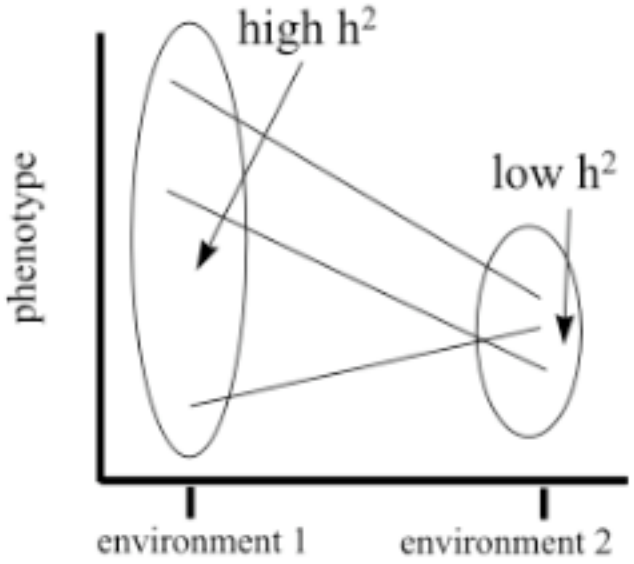

Figure 6.1. A hypothetical population with three genotypes characterized by distinct reaction norms along an environmental gradient. Where the reaction norms converge, the heritability of the trait is low; where they diverge, the heritability is high. Heritability, therefore, is a complex function of the environment and of the genetic constitution of the population. (The diagram assumes equal phenotypic variances in the two environments.)

any kind would be worth a dime. However, we know that this is not the real scenario, because we have plenty of evidence supporting the existence of some significant environmental effect on human behavior. The extreme nurturist hypothesis could instead be represented by reaction norms characterized by a steep slope, indicating a dramatic effect of the environment. They would also be close together in a bundle because of no average differences among genotypes. Even the staunchest supporter of liberal social policies (such as myself), however, will admit that this scenario is as unlikely and inconsistent with the scant data we have as the previous one. This is simply because the overwhelming majority of studies of reaction norms in other species have detected ample genetic variation for the ability of genotypes to respond to all sorts of environmental stimuli (Pigliucci 2001b, especially chapter 4).

What is the real shape of human reaction norms? I would *guess* (based on the typical situation in other animals) that some genotypes are almost not responsive to environmental changes, others respond significantly, and still others are dramatically

affected by the environment. Furthermore, the differences among genotypes are probably large in some environments but small in others. If this were the case, the truth would indeed lie somewhere in the middle, though such middle would obviously not be a simple average of genetic and environmental effects. However, the fact of the matter is that, again, we do not know. While we can confidently exclude both extreme scenarios, there are an infinite number of potential configurations in the middle. Since we cannot technically and ethically grow genetic replicates of human beings under controlled and diverse environmental conditions, we simply do not know what the patterns of plasticity for human cognitive traits are. Scientists should acknowledge this and move on. But the temptation is apparently too great for them not to sin.

A characteristic example of how difficult it is even for renowned biologists to get over what Feldman and Lewontin (1990) called "the heritability hang-up" can be found in works by sociobiologist Edward O. Wilson. In his *Consilience: the Unity of Knowledge* (1998) he explains his theory of gene-culture evolution, how genes and environments co-evolve to shape humanity's destiny. The chapter entitled "From Genes to Culture" is an odd mix of an enlightened vision of the problem and a stubbornly orthodox one, as if Wilson couldn't make up his mind. After admitting that "all biologists speak of the interaction between heredity and environment," he goes on to introduce the concept of reaction norm. He correctly points out that:

> "Redefined with the more precise concepts of genetics, nurturists can now be seen to believe that human behavioral genes have very broad norms of reaction, while hereditarians think the norms are relatively narrow. In this sense the difference between the two opinions is thus one of degree, not of kind. It becomes a matter

that can be settled and agreed upon empirically."

Wilson must know that the empirical way to settle this matter would be to selectively breed humans and raise them under controlled conditions, hardly something that can actually be done. Nevertheless, he expresses some optimism that the matter will eventually be settled through the accumulation of data while not providing a clue as to how this could possibly happen. Talk about "bad habits" ingrained in someone's writings about science!

Wilson also brings up the concept of heritability. Even though he expressly admits that there is a genotype-environment correlation that makes the use of heritability measures inadequate, and that heritability is 'flexible' (i.e., it depends on the environment), he calls these "peculiar twists." While coming so close to destroying (or at least greatly reducing) the very meaning and usefulness of heritability, the nature-nurture habit is apparently too strong, and Wilson ends up affirming that "the measure has considerable merit, and in fact is the backbone of human behavioral genetics." He proceeds by imagining scenarios in which we would be able to measure heritability in different cultural contexts and predicting that in some cases heritability would increase, in other it would decrease. Let me be clear about it: such predictions have no foundations whatsoever given current procedures in empirical quantitative genetics. The only way to make them in a reliable way is either to do the experiment, which is both ethically impossible and technically extremely difficult, or to understand exactly how genes and environments interact from a mechanistic standpoint. The latter objective is not as of now even within the range of the most powerful telescope that any scientist may hope to use to divine the future. Worse, immediately after providing the reasons for being extremely

cautious about heritability measures, Wilson says:

> "Nurturists have also traditionally thought that the heritability of intelligence and personality traits is low, while hereditarians have considered it to be high. That disagreement has largely been solved. In contemporary Caucasians of Europe and the Unites States at least, heritability is usually in midrange, with its exact value varying from one trait to another."

It is astounding to see this kind of reasoning on the part of a biologist who demonstrably knows better. As it is clear when one considers the environmental plasticity of heritability measures, it just does not make any sense to average heritability across environments. Since the Caucasians to whom Wilson is referring were raised under different environments, the resulting estimates of heritability of IQ are a hopelessly confounded hodgepodge of nature and nurture.

The environmental dependence of heritability is by no means a mere theoretical possibility. In many organisms in which the experiment can and has been done, a strong dependence of heritability on the environment has been found for many (albeit not all) traits. For example, the work of Mazer and Schick (1991) on wild radish clearly shows that the heritabilities of three traits (flowering time, petal area, and pollen volume) are highly dependent on the density of conspecifics, a crucial environmental parameter in plants. More importantly, there is no regular pattern describing such interdependence: heritability is lowest under high density for flowering time, but under low density for petal area, and under medium density for pollen volume. Therefore, one cannot even say that a particular environmental range is likely to yield high heritabilities while another is

associated with low heritabilities. It depends on the character (and, probably, on the population in which the study is conducted). This may be a frustrating aspect of biological reality, but ignoring it does not help, and–in the case of humans–may lead to disastrous social policies.

What are we to do?

What kind of evidence would settle the nature-nurture debate on the biological basis of human behavior? The answer is in fact very simple, and such an experiment has already been done, for example, by Cooper and Zubek (1958). On rats, that is. They compared the ability to avoid mistakes in running through a maze in two inbred, genetically distinct lines of rats. One line had been selected for high performance in the maze ("bright" line), the other for particularly low performance ("dull" line). When reared under a standard environment, comparable to the one in which the selection process occurred, the two lines showed a highly significant difference in their abilities. Cooper and Zubek, however, also reared individuals of the two lines in two other environments: a situation in which the cage was entirely devoid of visual and tactile stimuli ("poor" environment), and one in which the developing animals were exposed to brightly colored walls and toys ("enriched" environment). The results are simply stunning (**Figure 6.2**). Under the poor conditions, the bright rats performed as badly as the dull ones; furthermore, under the enriched environment the dull rats did as well as the bright ones! This translates into a high heritability and marked genetic differences in the standard environment, but no heritability under either extreme environment. The inescapable conclusion is that maze-running ability in these rat genotypes has a very plastic reaction

norm, and that different genotypes converge onto similar phenotypes at extreme environmental conditions. I am not suggesting here that this particular experiment is flawless. For one thing, only two genotypes were tested, and they were certainly not a representative sample of natural populations. Furthermore, the cognitive ability being tested was a relatively specific one, and its correlation to generalized intelligence (if there is such thing) remains to be ascertained. Finally, one could easily conceive of other (and more naturalistic) kinds of environments that may yield very different results. However, all these problems are also common to any human study published so far. The point I want to make is that this is the kind of data that would go a long way toward empirically answer the question of heritability and plasticity of human behaviors. Yet, we simply cannot expect this kind of data to appear any time soon in the scientific literature about humans, although similar experiments in other vertebrates would go much further than any piece of rhetoric or speculation.

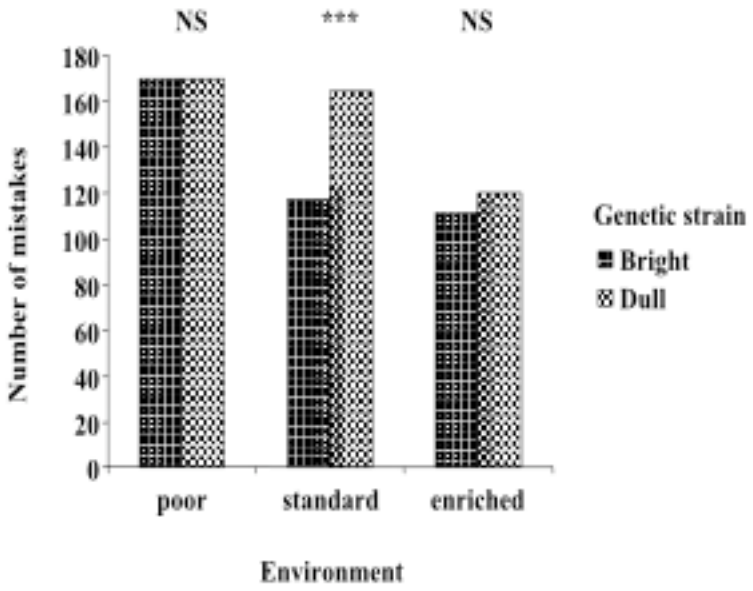

Figure 6.2. Effect of nurture (quality and stimulation properties of the environment) on "intelligence" (ability to avoid mistakes in mazes) in genetically distinct laboratory lines of rats. Notice how the genetic differences are significant (i.e., there is a significant heritability, indicated by the asterisks) only in the environment in which the lines were selected, but they disappear in the other two environments (NS=not significant). After data from Cooper and Zubek 1958.

So, how should scientists answer pressing requests from the public and policy makers when it comes down to our knowledge of intelligence and other behavioral traits in humans? I believe some answers and suggestions can be given, but only if accompanied by careful statements about the great limitations imposed by the impossibility of carrying out the proper experimental manipulations. While good science can be done by using only indirect and statistical approaches, we have to remember that the only way known so far to dissect the functionality of complex systems is by direct experimental manipulation.

The first and perhaps most important thing we can conclude from what we know of the biological basis of human behavioral traits is that they tend to be plastic. Therefore, regardless of any politically motivated rhetoric, educational programs do have a good chance to succeed. In fact, even if humans were genetically distinct in their abilities, and even if an improvement in environmental conditions will maintain such differences, it would still be worth funding educational programs. This is because a better quality of the environment would likely improve the performance of most genotypes (what in ecology is known as the "silver spoon effect": Sultan 1995).

The second fact that we can confidently state is that most human cognitive traits do have a genetic basis. It simply could not be otherwise, given that the structure of the brain is dictated in part by genes, and considering all we know about the effects on behavior of brain anomalies. This translates into a politically incorrect position in which genetic differences among races (Chapter 3), genders, and individuals may indeed exist. I don't see, however, how this acknowledgment would necessarily lead to discrimination and abuse. The first is a fact of nature, completely impervious to our judgments and hopes. The second is a matter of attitudes and policies, both of which can be changed. To translate one into the other is to commit the naturalistic fallacy(Hume 1739-40).

What scientists cannot and should not venture to say, however, is how the two previous conclusions can be combined in order to provide a satisfactory account of interactions of genes and environments in humans. This is indeed a crucial question, because educational approaches may be more or less fruitful depending on the precise shape of human reaction norms. The same can be said for policies concerned with curbing crime, or for a host of other fundamental and difficult decisions we have to make in our societies. Unfortunately, it should be clear by now that this is where the line must be drawn and the only honest answer a scientist can give is: I don't know. It is astounding to see how difficult the utterance of these simple words can be, a side effect of which is the publication of scientifically largely useless volumes such as the *Bell Curve* and its many rebuttals. The reasons that lead one to yield to the temptation of saying more than one's data and theories strictly allow are, of course, relatively easy to understand. Sociologists, psychologists, and philosophers of science have long pointed out that personal egos, social prestige, and financial rewards all play into this in a remarkably complex ensemble (Kuhn 1970; Longino 1990). The fact remains, however, that there are questions that science cannot answer (either at the moment, or in general). As Richard Lewontin put it in a similar context:

> "I must say that the best lesson our readers can learn is to give up the childish notion that everything that is interesting about nature can be understood. ... It might be interesting to know how cognition (whatever that is) arose and spread and changed, but we cannot know. Tough luck." (Lewontin 1998)

Learning to live with this conclusion actually empowers scientists, because by not

pretending to be omniscient they can enjoy the fruits of the single most effective tool humans have ever devised to understand the world: science itself. It also is the only decent thing to do.

Part III - Metaphors: good and bad ones

The final section of this work considers two instances of the use of metaphors in scientific discourse. One of these examples provides, I think, a quintessential illustration of how misleading metaphors can be in biology; the other example presents us with a somewhat more fecund (albeit by all means not devoid of pitfalls) metaphor, that has significantly furthered discussion in an important field in evolutionary biology.

My point, as stated in the Introduction, is of course not that scientists should avoid the use of metaphors or imagery. This would be impossible given that this sort of thinking aids seem to be ingrained in the way human beings in general approach complex problems. Reminders of this are everywhere around us, including indicative popular sayings in our culture ("a picture is worth a thousand words," "visualize this or that," etc.). Rather, the idea is to examine case studies in which metaphors have hindered or furthered scientific progress, to help both philosophers and scientists think about what may make for a good or bad metaphor and, hence, what sort of thinking patterns we may want to avoid or foster in our students, colleagues, and ourselves.

The first chapter of this cluster criticizes the metaphor that characterizes genes as "selfish" biological entities, popularized (though not originated) by Richard Dawkins in his famous 1976 book, *The Selfish Gene*. There is somewhat of a cottage industry in philosophy of science that flourished around criticism (and less frequently defense) of Dawkins' ideas, despite the fact that he never wrote a technical treatise about them. My contribution aims at attacking the metaphor by adopting the unusual tactic of illustrating

the nonsensical consequences that can be derived by applying it even more widely than Dawkins himself dared to suggest: if there are selfish genes, why not selfish triplets of DNA (the minimum coding region), or even selfish base pairs (the level at which physical recombination can no longer take place)? The criteria offered by Dawkins to identify *the* (alleged) target of selection do not allow us to discriminate among these possibilities.

The second and last chapter examines the often acrimonious controversy catalyzed by the publication of a paper that adopted a metaphor derived from the field of architecture to criticize the overly facile use of natural selection as an explanation for complex biological structures. While in the end I do not come strongly on the side of either the "panglossian" or the "spandrelist" extremes, I think that the now classic paper by Gould and Lewontin (1979) has done exactly what a good scientific paper should do: it has spurred a significant amount of debate, a critical reexamination of positions taken for granted, and a number of empirical studies that have taken into account the developments of the debate itself. This, of course, doesn't mean that the question of the relative contribution of constraints and selection to shaping phenotypic evolution is settled. That can only happen empirically. But the conceptual air has been sufficiently cleaned through what counts as one of the few philosophical analyses ever to directly make a significant dent into the thinking of practicing scientists.

Chapter 7: Against the selfish gene

"The folly of mistaking a paradox for a discovery, a metaphor for a proof, a torrent of verbiage for a spring of capital truths, and oneself for an oracle, is inborn in us." -Paul Ambroise Valery, 1871-1945

Perhaps one of the most controversial examples of widespread metaphors in modern evolutionary biology is the idea of "selfish genes" introduced by Richard Dawkins (1976) in his popular science book of the same title. I think the selfish gene metaphor has largely turned out to be a "bad" one, in the sense that it tends to mislead us into dangerous habits of thinking, while at the same time confining our imagery to a rather narrow view of the evolutionary process. I will try to substantiate this claim throughout this chapter.

Selfish genes, I will maintain, do not exist for the simple reason that genes do not exist. At least, not as the evolutionarily stable entities imagined by Dawkins to justify his extreme gene-centric view of the evolutionary process. This claim may seem excessive, but I will attempt to show that it is the necessary consequence of what we know about molecular and developmental genetics, and that such an outcome could have in fact been predicted since the time of the publication of *The Selfish Gene*, something that would have spared evolutionary biology a wasteful and acrimonious debate on the units of selection that has plagued the field for the last thirty years (Williams 1992).

Dawkins' central idea is a popularization of earlier work by Williams (1966) and Hamilton (1963). Both latter authors intended their work to be a refutation of the then

rampant fuzzy thinking about group selection (Wynne-Edwards 1962), and both later considerably retreated from the extreme gene's eye view still advocated by Dawkins and moved toward a more pluralistic multilevel theory of natural selection (Hamilton 1975; Williams 1992; see discussion in Sober and Wilson 1998, and in Keller 1999, especially chapter 1).

The selfish gene paradigm, I maintain here, can be attacked from two perspectives, one hierarchically above the gene level of analysis, the second one hierarchically lower. I will briefly summarize the arguments advanced by Sober and Wilson (1998) and earlier on by Sober and Lewontin (1982) to effectively deliver the attack from above, and I will then present novel arguments for the attack from below. Taken together, these criticisms of the selfish gene will show that—while indubitably important in any evolutionary theory—genes are neither central nor, *a fortiori*, the only targets of selection, and therefore promoters of evolutionary change.

Before we enter into the thicket of the discussion, let us briefly summarize Dawkins' views on the units of selection. While this is no easy task, given that his positions have shifted significantly over the years (especially from Dawkins 1976 to 1982) and have become more mainstream, Dawkins maintains that the only things that can be properly selected are *replicators*, i.e., entities that can be copied.[11] According to Dawkins, active germ-line replicators are the "optimons," i.e., the units that benefit from adaptations. The organism hosting the optimon is called a "vehicle," and the optimon obviously benefits from the survival and reproduction of its vehicle. That is how organism-level adaptations are explained in the selfish gene theory. Organisms are not seen by Dawkins as replicators for two reasons: first, because they are too large and

[11] It is worthwhile remembering that the distinction between replicators and interactors has been largely elaborated upon by Hull , and is ultimately based on Williams' concept of "evolutionary gene". However, this is not the place for a review of the history of the idea itself.

temporally unstable to be replicators (though his emphasis on temporal stability has later been reduced). In fact, even genomes or chromosomes are considered too unstable because of recombination and crossing-over. As we shall see below, it is peculiar that Dawkins fails to realize that this exact reasoning applies to genes as well and undermines their centrality in evolutionary biology theory. Second, organisms are not replicators because organisms are not copied, genes are. I will argue below that this is a largely incomplete view of what genes and organisms actually are and do.

The attack from above: individual, kin and group selection

Sober and Wilson (1988, chapter 2) have clearly and elegantly showed that all supposed alternatives to group selection, such as the theory of inclusive fitness (or kin selection, as it is often called) as well as classical individual selection, are special cases of a general formulation of the theory of natural selection due to Price (1970). Price invented a mathematical formalism that makes clear that selection can *always* be thought of as the outcome of processes acting at two (or more) hierarchical levels, for example (but not only) individual and group selection. The famous Price equation allows to calculate the change in gene frequency in the total population (i.e., across groups), ∏P, as:

$$\prod P = dp_{n'} + cov_n (s,p) / s_n$$

Where: dp is the change in gene frequency within the average group weighted by the size of the group after selection (n'), s is the benefit of the trait to the group, p is the frequency of altruists in the group, cov stands for covariance, is the average group benefit

weighted by the group size before selection (n). In essence, the first term measures within-group selection and the second term measures between-group selection. This equation can be extended to the general case of many hierarchical levels of selection, both below and above the individual, as well as to more than two groups (which was the extent of application of the original formulation).

That group (i.e., above-individual) selection is not only possible, but empirically demonstrable, is clear when one examines the cases that Williams and Hamilton themselves have considered, and that have been summarized in Sober and Wilson's book. For example, the evolution of virulence and the evolution of unequal sex ratios. In both cases the strategy formulated by Sober and Wilson (1988, chapter 3) considers first the two extreme scenarios of what presumably adaptive traits (such as sex ratios and degree of virulence) would look like if there were only individual or only group selection. Then one compares the observed range of the traits in question to estimate the relative importance of group vs. individual selection. To take the example of sex ratios, the expected value if only individual selection were occurring is 0.5 (maintained by frequency dependent selection in favor of the minority sex every time that deviations occur). If only group selection were happening, the expectation is a sex ratio as close to 100% females as it is made possible by the population structure and mating system of the organisms in question. Intermediate values (which need to be calculated for each case based on detailed knowledge of the biology of the species) would then represent prima facie evidence for mixed selective pressures acting upon the system.[12] Indeed, while unequal sex ratios were once considered very rare, several cases have been documented, and the phenomenon may be very common within certain groups of organisms, such as invertebrates (references in

[12] I am purposely avoiding the use of the "force" metaphor in evolution, which Sober freely employs, because I am convinced that recent criticisms of it are right on target.

Sober and Wilson 1998).

I do not wish to give here any further justification for expanding levels of selection above the gene, since this would essentially require to repeat the arguments and extensive literature citation that can be found in Sober and Wilson's book. Let me therefore turn my attention to a brief summary of the type of selection that does indeed go on directly at (or around) the gene level. Such discussion will make clear that I do not wish to deny the existence of selection directly on genes, but rather to emphasize the limits of the gene's eye view of evolution. The terrain will therefore be prepared for the extension of selection theory and of the criticism of the selfish gene *below* the gene level.

Selfish genetic elements

While selfish genes do not exist in the all-encompassing sense maintained by Dawkins, of course "selfish" genetic elements have been documented and are relatively well understood.[13] Yet, they play a much minor role than gene selectionists would like us to believe. Among the best examples of selfish "genes" are highly repetitive DNA (Fatyol et al. 2001, transposons (Miller et al. 2000), so-called oncogenes (Robinson et al. 2001), supernumerary chromosomes (Han et al. 2001), and segregation distorters (Kusano et al. 2003). Correspondingly, the best-known mechanisms used by selfish elements are transposition and meiotic drive (Lande and Wilkinson 1999).[14]

Highly repetitive DNA sequences, such as the human *Alu*, are very short (180 to

[13] Of course, one of the problems with this debate is the very use of the word "selfish." Within this context, I take this to mean an entity that can be replicated and that directly or indirectly affects its surrounding environment so to maximize the chances of its own replication—regardless of the consequences to either other similar entities or to the environment.

[14] For recent reviews on the role and persistence of selfish genetic elements in evolution see Hatcher 2000; Hurst and werren 2001.

280 bases), non-coding regions that can hardly be considered "genes" at all. Through errors of duplication, which their shortness and repetitiveness actually facilitates, these sequences can grow in number dramatically, up to the point of constituting a very large fraction of an organism's genome; millions of *Alu* are present in the human genome, amounting to about ten percent of the total (Kazazian and Moran 1998). Transposons are mobile genetic elements that are flanked by DNA sequences that make it particularly easy for them to self-excise from a chromosomal location and insert the DNA region into another location, either on the same or on another chromosome. Some transposons are capable of leaving copies of themselves behind; since often their DNA does not code for anything of use to the organism, these can rightly be thought of as selfish sequences. Oncogenes, or genes causing cancers, are of course nothing of the sort. Their normal function in the cell is usually associated with cell division and it is only when they are mutated or escape their normal regulatory sequences that such genes—as in the case of the human *myc* gene—actually cause tumors. Supernumerary chromosomes, sometime referred to as "B" chromosomes (as opposed to the normal complement, or "A" chromosomes) are large chunks of DNA, particularly common in plants and insects, which behave independently from the rest at meiosis, and that carry genes capable of altering the normal Mendelian ratio so to increase their chances of being passed into the germ line and the next generation. Finally, segregation distorters are genes that alter the normal course of meiotic divisions so that more copies of themselves are passed into the germ cells than they would if they were obeying standard Mendelian laws.

All these selfish elements have been demonstrated to be deleterious to the fitness of the host organism, which clearly shows that indeed natural selection can occur below the level of the organism. However, it is also equally clear that selection *at the organismal*

level poses a limit to the ability of these elements to be selfish. Just in the same way as a virulent parasitic strain sets off an arms race with its host, so do selfish genetic elements with the organisms they inhabit. Transposons and highly repetitive DNA sequences cannot replicate to the point of constituting the entire genome of the organism, or the latter would not be able to reproduce, thereby killing the too successfully selfish genes. Analogously, oncogenes are the exception rather than the rule, and one has to marvel at the high level of efficiency of the cellular mechanisms that have evolved to minimize the probability of a selfish cell breaking the ranks and causing a tumor in an organism. B chromosomes too have been demonstrated to increase in number within an individual, and in frequency within a population, only up to a certain point (Camacho et al. 2000). Beyond that the fitness cost to the organism is too high. In all these cases, selection at a hierarchically superior level to the selfish genetic element obviously puts a stop to the degree of selfishness going on at the lower level. In general, truly selfish genetic elements persist in populations only at very low frequencies, their further spread being limited by a variety of mechanisms, including population structure and intra-genomic conflicts (Hatcher 2000). Far from representing the ultimate success of the gene's eye view, selfish DNA clearly shows how a multilevel theory of natural selection is the only one capable of explaining what actually occurs in nature. Selection on individuals is very efficient at circumscribing selection on genes.

The reason for this overriding effect of individual-level selection is of course that genes are physically bound within organisms and are not capable of reproducing themselves outside of the organism or without help from the organism's molecular and cellular machinery. In a very important sense, organisms are groups of genes (and more), and it is group selection at this level that interacts with individual selection at the lower

(genic) level. The result, as in any compromise, is not entirely favorable to either side: selfish genetic elements are limited in their selfishness, but organisms are burdened by the existence of their genetic parasites and are therefore less efficient metabolically and in other ways in their competition with other organisms.

An example of a genetic phenomenon that is universal in the machinery of living organisms but is hardly emphasized by gene selectionists (because it does not fit the idea of every gene against each other) is epistasis (Wolf et al. 2000). Defined as gene-gene (or, better, gene product-gene product) interaction, it has been universally demonstrated to occur, and it is indeed a necessary consequence of the fact that genes code for proteins which themselves do not act alone but always in concert with other proteins, often in long and highly branched metabolic pathways. Epistasis is a perfect example of inter-genic cooperation that benefits the organism (without it, there would be no metabolism, and therefore no life as we know it). Selfish genes therefore have two "choices": to remain highly individualistic parasites (such as transposons and highly repetitive DNA), but be strictly limited in their ability to take over the organism; or cooperate to enhance the organism's fitness and therefore, indirectly, their own. It is highly likely that epistasis is a (universal) case of molecular "altruism" (if we wish to keep indulging in the widespread use of anthropocentric metaphors), because each gene would demonstrably do better in the short term just minding its own business and devoting all energies to its own reproduction. If genes were the only target of natural selection, epistasis would simply not make sense, and yet it is perhaps the only universal characteristic of every genetic system that ever appeared on earth after the origin of life itself. Because of epistasis, the standard trick of population genetics theory of calculating individual genes' fitnesses to use in mathematical models is at best an extreme simplification of what goes on. Single

gene fitness can be estimated only as an average across genetic backgrounds, but it is this sort of averaging that obscures, rather than clarify, what is really going on by creating fictitious targets (the individual genes) for natural selection to operate upon. This is so widespread in modern evolutionary thinking that Sober and Wilson give it the name of "averaging fallacy" (Sober and Wilson 1998, chapter 1).

It seems clear, therefore, that selfish genes do exist in a special sense, but that their power and evolutionary role is in fact much more limited than currently accepted by a large number of evolutionary biologists. Part of this limitation comes from genetic mechanisms themselves, such as epistasis, and part from the demonstrable existence of higher levels of selection, from the individual and above. But there are other, perhaps even more fundamental, reasons to discard the selfish gene metaphor, namely that genes themselves do not exist in the way the theory requires them to be.

The attack from below: what is a gene, anyway?

Commonsensical and largely accepted as the notion that genes are replicators is nowadays (but see Griffiths and Gray 1997), it is too simplistic given what we know now from molecular and developmental biology. To say the least, genes are certainly not the smallest and most stable physical units with phenotypic effects that replicate.[15] Besides the obvious (and important) question of what genes are *for* (see Chapter 2), which impinges on our understanding of what they are, it is not even clear how genes should be defined at the molecular level for a variety of reasons that I will briefly discuss below.

[15] As I mentioned above, Dawkins has more recently de-emphasized the stability criterion for talking about selfish genes. However, if stability doesn't matter, and if self-replicating units can be found *within* "genes," then it is not clear what makes genes *the* replicators and *the* targets of selection.

First, as it was discovered since the beginning of modern genetic research in the early 20th century, genes often undergo recombination. At the molecular level, this is a process that literally breaks a gene apart in at least two fragments, switches them with similar fragments from another gene, and stitches the pieces back together to regain a functional unit at the end of meiosis (Smith and Nicolas 2001). Because of this process, usually referred to as crossing over, genes cannot be considered stable entities over even rather short periods of time, quite apart from the other obvious source that damages their integrity on the long run, mutation (Nachman and Crowell 2000). If they are not long term entities, one of the distinctions between replicators and vehicles (the latter seen as temporary) ceases to exist, especially if one considers that some higher-level entities such as groups can exist essentially unchanged for much longer periods of time than genes (just think of social groups such as tribes of hunter-gatherers, or cultural groups such as the members of a given religion). Indeed, it is interesting to note that many selfish genetic elements exist (and probably persist) within "recombination-free" genomic regions, as is the case for meiotic drive genes in *Drosophila* (Palopoli and Wu 1996).

Second, as it has been apparent since the 1970s, genes are not "units" even in the functional sense, since many of them are made of a series of segments of two fundamental types: introns and exons (Chechetkin and Lobzin 1998). Exons are the segments that are actually retained as functional messenger RNA and translated to produce a protein when read by the ribosomal machines inside the cell. Introns are allegedly "spacers" that are spliced out after transcription, although occasionally evidence has turned out that they might have a regulatory role of some sort to play (Rose and Beliakoff 2000). So genes are neither historically nor functionally continuous units.

Third, there are plenty of other segments of DNA in a genome whose status as a

gene is uncertain, again highlighting that genes might be better described as a "family resemblance" concept (Wittgenstein 1953; Simon 1969; see also Chapter 4) rather than in an essentialist manner. For example, the recently quasi-completed human genome project has provided an estimate of about 30,000 genes in humans, much less than the 100,000 previously anticipated by evolutionary geneticists (Wolfsberg et al. 2001). However, this estimate is based on the classical kind of gene that actually codes for a protein, either with a structural or regulatory role. But what of the many up- and downstream short sequences that are known to regulate when and how a protein-producing gene is activated (Roosa et al. 2000)? These sequences do not "code" for anything per se, but mutations within their confines can have highly disruptive phenotypic effects, since the regulators determine when during development, in what cells and tissues, and in response to which environmental stimuli (if any) the protein-producing gene is to be expressed. As I mentioned above, one of the situations that lead to cancer is when "oncogenes" fall under the control of a wrong "promoter" sequence like the ones I am discussing here. In this case, is it the oncogene that is behaving "selfishly," is it the promoter sequence, or both?

Fourth, not only genes are made of introns and exons (in the case of protein-coding ones), but in fact one could argue that the "real" replicable units are the individual triplets coding for each amino acid, not the whole gene. Certainly triplets are more faithfully duplicated than genes because rarely a crossing over cuts within any given triplet, so at least triplets are more evolutionarily stable than genes (except for the occasional mutation, of course). Should we then consider genes as cooperatives (i.e., groups) of more or less altruistic triplets? And why stop at that level? Why not going one step further down and suggest that the "real" replicators are actually individual nucleotides? These are certainly more evolutionarily stable than triplets because they are

never broken up by crossing overs, and the only thing that can change them is a mutation, which is a much rarer event. In fact, there may be quite good evidence for the selfishness of individual bases, even though to my knowledge nobody has quite looked at it that way in the molecular literature. It is well-known that genomes have pockets of DNA that are highly enriched in AT (Adenine-Thiamin) CG (Cytosine-Guanine) base pairs (Hapgood et al. 2001). GC or AT-rich sequences are prone to produce more GC or AT pairs because of the way the DNA replicating and proofreading machineries work, so one can legitimately see these pockets as the outcome of inter-base competition for space in the genome, quite independently not only of the function of the organism, but even of the integrity of genes. Furthermore, recent research has clearly shown that short repetitive sequences do affect the phenotype dramatically *by simple virtue of their length*, not their base sequence. For example, the number of CAG repeats in a gene on the human chromosome 4 is directly proportional to the age of onset of Huntington's chorea (Gusella et al. 1996); in fact, the entire "gene" is made of a variable (among individuals) number of these short repeats! This means that short segments of DNA can have a function just as standard genes do. Why shouldn't *they* be the targets of selection?

If all of this talk of below-gene replicators sounds queer, it should. But not any more (and in fact significantly less) than talking about genes as the only units of replication and selection. The selfish gene as the true target of natural selection is a simplifying fiction of evolutionary genetic theory that—in the face of modern knowledge of molecular biology—should go the way of earlier simplifying fictions, like the one gene = one trait (introduced by Mendel: see Orel and Wood 2000) or the one gene = one enzyme (introduced by Beatle and Tatum: see Morange 1995) assumptions before it. The fact that most evolutionary biologists are still attached to the one-level explanation

offered by Dawkins and other selfish gene theorists is simply a matter of (bad) mental habit, not a reflection of what we actually know of genetics and evolution.

Dawkins' idea that genes and not organisms are copied is based on a simplistic understanding of information transfer in biology, and ignores some by now well-known facts (Roth 2001). Genes are not blueprints for organisms, they only contribute part of the information necessary to build a new individual. Try as you may, you will never succeed in making a living being out of a readout of its DNA.[16] The reason for this is that additional information is needed, for example the parental contribution in non-genetic material, which starts out the embryo's growth along the right spatial axes. Many epigenetic effects (such as tissue-tissue interactions and inductions) are only very remotely "controlled" by the genes (Newman and Muller 2001), as are the panoply of genotype-environment interactions that are also fundamental to the formation of any functional organism (Schlichting and Pigliucci 1998). Philosophers and biologists have started to crystallize all this into an alternative view of evolution appropriately termed Developmental Systems Theory (DST: Oyama 1985; Griffiths and Gray 1997; Oyama et al. 2001; Griffiths and Gray, in press), an attempt at providing an overarching conceptual framework that extends the classical neo-Darwinian focus on organisms, instead of reducing it to the level of the gene. According to DST, the genes are neither the only carriers of information, nor the only things that get replicated from one generation to the other, but just one of many necessary pieces of the evolutionary puzzle.

As perhaps the best understood example of the limits of genetic information in determining organismal structures, let us briefly consider the concept of homology. Homologous structures are those that are inherited from common ancestors, regardless of

[16] Which is why after the era of "genome" projects we are now talking of "proteome," "metabolome," and even "phenome."

their actual function in the descendants. For example, the upper limbs of all vertebrates are homologous despite the fact that some are used for terrestrial locomotion and some for flying. Following the gene-centric paradigm, it has always been assumed that homology at the developmental-morphological level had to correspond to homology at the genetic level. That is, biologists predicted that the same sort of genes would "control" similar sorts of structures. But research in molecular developmental genetics has shown that the reality is much more complex: some clearly non-homologous structures turned out to share the same genetic switches, while others that are equally clearly evolutionarily homologous do not (Treisman 1999). Roth (1988) has termed the process by which new genes get to take over the underpinning of homologous structures "genetic piracy." These findings have led to the vindication of an early proposal by van Valen (1982) that the continuity of information explaining homology is *not* genetic, but developmental. If this view is correct, it makes sense to say that homologous structures are "replicated" throughout evolution (and are the target of selection), regardless of their specific genetic basis. So much for genes as the only units of replication (and selection).

Toward a truly pluralistic theory of natural selection

The attack "from above" on the gene-centered paradigm summarized at the beginning of this chapter, together with my extension on that attack "from below" the gene hierarchical level, should make clear enough that evolutionary biology needs a truly pluralistic theory of natural selection. Pluralism has a bad reputation in science because of the dominant assumption that one unified explanation for every phenomenon in the world is not only possible, but probably at hand (Weinberg 1992b), at the least in the physical

sciences. However, as Dupré (1993) has convincingly argued, nature is much more likely to be significantly "disorderly," meaning that emergent properties at different levels of organization make a simple-minded ultra-reductionism a scientific and philosophical chimera. To paraphrase Dupré, the world is certainly made of material things, and larger things are always assembled from smaller ones. But all this implies is that studying the lower levels of organization helps us understanding *how* things at higher level do what they do, it doesn't necessarily aid us in predicting *what* specifically they will do. Emergent properties, once the scorn of the reductionist, are now the focus of attention in developing fields such as chaos (Stewart 2000; Pigliucci 2001c) and complexity(Bak and Chen 1991; Kauffman 1993; Sole` et al. 1999) theories, and concepts such as deterministic unpredictability and sensitivity to initial conditions are increasingly accepted as a natural property of complex, high-level systems of organization. In fact, leading evolutionary biologists have been proposing that major hitherto unexplained evolutionary transitions such as the origin of life and of societies are best thought of in terms of the emergence of higher levels of organization and of their level-specific properties (Maynard-Smith and Szathmary 1995). Why should the alleged centrality of the gene be the exception to multilevel explanations in science?

Naturally, for a truly multilevel or pluralistic theory to be scientifically useful it has to meet certain criteria. Let us consider as an example the list that Sober and Wilson (1998) developed to show that group selection can occur above the individual level *as a process distinct from (i.e., not reducible to) individual selection*: 1) There must be more than one group; 2) Groups must vary in some characteristic; 3) Groups with a certain trait must be more fit than groups without that trait (i.e., there has to be a relationship between the trait allegedly selected for and fitness); 4) Groups must not be too isolated

from each other because "progeny" of multiple groups must be able to mix or compete with each other to form new groups); and 5) The differential fitness of the groups must be strong enough to counter the differential fitness of individuals within groups (i.e., the outcome of any selection process will actually be a balance between selection at the lower and higher level(s)).

I maintain that from this perspective, just in the same way as organisms include groups of genes (as is readily admitted by mainstream evolutionary theorists, e.g., Buss 1987), and populations are groups of organisms, so genes can be thought of as groups of triplets or even individual bases, and are therefore no more fundamental than other levels in the hierarchy. In the same way as group selection allows the evolution of altruistic behavior in individuals, and individuals put a limit to the selfishness of genes through epistasis and other forms of regulation and interaction, so genes put a limit to the selfishness of their own constituent units.

To follow Sober and Wilson's criteria in the case, for example, of GC-rich regions: 1) Groups exist as discrete genomic regions with different amounts of GC pairs; 2) Genomic regions vary in the proportion of "altruistic" types, that is of bases that do not associate only with bases of the same kind (which would increase their own chances at reproduction), but with other kinds so to produce a functional gene (a group property); 3) Groups (i.e., genes) with more "altruistic" bases are more fit than groups without altruists for the simple reason that the first groups get replicated into future generations of organisms because they make an organism possible to begin with. Each organism can only sustain a certain amount of freeloading selfish bases, or it will pay too high a metabolic cost; 4) The groups (the genes) are certainly not completely isolated from each other. They interact and get reshuffled within an organism because of crossing over, and they

also get mixed by the process of sex in non-autogamic organisms; 5) From the fact that there are no organisms made only of GC-rich sequences we must infer that the differential fitness of genes made of high proportions of altruistic base pairs is high enough compared to the fitness of the selfish genetic elements to counterbalance selection at the individual base level. On the other hand, as remarked above, what makes this example much less fanciful than it might appear at first sight is that selection for selfish reproduction and extension of GC pairs is a likely candidate to explain the very existence of these genomic regions, the function of which has so far been highly debated.

I therefore suggest that the selfish gene—though not quite dead—needs to be drastically reduced to just one of many explanatory frameworks within evolutionary theory, and that the so-called gene's eye view represents just one of many levels at which natural selection can and does act, and cannot aspire to any privileged ontological role. Evolutionary theory must be pluralist at its core if it is to go beyond the limits of the neo-Darwinian paradigm as solidified toward the middle of the 20th century (Mayr and Provine 1980; Schlichting and Pigliucci 1998).

Chapter 8: The fall and rise of Dr. Pangloss: adaptationism and the Spandrels paper 20+ years later

"The world is emblematic. Parts of speech are metaphors because the whole of nature is a metaphor of the human mind." -Ralph Waldo Emerson, 1803-1882

In 1979, Gould and Lewontin published "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme" (Gould and Lewontin 1979). In it, they described "the adaptationist program" as an attempt to explain the existence and particular forms of any phenotypic trait as the result of natural selection. The paper's exotic title derived from the example with which the piece began, the second instance of metaphors to be examined in this dissertation, and one that has had—I think—an overall more beneficial impact (notwithstanding some limitations soon to be discussed) on evolutionary thinking than Dawkins' perhaps more popular analogy of the "selfish" gene discussed in Chapter 7.

Gould and Lewontin noted that the tapered spaces (the 'spandrels') between the archways supporting the domed roof of the basilica of St. Mark in Venice were beautifully decorated in a way that made admirable use of the triangular space. They claimed that while this space was put to great artistic use, it was an architectural byproduct of employing arches to support a domed room, not designed for that artistic use. The lesson they thought should be drawn from this was that when faced with a 'trait' that is being put to good use, one ought not to jump to the conclusion that that

particular use is the *reason* the trait is there. Gould and Lewontin argued that a particularly sloppy form of evolutionary thinking did just that, and that its practitioners went so far as assuming that any trait *had* to have a good use that explained its presence, even if none was obvious. Gould and Lewontin proceeded to criticize the assumptions that they thought such a program depended upon (**Box 8.1**), and laid out an alternative approach they felt better accounted for the complexities of the evolutionary process.

Over the more than twenty years since its publication, *Spandrels* has been much cited and criticized. Criticisms of it have ranged from the significant (for example, that the role of constraints in evolution is overstated: Rose and Lauder 1996), to the irrelevant (that the architectural feature they referred to is really called a 'pendentive,' not a 'spandrel': Mark 1996). Today, it is commonplace to pay lip service to the sorts of difficulties involved in demonstrating that a trait is 'adaptive' that Gould and Lewontin pointed out. More than two decades is a long enough time for the dust to have settled and ask to what extent the

Box 8.1. The objects of the attack: according to Gould and Lewontin, the adaptationist program is characterized by the telling of stories involving natural selection, which account for the presence and particular forms of traits by reference to their hypothesized adaptive significance.

Adaptationists, according to Gould and Lewontin, assume the following:

Ontological assumptions

- Organisms can be usefully considered as assemblages of traits, the adaptive nature of each of which can be considered independently of the others.
- Constraints on the power of natural selection can be assumed to be minor, so adaptation by natural selection is the proper null hypothesis for particular traits, either behavioral or physical.

Methodological practices

- Consistency between the observed trait and an explanatory story told in terms of natural selection is sufficient for the preliminary acceptance of the hypothesis that the trait is adaptive and evolved in just that way.
- The failure of one adaptive story leads immediately to the search for another story of the same sort.
- Any failure of particular traits to be optimal is accounted for by evoking 'tradeoffs' with other adaptive traits

Because of these features, adaptive stories are very easy to create, and very hard to falsify – hallmarks, Gould and Lewontin argue, of poor scientific hypotheses. Indeed, they claim that it is very difficult to ever reject the hypothesis that the trait is adaptive (in some way or other) within the context of the adaptationist program.

difficulties raised by the original paper can now be addressed. Of particular concern here is if and how conceptual and empirical advances have helped biologists to move away from adaptationist storytelling and towards testable hypotheses involving adaptation.

Whence Spandrels?

To begin with, it is worth thinking about what Gould and Lewontin's attack on the adaptationist program was motivated by. In large part, it was a reaction to two related phenomena. One was the rise of sociobiology (attributable in large part to E.O. Wilson's 1975 book by the same name), and the other was the popularization of the narrowly gene-centric approach favored by Dawkins (and defended explicitly in his 1976 book, *The Selfish Gene*; see Chapter 7). Those features which Gould and Lewontin claim describe the theory and practice of adaptationism bear directly upon the styles of reasoning used in these arenas.

As Lewontin has always been quick to point out, sociobiology has primarily been concerned with telling stories about (for example) human behavior that make such behavior out to be adaptive (Kaplan 2002). The charge is that adaptationists regard the *consistency* of the trait with the hypothesis that it is an adaptation shaped by natural selection as good evidence for the conclusion that it is. Arguments such as Dawkins', meant to demonstrate that the focus of selective forces are individual genes made visible to selection through their direct relationship with particular phenotypic features, were then an integral part of the adaptationist program. These assumptions, when applied to human behaviors, quickly yield conclusions with dramatic political and social implications (Lewontin 1992). The attack on adaptationism expressed in *Spandrels*, and which

Lewontin and Gould each pursued in many other works, would likely have been far less aggressive if all that were at stake was the adaptive significance of the variation in the color of snail-shells.

Rather than viewing organisms as collections of more or less optimized individual traits, Gould and Lewontin called for the acknowledgment of other possibilities and for predictions made on the basis of these alternative models to be compared (**Boxes 8.2** and **8.3**). They argued that this entailed recognizing the legitimacy of alternate approaches to evolution, in which constraints upon optimization through natural selection would play a more central role. Of course, Gould and Lewontin freely admitted that adaptation by natural selection has been and is a powerful force in shaping the phenotypic traits of organisms. The question they wished to raise was what kind of evidence is necessary to support the hypothesis that a trait is an adaptation formed by natural selection, and what evidence might point towards some other cause of the current state of the trait in question.

Box 8.2. Alternatives to adaptationism:

Gould and Lewontin suggested that at least the following alternative hypotheses be considered in attempts to understand the etiology of particular traits:

- No adaptation and no selection: the trait in question may be the result of genetic drift.
- Indirect selection: the trait was not the subject of selection – its features are the result of its association with another trait (which may or may not itself have been the subject of natural selection).
- Selection without adaptation: a trait may increase in frequency due to natural selection, but not be 'adaptive' as generally understood (Lewontin's example involves a resource-limited species and a genetic mutation which doubles fecundity: this does not increase the population's mean fitness, it only alters the population dynamics).
- Adaptation without selection: the trait may be adaptive, but not itself the product of selection explicitly for that particular form of the trait. Some types of phenotypic plasticity may be a special case of this, as is behavioral flexibility. Although these abilities may be selected for, the traits themselves (which are the result of the plasticity/flexibility) were not necessarily the result of selection for that form of the trait.
- Adaptation and selection, but no basis for distinguishing between adaptations: while the trait may be adaptive and have been selected for, there may be no way of distinguishing between different forms of a trait on the basis of their adaptive significance (the problem of multiple adaptive peaks).
- Adaptation and selection, but the particular adaptation represents a 'secondary' use of a trait already present for other (generally historical) reasons. This is the case to which Gould and Vrba 1982 later referred to as 'exaptation'.

Conceptual advances since the Spandrels

Since the publication of the *Spandrels* paper, theoretical evolutionary biology has made several advances that helped researchers to question a purely adaptationist approach to the study of phenotypic evolution. In particular, the role of constraints, tradeoffs, and costs in evolution has been widely discussed and generally acknowledged (Gould 1980b; Cheverud 1984; Maynard-Smith et al. 1985). Although the

Box 8.3. On constraints versus selection: which is the correct null hypothesis?

The problem with constraints, as Antonovics and van Tienderen (1991) insightfully remarked, is that it is difficult to envision what the null hypothesis is (see Chapter 5). Schlichting and Pigliucci (1998) discuss a potentially useful method that can be applied to test both spandrelism and panglossianism in some detail. This is based on the idea of **transitional probability matrices**. These matrices specify what the probability of a given evolutionary change is against the probabilities of alternative evolutionary pathways. Theoretical biologists are therefore forced to explicitly formulate their hypotheses in statistical terms. One could then in principle measure actual probabilities (e.g., from a phylogenetically informed data set) and compare the empirical data with the predictions by means of any matrix-comparison algorithm .

The examples below **(Fig. 8.1)** present possible transitional matrices for the simple case of three character states (A, B, and C) under neutral evolution or stabilizing selection. The entries in each matrix represent the probability of a given state to evolve from a certain initial condition (given in the columns) to another condition or to remain invariant (rows). In the case of neutral evolution, one might model the situation by assigning to any transition (including from a given state to the same state, i.e., no change) the same probability. Stabilizing selection could be modeled by a matrix that includes high probability of the phenotype to converge on one state (B in the example) and a low probability of any other state to evolve or being maintained. Notice that the sum of the probabilities along a given column must be one (i.e., there are no other possibilities aside the ones considered in the matrix).

1 NEUTRAL EVOLUTION

To \ From	A	B	C
A	0.33	0.33	0.33
B	0.33	0.33	0.33
C	0.33	0.33	0.33

2 DIRECTIONAL SELECTION (Favoring A)

To \ From	A	B	C
A	0.90	0.90	0.90
B	0.05	0.05	0.05
C	0.05	0.05	0.05

3 DISRUPTIVE SELECTION (Favoring A and C)

To \ From	A	B	C
A	0.45	0.45	0.45
B	0.10	0.10	0.10
C	0.45	0.45	0.45

4 STABILIZING SELECTION (Favoring B)

To \ From	A	B	C
A	0.05	0.05	0.05
B	0.90	0.90	0.90
C	0.05	0.05	0.05

Fig. 8.1-Transitional tables illustrating how complex null hypotheses concerning patterns of selection can be built in the case of multiple characters. See Box 8.3 for details.

terminology of constraints is still vague, most people seem to agree that at least genetic and developmental (sometimes referred to as epigenetic) constraints are a reality and can be both measured in practice and accounted for in theory (Barton and Partridge 2000; Wagner and Schwenk 2000; Fusco 2001; McGhee 2001; Schwenk and Wagner in press).

The common perception of the Fisherian idea that selection is an omnipowerful force that always brings a population to the maximum fitness peak available in the adaptive landscape (although Fisher's contribution was actually more subtle than that, and did not make use of the landscape metaphor: see Frank and Slatkin 1992) has been dealt a fatal blow by more recent advances in population genetics theory. Fisher's 'fundamental theorem of natural selection' (Fisher 1930), often perceived as the quintessential incarnation of the Panglossian paradigm, was originally developed for the special case of simple one locus-two allele systems (though it can be generalized for multiple alleles). However, even slightly more complex models such as two loci with epistasis, fecundity selection, linkage disequilibrium, and frequency dependence will often, albeit not necessarily, result in adaptive landscapes characterized by maladaptive evolution in which selection drives the population 'downhill' (Hartl and Clark 1989). More complex and realistic models have proven too difficult for mathematical tractability. The currently accepted generalization is that complex genotypic architectures are less, not more, amenable to being altered by natural selection (Cheverud 1988), that is, constraints are indeed a very common and inescapable feature of living systems (Baatz and Wagner 1997). The debate, however, still rages around how much genetic information is necessary to include in models of evolutionary trajectories, with population geneticists and optimality theorists discussing the possibility of a convergence of the two approaches (the so-called 'streetcar' model: Marrow and Johnstone 1996).

Another major (and slow in coming) conceptual shift in evolutionary biology theory that has contributed to a widespread questioning of the adaptationist program is the recognition of epigenesis as a central, and largely not understood, player in mediating the genotype-phenotype mapping function (Alberch 1991). Perhaps in part because of the recent advances in molecular developmental genetics (Wray 1994; Wilkins 2002), not an original component of the adaptationist program, the simplistic idea of selection acting more or less directly on genes is being set aside as a serious contender in the evolutionary arena. Even some sociobiologists have backed down from straight gene selection talk to embrace perhaps more vague, but certainly more realistic, 'epigenetic rules' (Wilson 1998). However, a new brand of sociobiology, renamed 'evolutionary psychology' seems to be retracing the whole path of mistakes of its ancestor making wild claims on the genetic basis of human behavior while ignoring the two decades of debate about the adaptationist program (Kaplan 2002).

These observations suggest that another way to avoid slipping into the adaptationist habit (suggested by Gould and Lewontin themselves) is to think not about the current state of the trait, but how, given its phylogenetic history and functional significance, that trait *should* be. Although a purely engineering approach is not very informative because it fails to account for historical pathways, once we take enough of the organism's basic developmental features into account, an analysis in terms of optimization theory can be revealing (Niklas 1997). This can then lead to the specification of appropriate hypotheses to test selective and constraint scenarios.

However, the role of constraints in evolution was probably oversold in the immediate aftermath of the *Spandrels*, perhaps as a result of the intuitive appeal of the architectural metaphor employed by Gould and Lewontin. This has led to an almost

comic proliferation of types of constraints, resulting into a rather confusing terminology and literature (Antonovics and van Tienderen 1991; van Tienderen and Antonovics 1994). Schlichting and Pigliucci (1998, chapter 6) have attempted to simplify and clarify the issue by proposing that all categories of constraints (with the exception of absolute constraints imposed by fundamental laws of physics) can be considered genetic or epigenetic in nature. This leaves selection and constraints as the two major deterministic players on the evolutionary stage, sometimes in opposition, other times working in concert to push a population in one direction or another of the adaptive landscape. It is this synthesis of constraints (spandrelism) and selection (panglossianism) that is key to a more sober and realistic understanding of phenotypic evolution.

The other factor that has been increasingly taken into consideration by theoretical evolutionary biologists in recent times is the fact that the environment (and therefore the adaptive landscape) is far from constant. While models of environmental heterogeneity had been proposed before the Spandrels paper (Levene 1953), the effects of environmental heterogeneity and the evolutionary strategies or outcomes that it can elicit have been investigated with a wider range of theoretical tools, including quantitative genetic and optimization models (e.g., Gabriel and Lynch 1992; Zhivotovsky et al. 1996). The general agreement is that the external environment poses as much of a limitation to adaptive evolution as internal (genetic-developmental, or epigenetic) constraints, and that they both constitute moving targets for natural selection (Schlichting and Pigliucci 1998). This, of course, does not mean that organisms do not evolve in response to selection, but the currently available theory leads us away from the idea of survival of the fittest and toward a model of survival of the "barely tolerable," analogous to the concept of 'satisficing' in foraging theory(Ward 1992).

Some of these ideas were actually expressed just before the publication of the *Spandrels* by Francois Jacob (1977), who presented a model of evolution as 'tinkering' (*bricoleur*, in the original French). According to Jacob, natural selection only works with the materials available and within the constraints present at a particular time in a particular place (see the metaphor of the builder using stones that fall from a cliffside in Darwin 1868). Jacob may in fact have struck closest to the balance between spandrelism and adaptationism advocated in this chapter.

Empirical advances since the Spandrels

Since the publication of the *Spandrels*, the roles played by constraints in evolution have been investigated empirically. We are now slowly gaining data on how common constraints actually are, and on how they evolve when studied in the context of phylogenetic hypotheses (Pigliucci and Preston in press). Of course, one must keep in mind that the fallacy of spandrelism is as easy to commit as that of adaptationism. Gould, in particular, has posited constraints in the land snail *Cerion* (Gould 1989), whereas a computer simulation of the shell morphospace has actually indicated that such space is essentially saturated, with no evidence of a limitation on form (Stone 1996).

The empirical study of constraints has been approached from different perspectives, but one of the most intriguing from an evolutionary standpoint focuses on research on the stability (or lack thereof) of genetic variance-covariance matrices. An example of this is the work of Conner and Via (1992) on natural selection on body size in the flour beetle *Tribolium*. They estimated the intensity and direction of selection, concluding that it would favor an increase in pupal weight and a decrease in width in male

beetles. However, they detected a positive genetic covariance between the two characters. When they used this information in a model predicting the likely evolutionary response of the population, they obtained a small change in pupal weight (despite the strong selection coefficient) and a large increase in width (despite selection acting in the opposite direction). The stability of variance-covariance matrices is currently a hotly disputed field of inquiry (e.g., Brodie III 1993; Shaw et al. 1995; Johnson and Wade 1996; Roff and Mousseau 1999; Roff 2000; Begin and Roff 2001; Phillips et al. 2001; Steppan et al. 2002).

Even though constraints exist, is it possible to somehow 'break' them? In other words, how stable are genetic covariances through evolutionary time? Several approaches have been used to answer this question. Steppan (1997), for example, has reconstructed the evolution of variance-covariance matrices using a phylogenetic hypothesis describing the historical branching of the leaf-eared mice of the genus *Phyllotis*. He concluded that the matrix describing the evolution of cranial characters is variable within species, and that this variation sorts out among taxa at higher phylogenetic levels. Steppan therefore concludes that there is 'no support [for] attempts to extrapolate structure to explain or predict macroevolutionary change'. However, experimental manipulations of genetic matrices by mutation-selection studies have also pointed towards the stability of at least some constraints in the crucifer plant *Arabidopsis thaliana* (Camara and Pigliucci 1999; Camara et al. 2000).

Immediately after the publication of the *Spandrels*, Lande and Arnold (1983) developed a conceptually straightforward way to measure natural selection and determine its magnitude and type. The idea is to use a regression analysis of the trait(s) of interest against a reasonable measure of fitness. Many variations and improvements on this theme

have been proposed since, including the use of genotypic rather than phenotypic data (Rausher 1992; Stinchcombe et al. 2002), as well as path analysis (Crespi and Bookstein 1989; Scheiner et al. 2000) and non parametric techniques Schluter and Nychka 1994). Of course, demonstrating selection currently acting on a character is not the same as demonstrating that the character is an adaptation (in the historical process sense of the term: Gould and Vrba 1982), yet it is a necessary step to build the case for current adaptive value.

Such a case can then be expanded by the use of manipulative experiments, both under controlled and field conditions. These experiments have a long and venerable history in evolutionary biology (Clausen et al. 1940), and can be used to bridge the gap between simply demonstrating the action of selection and actually identifying the causes underlying the observed selection gradients. Along similar lines, experimental evolution research has been conducted on a series of model organisms (Travisano et al. 1995; Travisano 1997). In these cases, one starts out with a base population that is subjected to a novel environment. The evolutionary trajectories that yield an increased adaptation to that environment are then followed and compared with predictions based on a priori functional and quantitative genetic models. All of these approaches are yielding more satisfactory insights into the dynamics and constraints of adaptive phenotypic evolution. This information is critical to assessing the actual balance between non-adaptive forces and selection in natural populations.

The new synthesis between spandrelism and adaptationism

The difficulty of assessing where the true middle lies between spandrelism and panglossianism is to be found in the fact that biology is partly an experimental science, but partly an historical one (Chapter 1). It is this latter component that makes it difficult, albeit not impossible, to deduce the underlying processes by observing the patterns they produce, because similar patterns can be the outcome of different processes. For example, I have mentioned the study of natural selection by the use of regression techniques. Even when successful, such techniques tell us little or nothing about the *nature* of the observed selection (Stratton and Bennington 1996). What aspects of the environment are exerting the selective force? How does variation in the genetic architecture of the traits under selection allow the population to respond? What other, unmeasured, traits might be driving the dynamics of the system? The answers to these questions do not come easily, and they can only be sought through a complex feedback between field studies and manipulative experiments, combined with mechanistic knowledge of the epigenetics of the traits themselves.

From the standpoint of theory, quantitative genetics and optimization theories have contributed substantially to our understanding of the interplay between selection and constraints. However, neither approach is without severe limitations (Pigliucci and Schlichting 1997). As far as quantitative genetics is concerned, several authors have demonstrated that the relationship between the observable genetic variance-covariance matrix and the underlying genetic architecture is not as simple as it was once thought (Houle 1991; Gromko 1995). One cannot deduce the existence (or non existence) of a constraint simply because a genetic correlation has been observed (or not). Optimization

theory has shown its own limits by demonstrating that the more complex and 'multitasking' a phenotype is, the more likely one is to find many alternative, equally 'optimal' designs (Niklas 1997). While these limits are arguably a fact of life, not a shortcoming of our theories, it is clear that our current mathematical tools are sufficient to highlight them, but not to provide a comprehensive theory of evolution on complex adaptive landscapes.

These, of course, are not reasons to despair, just inspiration for more work to come. However, the considerations submitted here do underscore not only that the world is neither made wholly of spandrels nor of panglossianisms, but that simply saying that the true answer lies in the proverbial middle is not really saying much. Unless one is prepared to roll up one's sleeves and thoroughly investigate that middle ground for decades to come.

Conclusions

So, what does all of this amount to? After having considered in some detail eight cases of normally unexamined philosophical baggage that accompanies research in biology, it is worth pausing to draw whatever general lessons can be evinced from the consideration of such case studies. Clearly, the message of this dissertation is not that scientists should avoid thinking by analogy, using metaphors, or making more or less explicit assumptions underlying their work. That would not only be counterproductive, given the positive power of these approaches, but indeed impossible to do because of how human beings naturally think.

Rather, the underlying message throughout the previous chapters has been that science has much to benefit from philosophical analysis, if only scientists would pay more attention to it, give more credits to philosophers, and perhaps even spend a bit more time reading the philosophical literature now and then. There are, of course, scientists that take philosophy seriously, or even engage personally in philosophical analyses. Most certainly Richard Lewontin is the quintessential example, but so are people like Stephen Gould or David Hull. Indeed, since I have started to frequent professional meetings devoted to philosophy of science I have discovered that several of my biologist colleagues go there as well, more often than not presenting their own papers in the pursuit of a somewhat parallel (and often not highly publicized to other biologists) career. Their papers, like this dissertation, tend to focus on the areas of philosophy of science that are most closely related to their own interests as empirical scientists, as it is logical that it should be. Some contributions by folks like Kevin de Queiroz could equally be published

in biology or philosophy journals, with little change in the emphasis of the points made or in the conceptual layout of the paper. And I mean this as a sincere compliment.

On the other hand, it would certainly be unfair -- and indeed downright counterproductive -- to ask the majority of scientists to become part-time philosophers, or to regularly devote part of their time and resources to philosophical inquiry, even if limited to the narrow confines of their own scientific interests. By the same token, of course, philosophers are not required to actually spend time in a laboratory and *do* research in science in order to be able to think about what scientists do. However, here the symmetry breaks down and a curious preconception has settled on us during the past couple of centuries: on the one hand, it is (rightly) expected that philosophers who wish to talk about science should devote a significant amount of time to *learn* science; on the other hand, it is somehow all right for scientists to go on with their daily business under the (philosophically highly untenable) assumption that science can do without an examination of its philosophy.

I am referring here, of course, to the well known -- and often hotly debated -- divide between the "two cultures," as C.P. Snow (1959) famously referred to them. The divide was not always in place, and it is in fact a rather recent feature of the human intellectual landscape. Until the 18th century, there really was no distinction between philosophy and science, the latter often referred to as "natural philosophy." People we think of today as scientists, such as Galilei or Newton, thought of themselves as philosophers. Moreover, some philosophers who lived in the age when science and philosophy were beginning to diverge, like Descartes, thought of themselves as scientists.[17] It was only when science began to turn into an increasingly specialized set of

[17] Descartes (1632) was more concerned with the acceptance of his theories of physics (which turned out to be wrong) than with his metaphysics, which he saw mostly as a vehicle to introduce and support his science.

disciplines, toward the end of the 19th century, that the two old companions parted ways completely. As late as 1859, Charles Darwin could devote a chapter of his *Origins of Species* to a refutation of William Paley's (1831) largely metaphysical arguments inferring intelligent design as the cause of the complexity of biological structures.

Those time have passed, and in an important sense the separation of science and philosophy has marked the maturation of both, albeit in very different senses. I am certainly not advocating an anachronistic reunification as the way to further progress. However, I hope that the examples we examined in some details strongly suggest that science can benefit from philosophical analysis. The opposite has certainly been true for some time, and is increasingly recognized by philosophers, notwithstanding the warnings and cautionary statements about scientism of scholars such as Sorell (1991). For example, discussions of ethics have greatly benefited from a variety of inputs originating from several fields of science. To mention but a few, the emerging field of the neurobiology of moral decision-making (e.g., Adams 2000; de Oliveira-Souza 2000), the comparative study of moral behavior in primates that are relatively closely related to humans (e.g., Katz 2000), the application of game theoretical modeling to the understanding of the dynamics characterizing ethical dilemmas (e.g., Sober and Wilson 1998; Skyrms 2000), and the comparative anthropology of moral practices and their consequences in different human societies (e.g., Turner 1998; Nie 2000). Entire areas of philosophizing, it seems, can no longer afford (if they ever did) to ignore science. This is not because science can solve or somehow subsume philosophical questions (as Wilson [1998] has incorrectly assumed in one of the most bold recent attempts by a scientist at "cultural colonization" of nonscientific fields). Rather, if philosophers wish to engage in discussions of matters of concern to the real world and to human beings, they are much better off doing that on the

basis of the best understanding of such world that we can afford, namely the one provided by science.

Scientists, however, have been much more recalcitrant than philosophers to accept the other side of the same coin. Perhaps affected by the *hubris* that understandably accompanies the rapid development of a new field that in short succession uncovers major principles underlying the laws of the physical universe, the diversity of biological organisms, and the nature of genetic information to name but a few, most scientists seem convinced that philosophy is of only passing interest, no more than a relict from the past, destined to be made irrelevant by the fast pace of further scientific discovery. But this would be a naive mistake, for a variety of reasons.

First, and most obviously, science simply cannot even begin to work unless it is willing to make an assortment of metaphysical and more **generally philosophical assumptions**. Some of these are so second nature to us that we hardly give it a thought. For example, we have to assume both *realism* (that there actually is a "real" world out there to be investigated) and *naturalism* (that that real world can be explained by using only natural laws, with no necessity for supernatural or metaphysical explanations). Closer to the everyday practice of the scientist, we trust in the working of induction as a fundamental way of making inferences about the natural world from our observations and experiments. None of this is justifiable empirically or from within the purview of science. It may not even been entirely justifiable philosophically, but philosophy is the discipline that is best positioned to analyze the validity and reasonability of such fundamental assumptions that science has to take for granted. Surely it is comforting for the scientist to know that somebody else is checking on the solidity of the foundations of the edifice on which science is based.

Second, recall my distinction in the Introduction among the three major purposes of philosophy of science: to inquire into the nature of science itself, to engage into a conceptual analysis of what scientists say and how they use their key terms, and to challenge the validity of specific scientific claims based on in-depth criticism of the strength of the connection between the available empirical evidence and the general conclusions that scientists wish to derive from it.

In terms of the **broad inquiry** that philosophers like Popper (1968), Kuhn (1970), or Lakatos (1977) have pursued, it seems at least a matter of intellectual curiosity for scientists to learn a bit about what they are doing as seen from the outside, from a level of meta-analysis that is made possible by philosophy but could not be undertook from within science itself. More importantly, though, an at least occasional reflection on how science proceeds, of its strengths and weaknesses, it seems to me should be part of the training of any graduate student in the sciences. A better understanding of why what one does works (or doesn't work) should make for better science in general.

When we get closer to specific practices within science, the relevance of philosophy to scientists increases accordingly. So the second level of inquiry, that of **conceptual analysis of the use of key terms in science** -- such as species, cause, natural selection, force, etc. -- should again at least occasionally be brought to the attention of the practicing scientist to see if and to what extent that sort of reflection may have practical consequences. As I have tried to argue, for example, in Chapter 4 (on species concepts), a lot of time and resources could have been better employed if scientists had paid more attention to Wittgenstenian ideas such as cluster concepts and how they may apply to the discussion of what species "really" are. When I have presented my ideas on this specific subject to biologist colleagues, an often-made remark

is that this is all very interesting, but "it will not help taxonomists in the field." True, nor is it meant to do that. However, a clearer conceptual understanding of what species are should have an impact on the usage of the term in the biological literature, and in particular in the theoretical framework that has developed around species concepts and the process of speciation. This, in turn, has very tangible effects on how empirical evolutionary biologists go about studying species and speciation. The taxonomic level, the pragmatic aspect of actually recognizing closely related species from each other, is the last in a long chain of interesting problems, and obviously the most remote from philosophical input. But this does not mean that philosophical analysis does not affect *any* of the other levels of concern to evolutionary biologists.

Finally, we get to what I have termed "**science criticism**" in the Introduction, the level of philosophical analysis most close to the interests of practicing biologists. Typical examples here, it will be recalled, are criticisms of specific research claims made by evolutionary psychologists or, in the case of my work in this dissertation, surrounding the nature-nurture debate about human cognitive abilities (Chapter 6). It has been my experience that science criticism is both where philosophers can make the most direct (though not necessarily most important) contributions to science, and where scientists tend to display the most skepticism or outright rejection of philosophical insights. Again, my direct experience with science colleagues helps understanding the situation here. The most common remarks about science criticism by scientists I have talked to are "what do they [the philosophers] know about it?" and "why couldn't a scientist do that as well?" It is important to ponder these questions in some detail.

The answer to the first question is that it simply reflects a prejudice: while occasionally one can catch philosophers of science who talk about specific scientific

issues without enough technical understanding of the issues themselves, this is becoming more and more rare. Philosophy of science is an increasingly specialized field, so much so that there are few people who even use the label "philosopher of science," or even "philosopher of physics." We are getting to the point of having researchers who work in philosophy of quantum mechanics, or of specialized fields of mathematics. This is of course a reflection of the fact that science itself keeps getting more and more technical and specialized, so that if somebody wishes to engage in conceptual analysis of a given scientific field, one has to be prepared to understand quite a bit of the intricacies of that field. Indeed, sometime the works of philosophers of quantum mechanics or mathematics are difficult to distinguish from theoretical papers written in the same fields by physicists or mathematicians. While this is not (yet?) true of biology, that field is also characterized now by such a vast and ever increasing literature that few philosophers dare engaging in it without a serious understanding of that literature.

The second question, "could this not have been done by a scientist?" has two simple answers: on the one hand, yes, it could have, because scientists presumably have the same general intelligence and ability to conceptualize that philosophers have. On the other hand, no, because scientists -- rather ironically -- are not explicitly trained in critical thinking at all.[18] Moreover, they simply don't have *the time* to engage in this sort of analysis because, unlike philosophers, they have to get experiments running and grants funded. This is simply the reality of how modern science works, regardless of any idealistic (or moralistic) vision of how it *should* work. Of course scientists engage in critical peer review of each other's papers, and of course every active laboratory runs a

[18] This point would lead us too far afield if pursued any further, but there is empirical evidence from the literature on science education that teaching science does not automatically translate into better critical thinking, as it is often assumed, and it could even be *detrimental* to one's critical thinking abilities: Ede 2000; Goode 2002; Walker and Hoekstra 2002.

journal club devoted to critically dissecting papers in that lab's field. But these activities tend to be focused on the technical details of the works being analyzed, and even when they get conceptual they assume or take for granted a lot more than a philosopher looking from the outside would. This localized approach to peer review works well most of the time, but occasionally we as scientists do miss the forest because we are paying too much attention to the individual trees. Philosophers can help by occasionally reminding us of what the entire forest looks like from a distance.

In the end, what I have been advocating explicitly in the Introduction and Conclusion to this work, and implicitly in my analysis of specific case studies in the main body of this dissertation, is that both philosophers and scientists have a lot to gain from interacting with each other, and ironically scientists are likely to benefit the most. This interaction should be fostered by a variety of means, including truly interdisciplinary research projects and teaching efforts, not in order to seek an artificial reconciliation between the "two cultures," but simply and pragmatically to combine forces to better understand the world in which we all live. After all, as philosopher Edmund Husserl wrote (in *The Crisis of European Sciences and Transcendental Philosophy*), "Merely fact-minded sciences make merely fact-minded people," to which we may juxtapose scientist Henri Poincare`'s conclusion (in *Science and Hypothesis*) that "Science is built up with facts, as a house is with stones. But a collection of facts is no more a science than a heap of stones is a house."

Literature Cited

Abelson, R.P. *Statistics as Principled Argument*. Hillsdale, NJ: Lawrence Erlbaum, 1995.

Ackermann, M., R. Bijlsma, A.C. James, L. Partridge, B.J. Zwaan, and S.C. Stearns. "Effects of Essay Conditions in Life History Experiments with *Drosophila Melanogaster*." *Journal of Evolutionary Biology* 14 (2001): 199-209.

Adams, J. *Gray Matter and Gray Areas: The Neuroscience of Morality* [web page]. brain.com, 2000 [cited 8/25/00. Available from www.brain.com.

Alberch, P. "From Genes to Phenotypes: Dynamical Systems and Evolvability." *Genetica* 84 (1991): 5-11.

Anderson, D.R., K.P. Burnham, and W.L. Thompson. "Null Hypothesis Testing: Problems, Prevalence, and an Alternative." *Journal of Wildlife Management* 64 (2000): 912-23.

Andreasen, R.O. "A New Perspective on the Race Debate." *British Journal of Philosophy of Science* 49 (1998): 199-225.

Antonovics, J., and P.H. van-Tienderen. "Ontoecogenophyloconstraints? The Chaos of Constraint Terminology." *Trends in Ecology and Evolution* 6 (1991): 166-68.

Appiah, K.A. "Race, Culture, Identity: Misunderstood Connections." In *Color Conscious: The Political Morality of Race*, edited by K.A. Appiah and A. Gutman, 30-105. Princeton, NJ: Princeton University Press, 1996.

Armstrong, D. "What Is Consciousness?" In *The Nature of Consciousness*, edited by N. Block, O. Flanegan and G. Guzeldere, 721-28. Cambridge, MA: MIT Press, 1997.

Asins, M.J. "Present and Future of Quantitative Trait Locus Analysis in Plant Breeding." *Plant Breeding* 121 (2002): 281-91.

Atchley, W.R., and B.K. Hall. "A Model for Development and Evolution of Complex

Morphological Structures." *Biological Review* 66 (1991): 101-57.

Baatz, M., and G.P. Wagner. "Adaptive Inertia Caused by Hidden Pleiotropic Effects." *Theoretical Population Biology* 51 (1997): 49-66.

Bacon, F. *Novum Organum*. Cambridge, England: Cambridge University Press, 1620 / 2000.

Bak, P., and K. Chen. "Self-Organized Criticality." *Scientific American* January (1991): 46-53.

Barton, N., and L. Partridge. "Limits to Natural Selection." *BioEssays* 22 (2000): 1075-84.

Basten, C.J., B.S. Weir, and Z.-B. Zeng. *Qtl Cartographer, Version 1.14*. Raleigh, NC: North Carolina State University, 2000.

Begin, M., and D.A. Roff. "An Analysis of G Matrix Variation in Two Closely Related Cricket Species, *Gryllus Firmus* and *G. Pennsylvanicus*." *Journal of Evolutionary Biology* 14 (2001): 1-13.

Berlin, I. "The Purpose of Philosophy." In *Categories and Concepts*, 1-11. New York: Viking Press, 1979.

Berry, A. , and M. Kreitman. "Molecular Analysis of an Allozyme Cline: Alcohol Dehydrogenase in *Drosophila Melanogaster* on the East Coast of North America." *Genetics* 134 (1993): 869-93.

Boyd, R.S., and S.N. Martens. "Nickel Hyperaccumulation in *Thlaspi Montanum* Var. *Montanum* (Brassicaceae): A Constitutive Trait." *American Journal of Botany* 85 (1998): 259-65.

Bradshaw, A.D. "Evolutionary Significance of Phenotypic Plasticity in Plants." *Advances in Genetics* 13 (1965): 115-55.

Bradshaw, A.D., and K. Hardwick. “Evolution and Stress - Genotypic and Phenotypic Components.” *Biological Journal of the Linnean Society* 37 (1989): 137-55.

Brodie III, E.D. “Homogeneity of the Genetic Variance-Covariance Matrix for Antipredator Traits in Two Natural Populations of the Garter Snake *Thamnophis Ordinoides*.” *Evolution* 47 (1993): 844-54.

Burian, R.M. “Comments on Complexity and Experimentation in Biology.” *Philosophy of Science* 64 (1997): S279-S91.

Buss, L.W. *The Evolution of Individuality*. Princeton: Princeton University Press, 1987.

Callahan, H. , and M. Pigliucci. “Shade-Induced Plasticity and Its Ecological Significance in Wild Populations of *Arabidopsis Thaliana*.” *Ecology* 83 (2002): 1965-80.

Callahan, H.S., C.L. Wells, and M. Pigliucci. “Light-Sensitive Plasticity Genes in *Arabidopsis Thaliana*: Mutant Analysis and Ecological Genetics.” *Evolutionary Ecology Research* 1 (1999): 731-51.

Camacho, J.P.M., T.F. Sharbel, and L.W. Beukeboom. “B-Chromosome Evolution.” *Philosophical Transactions of the Royal Society of London* B355 (2000): 163-78.

Camara, M., and M. Pigliucci. “Mutational Contributions to Genetic Variance/Covariance Matrices: An Experimental Approach Using Induced Mutations in *Arabidopsis Thaliana*.” *Evolution* 53 (1999): 1692-703.

Camara, M.D., C.A. Ancell, and M. Pigliucci. “Induced Mutations: A Novel Tool to Study Phenotypic Integration and Evolutionary Constraints in *Arabidopsis Thaliana*.” *Evolutionary Ecology Research* 2 (2000): 1009-29.

Cardinal, R.N., D.R. Pennicott, C.L. Sugathapala, T.W. Robbins, and B.J. Everitt. “Impulsive Choice Induced in Rats by Lesions of the Nucleus Accumbens Core.” *Science* 292 (2001): 2499-501.

Casal, J.J. “Coupling of Phytochrome B to the Control of Hypocotyl Growth in *Arabidopsis*.” *Planta* 196 (1995): 23-29.

Casal, J.J., and M.A. Mazzella. “Conditional Synergism between Cryptochrome 1 and Phytochrome B Is Shown by the Analysis of Phya, Phyb, and Hy4 Simple, Double, and Triple Mutants in *Arabidopsis*.” *Plant Physiology* 118 (1998): 19-25.

Casal, J.J., and H. Smith. “The Function, Action and Adaptive Significance of Phytochrome in Light-Grown Plants.” *Plant, Cell and Environment* 12 (1989): 855-62.

Cavalli-Sforza, L. L., P. Menozzi, and A. Piazza. *The History and Geography of Human Genes*. Princeton, NJ: Princeton University Press, 1994.

Cavalli-Sforza, L.L., and F. Cavalli-Sforza. *The Great Human Diasporas*. Reading, MA: Addison-Wesley, 1995.

Chamberlain, T.C. “The Method of Multiple Working Hypotheses.” *Science* 15 (1897): 92-96.

Chechetkin, V.R., and V.V. Lobzin. “Study of Correlations in Segmented DNA Sequences: Application to Structural Coupling between Exons and Introns.” *Journal of Theoretical Biology* 190 (1998): 69-83.

Cheverud, J.M. “The Evolution of Genetic Correlation and Developmental Constraints.” In *Population Genetics and Evolution*, edited by G. de Jong, 94-101. Berlin: Springer-Verlag, 1988.

———. “Quantitative Genetics and Developmental Constraints on Evolution by Selection.” *Journal of Theoretical Biology* 110 (1984): 155-71.

Cheverud, J.M., and E.J. Routman. “Epistasis and Its Contribution to Genetic Variance Components.” *Genetics* 139 (1995): 1455-61.

Chory, J., and J. Li. "Gibberellins, Brassinosteroids and Light-Regulated Development." *Plant, Cell and Environment* 20 (1997): 801-06.

Churchland, P. "Eliminative Materialism." In *Matter and Consciousness*, 43-49. Cambridge, MA: MIT Press, 1988.

Claridge, M.F., H.A. Dawah, and M.R. Wilson, eds. *Species: The Units of Biodiversity*. London: Chapman & Hall, 1997.

Clarke, G.M. "The Genetic Basis of Developmental Stability. I. Relationship between Stability, Heterozygosity and Genomic Coadaptation." *Genetica* 89 (1993): 15-23.

Clausen, J., D. Keck, and W.M. Hiesey. *Experimental Studies on the Nature of Plant Species. I. Effect of Varied Environment on Western North American Plants*. Washington, DC: Carnegie Institution, 1940.

Clegg, M.T., and M.L. Durbin. "Flower Color Variation: A Model for the Experimental Study of Evolution." *Proceedings of the National Academy of Science USA* 97 (2000): 7016-23.

Clemens, S. "Molecular Mechanisms of Plant Metal Tolerance and Homeostasis." *Planta* 212 (2001): 475-86.

Cohen, J. "The Earth Is Round ($P<0.05$)." *American Psychologist* 49 (1994): 997-1003.

Conner, J., and S. Via. "Natural Selection on Body Size in *Tribolium*: Possible Genetic Constraints on Adaptive Evolution." *Heredity* 69 (1992): 73-83.

Cooper, R.M., and J.P. Zubek. "Effects of Enriched and Restricted Early Environments on the Learning Ability of Bright and Dull Rats." *Canadian Journal of Psychology* 12 (1958): 159-64.

Crespi, B.J., and F.L. Bookstein. "A Path-Analytic Model for the Measurement of Selection on Morphology." *Evolution* 43 (1989): 18-28.

Cronbach, L.J. “Beyond the Two Disciplines of Scientific Psychology.” *American Psychologist* 30 (1975): 116-27.

Cronquist, A. “Once Again, What Is a Species?” In *Biosystematics in Agriculture*, edited by J.A. Ramberger, 3-20. Monclair, NJ: Allanheld and Osmun, 1978.

Culp, S. “Establishing Genotype/Phenotype Relationships: Gene Targeting as an Experimental Approach.” *Philosophy of Science* 64 (1997): S268-S78.

Damasio, A. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York, NY: Harcourt Brace & Co., 1999.

Darwin, C. *The Origin of Species by Means of Natural Selection: Or, the Preservation of Favored Races in the Struggle for Life*. New York, NY: A.L. Burt (1910), 1859.

———. *The Variation of Animals and Plants under Domestication*. New York, NY: D. Appleton and company, 1900, 1868.

Davies, P.S. “Troubles for Direct Proper Functions.” *Nous* 28 (1994): 363-81.

Dawkins, R. *The Extended Phenotype*. San Francisco, CA: Freeman, 1982.

———. *The Selfish Gene*. New York, NY: Oxford University Press, 1976.

de Oliveira-Souza, R. “The Moral Brain: A Functional Mri Study of Moral Judgment.” *Neurology* 54 (2000): A104.

de Queiroz, K. “The General Lineage Concept of Species and the Defining Properties of the Species Category.” In *Species: New Interdisciplinary Essays*, edited by R.A. Wilson, 49-89. Cambridge, MA: MIT Press, 1999.

———. “The General Lineage Concept of Species, Species Criteria, and the Process of Speciation: A Conceptual Unification and Terminological Recommendations.” In *Endless Forms: Species and Speciation*, edited by D.J. Howard and S.H. Berlocher, 57-75. Oxford, England: Oxford University Press, 1998.

———. “Phylogenetic Definitions and Taxonomic Philosophy.” *Biology and Philosophy* 7 (1992): 295-313.

Dennett, D. *Darwin's Dangerous Idea*. New York, NY: Simon & Schuster, 1995.

DeSalle, R., A. Templeton, I. Mori, S. Pletscher, and J.S. Johnson. “Temporal and Spatial Heterogeneity of Mtdna Polymorphisms in Natural Populations of *Drosophila Mercatorum*.” *Genetics* 116 (1987): 215-23.

Descartes, R. *Meditations*. New York: Liberal Arts Press, 1632 /1951.

Diamond, J. *Guns, Germs, and Steel*. New York, NY: Norton, 1997.

Dixon, P., and T. O'Reilly. “Scientific Versus Statistical Inference.” *Canadian Journal of Experimental Psychology* 53 (1999): 133-49.

Dorn, L.A., E.H. Pyle, and J. Schmitt. “Plasticity to Light Cues and Resources in *Arabidopsis Thaliana*: Testing for Adaptive Value and Costs.” *Evolution* 54 (2000): 1982-94.

Dudley, S.A., and J. Schmitt. “Testing the Adaptive Plasticity Hypothesis: Density-Dependent Selection on Manipulated Stem Length in *Impatiens Capensis*.” *The American Naturalist* 147 (1996): 445-65.

Duncan, J., R.J. Seitz, J. Kolodny, D. Bor, H. Herzog, and et al. “A Neural Basis for General Intelligence.” *Science* 289 (2000): 457-60.

Dupré, J. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, MA: Harvard University Press, 1993.

Ede, A. “Has Science Education Become an Enemy of Scientific Rationality?” *Skeptical Inquirer*, Jul/Aug 2000, 48-51.

Entine, J. *Taboo: Why Black Athletes Dominate Sports and Why We're Afraid to Talk About It*. New York, NY: Public Affairs, 2000.

Fatyol, K., K. Illes, D.C. Diamond, C. Janish, and A.A. Szalay. “Mer22-Related Sequence Elements Form Pericentric Repetitive DNA Families in Primates.” *Molecular and General Genetics* 262 (2001): 931-39.

Fedlman, J. “Minimization of Boolean Complexity in Human Concept Learning.” *Nature* 407 (2000): 630-33.

Feldman, M.W., and R.C. Lewontin. “The Heritability Hang-Up.” *Science* 190 (1990): 1163-68.

Fisher, R.A. *The Genetical Theory of Natural Selection*. Oxford: Clarendon, 1930.

Frank, S.A., and M. Slatkin. “Fisher's Fundamental Theorem of Natural Selection.” *Trends in Ecology & Evolution* 7 (1992): 92-95.

Fusco, G. “How Many Processes Are Responsible for Phenotypic Evolution?” *Evolution & Development* 3 (2001): 279-86.

Futuyma, D. *Evolutionary Biology*. Sunderland, MA: Sinauer, 1998.

Gabriel, W., and M. Lynch. “The Selective Advantage of Reaction Norms for Environmental Tolerance.” *Journal of Evolutionary Biology* 5 (1992): 41-59.

Ganfornina, M.D. , and D. Sanchez. “Generation of Evolutionary Novelty by Functional Shift.” *BioEssays* 21 (1999): 432-39.

Gazzaniga, M.S. “Cerebral Specialization and Interhemispheric Communication. Does the Corpus Callosum Enable the Human Condition?” *Brain* 123 (2000): 1293-326.

———. *The Mind's Past*. Berkeley, CA: University of California Press, 1998.

Gergen, K.J. “Social Psychology as History.” *Journal of Personality and Social Psychology* 26 (1973): 309-20.

Giray, E. “An Integrated Biological Approach to the Species Problem.” *British Journal for the Philosophy of Science* 27 (1976): 317-28.

Givnish, T.J. “On the Adaptive Significance of Leaf Height in Forest Herbs.” *The American Naturalist* 120 (1982): 353-81.

Gleason, J.M., E.C. Griffith, and J.R. Powell. “A Molecular Phylogeny of the *Drosophila Willistoni* Group: Conflicts between Species Concepts?” *Evolution* 52 (1998): 1093-103.

Godfrey-Smith, P. “A Modern History Theory of Functions.” *Nous* 28 (1994): 344-62.

Goode, E. “Education, Scientific Knowledge, and Belief in the Paranormal.” *Skeptical Inquirer*, Jan/Feb 2002, 24-27.

Gotelli, N.J., and G.R. Graves. *Null Models in Ecology*. Chicago: University of Chicago Press, 1996.

Gould, S.J. “Covariance Sets and Ordered Geographic Variation in *Cerion* from Aruba, Bonaire and Curacao: A Way of Studying Nonadaptation.” *Systematic Zoology* 33 (1984): 217-37.

———. “A Developmental Constraint in *Cerion*, with Comments on the Definition and Interpretation of Constraint in Evolution.” *Evolution* 43 (1989): 516-39.

———. “The Evolutionary Biology of Constraint.” *Daedalus* 109 (1980): 39-52.

———. *The Mismeasure of Man*. New York: W.W. Norton & Co., 1996.

———. “The Promise of Paleobiology as a Nomothetic, Evolutionary Discipline.” *Paleobiology* 6 (1980): 96-118.

———. *The Structure of Evolutionary Theory*. Cambridge, MA: Harvard University Press, 2002.

Gould, S.J., and R.C. Lewontin. “The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme.” *Proceedings of the Royal Society of London* B205 (1979): 581-98.

Gould, S.J., and E.S. Vrba. “Exaptation - a Missing Term in the Science of Form.” *Paleobiology* 8 (1982): 4-15.

Grant, V. “Evolution of the Species Concept.” *Biol. Zent. bl.* 113 (1994): 401-15.

Gray, R. “Death of the Gene: Developmental Systems Strike Back.” In *Trees of Life*, edited by P. Griffiths: Kluwer Academic Publisher, 1992.

Greene, J.D., R.B. Somerville, L.E. Nystrom, J.M. Darley, and J.D. Cohen. “An Fmri Investigation of Emotional Engagement in Moral Judgment.” *Science* 293 (2001): 2105-08.

Gregson, R.A.M. “Signs of Obsolescence in Psychological Statistics: Significance Versus Contemporary Theory.” *Australian Journal of Psychology* 49 (1997): 59-63.

Griffiths, P., and R.D. Gray. “The Developmental Systems Perspective: Organism-Environment Systems as Units of Development and Evolution.” In *The Evolutionary Biology of Complex Phenotypes*, edited by M. Pigliucci and K. Preston. Oxford, England: Oxford University Press, In press.

Griffiths, P.E. “Adaptive Explanation and the Concept of a Vestige.” In *Trees of Life*, edited by P. Griffiths: Kluwer Academic Publisher, 1992.

———. “Functional Analysis and Proper Functions.” *British Journal for the Philosophy of Science* 44 (1993): 409-22.

Griffiths, P.E., and R.D. Gray. “Replicator Ii - Judgment Day.” *Biology and Philosophy* 12 (1997): 471-92.

Gromko, M.H. “Unpredictability of Correlated Response to Selection: Pleiotropy and Sampling Interact.” *Evolution* 49 (1995): 685-93.

Gusella, J.F., S. McNeil, F. Persichetti, J. Srinidhi, A. Novelletto, E. Bird, P. Faber, J.-P. Vonsattel, R.H. Myers, and M.E. McDonald. “Huntington's Disease.” *Cold Spring*

Harbor Symposia on Quantitative Biology 61 (1996): 615-26.

Halliday, K.J., M. Koornneef, and G.C. Whitelam. "Phytochrome B and at Least One Other Phytochrome Mediate the Accelerated Flowering Response of *Arabidopsis Thaliana* L. To Low Red/Far-Red Ratio." *Plant Physiology* 104 (1994): 1311-15.

Hamilton, W.D. "The Genetical Evolution of Social Behavior." *The American Naturalist* 97 (1963): 31-33.

———. "Innate Social Aptitudes of Man: An Approach from Evolutionary Genetics." In *Biosocial Anthropology*, edited by R. Fox, 133-55. New York, NY: John Wiley & Sons, 1975.

Han, Y.N., X.G. Liu, U. Benny, H.C. Kistler, and H.D. VanEtten. "Genes Determining Pathogenicity to Pea Are Clustered on a Supernumerary Chromosome in the Fungal Plant Pathogen *Nectria Haematococca*." *Plant Journal* 25 (2001): 305-14.

Hapgood, J.P., J. Riedemann, and S.D. Scherer. "Regulation of Gene Expression by Gc-Rich DNA Cis-Elements." *Cell Biology International* 25 (2001): 17-31.

Hartl, D.L., and A.G. Clark. *Principles of Population Genetics*. Sunderland, MA: Sinauer, 1989.

Hatcher, M.J. "Persistence of Selfish Genetic Elements: Population Structure and Conflict." *Trends in Ecology & Evolution* 15 (2000): 271-77.

Herrnstein, R.J., and C. Murray. *The Bell Curve: Intelligence and Class Structure in American Life*. New York, NY: Free Press, 1994.

Hey, J. *Genes, Categories and Species*. Oxford, England: Oxford University Press, 2001.

———. "The Mind of the Species Problem." *Trends in Ecology & Evolution* 16 (2001): 326-27.

Hilborn, R., and M. Mangel. *The Ecological Detective: Confronting Models with Data*.

Princeton, NJ: Princeton University Press, 1997.

Hirschfeld, L.A. “Natural Assumptions: Race, Essence, and Taxonomies of Human Kinds.” *Social Research* 65 (1998): 331-49.

Hoberman, J. “Totem and Taboo: The Myth of Race in Sports.” *Skeptic* 2000, 35-37.

Hoffmann, A.A., R. Hallas, C. Sinclair, and L. Patridge. “Rapid Loss of Stress Resistance in *Drosophila Melanogaster* under Adaptation to Laboratory Culture.” *Evolution* 55 (2001): 436-38.

Holliday, R. “From DNA Modification to Epigenetics.” *Journal of Genetics* 77 (1998): 65-69.

Horvath, C.D. “Phylogenetic Species Concept: Pluralism, Monism, and History.” *Biology and Philosophy* 12 (1997): 225-32.

Houle, D. “Genetic Covariance of Fitness Correlates: What Genetic Correlations Are Made of and Why It Matters.” *Evolution* 45 (1991): 630-48.

Howard, D.J., and S.H. Berlocher, eds. *Endless Forms: Species and Speciation*. Oxford, England: Oxford University Press, 1998.

Howson, C., and P. Urbach. *Scientific Reasoning: The Bayesian Approach*. La Salle, IL: Open Court, 1991.

Hubbell, S.P. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princenton, NJ: Princenton University Press, 2001.

Huelsenbeck, J.P., B. Rannala, and J.P. Masly. “Accomodating Phylogenetic Uncertainty in Evolutionary Studies.” *Science* 288 (2000): 2349-50.

Hull, D. “The Effect of Essentialism on Taxonomy -- Two Thousand Years of Stasis.” *British Journal for the Philosophy of Science* 15/16 (1965): 314-26 / 1-18.

———. *Science & Selection: Essays on Biological Evolution & the Philosophy of Science*.

New York, NY: Cambridge University Press, 2000.

Hull, D.L. “Species, Subspecies, and Races.” *Social Research* 65 (1998): 351-67.

Hume, D. *Enquiry Concerning Human Understanding*. Chicago, IL: Gateway Editions, 1748 / 1956.

———. *A Treatise of Human Nature*. New York, NY: Clarendon Press, 1739-40 / 1978.

Hurst, G.D.D., and J.H. Werren. “The Role of Selfish Genetic Elements in Eukaryotic Evolution.” *Nature Reviews Genetics* 2 (2001): 597-606.

Hurst, L.D. “The Evolution of Genomic Anatomy.” *Trends in Ecology and Evolution* 14 (1999): 108-12.

Jablonka, E., M. Matzke, D. Thieffry, and L. Van Speybroeck. “The Genome in Context: Biologists and Philosophers on Epigenetics.” *BioEssays* 24 (2002): 392-94.

Jacob, F. “Evolution and Tinkering.” *Science* 196 (1977): 1161-66.

Jefferys, W.H., and J.O. Berger. “Sharpening Ockam's Razor on a Bayesian Strop.” *American Scientist* 1992, 64-72.

Jiggins, C.D., R.E. Naisbit, R.L. Coe, and J. Mallet. “Reproductive Isolation Caused by Colour Pattern Mimicry.” *Nature* 411 (2001): 302-05.

Johnson, N.A., and M.J. Wade. “Genetic Covariances within and between Species: Indirect Selection for Hybrid Inviability.” *Journal of Evolutionary Biology* 9 (1996): 205-14.

Jordan, G.J., B.M. Potts, P. Chalmers, and R.J.E. Wiltshire. “Quantitative Genetic Evidence That the Timing of Vegetative Phase Change in *Eucalyptus Globulus* Ssp Globulus Is an Adaptive Trait.” *Australian Journal of Botany* 48 (2001): 561-67.

Joron, M, I.R. Wynne, G. Lamas, and J. Mallet. “Variable Selection and the Coexistence of Multiple Mimetic Forms of the Butterfly *Heliconius Numata*.” *Evolutionary*

Ecology 13 (1999): 721-54.

Kaplan, J. "Historical Evidence and Human Adaptations." *Philosophy of Science* 69 (2002): S294-S304.

———. *The Limits and Lies of Human Genetic Research: Dangers for Social Policy*. New York, NY: Routledge, 2000.

Kaplan, J., and M. Pigliucci. "Genes 'for' Phenotypes: A Modern History View." *Biology and Philosophy* 16 (2001): 189-213.

Karan, D. , and R. Parkash. "Desiccation Tolerance and Starvation Resistance Exhibit Opposite Latitudinal Clines in Indian Geographical Populations of *Drosophila Kikkawai*." *Ecological Entomology* 23 (1998): 391-96.

Kass, R.E., and A.E. Raftvery. "Bayes Factors." *Journal of the American Statistical Association* 90 (1995): 773-95.

Katz, L.D., ed. *Evolutionary Origins of Morality: Cross-Disciplinary Perspectives*. Exeter, UK: Imprint Academic, 2000.

Kauffman, S.A. *The Origins of Order*. New York: Oxford University Press, 1993.

Kazazian, H.H., and J.V. Moran. "The Impact of Li Retrotransposons on the Human Genome." *Nature Genetics* 19 (1998): 19-24.

Keightley, P.D., and H. Kacser. "Dominance, Pleiotropy and Metabolic Structure." *Genetics* 117 (1987): 319-29.

Keita, S.O.Y. , and R.A. Kittles. "The Persistence of Racial Thinking and the Myth of Racial Divergence." *American Anthropologist* 99 (1997): 534-44.

Keller, L., ed. *Levels of Selection in Evolution*. Princeton, NJ: Princeton University Press, 1999.

King, R.C., and W.D. Stansfield. *A Dictionary of Genetics*. New York, NY: Oxford

University Press, 1990.

Kingsolver, J.G., H.E. Hoekstra, J.M. Hoekstra, D. Berrigan, S.N. Vignieri, C.E. Hill, A. Hoang, P. Gibert, and P. Beerli. "The Strength of Phenotypic Selection in Natural Populations." *American Naturalist* 157 (2001): 245-61.

Kinley, T.A., and C.D. Apps. "Mortality Patterns in a Subpopulation of Endangered Mountain Caribou." *Wildlife Society Bulletin* 29 (2001): 158-64.

Kitcher, P. *The Advancement of Science: Science without Legend, Objectivity without Illusions*. New York, NY: Oxford University Press, 1995.

———. "Battling the Undead: How (and How Not) to Resist Genetic Determinism." In *Thinking About Evolution: Historical, Philosophical, and Political Perspectives*, edited by R. Singh, C. Krimbas, D. Paul and J. Beatty, 396-414. Cambridge, England: Cambridge University Press, 2001.

———. "Species." *Philosophy of Science* 51 (1984): 308-33.

Kliebenstein, D., A. Figuth, and T. Mitchell-Olds. "Genetic Architecture of Plastic Methyl Jasmonate Responses in *Arabidopsis Thaliana*." *Genetics* 161 (2002): 1685-96.

Koornneef, M., L.W.M. Dellaert, and J.H. van-der-Veen. "Ems- and Radiation-Induced Mutation Frequencies at Individual Loci in *Arabidopsis Thaliana* (L.) Heynh." *Mutation Research* 93 (1982): 109-23.

Kuhn, T. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1970.

Kusano, A., C. Staber, H.Y.E. Chan, and B. Ganetzky. "Closing the (Ran)Gap on Segregation Distortion in *Drosophila*." *BioEssays* 25 (2003): 108-15.

Lakatos, I. *The Methodology of Scientific Research Programmes*. Cambridge, England:

Cambridge University Press, 1977.

Lande, R, and G.S. Wilkinson. "Models of Sex-Ratio Meiotic Drive and Sexual Selection in Stalk-Eyed Flies." *Genetical Research* 74 (1999): 245-53.

Lande, R., and S.J. Arnold. "The Measurement of Selection on Correlated Characters." *Evolution* 37 (1983): 1210-26.

Larson, E.J., and L. Witham. "Scientists Are Still Keeping the Faith." *Nature* 386 (1997): 435-36.

Lawton, J.H. "Are There General Laws in Ecology?" *Oikos* 84 (1999): 177-92.

Levene, H. "Genetic Equilibrium When More Than One Ecological Niche Is Available." *The American Naturalist* 87 (1953): 331-33.

Levin, D.A. "Plant Outliers: An Ecogenetic Perspective." *The American Naturalist* 145 (1995): 109-18.

Levins, R., and R.C. Lewontin. *The Dialectical Biologist*. Cambridge, MA: Harvard University Press, 1985.

Levitt, N. "Why Professors Believe Weird Things: Sex, Race, and the Trials of the New Left." *Skeptic* 1998, 28-35.

Lewontin, R. C., S. Rose, and L. J. Kamin. *Not in Our Genes*. New York, NY.: Pantheon, 1984.

Lewontin, R.C. "The Analysis of Variance and the Analysis of Causes." *American Journal of Human Genetics* 26 (1974): 400-11.

———. *Biology as Ideology: The Doctrine of DNA*. New York, NY: HarperPerennial, 1992.

———. "The Evolution of Cognition: Questions We Will Never Answer." In *Methods, Models, and Conceptual Issues. An Invitation to Cognitive Science*, edited by D.

Scarborough and S. Sternberg, 107-32. Cambridge, MA: MIT Press, 1998.

———. *The Genetic Basis of Evolutionary Change*. New York, NY: Columbia University Press, 1974.

———. "Is Nature Probable or Capricious?" *BioScience* January (1966): 25-27.

Loftus, G.R. "A Picture Is Worth a Thousand *P* Values: On the Irrelevance of Hypothesis Testing in the Microcomputer Age." *Behavior Research Methods, Instruments & Computers* 25 (1993): 250-56.

Longino, H.E. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, NJ: Princeton University Press, 1990.

Lopez-Juez, E., M. Kobayashi, A. Sakurai, Y. Kamiya, and R.E. Kendrick. "Phytochrome, Gibberellins, and Hypocotyl Growth." *Plant Physiology* 107 (1995): 131-40.

Lorenz, K. *On Aggression*. New York, NY: Harcourt, Brace and World, Inc., 1963.

Loudet, O., S. Chaillou, C. Camilleri, and D. Bouchez. "Bay-0 X Shahdara Recombinant Inbred Line Population: A Powerful Tool for the Genetic Dissection of Complex Traits in *Arabidopsis*." *Theoretical and Applied Genetics* 104 (2002): 1173-84.

Lovley, D.R. "Anaerobes into Heavy Metal: Dissimilatory Metal Reduction in Anoxic Environments." *Trends in Ecology and Evolution* 8 (1993): 213-17.

Lu, G., D.J. Basley, and L. Bernatchez. "Contrasting Patterns of Mitochondrial DNA and Microsatellite Introgressive Hybridization between Lineages of Lake Whitefish (*Coregonus Clupeaformis*); Relevance for Speciation." *Molecular Ecology* 10 (2001): 965-85.

MacEachern, S. "Genes, Tribes, and African History." *Current Anthropology* 41 (2000): 357-84.

Magwene, P.M. “New Tools or Studying Integration and Modularity.” *Evolution* 55 (2001): 1734-45.

Malakoff, D. “Bayes Offers a 'New' Way to Make Sense of Numbers.” *Science* 286 (1999): 1460-64.

Margolis, E., and S. Laurence, eds. *Concepts: Core Readings*. Cambridge, MA: MIT Press, 1999.

Mark, R. “Architecture and Evolution.” *American Scientist*, July-August 1996, 383-89.

Marrow, P., and R.A. Johnstone. “Riding the Evolutionary Streetcar: Where Population Genetics and Game Theory Meet.” *Trends in Ecology and Evolution* 11 (1996): 445-46.

Martinez, D.E., and J. Levington. “Adaptation to Heavy Metals in the Aquatic Oligochaete *Limnodrilus Hoffmeister*: Evidence for Control by One Gene.” *Evolution* 50 (1996): 1339-43.

Mas, P., P.F. Devlin, S. Panda, and S.A. Kay. “Functional Interaction of Phytochrome B and Cryptochrome 2.” *Nature* 408 (2000): 207-11.

Mathews, S., and R.A. Sharrock. “Phytochrome Gene Diversity.” *Plant, Cell and Environment* 20 (1997): 666-71.

Matloff, N.S. “Statistical Hypothesis Testing: Problems and Alternatives.” *Environmental Entomology* 20 (1991): 1246-50.

Matos, M., M.R. Rose, M.T. Rocha Pite`, C. Rego, and T. Avelar. “Adaptation to the Laboratory Environment in *Drosophila Subobscura*.” *Journal of Evolutionary Biology* 13 (2000): 9-19.

Matthen, M., and A. Ariew. “Two Ways of Thinking About Fitness and Natural Selection.” *Journal of Philosophy* 49 (2002): 55-83.

Maxwell, N. “Induction, Simplicity and Scientific Progress.” *Scientia* 114 (1979): 629-74.

Mayden, R.L. “A Hierarchy of Species Concepts: The Denouement in the Saga of the Species Problem.” In *Species: The Units of Biodiversity*, edited by M.F. Claridge et al., 381-424: Chapman & Hall, 1997.

Maynard-Smith, J., R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup, and L. Wolpert. “Developmental Constraints and Evolution.” *The Quarterly Review of Biology* 60 (1985): 265-87.

Maynard-Smith, J., and E. Szathmary. *The Major Transitions in Evolution*. Oxford, England: Oxford University Press, 1995.

Mayr, E. *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge, MA: Belknap Press, 1982.

———. *Introduction to Darwin's on the Origin of Species*. Cambridge, MA: Harvard University Press, 1964.

———. “Speciation Phenomena in Birds.” *American Naturalist* 74 (1940): 249-78.

———. “What Is a Species, and What Is Not?” *Philosophy of Science* 63 (1996): 262-77.

Mayr, E., and W.B. Provine. *The Evolutionary Synthesis. Perspectives on the Unification of Biology*. Cambridge, MA: Harvard University Press, 1980.

Mazer, S.J., and C.T. Schick. “Constancy of Population Parameters for Life-History and Floral Traits in *Raphanus Sativus* L. Ii. Effects of Planting Density on Phenotype and Heritability Estimates.” *Evolution* 45 (1991): 1888-907.

McGhee, G.R. “Exploring the Spectrum of Existent, Nonexistent and Impossible Biological Form.” *Trends in Ecology & Evolution* 16 (2001): 172-73.

McIntosh, R.P. *The Background of Ecology: Concept and Theory*. Cambridge, England: Cambridge University Press, 1985.

McKitrick, M.C., and R.M. Zink. “Species Concepts in Ornithology.” *Condor* 90 (1988): 1-14.

McPeek, M.A., and G.A. Wellborn. “Genetic Variation and Reproductive Isolation among Phenotypically Divergent Amphipod Populations.” *Limnology and Oceanography* 43 (1998): 1162-69.

Meehl, P.E. “Theoretical Risks and Tabular Asterisks: Sir Karl, Sir Ronald, and the Slow Progress of Soft Psychology.” *Journal of Consulting and Clinical Psychology* 46 (1978): 806-34.

Miller, B.L., W.W. Seeley, P. Mychack, H.J. Rosen, I. Mena, and K. Boone. “Neuroanatomy of the Self: Evidence from Patients with Frontotemporal Dementia.” *Neurology* 57 (2001): 817-21.

Miller, W.J., A. Nagel, J. Bachmann, and L. Bachmann. “Evolutionary Dynamics of the Sgm Transposon Family in the *Drosophila Obscura* Species Group.” *Molecular Biology and Evolution* 17 (2000): 1597-609.

Millstein, R.L. “Are Random Drift and Natural Selection Conceptually Distinct?” *Biology and Philosophy* 17 (2002): 33-53.

Mishler, B.D., and R.N. Brandon. “Individuality, Pluralism, and the Phylogenetic Species Concept.” *Biology and Philosophy* 2 (1987): 397-414.

Mishler, B.D., and M.J. Donoghue. “Species Concepts: A Case for Pluralism.” *Systematic Zoology* 31 (1982): 491-503.

Moir, A., and D. Jessel. *Brain Sex: The Real Difference between Men and Women*. New York, NY: Delta, 1992.

Moller, A.P., and M.D. Jennions. “Testing and Adjusting for Publications Bias.” *Trends in Ecology & Evolution* 16 (2001): 580-86.

Monod, J. *Chance and Necessity; an Essay on the Natural Philosophy of Modern Biology*. New York, NY: Knopf, 1971.

Morange, M. "A Brief Overview of Molecular Biology." *Biofutur* 142 (1995): 15-18.

Munjal, A.K., D. Karan, P. Gibert, B. Moreteau, R. Parkash, and J.R. David. "Thoracic Trident Pigmentation in *Drosophila Melanogaster*: Latitudinal and Altitudinal Clines in Indian Populations." *Genetics Selection Evolution* 29 (1997): 601-10.

Murray, C. *Iq Will Put You in Your Place* [web page]. 1998 [cited 11/7 1998]. Available from www.eugenics.net.

Nachman, M.W., and S.L. Crowell. "Estimate of the Mutation Rate Per Nucleotide in Humans." *Genetics* 156 (2000): 297-304.

Neff, M.M., and E. Van-Volkenburgh. "Light-Stimulated Cotyledon Expansion in *Arabidopsis* Seedlings: The Role of Phytochrome B." *Plant Physiology* 104 (1994): 1027-32.

Newberg, A., and E. D'Aquili. *Why God Won't Go Away: Brain Science & the Biology of Belief*. New York, NY: Ballantine, 2001.

Newman, S.A., and G.B. Muller. "Epigenetic Mechanisms of Character Origination." In *The Character Concept in Evolutionary Biology*, edited by G.P. Wagner, 559-80. San Diego, CA: Academic Press, 2001.

Nie, J.-B. "The Plurality of Chinese and American Medical Moralities: Toward an Interpretive Cross-Cultural Bioethics." *Kennedy Institute of Ethics Journal* 10 (2000): 239-60.

Niklas, K.J. "Adaptive Walks through Fitness Landscapes for Early Vascular Land Plants." *American Journal of Botany* 84 (1997): 16-25.

Nixon, K.C., and Q.D. Wheeler. "An Amplification of the Phylogenetic Species

Concept.” *Cladistics* 6 (1990): 211-23.

Noor, M.A.F. “Is the Biological Species Concept Showing Its Age?” *Trends in Ecology & Evolution* 17 (2002): 153-54.

Okasha, S. “The Underdetermination of Theory by Data and the "Strong Programme" in the Sociology of Knowledge.” *International Studies in the Philosophy of Science* 14 (2000): 283-97.

Orel, V., and R.J. Wood. “Essence and Origin of Mendel's Discovery.” *Comptes Rrendus de l'Academie des Sciences Serie III-Sciences de la Vie-Life Sciences* 323 (2000): 1037-41.

Otte, D., and J.A. Endler, eds. *Speciation and Its Consequences*. Sunderland, MA: Sinauer, 1989.

Otto, S.P., and C.D. Jones. “Detecting the Undetected: Estimating the Total Number of Loci Underlying a Quantitative Trait.” *Genetics* 156 (2000): 2093-107.

Oyama, S. *The Ontogeny of Information*. Cambridge, England: Cambridge University Press, 1985.

Oyama, S., P.E. Griffiths, and R.D. Gray, eds. *Cycles of Contingency: Developmental Systems Theory and Evolution*. Cambridge, MA: MIT Press, 2001.

Pal, C., and I. Miklos. “Epigenetic Inheritance, Genetic Assimilation and Speciation.” *Journal of Theoretical Biology* 200 (1999): 19-37.

Paley, W. *Natural Theology: Or, Evidences of the Existence and Attributes of the Deity, Collected from the Appearances of Nature*. Boston, MA: Gould, Kendall, and Lincoln, 1831.

Palmateer, A.J., M.E. Schmidt, S.R. Stetina, and J.S. Russin. “Temperature Effects on Race Determination in *Heterodera Glycines*.” *Journal of Nematology* 32 (2000): 349-

55.

Palopoli, M.F., and C.I. Wu. “Rapid Evolution of a Coadapted Gene Complex: Evidence from the Segregation Distorter (Sd) System of Meiotic Drive in *Drosophila Melanogaster*.” *Genetics* 143 (1996): 1675-88.

Peters, R.H. *A Critique for Ecology*. Cambridge, England: Cambridge University Press, 1991.

Phillips, P.C. “From Complex Traits to Complex Alleles.” *Trends in Genetics* 15 (1999): 6-8.

Phillips, P.C., M.C. Whitlock, and K. Fowler. “Inbreeding Changes the Shape of the Genetic Covariance Matrix in *Drosophila Melanogaster*.” *Genetics* 158 (2001): 1137-45.

Pier, G.B., M. Grout, T. Zaide, G. Meluleni, S.S. Mueschenborn, G. Banting, R. Ratcliff, M.J. Evans, and W.H. Colledge. “Salmonella Typhi Uses Cftr to Enter Intestinal Epithelial Cells.” *Nature* 393 (1998): 79-82.

Pigliucci, M. “Are Ecology and Evolutionary Biology "Soft" Sciences?” *Annales Zoologici Finnici* 39 (2002): 87-98.

———. “Chaos and Complexity: Should We Be Skeptic?” *Skeptic* 2001, 62-70.

———. *Denying Evolution: Creationism, Scientism and the Nature of Science*. Sunderland, MA: Sinauer, 2002.

———. “Design Yes, Intelligent No: A Critique of Intelligent Design Theory and Neo-Creationism.” *Philosophy Now*, June/July 2001, 26-29.

———. “From Molecules to Phenotypes? The Promise and Limits of Integrative Biology.” *Basic and Applied Ecology* (In press).

———. “How Organisms Respond to Environmental Changes: From Phenotypes to

Molecules (and Vice Versa)." *Trends in Ecology and Evolution* 11 (1996): 168-73.

———. *Phenotypic Plasticity: Beyond Nature and Nurture*. Baltimore, MD: Johns Hopkins University Press, 2001.

———. "Science and Religion." In *The Skeptic Encyclopedia of Pseudoscience*, edited by M. Shermer, 443-54. Santa Barbara, CA: ABC Clio, 2002.

———. "Species as Family Resemblance Concepts: The (Dis-)Solution of the Species Problem?" *BioEssays* (In press).

Pigliucci, M., and J. Kaplan. "The Fall and Rise of Dr. Pangloss: Adaptationism and the Spandrels Paper 20 Years Later." *Trends in Ecology & Evolution* 15 (2000): 66-70.

Pigliucci, M., and K. Preston, eds. *The Evolutionary Biology of Complex Phenotypes*. Oxford: Oxford University Press, In press.

Pigliucci, M., and C.D. Schlichting. "On the Limits of Quantitative Genetics for the Study of Phenotypic Evolution." *Acta Biotheoretica* 45 (1997): 143-60.

Pigliucci, M., and J. Schmitt. "Genes Affecting Phenotypic Plasticity in *Arabidopsis*: Pleiotropic Effects and Reproductive Fitness of Photomorphogenic Mutants." *Journal of Evolutionary Biology* 12 (1999): 551-62.

Pinker, S. *The Blank Slate: The Modern Denial of Human Nature*: Viking Press, 2002.

———. *How the Mind Works*. New York, NY: W.W. Norton, 1997.

Platt, J.R. "Strong Inference." *Science* 146 (1964): 347-53.

Popper, K.R. *Conjectures and Refutations: The Growth of Scientific Knowledge*. New York, NY: Harper & Row, 1968.

Prawitz, D. "Progress in Philosophy." In *The Idea of Progress*, edited by A. Burgen, P. McLaughlin and J. Mittelstab. Berlin: Walter de Gruyter, 1997.

Price, G.R. "Selection and Covariance." *Nature* 277 (1970): 520-21.

Prokudina, E.I., R.Y. Valeev, and R.M. Tchuraev. “A New Method for the Analysis of the Dynamics of the Molecular Genetic Control Systems. Ii. Application of the Method of Generalized Threshold Models in the Investigation of Concrete Genetic Systems.” *Journal of Theoretical Biology* 151 (1991): 89-110.

Qin, M., R. Kuhn, S. Moran, and P.H. Quail. “Overexpressed Phytochrome C Has Similar Photosensory Specificity to Phytochrome B but a Distinctive Capacity to Enhance Primary Leaf Expansion.” *Plant Journal* 12 (1997): 1163-72.

Quail, P.H. “The Phytochromes: A Biochemical Mechanism of Signaling in Sight?” *BioEssays* 19 (1997): 571-79.

Quail, P.H., M.T. Boylan, B.M. Parks, T.W. Short, Y. Xu, and D. Wagner. “Phytochromes: Photosensory Perception and Signal Transduction.” *Science* 268 (1995): 675-80.

Queitsch, C., T.A. Sangster, and S. Lindquist. “Hsp90 Buffers Genetic Variation, Environmental Responses, and Maintains Developmental Stability.” *Nature* (2002).

Ramachandran, V.S., and S. Blakeslee. *Phantoms in the Brain: Probing the Mysteries of the Human Mind*. New York, NY: William Morrow, 1998.

Raup, D.M., and S.J. Gould. “Stochastic Simulation and Evolution of Morphology - Towards a Nomothetic Paleontology.” *Systematic Zoology* 23 (1974): 305-22.

Rausher, M.D. “The Measurement of Selection on Quantitative Traits: Biases Due to Environmental Covariances between Traits and Fitness.” *Evolution* 46 (1992): 616-26.

Reed, J.W., A. Nagatani, T.D. Elich, M. Fagan, and J. Chory. “Phytochrome a and Phytochrome B Have Overlapping but Distinct Functions in *Arabidopsis* Development.” *Plant Physiology* 104 (1994): 1139-49.

Reed, J.W., P. Nagpal, D.S. Poole, M. Furuya, and J. Chory. “Mutations in the Gene for

the Red/Far-Red Light Receptor Phytochrome B Alter Cell Elongation and Physiological Responses Throughout *Arabidopsis* Development." *The Plant Cell* 5 (1993): 147-57.

Rehfeldt, G.E., and L.A. Gallo. "Introduction of Ponderosa Pine and Douglas-Fir to Argentina - Using Quantitative Traits for Retrospective Identification and Prospective Selection of Provenances." *New Forests* 2 (2001): 35-44.

Rice, W.R. "Analyzing Tables of Statistical Tests." *Evolution* 43 (1989): 223-25.

Ridley, M. "The Cladistic Solution to the Species Problem." *Biology and Philosophy* 4 (1989): 1-16.

Rieppel, O. "Structuralism, Functionalism, and the Four Aristotelian Causes." *Journal of the History of Biology* 23 (1990): 291-320.

Roff, D. "The Evolution of the G Matrix: Selection or Drift?" *Heredity* 84 (2000): 135-42.

Roff, D.A., and T.A. Mousseau. "Does Natural Selection Alter Genetic Architecture? An Evaluation of Quantitative Genetic Variation among Populations of *Allenomobius Socius* and *A. Fasciatus*." *Journal of Evolutionary Biology* 12 (1999): 361-69.

Rogers, A.R. "Genetic Evidence for a Pleistocene Population Explosion." *Evolution* 49 (1995): 608-15.

Roosa, J.R., C. Gervasi, and B.G. Szaro. "Structure, Biological Activity of the Upstream Regulatory Sequence, and Conserved Domains of a Middle Molecular Mass Neurofilament Gene of *Xenopus Laevis*." *Molecular Brain Research* 82 (2000): 35-51.

Rose, A.B., and J.A. Beliakoff. "Intron-Mediated Enhancement of Gene Expression Independent of Unique Intron Sequences and Splicing." *Plant Physiology* 122 (2000): 535-42.

Rose, M.R., and G.V. Lauder. “Post-Spandrel Adaptationism.” In *Adaptation*, edited by M.R. Rose and G.V. Lauder, 1-8. San Diego, CA: Academic Press, 1996.

Rosenberg, A. “How Is Biological Explanation Possible?” *British Journal for the Philosophy of Science* 52 (2001): 735-60.

Roth, L. “The Biological Basis of Homology.” In *Ontogeny and Systematics*, edited by C.J. Humphries, 1-26. New York: Columbia University Press, 1988.

Roth, V.L. “Character Replication.” In *The Character Concept in Evolutionary Biology*, edited by G.P. Wagner, 81-108. San Diego, CA: Academic Press, 2001.

Rudge, D. W. “A Bayesian Analysis of Strategies in Evolutionary Biology.” *Perspectives on Science* 6 (1998): 341-60.

Ruse, M. “Definitions of Species in Biology.” *British Journal for the Philosophy of Science* 20 (1969): 97-119.

Rushton, J.P. *The Mismeasure of Gould* [web page]. 1998 [cited. Available from www.eugenics.net.

Rutherford, S.L., and S. Lindquist. “Hsp90 as a Capacitor for Morphological Evolution.” *Nature* 396 (1998): 336-42.

Sauer, N.J. “Forensic Anthropology and the Concept of Race: If Races Don't Exist, Why Are Forensic Anthropologists So Good at Identifying Them?” *Social Science and Medicine* 34 (1992): 107-11.

Schat, H., R. Vooijs, and E. Kuiper. “Identical Major Gene Loci for Heavy Metal Tolerances That Have Independently Evolved in Different Local Populations and Subspecies of *Silene Vulgaris*.” *Evolution* 50 (1996): 1888-95.

Scheiner, S.M., R.J. Mitchell, and H.S. Callahan. “Using Path Analysis to Measure Natural Selection.” *Journal of Evolutionary Biology* 13 (2000): 423-33.

Schenk, P.M., K. Kazan, I. Wilson, J.P. Anderson, T. Richmond, S.C. Somerville, and J.M. Manners. "Coordinated Plant Defense Responses in *Arabidopsis* Revealed by Microarray Analysis." *Proceedings of the National Academy of Science USA* 97 (2000): 11655-60.

Schick, T.-Jr., ed. *Readings in the Philosophy of Science: From Positivism to Postmodernism*. Mountain View, CA: Mayfield, 2000.

Schlichting, C.D., and M. Pigliucci. "Gene Regulation, Quantitative Genetics and the Evolution of Reaction Norms." *Evolutionary Ecology* 9 (1995): 154-68.

———. *Phenotypic Evolution, a Reaction Norm Perspective*. Sunderland, MA: Sinauer, 1998.

Schluter, D., and D. Nychka. "Exploring Fitness Surfaces." *The American Naturalist* 143 (1994): 597-616.

Schmalhausen, I.I. *Factors of Evolution. The Theory of Stabilizing Selection*. Chicago: University of Chicago Press, 1949.

Schmitt, J. "Is Photomorphogenic Shade Avoidance Adaptive? Perspectives from Population Biology." *Plant, Cell and Environment* 20 (1997): 826-30.

Schwenk, K., and G.P. Wagner. "The Relativism of Constraints on Phenotypic Evolution." In *The Evolutionary Biology of Complex Phenotypes*, edited by M. Pigliucci and K. Preston. Oxford: Oxford University Press, In press.

Sgro`, C.M., and L. Partridge. "Laboratory Adaptation of Life History in *Drosophila*." *American Naturalist* 158 (2001): 657-58.

Shaw, F.H., R.G. Shaw, G.S. Wilkinson, and M. Turelli. "Changes in Genetic Variances and Covariances: **G** Whiz!" *Evolution* 49 (1995): 1260-67.

Shermer, M. "Blood, Sweat, and Fears." *Skeptic* 2000, 44-53.

Shinomura, T., A. Nagatani, J. Chory, and M. Furuya. “The Induction of Seed Germination in *Arabidopsis Thaliana* Is Regulated Principally by Phytochrome B and Secondarily by Phytochrome A.” *Plant Physiology* 104 (1994): 363-71.

Shipley, B. *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference*. Cambridge, England: Cambridge University Press, 2000.

Shirley, M.D.F., and R.M. Sibly. “Genetic Basis of a between-Environment Trade-Off Involving Resistance to Cadmium in *Drosophila Melanogaster*.” *Evolution* 53 (1999): 826-36.

Shoemaker, J.S., I.S. Painter, and B.S. Weir. “Bayesian Statistics in Genetics: A Guide for the Uninitiated.” *Trends in Genetics* 15 (1999): 354-58.

Shrader-Frechette, K.S. , and E.D. McCoy. *Method in Ecology: Strategies for Conservation*. Cambridge, England: Cambridge University Press, 1994.

Silva, P.J.N., and D.E. Dykhuizen. “The Increased Potential for Selection of the *Lac* Operon of *Escherichia Coli*.” *Evolution* 47 (1993): 741-49.

Simon, M.A. “When Is a Resemblance a Family Resemblance?” *Mind* 78 (1969): 408-16.

Simpson, G.G. “Criteria for Genera, Species, and Subspecies in Zoology and Paleontology.” *Annals of the New York Academy of Sciences* 44 (1943): 145-78.

———. *Principles of Animal Taxonomy*. New York, NY: Columbia University Press, 1961.

Simpson, G.G., and C. Dean. “*Arabidopsis*, the Rosetta Stone of Flowering Time?” *Science* 296 (2002): 285-89.

Sinsheimer, J.S., J.A. Lake, and R.J.A. Little. “Bayesian Hypothesis Testing of Four-Taxon Topologies Using Molecular Sequence Data.” *Biometrics* 52 (1996): 193-210.

Skyrms, B. “Game Theory, Rationality and Evolution of the Social Contract.” *Journal of Consciousness* 7 (2000): 269-84.

Smith, H. “Light Quality, Photoreception, and Plant Strategy.” *Annual Review of Plant Physiology* 33 (1982): 481-518.

Smith, K.N., and A. Nicolas. “Recombination at Work for Meiosis.” *Current Opinion in Genetics & Development* 8 (2001): 200-11.

Smith-Gill, S.J. “Developmental Plasticity: Developmental Conversion Versus Phenotypic Modulation.” *American Zoologist* 23 (1983): 47-55.

Sneath, P.H.A. “Phenetic Taxonomy at the Species Level and Above.” *Taxon* 25 (1976): 437-50.

Sneath, P.H.A., and R.R. Sokal. *Numerical Taxonomy*. San Francisco: W.H. Freeman & Co., 1973.

Snow, C.P. *The Two Cultures and the Scientific Revolution*. Cambridge, England: Oxford University Press, 1959.

Sober, E. *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. Cambridge, MA: MIT Press, 1984.

———. *Philosophy of Biology*. Boulder, CO: Westview Press, 1993.

Sober, E., and R.C. Lewontin. “Artifact, Cause and Genic Selection.” *Philosophy of Science* 49 (1982): 157-80.

Sober, E., and D.S. Wilson. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press, 1998.

Sokal, R.R., and F.J. Rohlf. *Biometry*. New York, NY: Freeman & Co., 1995.

Sole`, R.V., S.C. Manrubia, M. Benton, S. Kauffman, and P. Bak. “Criticality and Scaling in Evolutionary Ecology.” *Trends in Ecology & Evolution* 14 (1999): 156-60.

Sollars, V., X. Lu, L. Xiao, X. Wang, M.D. Garfinkel, and D.M. Ruden. “Evidence for an Epigenetic Mechanism by Which Hsp90 Acts as a Capacitor for Morphological Evolution.” *Nature Genetics* (2003).

Sorell, T. *Scientism: Philosophy and the Infatuation with Science*. London: Routledge, 1991.

———. *Scientism: Philosophy and the Infatuation with Science*. London: Routledge, 1991.

Splitter, L.J. “Species and Identity.” *Philosophy of Science* 55 (1988): 323-48.

Stanley, S.M. “A Theory of Evolution above the Species Level.” *Proceedings of the National Academy of Science USA* 72 (1975): 646-50.

Stephens, J.C., D.E. Reich, D.B. Goldstein, H.D. Shin, M.W. Smith, and et-al. “Dating the Origin of the Ccr5-Delta32 Aids-Resistance Allele by the Coalescence of Haplotypes.” *American Journal of Human Genetics* 62 (1998): 1507-15.

Steppan, S.J. “Phylogenetic Analysis of Phenotypic Covariance Structure. Ii. Reconstructing Matrix Evolution.” *Evolution* 51 (1997): 587-94.

Steppan, S.J., P.C. Phillips, and D. Houle. “Comparative Quantitative Genetics: Evolution of the G Matrix.” *Trends in Ecology & Evolution* 17 (2002): 320-27.

Sterelny, K. “The Nature of Species.” *Philosophical Books* 35 (1994): 9-20.

Sterelny, K., and P. Kitcher. “The Return of the Gene.” *Journal of Philosophy* 85 (1988): 339-61.

Stewart, I. “The Lorenz Attractor Exists.” *Nature* 406 (2000): 948-49.

Stinchcombe, J.R., M.T. Rutter, D.S. Burdick, P. Tiffin, M.D. Rausher, and R. Mauricio. “Testing for Environmentally Induced Bias in Phenotypic Estimates of Natural Selection: Theory and Practice.” *American Naturalist* 160 (2002): 511-23.

Stone, G., R. Atkinson, A. Rokas, G. Csoka, and J.L. Nieves-Aldrey. “Differential

Success in Northwards Range Expansion between Ecotypes of the Marble Gallwasp *Andricus Kollari*: A Tale of Two Lifecycles." *Molecular Ecology* 10 (2001): 761-78.

Stone, J.R. "Computer-Simulated Shell Size and Shape Variation in the Caribbean Land Snail Genus *Cerion*: A Test of Geometrical Constraints." *Evolution* 50 (1996): 341-47.

Stratton, D.A., and C.C. Bennington. "Measuring Spatial Variation in Natural Selection Using Randomly-Sown Seeds of *Arabidopsis Thaliana*." *Journal of Evolutionary Biology* 9 (1996): 215-28.

Su, W., and S.H. Howell. "The Effects of Cytokinin and Light on Hypocotyl Elongation in *Arabidopsis* Seedlings Are Independent and Additive." *Plant Physiology* 108 (1995): 1423-30.

Sultan, S.E. "Phenotypic Plasticity and Plant Adaptation." *Acta Botanica Neederlandica* 44 (1995): 363-83.

Tattersall, I. *Becoming Human*. New York, NY: Harcourt Brace & Company, 1998.

———. *The Fossil Trail: How We Know What We Think We Know About Human Evolution*. New York, NY: Oxford University Press, 1995.

———. "Once We Were Not Alone." *Scientific American*, January 2000, 56-62.

Templeton, A.R. "Human Races: A Genetic and Evolutionary Perspective." *American Anthropologist* 100 (1999): 632-50.

———. "The Meaning of Species and Speciation: A Genetic Perspective." In *Speciation and Its Consequences*, edited by D. Otte and J.A. Endler, 3-27, 1989.

Templeton, A.R., S.D. Maskas, and M.B. Cruzan. "Gene Trees: A Powerful Tool for Exploring the Evolutionary Biology of Species and Speciation." *Plant Species Biology* 15 (2000): 211-22.

Tooby, J. , and L. Cosmides. “The Innate Versus the Manifest: How Universal Does Universal Have to Be?” *Behavioral and Brain Sciences* 12 (1989): 36-37.

———. “On the Universality of Human Nature and the Uniqueness of the Individual: The Role of Genetics and Adaptation.” *Journal of Personality* 58 (1990): 17-67.

Travisano, M. “Long-Term Experimental Evolution in *Escherichia Coli*. Vi. Environmental Constraints on Adaptation and Divergence.” *Genetics* 146 (1997): 471-79.

Travisano, M., J.A. Mongold, A.F. Bennett, and R.E. Lenski. “Experimental Tests of the Roles of Adaptation, Chance, and History in Evolution.” *Science* 267 (1995): 87-90.

Treisman, J.E. “A Conserved Blueprint for the Eye?” *BioEssays* 21 (1999): 843-50.

Tukey, J.W. “Analyzing Data: Sanctification or Detective Work?” *American Psychologist* 24 (1969): 83-91.

Turesson, G. “The Genotypical Response of the Plant Species to the Habitat.” *Hereditas* 3 (1922): 211-350.

Turkheimer, F. , K. Pettigrew, L. Sokoloff, C.B. Smith, and K. Schmidt. “Selection of an Adaptive Test Statistic for Use with Multiple Comparison Analyses of Neuroimaging Data.” *Neuroimage* 12 (2000): 219-29.

Turner, L. “An Anthropological Exploration of Contemporary Bioethics: The Varieties of Common Sense.” *Journal of Medical Ethics* 24 (1998): 127-33.

van Tienderen, P.H., and J. Antonovics. “Constraints in Evolution: On the Baby and the Bath Water.” *Functional Ecology* 8 (1994): 139-40.

van Tienderen, P.H., I. Hammad, and F.C. Zwaal. “Pleiotropic Effects of Flowering Time Genes in the Annual Crucifer *Arabidopsis Thaliana* (Brassicaceae).” *American Journal of Botany* 83 (1996): 169-74.

van Valen, L. “Ecological Species, Multispecies, and Oaks.” *Taxon* 25 (1976): 233-39.

van Valen, L.M. “Homology and Causes.” *Journal of Morphology* 173 (1982): 305-12.

———. “Species, Sets, and the Derivative Nature of Philosophy.” *Biology and Philosophy* 3 (1988): 49-66.

Vekemans, X., and C. Lefebvre. “On the Evolution of Heavy-Metal Tolerant Populations in *Armeria Maritima*: Evidence from Allozyme Variation and Reproductive Barriers.” *Journal of Evolutionary Biology* 10 (1997): 175-92.

Via, S. “Genetic Constraints on the Evolution of Phenotypic Plasticity.” In *Genetic Constraints on Adaptive Evolution*, edited by V. Loeschcke, 47-71. Berlin: Springer-Verlag, 1987.

Via, S., and R. Lande. “Genotype-Environment Interaction and the Evolution of Phenotypic Plasticity.” *Evolution* 39 (1985): 505-22.

Vicente, J.G., J. Conway, S.J. Roberts, and J.D. Taylor. “Identification and Origin of *Xanthomonas Campestris* Pv. Campestris Races and Related Pathovars.” *Phytopathology* 91 (2001): 492-99.

Waddington, C.H. “Canalization of Development and the Inheritance of Acquired Characters.” *Nature* 150 (1942): 563-65.

———. “Genetic Assimilation.” *Advances in Genetics* 10 (1961): 257-90.

Waddle, D.M. “Matrix Correlation Tests Support a Single Origin for Modern Humans.” *Nature* 368 (1994): 452-54.

Wade, P.R. “Bayesian Methods in Conservation Biology.” *Conservation Biology* 14 (2000): 1308-16.

Wagner, G., ed. *The Character Concept in Evolutionary Biology*. San Diego, CA: Academic Press, 2001.

Wagner, G.P., and K. Schwenk. “Evolutionarily Stable Configurations: Functional Integration and the Evolution of Phenotypic Stability.” *Evolutionary Biology* 31 (2000): 155-217.

Walker, W.R., and S.J. Hoekstra. “Science Education Is No Guarantee of Skepticism.” *Skeptic* 2002, 24-27.

Wang, R.L., J. Wakeley, and J. Hey. “Gene Flow and Natural Selection in the Origin of *Drosophila Pseudoobscura* and Close Relatives.” *Genetics* 147 (1997): 1091-106.

Ward, D. “The Role of Satisficing in Foraging Theory.” *Oikos* 63 (1992): 312-17.

Waters, C.K. “Tempered Realism About the Force of Selection.” *Philosophy of Science* 58 (1991): 553-73.

Weinberg, S. “Against Philosophy.” In *Dreams of a Final Theory*, 166-90. New York: Pantheon, 1992.

———. *Dreams of a Final Theory*. New York: Pantheon, 1992.

Weinig, C. “Limits to Adaptive Plasticity: Temperature and Photoperiod Influence Shade-Avoidance Responses.” *American Journal of Botany* 87 (2000): 1660-68.

Weis, A.E., and W.L. Gorman. “Measuring Selection on Reaction Norms: An Exploration of the *Eurosta-Solidago* System.” *Evolution* 44 (1990): 820-31.

Wells, C., and M. Pigliucci. “Heterophylly in Aquatic Plants: Considering the Evidence for Adaptive Plasticity.” *Perspectives in Plant Ecology Evolution and Systematics* 3 (2000): 1-18.

Whewell, W. *The Philosophy of the Inductive Sciences, Founded Upon Their History*. London: J.W. Parker, 1840.

Whitelam, G.C., and P.F. Devlin. “Light Signalling in *Arabidopsis*.” *Plant Physiology and Biochemistry* 36 (1998): 125-33.

Wilkins, A.S. *The Evolution of Developmental Pathways*. Sunderland, MA: Sinauer, 2002.

———. "Why the Philosophy of Science Actually Does Matter." *BioEssays* 23 (2001): 1-2.

Williams, G.C. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton: Princeton University Press, 1966.

———. *Natural Selection: Domains, Levels and Challenges*. New York, NY: Oxford University Press, 1992.

Wills, C. "When Did Eve Live? An Evolutionary Detective Story." *Evolution* 49 (1995): 593-607.

Wilson, E.O. *Consilience: The Unity of Knowledge*. New York, NY: Knopf : Distributed by Random House, 1998.

———. *Sociobiology: The New Synthesis*. Cambridge, MA: Harvard University Press, 1975.

Windelband, W. "History and Natural Science." *History and Theory* 19 (1894 / 1980): 169-85.

Wittgenstein, L. *Philosophical Investigations*. New York, NY: Macmillan, 1953 / 1973.

Wolf, J.B., E.D. Brodie-III, and M.J. Wade. *Epistasis and the Evolutionary Process*. Oxford, England: Oxford University Press, 2000.

Wolford, G., M.B. Miller, and M. Gazzaniga. "The Left Hemisphere's Role in Hypothesis Formation." *Journal of Neuroscience* 20 (2000): 1-4.

Wolfsberg, T.G. , J. McEntyre, and G.D. Schuler. "Guide to the Draft Human Genome." *Nature* 409 (2001): 824 - 26.

Wray, G.A. "Developmental Evolution: New Paradigms and Paradoxes." *Developmental Genetics* 15 (1994): 1-6.

Wu, R. “The Detection of Plasticity Genes in Heterogeneous Environments.” *Evolution* 52 (1998): 967-77.

Wu, S.-H., and J.C. Lagarias. “The Phytochrome Photoreceptor in the Green Alga *Mesotaenium Caldariorum*: Implication for a Conserved Mechanism of Phytochrome Action.” *Plant, Cell and Environment* 20 (1997): 691-99.

Wynne-Edwards. *Animal Dispersion in Relation to Social Behavior*. Edinburgh: Oliver & Boyd, 1962.

Zhivotovsky, L.A., M.W. Feldman, and A. Bergman. “On the Evolution of Phenotypic Plasticity in a Spatially Heterogeneous Environment.” *Evolution* 50 (1996): 547-58.

Vita

Massimo Pigliucci was born in Monrovia (Liberia) on 16 January 1964, and is an Italian citizen. He has a doctorate in genetics from the University of Ferrara (Italy, 1989) and a PhD in Botany from the University of Connecticut (1994). He has carried out post-doctoral research in evolutionary ecology at Brown University and is currently Associate Professor of Ecology & Evolutionary Biology at the University of Tennessee. He is has been associated with the Department of Philosophy at the same University since Fall 2000, where he has pursued a Ph.D. in philosophy of science under the guidance of Dr. Jonathan Kaplan, which has led to this dissertation.

Dr. Pigliucci's fields of research in biology include plant ecological and evolutionary genetics, experimental and theoretical approaches to the study of organismal responses to environmental change (nature vs. nurture), and the extent of limits and constraints on natural selection.

In philosophy, Pigliucci's interests include philosophy of science and the relationship between science and philosophy as far as the pursue of joint research programs is concerned. He is a member of the American Philosophical Association and of the Philosophy of Science Association. His essays and lectures can be found at www.genotype-environment.org.