ASSESSING THE POSSIBILITY OF A FUNCTIONALLY DISCONTINUOUS BIOLOGICAL PARADIGM

A Thesis

by

JAMES WILLIAM SCHROEDER

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

December 2005

Major Subject: Philosophy

ASSESSING THE POSSIBILITY OF A FUNCTIONALLY DISCONTINUOUS BIOLOGICAL PARADIGM

A Thesis

by

JAMES WILLIAM SCHROEDER

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

Approved by:

Co-Chairs of Committee,	Roger Sansom
	James Womack
Committee Member,	Walter Bradley
Head of Department,	Robin Smith

December 2005

Major Subject: Philosophy

ABSTRACT

Assessing the Possibility of a Functionally

Discontinuous Biological Paradigm. (December 2005)

James William Schroeder, B.S., Texas A&M University

Co-chairs of Advisory Committee: Dr. Roger Sansom
Dr. James Womack

This project sets as its goal the development of an Intelligent Design paradigm that makes falsifiable predictions. According to Karl Popper, such falsifiability is a key component of scientific theories. To accomplish this, two hypothetical historical narratives are first outlined based on guided processes and the design points they predict.

A biochemical approach to characterizing organisms then defines a protein's global functional limits as determining the set of amino acids that allow it to successfully perform its functions in any situation. The local functional limits restrict this potential substitution set to only those proteins viable within an individual genetic background.

Proteins are referred to as the first-order of specified complexity because a protein's gene is the fundamental unit of inheritance. Other orders of specified complexity are described culminating in the organism level, which is the fundamental unit of selection.

Each phylogenetic tree within the two intelligent design scenarios is founded by an original group or archetype. The descendants of this archetype are known as the archetype's genus. Speciation events within the genus are brought about by a slow process called co-adapted drift that creates distinct species through functional incompatibilities.

A theory of natural selection is developed that attempts to characterize the relationship between the gene and the organism. Natural selection in this sense is described as a preservation mechanism that selects against deleterious phenotypes instead of selecting for beneficial ones.

Finally, a practical methodology is developed that begins by determining the history of a gene in a given species by the symmetrical causal relationships of the alleles and the species allelic distribution. The original alleles in this species and their local functional limits are then compared with those of analogous genes in similar species to determine if these species were functionally compatible at that time. The two Intelligent Design paradigms predict patterns of incompatibilities, or design points, where guided actions were involved. This is a falsifiable prediction that raises the status of these paradigms in a Popperian sense.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Recently, a new generation of dissension has arisen against a naturalistic view of the universe and orthodox Darwinian evolution. This group has become known as the Intelligent Design movement. Those within this movement come from varying disciplines and hold varying reasons for their participation. However, one common conviction can be said to unite them in their position on biological issues. This conviction is that the solely unguided processes of naturalistic evolution do not adequately explain the origin and diversity of life. This leads them to broadly conclude that an intelligent designer lies behind the history of life on Earth at least at some point. Whether it is possible to discover this agent's identity, or what intentions[1] this agent may have held in designing life are not questions that are pursued. Thus the main thrust of the Intelligent Design position is not an argument for a specific type of designer (i.e. God, aliens, etc.) or design story (i.e. Genesis, directed panspermia, etc.). Rather, the position is an argument for design per se.

A major difficulty raised against Intelligent Design has been their lack of a positive paradigm. Michael Ruse has commented that denying Darwinian evolution does not prove Creationism, rather one must have his own evidence in favor of his own theory.[2] While the Creationism Ruse is referring to is not the same thing as the Intelligent Design that I am talking about (for example, in exegeting and relying on particular biblical claims about history), the charge is still relevant to Intelligent Design in virtue of its attempt to present an alternative to Darwinian evolution. Intelligent Design arguments are often accused of

---

This thesis follows the style of the MLA Style Manual.

not being falsifiable or even making specific claims that could be falsified.[3] One of Karl Popper's contributions to the discussion of valid theory formulation is his emphasis on a theory's falsifiability as a means (although not guaranteed) of avoiding pseudoscience.[4] Such falsifiability is a feature of empirically-testable predictions that are logically deduced from the systematic form of new theories.[5] Although Popper's views have fallen on somewhat hard times, clearly unfalsifiability is a scientific vice. One quality that a scientific theory can display that avoids this problem, according to Popper, is found in its formulation of predictions. These should not be predictions that could yield multiple or ambiguous results that could be massaged into conforming with the theory in question.[6] Ideally, he states that a good theory is able to make 'risky' predictions, which if did not obtain, more or less refutes the theory in question.[7] According to Popper's criteria then, a way to rectify the non-falsifiability charge against an Intelligent Design scenario is to formulate risky predictions out of its structure that have the potential to disconfirm it. Such a formulation, however, only adds to the theory's status qua theory, it does not in itself say anything of the theory's actual confirmation or disconfirmation.

Design arguments are often accused of not being falsifiable because they are often attempting to explain historical dissimilarities among groups of organisms resulting from design. Such dissimilarities within the tree of life are taken to be examples of a positive hallmark of biological design. Guided processes at points of design, the design theorist claims, can more adequately explain these dissimilarities between groups of organisms. In contrast, the position taken by naturalistic evolution argues that these alleged dissimilarities are still explainable (or possibly can be) by unguided, natural processes.

To avoid this difficulty, design theorists could develop rigorous methods to specify the kind of dissimilarity between closely related species that they expect to find. This concerns two main aspects of biology. The first aspect of this strategy concerns the physiological and biochemical composition of living organisms themselves. If phylogenetic dissimilarities are to be taken seriously, it must be (reasonably) shown how individual organisms could not have endured the physiological changes needed to bridge these alleged functional dissimilarities. The second aspect of this strategy deals with the histories of specific groups of organisms. Many current methods of reconstructing these histories utilize quantifiable molecular remnants in extant groups of organisms. True dissimilarities in the history of life would be evidenced (or potentially evidenced) in these quantifiable molecular records. Design theorists need to develop methods to discern such evidence and then demonstrate that this evidence actually exists.

This project begins by considering some of the positive arguments for biological Intelligent Design that have surfaced during recent years. These arguments, I believe, can serve as crucial planks in the foundations of an Intelligent Design biological paradigm. I will then attempt to utilize aspects of these positive arguments (provisionally assuming they are correct) to develop the framework of an Intelligent Design paradigm following the above strategy. This paradigm exists as a positive assertion that is independent of negative references to naturalistic evolution.[8] It is also explanatory in its ability to address multiple physiological aspects of biological organisms, and capable of making testable, falsifiable predictions. Chapter V will discuss a methodology for testing these predictions and confirming or denying them. This methodology, I will argue, raises the status of the

theories making these predictions to scientific theories in a Popperian sense because of the character of the predictions they make.

TWO GENERAL FRAMEWORKS

We can distinguish between naturalistic evolution and biological design per se by delineating between guided versus unguided processes as causal mechanisms in species creation. Natural selection will only favor those traits that are presently advantageous. It merely operates on a breeding population's current set of variation within its present environmental conditions. Even if a presently disadvantageous mutation could help an organism in the future, that mutation will not be presently selected based on current function. Thus it would not likely be proliferated in the organism's gene pool for future use.[9] Whatever future advantage it may afford the organism would be irrelevant to the unguided natural processes that presently select for it. This summarizes what I mean in saying that naturalistic evolution proceeds by only unguided natural processes.

In contrast, in this project I will present two Intelligent Design frameworks that utilize some form of guided processes as the means by which design per se could be manifested in biological organisms. The first of these frameworks, *intervening-maintenance*, posits both unguided and guided processes. We can consider a hypothetical situation where mutant traits arise in a species that proves presently disadvantageous. These traits will, however, be advantageous in future conditions. In this case, the guided, goal-oriented process of a designer's intervening maintenance of these traits is presented as a possible explanation of the trait's proliferation. This process acts contrary to the

perpetual natural selective pressures encountered. This scenario is similar to a biologist's preservation of an organism's feature through artificial selection. This artificial selection acts despite whether that feature could otherwise persist in the organism's natural environment. The designer could thus maintain traits by this periodic intervention because it can plan for future situations where that trait would be advantageous. The large-scale artificial preservation[10] of novel traits in this way could explain the transcendence of functional dissimilarities between groups of related organisms.

Biology within a design framework is not restricted to the gradual, intervening process of the above scenario as a means of species generation. A creative action of a designer could also be manifested at single time-points creating entire organisms without precursors. These creative actions would occur at the origin of a given species or of the original life form (if monophyly was the case). This second framework, the *initial-construction* scenario, is similar to the manufacturing procedures of an automobile factory. Such a factory needs to exert a concerted construction effort to build a fully functional car. Only after this construction is the car functionally capable of standing up to the scrutiny of its "natural environment", the street. The automobile manufacturers continue to construct the integrated vehicle throughout this period because of their ability to foresee this end result. This is despite the inability of the car's individual parts to act in isolation in serving the same functions the completed car can. The mechanisms of this initial-construction scenario act in the absence of selective pressures. On the contrary, in the intervening-maintenance scenario described above, organisms are constantly subjected to natural selective pressures during the designing process, against which the designer is acting.

A common historical aspect shared by these two scenarios is a way in which dissimilarities between species would be manifested. In the intervening-maintenance scenario, the unnaturally maintained traits would have physically descended from naturally maintained precursor traits. Despite a physically uninterrupted descent in this case, this guided maintenance yields an example of what I will call functional dissimilarity. A functional dissimilarity exists between two species when the points along the simplest gradual sequence of mutations that would transcend the dissimilarity would result in severe loss of function. Thus, when there is not a direct gradualistic descent by natural selection between a precursor of a trait and the novel form of the trait - as in the case of intervening-maintenance - those two traits are functionally dissimilar. In the initial-construction scenario, none of a newly constructed species' traits have physical precursors. Thus these traits are also functionally dissimilar between themselves and any would-be precursor traits. In both scenarios, any proposed functional dissimilarities discovered by a researcher represent a proposed point of design. In the remainder of this project, I will refer to proposed functional dissimilarities as proposed design points.

POSITIVE INTELLIGENT DESIGN ARGUMENTS

One of the positive arguments put forward in support of Intelligent Design has been Michael Behe's argument for *irreducible complexity* in biological structures. Behe defines an irreducibly complex system as:

> …a single system composed of several well-matched, interacting parts that contribute to the basic function, wherein the removal of any one of the parts causes the system to effectively cease functioning.[11]

In his book *Darwin's Black Box*, Behe gives multiple examples of what he claims are irreducibly complex systems found throughout molecular biology. These examples include the mammalian blood clotting system, cilia, intracellular transport, the bacterial flagellum, and the immune system.[12] Let us now take the example of the bacterial flagellum. David DeRosier, who specializes in the study of the bacterial flagellum, describes it as follows:

> More than actomyosin or tubulokinesin, the bacterial flagellum of *Salmonella typhimurium* is the analogue of a man-made mechanical system. Its heart is a 15,000 revolutions per minute, reversible rotary motor powered by the proton-motive gradient across the cell's inner membrane. Each revolution consumes about 1000 protons. A drive shaft, held by a bushing in the outer membrane, transmits torque across the cell's envelope. Attached to the drive shaft, a universal joint enables the motor to drive the propeller, even when the drive shaft and propeller are not co-linear. A short junction joins the propeller to the drive shaft. The propeller, a long left-handed corkscrew, converts torque to thrust. A cap sits at the cell distal end of the filament. By electron microscopy, the motor associated parts and the bushing are seen to be rings of subunits, where as the drive shaft appears to be a helical assembly of subunits. About four dozen genes are needed to build the flagellum. Some are required for regulation of synthesis; some for export and assembly; some for the structure itself, and a few are of unknown function. Nineteen different proteins are known to be part of the flagellar structure; it is thought that there may be additional components.[13]

Behe argues that his examples of irreducibly complex systems, such as the bacterial flagellum, exist contrary to the unguided generative abilities presented by Darwinian stepwise mutation models. William Dembski argues that Behe's irreducible complexity examples therefore logically rule out direct Darwinian pathways of stepwise mutation as potential candidates for the origin of these systems.[14] To illustrate these claims, we can consider the previous discussion of unguided natural processes and disadvantageous novel mutations. Behe argues that an operational bacterial flagellum minimally requires multiple

proteins acting in coordination. Each of these proteins can only be selected for based on the fully operational flagellum's function. Therefore, he claims that if any of these proteins arose before the fully operational flagellum could be selected for then they could not have been selected for based on the flagellar function. The flagellum's function as a whole represents the functional advantage of each of its constituent proteins as well. Any future use of these constituent proteins in a fully operational flagellum would play no part in their unguided preservation before the flagellar function is realized. Dembski argues that there are two remaining explanations of the formation of irreducibly complex systems like the bacterial flagellum. The first is some version of indirect Darwinian mutation pathways. The second is simply the system's appearance in its fully functioning form in part by some goal-oriented action. For this project, I will develop the latter position. Whether or not the characterization of Behe's examples turns out to be empirically correct is another issue. These examples of irreducible complexity suggest a form of dissimilarity by which an Intelligent Design paradigm could be developed.

A second positive argument presented by design theorists is *specified complexity* developed by William Dembski.[15] Dembski argues that there are three criteria to establish specified complexity. The fulfillment of these criteria, he suggests, warrants a design inference. The first criterion is *contingency*. Contingency first dictates that the object in question must conform to the fundamental regularities that brought it about. However, it must exist as one among many possible outcomes of those regularities. As Dembski describes it:

> By being compatible with but not required by the regularities involved in its production, an object, event, or structure becomes irreducible to any

> underlying physical necessity…The method applies quite generally: the position of Scrabble pieces on a Scrabble board is irreducible to the natural laws governing the motion of Scrabble pieces…the sequencing of DNA bases is irreducible to the bonding affinities between the bases; and so on.[16]

His second criterion is an object's *complexity*. He argues that this criterion is also necessary to avoid improperly treating any probable chance event as a product of design. This complexity criterion therefore is stated in terms of probability: the greater the complexity, the smaller the probability. He writes:

> Thus to determine whether something is sufficiently complex to underwrite a design inference is to determine whether it has sufficiently small probability.[17]

Despite fulfilling these first two criteria an object could still exist as the product of an unguided event. In addressing Dembski's third criterion, *specificity*, we can consider a mountainside. With the combination of wind, erosion, and gravity, one expects to find a given mountainside with any multitude of possible shapes that its rocks could collectively form. All of these shapes are as equally unlikely. However, when confronted with a mountainside exhibiting the likenesses of Abraham Lincoln, George Washington, Thomas Jefferson, and Teddy Roosevelt, one rightly concludes that the mountain is the product of design, not the result of unguided natural forces. Dembski argues that these images fit his third criterion for a design inference by being specified to an independent pattern (the likeness of these former presidents). This pattern is both prior to, and independent of the production of the object in question - the faces on the mountainside. By applying this criterion, he believes one can avoid mistakenly classifying either random events or *ad hoc* patterns as genuine inferences to design.[18]

PROJECT GOALS

In Chapter II, I will begin developing my position by examining proteins as examples of specified complexity. In doing so, I am not attempting to argue that proteins are products of an intelligent designer simply because they fit Dembski's criteria of specified complexity. In other words, I am not arguing specifically for Dembski's conclusions about specified complexity. After all, within an unguided naturalistic evolutionary framework, these proteins could be suggested as being, as Dawkins would say, "designed" by the blind watchmaker of natural selection. Rather I am simply using the convention of specified complexity and its criteria to define protein sequence and function as: 1) contingent, 2) suitably complex, and 3) specified to an independent pattern. The protein's normal functions with other molecules within its host organism determine this independent pattern. An example of this is the specified complexity of a bacterial flagellum's constitutive protein. This protein is contingent in that it could have had a different sequence. It is also complex in that it contains multiple hundred amino acids. Finally, it exhibits specificity to an independent pattern, namely, that defined by its role in the formation, structure, or operation of the flagellum.

Following the discussion of protein specified complexity in the Chapter II, I will take a bottom-up approach in assessing the remaining project goals. This approach will address two aspects of the previously described strategy to avoid the problem of unfalsifiability.

The first aspect of the bottom-up approach consists of defining larger systems as outcomes of the integrated functions of their constitutive systems. The fundamental

starting point of defining the feature is the biochemical level. Below this level, there can be no further appeal to bioinformatic chemistry. From this, a specific definition of speciation will be developed where potential for mutational variation must remain within the functional constraints of integrated biological systems. Natural selection will then be addressed in relation to these ideas.

The second aspect of this approach consists in defining potential ancestral or future variations of extant biological systems by the functional limits of extant systems. This will be done by developing a methodology for empirically assessing their limitations. One possible conclusion of this methodology is that these limitations could not have been naturally transcended while retaining an advantageous function by direct pathways. This is evidence of a possible design point predicted by these two Intelligent Design paradigms. An investigation leading to a pattern of such conclusions could then provide some evidence in favor of design.

This aspect of the investigation model parallels the investigation model of Big Bang cosmology. The Big Bang event existed as an event in the past that was both unique in history as well as outside of any possible physical investigation.[19] It can be reliably inferred though as a historical event based upon the convincing evidence its effects provide. Similarly, I propose that this bottom-up approach of inferring phylogenetic history based upon the functional and integrative considerations of extant biological systems could be sufficient to establish their possible ancestral forms. This holds despite any inability to infer an unguided (and predictable) causal history prior to the system's emergence. The methodology developed by this approach is capable of roughly

determining any would-be points of origin of individual species and biological features. As I will argue later, the Intelligent Design scenarios described above predict a general pattern of proposed design points of a species' constitutive features at the time of its geological appearance. Hence large-scale application of this methodology can serve as a Popperian falsification test of such Intelligent Design scenarios (despite its inability to conclusive rule out naturalistic evolution scenarios). I will not enter into a philosophical or theological discussion of these positions past what an empirical inference allows.

NOTES

1. By 'intentions' here I mean a designer's intentions in creating life in the first place (subjective intentions such as satisfying this designer's desire to create). I do not mean a designer's intentions for the mechanistic composition of the design. The discipline of biology has taken upon itself the task of determining the physical nature of biological organisms as they operate in their natural environments and reproduce. This task is distinct from peripheral intentions an agent has in creating life, and is also the task this project is concerned with.

2. Michael Ruse, "Creation-Science Is Not Science," Philosophy of Science, ed. Martin Curd (New York: Norton & Company, 1998) 38-47.

3. Robert Pennock, "DNA by Design?," Debating Design, ed. Michael Ruse (Cambridge: Cambridge UP, 2004) 130-148.

4. Karl Popper, "Science: Conjectures and Refutations," Philosophy of Science, ed. Martin Curd (New York: Norton & Company, 1998) 3-10.

5. Karl Popper, "The Problem of Induction," Philosophy of Science, ed. Martin Curd (New York: Norton & Company, 1998) 426-432.

6. Karl Popper, "Science: Conjectures and Refutations," Philosophy of Science, ed. Martin Curd (New York: Norton & Company, 1998) 3-10.

7. Karl Popper, "Science: Conjectures and Refutations," Philosophy of Science, ed. Martin Curd (New York: Norton & Company, 1998) 3-10.

8. By this, I mean that this paradigm should be self-contained, such that its internal cohesion would not be altered if unguided Darwinian evolution was never proposed. To be sure, Darwinian evolution exhibiting no true dissimilarities is mentioned in this project. However, this is primarily to illustrate a point or serve as a point of comparison.

9. This is not to say that it is impossible for the trait to be maintained. It only states that its proliferation is contrary to its selective advantage that its current function provides.

10. I mean "artificial preservation" in the sense that human domestic animal breeding programs exercise artificial preservation. That is, traits are often selected and preserved based upon the goals of the breeder, and not based upon the trait's immediate functional advantage.

11. Michael Behe, Darwin's Black Box (New York: Free Press, 1996) 39.

12. Michael Behe, Darwin's Black Box (New York: Free Press, 1996) 39.

13. David DeRosier, "Spinning Tails," Curr Opin Struct Biol. 5.2 (1995): 187-93.

14. William Dembski, "Evolution's Logic of Credulity: An Unfettered Response to Allen Orr." 2002. Online posting. Design Inference Website, 2002 Postings Section. http://www.designinference.com/. October 2003.

15. William Dembski, No Free Lunch (Lanham: Rowman & Littlefield, 2002) 6-15.

16. William Dembski, No Free Lunch (Lanham: Rowman & Littlefield, 2002) 8.

17. William Dembski, No Free Lunch (Lanham: Rowman & Littlefield, 2002) 9.

18. William Dembski, No Free Lunch (Lanham: Rowman & Littlefield, 2002) 9-12.

19. This is simply because, on one view, the event itself preceded the origin of the physical laws that physicists use to investigate it.

CHAPTER II

SPECIFIED COMPLEXITY AS A POINT OF DEPARTURE

This chapter starts by defining and discussing the ultimate function of biological organisms. It then discusses the particular functions of an organism's individual systems contributing to that ultimate function. This methodology can be applied to protein function. The precise reasons why proteins constitute the smallest complete unit will be addressed in Chapter III. In this chapter, I will introduce two kinds of protein functional limits. The first will be as global functional limits. These limits define the maximum possible permutations of a protein that its particular functions permit. The second will be as the local functional limits. These limits describe a protein's allowable variants within a given organism's genetic background. From these starting points, the broader issues of the project can then be addressed using this bottom-up approach.

FUNCTIONS

A biological feature's *functions* are the operations that it is both specified for and naturally performs. This specification can be either chemical or physical. For example, a hormone receptor is chemically specified to bind its hormone counterpart. Likewise, a pig's femur bone is physically specified to structurally support its leg. However, it could be argued that hormone receptors are also specified for hormone mimetics. Pig femur bones similarly could be said to be specifically useful for occupying rambunctious canines. Additionally, both of these counterexamples could be said to occur naturally. Therefore, a crucial third criterion is needed to make this notion of functioning meaningful for

characterizing biological systems. This requirement is the feature's contribution to its organism's ultimate function.

Here, we will define an organism's *ultimate function* as the ability to survive and reproduce its genes. The *particular functions* an organism's features perform are those required for this ultimate function (redundancy notwithstanding). A feature's particular functions stand in opposition to any non-particular functions it can perform. These non-particular functions are those that do not contribute to the organism's ultimate function. The hormone receptor binding a hormone mimetic is then one of its non-particular functions. Binding this mimetic is not aiding its particular function facilitated by binding its true hormone. It is possibly even inhibiting it.[1]

A *system*, in this project, is any interacting group of biological components that performs particular functions. A system can then describe groups of cells, groups of proteins, or even proteins themselves. Taken as a system, a protein is a group of interacting amino acids that performs its particular functions. This term 'system' in this sense can apply to any group of components within an organism's *physiological hierarchy*. An organism's physiological hierarchy describes the ladder of defined systems arranged from smallest (proteins) to largest (the organism itself). Each system in this hierarchy encompasses the smaller systems it is composed of. For example, a cell (the larger system) contains protein networks (the smaller systems) inside of it. Similarly, a protein network (the larger system) includes proteins (the smaller systems) within itself. The *ultimate system* in this hierarchy is the organism itself. As its particular function, the ultimate

system (the organism) performs its ultimate function. For this project, all other systems are considered only with regard to their particular functions.

A system's structural specificity is determined by its function. Additionally, it would not carry out its function without one of a set of specified structures. A system's particular functions rely upon its structural specificity. Recall that specificity is Dembski's third criteria for specified complexity. I am assuming the other two criteria, contingency and complexity, when speaking of biological systems. By meeting these three criteria, I am taking these systems to exhibit specified complexity. This specified complexity is directed towards these systems' particular functions.

This specificity of biological systems has two elements. The first element is the 'parts list' of components involved in the system. For all systems other than the smallest in the physiological hierarchy, these 'components' can also be systems in their own right. An organelle is then a cellular component even though it is a system in itself. The second element is the ways in which these parts are coordinated together. The appropriate coordination of these components allows the system to perform its particular functions.

FUNCTIONAL ROLES

Functional roles within a biological system represent individual tasks contributing to the system's particular function. The total list of functional roles of a system constitutes the system's *specified complexity core*. Within each functional role, there may be one or more component that can perform that role. A component's particular function is a way of

describing what functional role it plays within the larger system. An organelle's particular function, for example, reflects the functional role it plays in the cell.

Component redundancies that exist within a functional role can take two forms. First, multiple components can be expressed along side each other within one functional role. Here, one component's malfunctioning might not be detrimental to the filling of its functional role. Second, redundant components within a functional role can be utilized in different circumstances. Such is the case when alternative transcription factors activate DNA transcription.

As a system, a protein's individual amino acid positions are normally seen as functional roles within its specified complexity core.[2] Protein specified complexity cores will allow loose specificity in most functional roles (perhaps two-thirds). That is, substitutions here still allow those roles' contributions to the protein's particular function. A protein's *functional limits* delineate between this set of allowable substitutions and deleterious substitutions. Deleterious substitutions are those that inhibit the particular functioning of that protein or another protein or system.

Systems of proteins are also consortiums of functional roles that combine to perform the functions of the system. As with individual proteins, there is the possibility of loose specificity within these systems' functional roles. However, distinct component substitutions are much rarer in protein networks than in individual proteins. There are two reasons for this. One is the higher degree of specificity proteins, the components of protein networks, have over amino acids, the components of proteins. Amino acids are, after all,

much more functionally versatile than proteins, relatively speaking. The second reason is the greater energetic cost of synthesizing proteins.

A system's *system limits* describe how malleable its components' specificities are within the system's functional roles. Similar to a protein's functional limits (as described below), system limits can be considered both locally and globally. The *local system limits* describe the possible compositions a system can undergo in order to perform its particular functioning within the confines of a specific genetic background. This would be, for example, the genetic background from an individual organism with its peculiar variants of systems and proteins. The *global system limits* describe any possible composition the system can assume while still performing its particular function within any genetic background. Any given local system limits is then a subset of the global system limits.

These local system limits define the role specificities by dictating the roles' particular functioning criteria within a specific genetic background. There are two of these criteria that were discussed in the previous section. One criterion is the general structural composition each component must conform to. This general composition is described by the proteins' functional limits, if the components considered are proteins. If the components are protein systems or larger systems, this composition is that system's own system limits. System compositions can then be partly reduced to their constitutive proteins functional limits via this first criterion. However, system limits also include a second criterion that applies to the system as a whole. This criterion determines where and when components are expressed and integrated, and their expression levels.

TYPES OF BIOLOGICAL SYSTEMS

An individual component within a functional role may or may not be necessary for that role. For example, a redundant component could be substituted with a second component sharing the common role. Some functional roles might themselves be unnecessary for the particular functioning of a system. Such roles might serve as merely an additive enhancement to the system. The remaining roles are those essential for the system to perform its particular functions. A system's functional specificity is then irreducible to any one of these essential functional roles. Furthermore, the system's particular functioning cannot be reduced beyond that full set of essential roles.

There are two general categories that describe these features of biological systems. Each general category contains two possible options within it:

Concerning functional roles -

1) Systems that contain crucial and non-crucial functional roles
2) Systems that contain only crucial functional roles

Concerning components within functional roles -

3) Systems whose functional roles are fulfilled by more than one component
4) Systems whose functional roles are fulfilled by only one component

Table 2.1

Types of Biological Systems.

	1) Crucial and non-crucial functional roles	***2) Crucial functional roles only***
3) Functional roles fulfilled by more than one component	**Type A**: System contains an irreducibly complex core with component redundancy	**Type B**: System with irreducibly complex set of functional roles
4) Functional roles fulfilled by only one component	**Type C**: System with an invariant irreducibly complex core	**Type D**: Irreducibly complex system

Biological systems can contain any combination of options within these two general categories. These possible combinations are exhibited in Table 2.1. The four combinations yielded will be referred to as system types A, B, C, and D. Each type requires at least a minimum set of functional roles that are crucial for the system's functioning. In types A and C, the set of roles crucial for the system's function is called the system's *irreducibly complex core*. This core is a subset of the system's specified complexity core. It refers only to the system's functional roles, not to its component diversity or invariability. Systems of type A include most, if not all proteins. Functional roles in these proteins (their amino acid positions) often are partially specific, allowing multiple amino acid substitutions. Many of these amino acid positions are not crucial for the protein's functioning either. A type C system example is a protein system containing a unique enhancement that is not crucial for the system's functioning. Even if this component is lost, the system can function normally without the component being replaced.

The remaining systems are of types B and D - those that contain only functional roles crucial to the system's functioning. In other words, these systems cannot be reduced beyond this set of roles without malfunctioning. This, however, is all that can be said of systems of type B. Type B systems might include components within a role that adapt to varying external circumstances. An example is the various transcription factors that activate DNA transcription, as noted earlier. While the transcription factors' role is crucial, the individual transcription factors filling this role are diverse. Occasionally, individual proteins and protein domains could themselves be examples of this type of system. In these cases, their amino acid positions might both be crucial for the protein's functioning, and allow variability. A system of type D is an irreducibly complex system such as Behe describes. These systems have only functional roles crucial for the system's functioning and unique components participating in each of these roles.

GLOBAL AND LOCAL FUNCTIONAL LIMITS

Recall that a protein's functional limits determine its allowed set of amino acid substitutions. These functional limits can be applied to proteins in two different ways. The first of these will be referred to as a protein's *global functional limits.* The global limits are the less restrictive of the two applications regarding the allowed substitutional diversity. Global functional limits define any potential variations proteins can assume while maintaining their particular functions. They do so by defining the set of all possible amino acid substitutions that permit these particular functions. Our knowledge of this substitution set is first informed by the substitutional diversity exhibited in a protein's

extant variants. This set is enhanced by considering novel artificial variants conducive with the protein's particular functioning. Such novel variants are those produced both *in vitro* and in silico.

		T				*D*												*T*		
		R				*K*	*G*						*P*					*D*		
		L			*L*	*S*	*N*		*R*				*T*					*Q*		*R*
P		**E**	*L*		**Q**	*A*	**Q**		**L**		*D*		**C**	*K*			*G*	**V**		*A*
V	**A**	**P**	**P**	**P**	**V**	**V**	**L**	**L**	**P**	**D**	**V**		**E**	**T**	**P**		**S**	**E**	**E**	**D**	**C**
1 –	2 –	3 –	4 –	5 –	6 –	7 –	8 –	9 –	10 –	11 –	12 –		13 –	14 –	15 –		16 –	17 –	18 –	19 –	20

Figure 2.1. Specified Complexity Core of Hypothetical Protein. (Italicized letters are amino acids within the parameters of the global functional limits. Bold letters are amino acids within the parameters of the global functional and local functional limits).

The second application of protein functional limits is the protein's *local functional limits*. These limits are defined relative to the protein's interactions within an individual organism's genetic background. This background is that individual organism's peripheral protein variants and biochemicals, with which the given protein must interact. Local functional limits describe the amino acid set that allows the protein's particular functions within this individual organism (Figure 2.1). Variants within an individual organism's local limits are those that can operate within its genetic background. This is contrasted to the full range of substitutions allowed by the protein's global functional limits. This full range pertains to any potential genetic background that preserves the protein's particular

functioning. A protein's local functional limits' substitution set is then a subset of its global functional limits' substitution set.

A protein's global functional limits could also refer to ways it functions in multiple systems. In these cases, investigators might wish to delineate between the protein's interactions in these different systems. Local functional limits could then be characterized relative to specific protein domains, active sites, etc. A particular protein variant could then exist within system A's local functional limits, yet outside system B's. The variation between these local limits should typically apply to the protein regions performing the distinct interactions.

Confirming candidate protein variants as genuine representatives of a specific protein is done in two ways. First, true variants are identified by their compositional conformance to their protein's global functional limits. Second, factors independent of composition also aid in verifying these true variants. This includes a holistic accounting of the protein's functional role(s) and its gene's locality in the genome. The functional roles of proteins within a system include their assisting the functions of peripheral proteins, biochemicals, and systems. This assistance is those proteins' particular functions. Having this proper coordination of all molecules within a system of proteins allows that system's particular functioning. A protein's expression must then coincide with the presence of the other components within its system. The genetic locus of all true variants of a protein should also be uniform within a species.[3] These two criteria for determining variants - composition and functional role - are also forms of the previous biological system criteria. These criteria, as stated earlier, are the 'parts list', and coordination of these parts.

Candidate variants must meet each of these pre-established criteria to be considered true variants of a protein.

While substitutions within a protein's global functional limits exist in the 'properly functioning' category, those outside exist in the 'malfunctioning' category. However, many substitutions will possibly not appear to fall neatly into these functioning/malfunctioning categories. Rather, they might seem to fall somewhere in between. The extent and relevance of this is an empirical issue, not one to be resolved here. For the present project, there are three main categories that all variants can be put into. The first two categories contain the *quantitatively inferior* and *quantitatively superior variants*. Both of these types of variants exist within their protein's functional limits. Thus, these variant types both exist in the 'properly functioning' category. This means that neither variant type is always consistently selected against were the organism dependent upon them. As their name implies, quantitatively inferior variants have a decreased functional performance compared to quantitatively superior variants. This being the case, these variants can be consistently selected against when in competition with superior variants. These two types of variants will be developed further in Chapter IV. Both types, though, are distinct from the third variant type - those outside the protein's functional limits. This third group is the *malfunctioning variants*. Variants of this third type are consistently selected against when functionally relied upon. Variants of the first two types are not always consistently selected against. These variant types could apply whether or not an investigator could definitively assign an individual variant to them.

Global functional limits are defined relative to a species. This is because a species is the largest group that can initially guarantee an interbreeding gene pool. This interbreeding gene pool requires a given protein to have uniform global functional limits for functional compatibility. However, two species could demonstrate a common ancestry, with the result that their analogous proteins have equivalent global functional limits. Information gleaned from variants in one species could then enhance an investigator's knowledge of the protein's global functional limits in both. This will be discussed further in Chapter V.

This could imply a revival of the distinction between analogous and homologous proteins across species. This distinction has become understandably blurred as a logical outcome of a continuous common lineage perspective. What would the results be of a common phylogenetic lineage? Proteins from similar species, appearing analogous in composition and function were likely derived from a common protein. Hence these similar proteins from similar species are homologous to each other.

By examining each species' proteins individually, these proteins can be characterized more precisely. This characterization is much more laborious though. If proteins from different species truly are analogous, their individual characterizations should be mostly equivalent, simplifying these investigations. This is due to the similar results that are yielded from their individual characterization requirements. These criteria include their global functional limits and functional roles exhibited within extant species members. The criteria are examined for each protein and then clarified by artificial perturbation, as discussed above. Proteins that are homologous across species will also

possess the same characterization for these criteria. An example of this is a cytochrome c protein from two species that descended from a common ancestral protein. Further characterization of homologous proteins will be developed in Chapter V.

COMPONENTS OF THE GENETIC BACKGROUND

An individual's genetic background is composed of two different types of components. These are those that are compositionally immutable and those that are compositionally mutable. The compositionally immutable components are molecules whose precise chemical composition defines their functional properties. Thus this composition fixes their potential functioning in a biological system. These molecules include simpler organic molecules (non-biopolymers) whose compositional modification alters their chemical properties. Altering these chemical properties disables their filling their respective functional roles in their system. Because these molecules are immutable, proteins interacting with them have only one set of functional limits. The local functional limits of these protein regions are thus equivalent to these region's global functional limits. These immutable molecules then partly ground the mutational potential of proteins they interact with. This applies at least insofar as these molecules' specific interactions with those proteins are concerned.

Compositionally mutable molecules are important because they are what impose varying local functional limits on proteins. In this, they allow for diverse local functional limits subsumed under the global functional limits. These molecules include other proteins primarily, but also include other biopolymers like DNA and RNA. These three

macromolecules fall under the ‘mutable component’ category because they can alter their sequential arrangement. The mutation of this sequential arrangement can also alter the interactive properties of these molecules. These altered interactive properties may or may not be deleterious to particular functioning though. This variation in interactive properties allows for different local functional limits in different individual organisms. Each individual organism could then potentially have its own local functional limits for each protein it carries. However, a protein’s local functional limits throughout all the species’ individuals should at least partially overlap. If not, some variants would be incompatible with the genetic backgrounds of other individuals within that species. This could render two individuals carrying mutually incompatible variants incapable of interbreeding and producing viable offspring. This applies insofar as this protein and its system are expressed and relied upon within their functional roles.

		T			*D*												*T*			
		R			*K*	*G*					*P*						*D*			
		L		*L*	**S**	*N*		*R*			*T*						*Q*		*R*	
P		**E**	*L*	**Q**	**A**	**Q**		**L**		*D*	*C*	*K*		*G*			*V*		*A*	
V	**A**	**P**	**P**	**P**	**V**	**V**	**L**	**L**	**P**	**D**	**V**	**E**	**T**	**P**	**S**	**E**	**E**	**D**	**C**	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	

Figure 2.2. Specified Complexity Core Sub-Field *A* of Hypothetical Protein Defined by Local Functional Limits *A*. (Codon positions 12 (V), 13 (E), and 18 (E) variations are functionally linked. Italicized letters = amino acids within the parameters of the global functional limits. Bold letters = amino acids within the parameters of the global functional and local functional limits).

		T			*D*													*T*		
		R			*K*	*G*					*P*							*D*		
		L		*L*	**S**	*N*		*R*			*T*							*Q*		*R*
P		**E**	*L*	**Q**	**A**	**Q**		**L**		**D**	**C**	*K*		*G*				**V**		*A*
V	**A**	**P**	**P**	**P**	**V**	**V**	**L**	**L**	**P**	**D**	*V*	*E*	**T**	**P**	**S**	**E**	*E*	**D**	**C**	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	

Figure 2.3. Specified Complexity Core Sub-Field *B* of Hypothetical Protein Defined by Local Functional Limits *B*. (Codon positions 12 (D), 13 (C), and 18 (V) variations are functionally linked. Italicized letters = amino acids within the parameters of the global functional limits. Bold letters = amino acids within the parameters of the global functional and local functional limits).

SUB-DOMAINS AND COMPONENT SUBSTITUTION

In some cases, multiple variants of a protein within a species display distinct amino acid combinations. These combinations consist of different amino acid positions that are *functionally linked*. This differs from typical protein variant characterizations that exist in any combination that allows particular functioning. Investigators could represent combinations that are functionally linked with multiple distinct *sub-fields* of substitution sets. Each of these sub-fields maintain their own signature local functional limits (Figures 2.2, 2.3). Although allowing distinct substitution sets, these sub-field functional limits are still restricted by the same peripheral molecules. Thus, these variants retain their similarity because they must each conform to these external constraints. If they do not, then they forfeit their particular functioning.

Functionally linked variants do not bring into question the overall validity of protein functional limits. Rather, they are a further testimony to them. If nothing else, this demonstrates a more subtle level of integration existing within each protein sub-field.

Historically, the original variants representing these sub-fields could have spawned all subsequent variants within their respective sub-fields. This would have happened through stepwise sequence divergence that was restricted within each sub-field. Each divergent variant would have maintained these functionally linked combinations within its sub-field's local functional limits. If they did not, their particular functioning would have been compromised unless it was somehow compensated for. Chapter V will address this issue of constrained sequence divergence further.

As discussed previously, two or more proteins could perform in one functional role. In these cases, the local functional limits of the regions governing the shared function(s) apply equally to both. This does so in a qualified sense though. The relationship of these regions' functional limits is similar to the relationship between distinct protein sub-fields. Both are determined by their means of interacting with peripheral molecules. Past this, they may each maintain functionally linked substitution sets.

CONCLUSION

This chapter introduced the notions of functional roles, particular functioning, and functional limits. A functional role is a task that an organism must perform to survive and reproduce its genes. Likewise, the individual tasks the organism's systems must perform are called the systems' functional roles. The individual components within these systems are said to perform their particular functions when they fill their systems' functional roles. Similarly, these systems perform their particular functions when they fill the organism's functional roles. When systems of proteins are discussed, their individual constituent

proteins' particular functions perform the systems' functional roles. These proteins' particular functioning is constrained by their global and local functional limits. A protein's functional limits determine which compositions will allow its variants to perform their particular functioning. The development of these three concepts will serve as a foundation for the remainder of this project.

NOTES

1. This does not, however, refer to systems in which an endogenous hormone and endogenous antagonist both naturally act on a hormone receptor in regulating the downstream effects of receptor activity. In this case, both the hormone and the antagonist serve particular functions and the ultimate function of the organism.

2. I say normally because it might often prove more practical to consider a protein's modular domains separately. These domains are then seen as possessing specified complexity cores of their own.

3. Chromosomal abnormalities notwithstanding.

CHAPTER III

CUMULATIVE FUNCTIONAL INTEGRITY AND SPECIES EMERGENCE

The previous chapter developed the notion of protein specified complexity in two ways. The first way was by way of the protein's global functional limits. These global limits describe all possible permutations the protein could take while remaining a viable variant. A protein's local functional limits describe a subset of possible variants that could function successfully within an individual organism. In this chapter, these ideas will now be used as the grounding of a view of biological systems' integration. The notion of the physiological hierarchy will first be expanded upon. This hierarchy systematically describes the internal architecture of organisms' biochemical and physiological systems. This hierarchy will then be used in developing a theory of speciation. This theory will argue for reproductive isolation as an outcome of mutually incompatible fluctuations within this integrated hierarchy. These ideas support the larger project goals by furthering the notion of functional limitations resulting from integration on multiple levels.

EXTRINSIC AND INTRINSIC SPECIFIED COMPLEXITY

There can be said to be two types of specified complexity that apply to biological systems. Here we will use proteins as examples of such systems. The first type of specified complexity concerns the proteins themselves. This type refers to the internal structures proteins must possess in accordance with their functional limits. We will call this the proteins' *intrinsic specified complexities*. Each of these proteins can also be said to

participate in the second type of specified complexity. This second type is their *extrinsic specified complexity* exhibited by their protein network as a whole.

These two terms refer to the relative levels of integration when discussing a given biological system. The internal integration of any given system is always that system's intrinsic specified complexity. Likewise, extrinsic specified complexity refers to a system's participation in the structure of the larger systems that contains it. For example, a bacterial flagellar protein's extrinsic specified complexity refers to its role in the coordinated protein system producing the flagellar function. It is this integrated net function Behe claims would effectively cease were any essential constitutive proteins removed. Likewise, each flagellar protein's intrinsic specified complexity is determined by the integration of its constitutive components - amino acids. This intrinsic/extrinsic specified complexity characterization is not limited to relationships between proteins and protein systems though. This will be discussed further below.

ORDERS OF SPECIFIED COMPLEXITY

Proteins are the base level, or *first-order* of specified complexity. The specific reasons why proteins are granted this position will be discussed in the next section. A protein network then exists as the *second-order* of specified complexity. Second-order systems are groups of first-order components (proteins) in conjunction with other pertinent elements. One of these elements concerns the particular compositions of these constitutive proteins and their interactions together. A second is the instances of transcription activation of the proteins' genes and their expression levels. Additionally, the structures

and interactions of any non-specifically complex organic molecules play roles in second-order systems. The specified complexities of second-order systems are collectively defined by each of these elements of its structure. These elements are considered insofar as they appear in the system's viable forms *in vivo*.

An even larger system, say an organelle, could be referred to as a *third-order* of specified complexity. As the second-order is extrinsically specifically complex to the first-order, so the third-order is extrinsic to the second-order. The specified complexity of second-order systems within third-order systems is that second-order system's intrinsic specified complexity. The third-order consists of groups of coordinated second-order systems, and possibly other non-specifically complex components. Additionally, the second-order system considerations mentioned above, such as expression times of constitutive components, also apply to third-order systems.

These orders of integration extend out to include larger systems in an organism's physiology. Each higher-order system relies upon the individual structures and collective integration of its constitutive lower-order components. These components include lower-order systems as well as non-specifically complex components within the system. A system's collective integration includes the appropriate sequence of component expression and utilization, expression localization, and expression levels.

Larger systems, which could perhaps be on the cellular level, constitute the *fourth-order* of specified complexity. This fourth-order consists of groups of third-order components. A *fifth-order* exists at the level of perhaps tissues, and so on up to the organism level. The organism as a whole exists as the final system in the physiological

hierarchy. Each order thus contains what we will call a *cumulative functional integrity*. This is summarized as a system's overall specified complexity, both exhibited in its own intrinsic specified complexity as well as that of the lower-order systems it contains. Where redundancy within a system's functional role exists, the immediately utilized component is considered part of the system's cumulative functional integrity. A system's cumulative functional integrity is then a dynamic application. It does not apply to a system's *potential* composition, as housed in the genome, but refers to its actual composition at a given moment. A system's potential composition is its potential cumulative functional integrity. The physiological hierarchy described here, though, is not intended to imply a definitive distinction between the orders of specified complexity. Indeed, the characterization of the orders of specified complexity given here is, in a sense, only instrumental. This will be addressed further in the next section.

A system's cumulative functional integrity defines the mutational limits of viable variants of that system. These *system limits* describe the need for the system's components to maintain an internal harmony for the system to function properly. System limits are the higher-order system counterpart to protein functional limits in this way. Indeed, they are, in part, extensions of the functional limits of their constitutive protein components. System limits are direct extensions of protein functional limits for second-order systems, which only contain first-order components. For third-order systems or higher, protein functional limits apply indirectly via their constitutive component systems' cumulative functional integrities. Variability in protein functional limits allows much of the local variability within the system limits. System limits also define the mutational boundaries of

any potential ancestral forms of the system. Thus, system limits restrict the possible structures from which a given system could have developed. This is similar to a protein's global functional limits defining the compositional boundaries of both its extant and potential ancestral variants. System limits are, primarily, only approximations of the limits of the systems they address. This is because most orders of specified complexity are only instrumental and any characterization of their limits is, however accurate, only partial.

An extended application of orders of specified complexity culminates in defining the cumulative functional integrity of an entire organism. As with any biological system, whole organisms contain at least an irreducibly complex core of functional roles. When a species originally emerged, it would have existed in its completed final form with regards to its cumulative functional integrity. This is because a whole organism's cumulative functional integrity fulfills each of its essential functional roles. The organism cannot function successfully without the fulfillment of these essential functional roles. This is not to say that these initial organisms' potential composition was identical to later potential compositions. This potential composition could have changed by loosing or gaining traits over time. At the very minimum though, a species' initial organisms contained a full and successfully functioning cumulative functional integrity.

ON THE LOWEST ORDER

First-order systems are significant because their structures ground the cumulative functional integrity of the systems in the remaining orders. As mentioned previously, the protein level is the first-order of specified complexity. The reason for this is not because

proteins are the smallest components of biological systems. Carbon, nitrogen, and other atoms utilized in biological systems are certainly smaller than proteins. Indeed, their quantum-level components are smaller still if physical size is the measure considered. The reason is also not because proteins are the smallest biochemical components. Proteins themselves are, after all, composed of amino acids, which are biochemicals as well. There are, however, compelling reasons to consider proteins as the first-order of specified complexity.

The first reason is that first-order proteins' compositions are irreducible to their chemical components. In other words, first-order components (amino acids) have no cumulative functional integrity of their own because they are not bioinformatically expressed by the organism's genome. Instead they are chemically produced by the true expressed products of an organism's genome - proteins. This distinguishes proteins from higher-order systems, such as the second-order system of which it is a member. Second-order systems by definition include the intrinsic specified complexities of proteins in their cumulative functional integrities. Thus, from the composition and coordination of these components, the second-order system's structure and particular functions are achieved. However, merely the presence of the correct amino acids does not guarantee a protein's proper composition. The proper sequencing of these amino acids is required as well. This sequencing is not, however, inherent in the chemistry of these amino acids. Instead, it is imposed by a parallel source - the sequence of the allele coding for that protein. This source is parallel in that it determines the protein's composition, yet is not itself part of that composition. It is upon this first-ordered sequencing that all higher-orders are given form.

Another reason is that proteins are the functional representatives of alleles that encode them. By "functional representatives" I mean that the selection criteria for alleles concern the particular functioning of those alleles' proteins. The allele itself has no particular function other than to produce viable proteins. If an allele expresses a malfunctioning protein, that allele could be selected against because its protein failed to perform its particular functions. Alternatively, alleles can produce properly functioning proteins that permit those alleles to be inherited based on their proteins' functional merits. This direct connection between proteins and their alleles is also important because it demonstrates why alleles are the fundamental units of inheritance. As stated above, proteins are the lowest order of bioinformatic chemistry, and here we see that proteins are functional representatives of alleles. Thus, alleles are functionally represented by the lowest order of bioinformatics chemistry. The fundamental bioinformatic contribution from a parent to an offspring is then found in the alleles.

Another order of specified complexity exists that is in one sense more foundational than the first-order. However, this order exists as a special case in the physiological hierarchy. I will call this order the *mediator-order*. The mediator-order includes molecules acting in the translation machinery between DNA, RNA, and proteins. Several types of organization unique to this order are found. One deals with an allele's codons within its messenger RNA (mRNA) transcribed from DNA. Each three-nucleotide codon codes for a specific amino acid and/or processing signal for translation from the mRNA into protein. These codon patterns can be said to be found in transfer RNAs (tRNAs) as well. Another type of information can also be demonstrated in tRNAs. These molecules

contain specific sequences that cause their unique secondary structures. The correlation between the tRNA's correct codon compliment and its ability to bind with the proper amino acid demonstrates a sequential link between these two aspects of its composition. Other types of molecules, such as ribosomal RNA (rRNA), and sequencing characterizations could also exist at this level.

This mediator-order information is a special case though for several reasons. First, its sequencing is not physically evident in higher-order systems like protein sequencing is. A protein is, after all, composed solely of amino acids, revealing no direct trace of its nucleotide codon origins. Additionally, higher-orders built off of proteins are subsequently not composed of codon ordering either. Nor do higher-orders contain any direct physical trace of these codons as causal elements. A protein's amino acids are indirect evidence of their respective codons only through one's prior knowledge of those codon/amino acid associations.

ON THE HIGHEST ORDER

The scenario to this point begins with proteins as the foundational order of the physiological hierarchy. As first-order components, proteins contribute in determining the structure and particular functioning of second-order systems. These second-order systems then in turn contribute in determining third-order systems and so on. What then exists as the final order of specified complexity? As we have already seen, the physiological hierarchy extends out at least as far as the organism-level. Additionally, every component in an organism's cumulative functional integrity is selected for at the organism level. For

example, a well-functioning heart is not independently selected for in an organism with malfunctioning lungs, despite the heart's fitness level. This is because an organism's heart cannot be selected for independently of its lungs. Likewise, systems are not typically selected independently of other systems in an organism. A possible exception to this is genes that cheat meiosis and whose selection status is not as clearly tied to the organism-level. Organisms themselves can be selected for as units independent of other organisms though. So, just as the gene/protein-level is the fundamental unit of inheritance, so the organism-level is the fundamental unit of selection. It is possible to appeal to the population or ecological levels for higher-orders of specified complexity. Although a characterization of that sort could prove fruitful, it extends beyond the scope of this project. The organism level then at least exists as the lowest order of selection. This is sufficient for present purposes as the final order of the physiological hierarchy.

The characterization of ascending orders of specified complexity is an instrumental approach to understanding the systematic integration of biological organisms. Specifically, the intermediate orders existing between the lowest and highest orders are merely instrumental. The highest and lowest orders are the only two orders that are said to be 'real'. The protein-level is said to be so because proteins' genes are the fundamental units of inheritance. The organism-level is said to be so because organisms are the fundamental units of selection. Indeed, with an exhaustive knowledge of an organism, one could speak of the physiological hierarchy simply by the collective interactions of its molecules. No human investigator obviously has this privileged perspective though. It would therefore behoove researchers to utilize the physiological hierarchy as a tool for

investigating biological systems. They should, however, proceed with the understanding that the intermediate orders are only conventional.

OUTLINE OF INTELLIGENT DESIGN SCENARIOS

Recall the earlier discussion of the two potential Intelligent Design scenarios of intervening-maintenance and initial-construction. These two scenarios differed from naturalistic evolution by their inclusion of guided processes. Here, guided processes occur in addition to the unguided (and predictable) processes that alone drive naturalistic evolution. In the intervening-maintenance scenario, species generation can occur by the artificial selection of existing species (or individual organisms) by an intelligent designer. These guided selection processes manipulate certain aspects of the unguided natural processes until a new species' body plan is established. At this point, the new species is functionally capable of successfully proliferating within its environment's selective constraints without artificial guiding support. This new species' organisms then propagate under the influence of unguided natural processes within their natural environment. Through time, these organisms' descendents could adapt within the system limits of their varying systems as natural pressures allow. This continues until the designer again employs guided processes to modify them into a still different species.[1] The period in the species' lifecycle when it is exposed to solely unguided natural processes begins at its initial release from the designer's guided processes. This period ends when guided processes again manipulate the species to form a new species.

In the initial-construction scenario, the appearance and proliferation of species follows a somewhat similar pattern. The species in this scenario, though, are not modified from presently existing organisms. Instead, they are constructed new from non-biological raw materials. However, they are only capable of surviving under their natural environment's pressures after existing in their completed final form. Species individuals then pass on their alleles based on the functional performance of these alleles' proteins within this environment.

CO-ADAPTATION AND SPECIATION

The period after the species' release from the direct influences of guided processes can be described similarly for both scenarios. First a new species appears in the environment in its *original form*. All alleles the new species carries are likewise in their original forms. A species' original form could have equally been derived through either one of the two Intelligent Design scenarios. We can call this original group of organisms the *archetype*. The archetype, as a group, then proceeds to the normal period of their existence. Archetype individuals engaging in their archetype's typical behaviors are what characterize this period. This includes archetype individuals interacting together and with their natural environment, and producing the next generations according to unguided processes. During this period, archetype individuals are subjected to mutation and selection based upon these unguided natural processes. Many allelic results of these processes will likewise be passed on to their descendents. As time passes, the successive descendents are subjected to further mutation and natural selection. The resulting novel

alleles then become increasingly deviated from their original forms insomuch as their global functional limits allow. Mutated alleles exceeding their protein's local functional limits are malfunctioning, and thus maladaptive mutations. If these maladapted alleles code for essential proteins, natural selection will purge them from the gene pool.[2] Several cases could exist for maladapted alleles that are not purged. The functional role(s) those alleles' proteins fulfill could be non-essential. In other instances, a redundancy for a protein's functional role could compensate for that protein's malfunctioning. It is also possible that an allele's phenotype is not consistently expressed, such as with recessive alleles (this will be discussed further in Chapter IV).

Some protein variants could be produced that deviate from their original form(s), yet still exist within their original local functional limits. Likewise, variants could be produced that are outside of these original limits yet exist within some newly redefined local functional limits (as described in Chapter II). As a result, these novel variants now tolerate different, yet overlapping, substitution sets from the proteins they interact with (Figure 3.1). Peripheral proteins interacting with this first protein now themselves have new local functional limits insofar as the first protein influences their limits. This allows viable novel mutations of peripheral proteins that were previously outside of their local functional limits. These new peripheral protein variants then in turn tolerate further mutations of the first protein. This includes mutations that were previously outside of the first protein's local functional limits. Additionally, other proteins interacting with these peripheral proteins also possess redefined local functional limits. This will then allow some of their previously malfunctioning variants to become viable as well. And thus

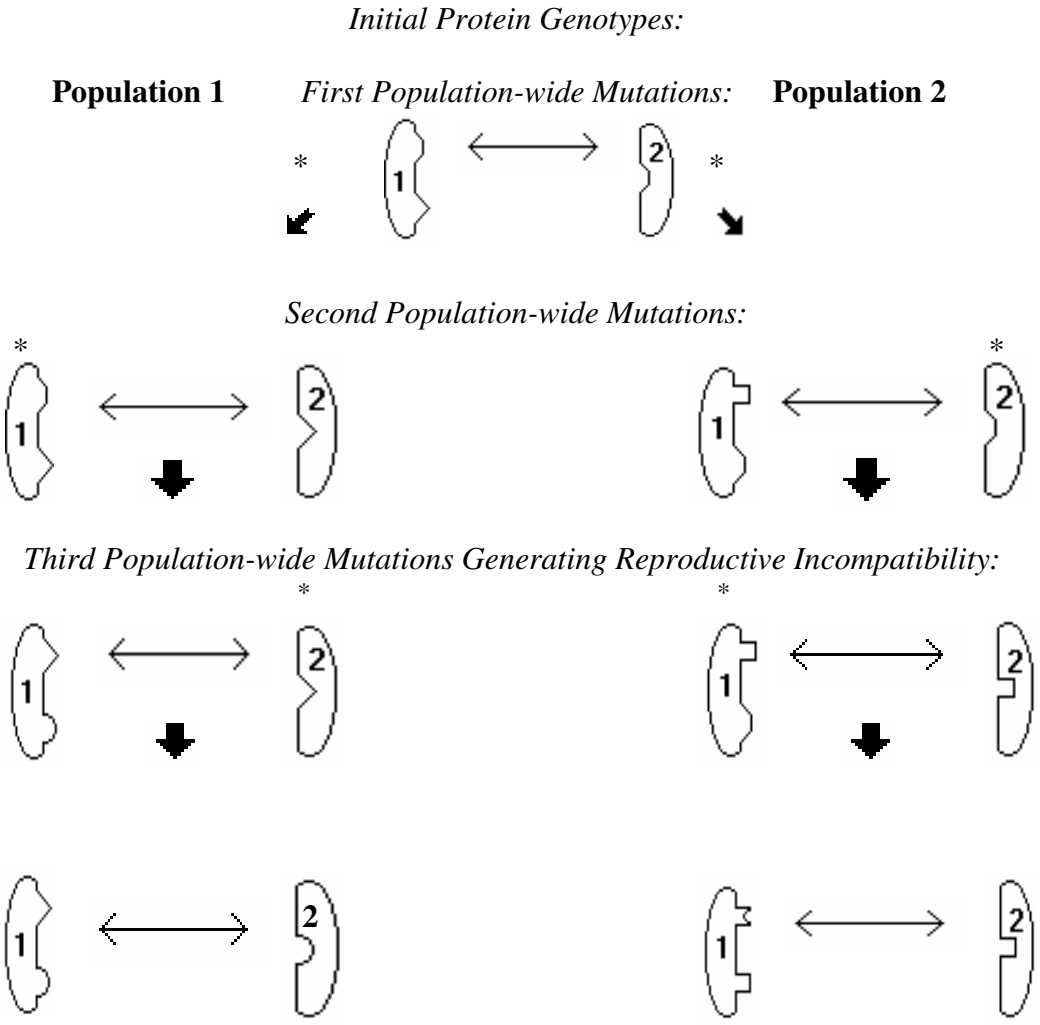

Figure 3.1. A Demonstration of a Speciation Event Between Co-adapted Proteins in Two Populations Descended From a Common Population. From a Single Ancestral Population Arise Two Offspring Populations That are Reproductively Incompatible With Each Other After Three Mutation Events. Each Mutation Event is Highlighted With a "*".

the cycle goes on. This cascading effect of mutation and redefining of local functional limits I will call *co-adaptive drift*. Co-adaptive drift combines the actions of random drift, occurring within functional limits, with natural selection, which constantly patrols these functional limits. The creation of novel variants by the exploitation of allelic diversity allowed by a protein's local functional limits is itself an action of random drift. As alleles randomly drift within their local functional limits, they are redefining the local functional limits of their neighboring proteins. In doing this, they are in fact redefining the selective constraints on those neighboring proteins.

As the archetype's genealogy descends through time, its genetic makeup slowly changes in accordance with this co-adaptive drift. Central to this co-adaptive drift is the push-and-pull interrelation of functional mutations among the protein components of systems. This interaction is continuously defining and redefining the local functional limits of proteins.

One might foresee this constant co-adaptive drift allowing higher-order systems to evolve through a near limitless number of compositional avenues. These system compositions, it could be said, could conceivably adapt to assume an equally limitless number of functional roles. However, this overlooks a basic characteristic of such systems. These systems are functionally accountable to the organism to perform some particular function. This is insofar as these systems are part of the cumulative functional integrity of a higher-order system. It is on the basis of this accountability that these system variants are selected for or against. The same applies to these systems' components, and ultimately to their constitutive proteins as well. The functional mutation capacity of a

system's components is limited to variants that contribute to the system's successful functioning as a whole. Or in other words, they are simply restricted by their global and local functional limits. Likewise, any higher-order systems' functional mutation capacities are limited to versions that allow the successful functioning of systems to which they belong. The cycle of co-adaptive drift via redefined local functional limits will thus be bounded. This is due to each protein's necessity to conform to its respective global functional limits. These global limits do, after all, constrain the protein to any potential functionally operative variants.

On the population level, co-adaptive drift could ultimately produce general *reproductive incompatibilities* between populations of archetype's descendents. This results from the drifting apart of the respective local functional limits of multiple proteins between these descendent populations. A general *functional incompatibility* for these proteins occurs if their respective local functional limits drifted apart until they were non-overlapping. Enough functional incompatibilities between proteins of one population and their synonymous systems in the other could create a reproductive incompatibility between these populations. These functional incompatibilities could eventually create multiple new species, or generally reproductively-isolated populations, within the archetype's *genus*. A genus in these Intelligent Design paradigms includes all organisms descended from one archetype. These new species still maintain the archetype's body plan. However, their co-adapted genetic constitutions have realized mutually non-overlapping areas within the system limits of the archetype's body plan.

New species are only generally reproductively incompatible because many variants with overlapping local functional limits are not entirely eliminated. Thus, when such proteins are expressed in hybrids of the two populations, they can function successfully. Wild type variants of these proteins would be functionally incompatible in hybrid organisms. This theory then predicts that on occasion, hybrids between two species within a genus could be reproductively successful. This only occurs with the correct assortment of alleles. These hybrids could possibly even display a superior fitness to the originals (depending on the environment).

A functional incompatibility between two species in a few essential protein systems would likely be sufficient for reproductive incompatibility. Wallace though has offered a stronger version of a form of genomic co-adaptation. He suggests that this extends out to include the co-adaptation of an organism's entire genome.[3] This could also be concluded from our current development of protein functional limits. If this treatment of functional compatibility is correct, conclusions from these principles should apply to all natural protein systems, not simply to a few examples. One of those conclusions is that entire genomes of organisms would be collectively co-adapted. For our present purposes, genomes are the source of an organism's cumulative functional integrity. If two archetype populations co-adaptively drifted apart over a sufficient time period, their co-adapted genomes yielding new species would not be surprising.

CONCLUSION

This chapter developed both the notions of co-adaptive drift and cumulative functional integrity within organisms' physiological hierarchy. These principles lend support to a holistic view of biology espoused by the Intelligent Design frameworks presented here. These two notions also address the primary issues pursued in this project. Cumulative functional integrity addresses the first project issue by explicitly defining larger systems as components of smaller, quantifiable systems. This is not merely as a collection of parts though. It is instead as an integrated whole wherein lower-order systems fulfill functional roles of higher-order systems. The entirety of these relationships is the organism's physiological hierarchy. Co-adaptive drift is also a theory of integration. The notion of co-adaptive drift relies upon protein functional limits and system limits within the physiological hierarchy. Cumulative functional integrity and co-adaptive drift also address the second primary issue of this project. This issue concerns describing potential ancestral or future variants of biological systems by examining extant versions of those systems. Both notions do this by an extended application of protein functional limits, culminating in the description of system limits. Taken by themselves, both types of limits provide a rough sketch of possible ancestral and future variants of proteins and systems. Chapter IV will examine natural selection in relation to the underlying theme of system's integration developed in this chapter.

NOTES

1. An alternative to this phase of intervening-maintenance is a continuation, even a perpetual continuation of guided processes throughout any stage of life. Additionally, studies of historical biology would be quite unreliable if there was never a time when natural processes were the sole perpetuators of an organism's physiological makeup. This is because the predictability that natural processes acting in isolation provide are absent in that case. Likewise, there would surely exist many, if not most, organisms on the planet exhibiting gene products that contribute to partially-formed biochemical systems. Although this is not impossible, it certainly does not seem to be what modern biology is discovering.

2. This will only fully apply to dominant alleles. This issue will be discussed more in the Chapter IV.

3. Bruce Wallace, "Coadaptation Revisited," Journal of Heredity 82.2 (1991): 89-95.

CHAPTER IV

NATURAL SELECTION REVISITED

In each of the two Intelligent Design scenarios presented here, archetypes originate through one of two types of guided processes. After this origin, the archetypes and their descendents are subjected to purely unguided, natural processes for the duration of their genus' existence. One of the unguided processes that genera will be primarily influenced by is natural selection. In this chapter we will consider how protein functional limits and system limits relate, through natural selection, to fitness levels of allelic and polygenic phenotypes.

THE MECHANISM OF NATURAL SELECTION

Natural selection, as a mechanism, preserves adaptive phenotypes and eliminates maladaptive ones. These actions transpire based upon those phenotypes' functional merits. This occurs on the organism-level because the organism is the fundamental unit of selection. Natural selection is then first a discussion about whole organisms living, reproducing, and dying based upon their collective functional merits. Organisms though do not exhibit multiple fitness levels reflecting the multiple fitness levels of each individual allele they carry. Instead, whole organisms will either survive and reproduce, or they will not. Whatever their fate is, is the fate of all of the genes that individual organism carries.

Notice also that phenotypes, and not necessarily their alleles, are either preserved or eliminated by the actions of natural selection. As was said previously though, alleles are

the fundamental unit of inheritance, not phenotypes. Genes cannot themselves be selected for or against apart from consideration of their protein's function. Thus, to say a gene is selected for based on function is to say that its phenotype (i.e. its expressed protein) is being selected. Genes can then only be acted upon by natural selection, via their functional merits, when their varying phenotypes are expressed.[1]

This assessment precludes alleles contributing to polygenic or polyallelic traits from being consistently selected. While these alleles are individually necessary for their traits, they are not sufficient for those phenotypes' expression. Traits such as these are polygenic traits and recessive and other polyallelic traits. Recessive phenotypes are polyallelic because they are only expressed when both copies of that allele are present. This is the sole time recessive allele phenotypes are acted upon by natural selection. Thus, due to this polyallelic nature, these are the only times recessive alleles are selected themselves. Traits resulting from codominant alleles are also polyallelic in this sense because both alleles are necessary to yield the codominant phenotype.

Genes of polygenic traits are equally subjected to natural selection only when those traits are expressed.[2] An example of this kind of trait is the extra set of wings on four-winged fruit flies. This phenotype in fruit flies is the product of three specific alleles in distinct genes.[3] All three alleles in these respective genes must be present in order for the phenotype to be expressed.

Even if a polygenic phenotype was eliminated in one generation (for example), its constitutive alleles would not be. This is because the individual genotypes are insufficient to produce the polygenic phenotype. Thus, due to this hybrid composition, specific

polygenic traits will only appear sporadically in a population or species. They will not be consistently inherited directly from one generation to the next (in sexual species). Dominant alleles alone are sufficient to simply consistently produce their phenotypes whenever they exist in organisms. The phenotypes of dominant alleles' appear then to be the only traits that can be simply consistently passed on in sexual species. Dominant alleles that can participate in codominant phenotypes are, however, an exception to this because those alleles can participate in two or more different phenotypes.[4]

So far, I have suggested that selection occurs on the organism-level via expressed phenotypes. From this we could assess whether natural selection selects for advantageous phenotypes or against disadvantageous ones. This requires determining situations where an allele's protein's functional merits consistently afford the organisms carrying it a functional advantage. This applies insofar as the function of that allele's protein is concerned.

Table 4.1

Viable and Defective Essential Proteins Compared to Different Genetic Backgrounds.

	Fit Genetic Background	**Unfit Genetic Background**
Viable Protein	Organism can survive	Organism cannot survive
Defective Protein	Organism cannot survive	Organism cannot survive

We can then develop a scenario such as that displayed in Table 4.1. Here, the functional merits of two different dominant alleles are compared against different genetic backgrounds. One of these alleles produces a viable essential protein within an essential protein network. The protein is viable in that it exists within its local functional limits. The other allele produces a defective essential protein in another essential protein network. The viable protein's expression within an organism's otherwise fit genetic background allows that organism the ability to survive. The same viable protein expressed in an organism with an unfit genetic background does not yield a viable organism. This organism is already unfit because of its genetic background. Adding this viable protein will not change that fact. The viable protein in this case is necessary for the organism's survival, although it is not sufficient to ensure this survival potential. A defective protein expressed within organisms with either genetic background would not permit survival. The defective protein ensures this lack of fitness in both cases. Thus wherever this defective protein's allele exists, it can cause an equally detrimental effect upon that organism. This defective protein is not necessary for the organism's death, but it is sufficient for it.

It appears then that maladapted phenotypes are the only instances where phenotypes are simply consistently subjected to selective pressures. Likewise, the only times alleles are consistently subjected to such pressures is when they are dominant. It appears then that natural selection only selects simply consistently against alleles.

Furthermore, the only alleles that will be consistently subjected to natural selection over time are deleterious dominant alleles not participating in codominant phenotypes.

This conclusion could equally be drawn from the implications of cumulative functional integrity developed in Chapter III. A main implication of cumulative functional integrity highlights the integration of each order of specified complexity. Indeed, a disabled system causes the malfunction of each higher-order system that included it in its cumulative functional integrity. Malfunctioning systems within an organism's cumulative functional integrity also cause the entire organism's malfunctioning. In other words, the organism would be unfit for reproduction or perhaps even survival. Thus, what is fundamentally advantageous for an organism is having a *harmonious* cumulative functional integrity. A harmonious cumulative functional integrity is one that allows the system's particular function. Thus, that system's components are properly integrated together, and consequently its constituent components (also systems) are as well. Systems with disharmonious cumulative functional integrities possess internal disintegration or malfunctioning constituent systems. Individual phenotypes can only potentially offer an advantage (beyond mere survival without competition) when their organism's cumulative functional integrity is harmonious.

THE OUTCOMES OF NATURAL SELECTION

Two different outcomes can come from this one mechanism of deleterious phenotypes being selected against. The first deals with alleles that have a direct correlation between their existence and their phenotype's existence. As discussed earlier, dominant

alleles not participating in codominant phenotypes are the only alleles that fall into this category. Since essential deleterious dominant alleles of this type are consistently selected against, it is only a matter of time before they go extinct. Natural selection's first outcome is then to effectively purge the species of these essential deleterious dominant alleles. This outcome is a tremendous means of preserving a population/species' fidelity against the onset of genetic diseases.

A second outcome of natural selection concerns the elimination of deleterious polygenic and polyallelic phenotypes composed of sets of otherwise viable alleles. These phenotypes are eliminated in the individuals expressing them. However, the individual alleles that contribute to these phenotypes can persist in the gene pool. In doing so they could continue to serve as constituents of different, non-deleterious phenotypes. In this way, the species can purge deleterious polygenic and polyallelic phenotypes as they arise. The diversity of alleles that combined to create these phenotypes could be preserved though. This preservation has two consequences. First, it makes it possible for the viable phenotypes that these alleles are constituents of to be preserved. Second, should a situation arise where these deleterious polygenic and polyallelic phenotypes become advantageous, they can still possibly be expressed. Eliminating deleterious polygenic and polyallelic phenotypes also benefits the species by preserving and promoting a viable genetic diversity. This is assuming the eliminated alleles are not removed by random drift. It also helps by maintaining an advantageous adaptive malleability to different circumstances with this diversity.

Natural selection's effects on polygenic and polyallelic traits apply to deleterious recessive alleles in a somewhat different manner. For at no time are polyallelic phenotypes of expressed deleterious recessive alleles advantageous under normal environmental conditions. Otherwise these alleles would not themselves be classified as deleterious. Natural selection does benefit the species here by removing two of these alleles along with this organism. Through this action, the frequency of this allele will be kept to a minimum in the species. This allows for the least possible number of individuals homozygous for this allele. However, recessive phenotypes cannot be put into the same category of the polygenic and polyallelic phenotypes discussed above. This is because any time the functional effects of these recessive alleles are manifest results in the deleterious phenotype. Individual alleles contributing to the first type of deleterious polygenic and polyallelic phenotypes can still contribute to beneficial phenotypes. Deleterious recessive alleles cannot. Thus, the preservation of otherwise viable alleles still benefits a species' genetic diversity and potential for adaptation. The preservation of deleterious recessive alleles does not do this. That is, at least not in *those* environmental conditions.

A deleterious recessive allele's protein can be outside of its local, and possibly its global functional limits. This composition is what can make it deleterious with respect to its second-order system. A special case could exist, though, that allows these proteins' composition to be viable. Here, a recessive allele's protein could be outside of its local limits, yet remain within its global limits. Its peripheral proteins could then mutate such that they redefine the deleterious allele's local functional limits. This would make the phenotype of this protein viable again within that new genetic background. This cannot

happen if the allele is outside of its global functional limits though. This is one instance where deleterious recessive alleles can yet contribute to their species' adaptive malleability. Otherwise, the preservation of these recessive deleterious alleles appears to provide little utility in a species' gene pool. The preservation of such alleles provides a useful tool for those studying historical biology though. This tool will be discussed more in Chapter V.

QUALITATIVE AND QUANTITATIVE MUTATIONS

A protein's local functional limits represent *qualitative* boundaries that define a protein's functional capabilities. This boundary delineates where the second-order protein network can function or malfunction based on that protein's functional performance. This applies if this protein is unique to its functional role in that system. There is also another way a protein's functional performance can be characterized that has not yet been considered. Multiple protein variants could exist within their local functional limits, yet still maintain varying, yet viable, fitness standings. In this scenario, a variant exhibiting superior fitness over a second variant is *quantitatively superior*. The second variant here is the *quantitative inferior* of the first. The term "quantitative" in this case captures the notion that both types of protein are qualitatively equal. They are equal in that they both reside within their local functional limits. However, they are still distinct in the degree of efficiency by which they perform their particular functions, one being relatively superior to the other.

A mutation substitution cannot qualitatively enhance a viable protein variant's particular functioning. Viable variants can already sufficiently perform their particular functions, and therefore having no qualitative room to improve. This is the case whether or not the variant still has quantitative room to improve. This can be illustrated by considering a car. If a car is to be legally permitted to drive in the state of Texas, this car must meet several basic functional requirements. These are fundamental requirements that allow it to operate successfully on any road. These requirements include a competent four-wheel frame, functional exhaust system, and so on. Additionally, a car must meet various safety requirements dictated by the state. These include well-functioning turn signals, windshield wipers, seat belts, and so on. By these standards, old beat-up Pintos barely passing the safety test are qualitatively equal to souped-up Ferraris in the eyes of a safety inspector. In this instance, the inspector is acting as the selection mechanism. That is, he facilitates what is and is not allowed to 'survive' on the street (its natural environment). The fulfillment of each requirement, however, merely allows a car to be legally allowed onto the road. Past these qualitative standards, the efficiency and magnitude of a car's features can still be quantitatively enhanced. They could even be dramatically enhanced – hence the Ferrari. This improvement though is distinct from any qualitative threshold. A quantitative improvement is useful only after the fulfillment of these qualitative standards. Thus a Pinto that passes its safety inspection will be allowed onto the street. However, a Ferrari with a broken taillight would not. Despite its quantitative enhancements, not fulfilling the safety inspection standards removes this Ferrari from its natural environment

(roads). Similarly, proteins must fulfill their qualitative functional standards to have the sustained option of quantitatively improving their particular functions' efficiencies.

STAGES OF EXTINCTION

Qualitatively inferior phenotypes are those that will not allow organisms carrying them to survive, were they relied upon. However, a phenotype could meet its qualitative requirements, yet be quantitatively inferior to another phenotype. Qualitatively adequate phenotypes can then be divided into two categories: quantitatively superior and quantitatively inferior phenotypes. Organism's carrying quantitatively inferior phenotypes may not be able to compete with organism's carrying their quantitatively superior counterparts. They can still survive, however, in the absence of competition from that quantitatively superior phenotype at least. A particular phenotype's relative fitness can change such that it alternates between these qualitative and quantitative categories. To understand this better, we will consider a way to categorize these oscillating phenotypic selection statuses. This characterization will organize phenotypic selection statuses into three *stages of extinction*. These stages will reflect the means by which their phenotypes are being selected against.

The qualitative requirement of phenotypes describes the phenotype's relationship to its environment. This relationship can be either internal, by its physiological interactions, or external, by its interactions with the external environment. The quantitative requirement, though, describes the phenotype's fitness compared to competing variations of that phenotype in other organisms. Quantitatively inferior phenotypes are then only

deficient when out-competed by superior phenotypes. This competition concerns everything from resources and mates to all other relevant aspects of survival and propagation. In the absence of organisms carrying quantitatively superior phenotypes, organisms carrying inferior phenotypes can consistently propagate successfully. This is insofar as that inferior phenotype's functional merits and contributions are considered.

We have considered two types of phenotypes that allow organisms carrying them the capability to consistently survive and reproduce. These are quantitatively superior phenotypes under all normal circumstances, and quantitatively inferior phenotypes when their superior counterparts are absent. The category containing these types of phenotypes we will call *extinction-stage one*. Phenotypes of this stage should constitute the bulk of extant phenotypes of any random sampling. This stage describes phenotypes that are consistently functionally successful. In fact, this is the only stage that contains phenotypes that are not consistently selected against based upon their functional merits. The remaining extinction-stages, however, do describe phenotypes that will be selected against in varying circumstances.

The second category, *extinction-stage two*, results from relationships between quantitatively inferior and superior phenotypes. In other words, this category concerns competition between different organisms over resources, mating rights, etc. This stage includes quantitatively inferior phenotypes existing in the presence of their quantitatively superior phenotypes. These phenotypes will be consistently selected against in the presence of their superior counterparts. Though, in the absence of these superior

phenotypes, the inferior phenotypes exist in extinction-stage one. Within that stage, they could consistently and successfully propagate based upon their functional merits.

An example of extinction-stage two is certain instances of acquired antibiotic resistance in bacteria.[5] The development of resistance to antibiotics in bacteria often occurs because of a mutated protein target for the antibiotic. The result is a reduced affinity for the target, which decreases the antibiotic's opportunity to kill the bacteria. This resistance is most often accompanied by a fitness cost to the bacteria though. Resistant bacteria are able to survive, but are often out-competed by the wild type strains in the absence of the antibiotic. In the antibiotic's absence, the resistant bacteria's phenotype is in extinction-stage two while wild type phenotypes are in extinction-stage one. When the antibiotic is present though, these phenotypes switch their respective extinction-stage assignments. The resistant phenotype then enters extinction-stage one, becoming the quantitatively superior phenotype.

If extinction-stage two phenotypes continue to persist in this stage, they could be selected out of the gene pool. This is possible if their stage one counterparts are in constant competition with them. The consistent elimination of polygenic phenotypes will not definitively alter their constituent alleles' frequencies in the species' gene pool. This is because polygenic phenotypes are not consistently inherited. It could, however, lower the frequencies of these constituent alleles. Phenotypes of essential dominant alleles continuing to persist in extinction-stage two, however, will likely be eliminated over time. This occurs if these phenotypes remain in constant competition with superior, stage one phenotypes.

Table 4.2

Quantitatively Superior and Inferior Alleles Compared Against Different Genetic Backgrounds.

	Fit Genetic Background	**Unfit Genetic Background**
Quantitatively Superior Phenotype	Organism survives and procreates more	Organism does not survive and procreate
Quantitatively Inferior Phenotype	Organism survives and procreates less	Organism does not survive and procreate

One might argue that stage two phenotypes are then not selected against. Rather, it is the superior phenotypes that are being selected for, which in turn eliminates the inferior phenotypes. To rectify this we can again consider a situation outlined in Table 4.2, similar to that described in Table 4.1. In situations where both phenotypes are in constant competition, the quantitatively inferior phenotype should be consistently selected against. This applies under circumstances that permit each phenotype to remain it its respective extinction-stage. The superior phenotype though only affords its organism an advantage if that organism's genetic background is fit otherwise. Thus, what is consistently subjected to natural selection in this situation is the inferior phenotype of extinction-stage two, which is selected against.

Extinction-stage three is the final category of phenotypic selection. This stage contains phenotypes that cannot successfully survive and reproduce in any natural scenario. These phenotypes are either essential protein variants existing outside their local

functional limits, or essential higher-order systems existing outside their system limits. These phenotypes will then cause a malfunction in the cumulative functional integrity of whichever higher-orders contain them. This malfunction extends on up to the organism-level, if these phenotypes were essential. Organisms carrying them are then incapable of surviving/reproducing because of the internal malfunction inherent to these phenotypes. Essential phenotypes in this stage should always be eliminated over time, if not immediately, even when there is no competition.

THE BENEFITS OF NATURAL SELECTION

Natural selection so construed plays a vital role in the two Intelligent Design scenarios developed here. First, as a preservation mechanism, natural selection eliminates harmful alleles by eliminating their harmful phenotypes. This is consistently true of dominant deleterious alleles not contributing to codominant phenotypes because of the sufficiency of the allele's presence to produce its deleterious phenotype.

This is not the case with phenotypes of deleterious recessive alleles, other deleterious polyallelic phenotypes, and more complex polygenic phenotypes. This is because of the insufficiency of their individual allele's presence to produce these deleterious phenotypes. By eliminating these phenotypes when they do occur, natural selection actively suppresses this stock of recessive alleles and alleles contributing to other deleterious polyallelic phenotypes. This aids in preventing more long-term harm throughout the gene pool in the future. Individual alleles of more complex polygenic phenotypes will be less suppressed by natural selection when more genes are involved in

the phenotype. For example, alleles of deleterious polygenic phenotypes with three genes will occur together and be eliminated more often than alleles of phenotypes with four or five genes. An organism's likelihood of possessing the correct allelic combination and expressing these deleterious phenotypes increases exponentially as more genes become involved. Very quickly the number of genes involved in these polygenic phenotypes will dictate that the likelihood of the phenotype occurring is nearly equivalent for all individuals in a population (or species).

Natural selection also serves to minimize deleterious effects in a gene pool while preserving a sizable allelic diversity. It does so by suppressing alleles more likely to cause deleterious phenotypes, while only mildly affecting those less likely to cause harm. This allelic diversity will allow for multiple variants of polygenic traits at any given time by recombination. Additionally, it maintains avenues for mutation to create new alleles. These new alleles could in turn propagate co-adaptive drift, which allows further allelic diversity to be actualized by mutation.

CONCLUSION

The characterization of natural selection developed here serves several ends for this project. Being approached from a systems integration perspective, it expands the development of cumulative functional integrity presented in the previous chapter. In this, natural selection can explain phenotypes' selection statuses as they relate to an organism's cumulative functional integrity. This then explains these phenotypes' selection statuses in regards to their level of selection – the organism. As a result, this characterization of

natural selection is also relevant to the proposal of co-adaptive drift. Here, natural selection describes the elimination of systems and proteins that are unharmonious with an organism's cumulative functional integrity. By eliminating unharmonious phenotypes while preserving a harmonious allelic diversity, harmonious co-adaptive drift can be perpetuated. The selection status of alleles of polyallelic phenotypes provided here will also be useful in the historical methodology developed in Chapter V.

NOTES

1. Instances exist, however, where two closely linked genes are typically passed on as a unit. Here, the action of natural selection on one gene affects the fate of the other nearly as often. However, if natural selection primarily acts by eliminating deleterious traits and not by preserving adaptive ones (as will be claimed), this can prove more detrimental to both genes. For this now means that one gene could be eliminated either based upon its own functional merits or those of its linked gene. If one gene is maladaptive and essential, it should not then be preserved if its partner is viable. On the contrary, both are now susceptible to selective elimination.

2. This does not necessarily apply to genes that can contribute to multiple systems and phenotypes.

3. Edward Lewis, "A Gene Complex Controlling Segmentation in Drosophila," Nature 276 (1978): 565-570.

4. Alleles that are partially dominant could also be classified as exceptions in this sense.

5. Dan Andersson, "The Biological Cost of Antibiotic Resistance," Curr Opin Microbiol 2 (1999): 489-493.

CHAPTER V

FUNCTIONAL MOLECULAR PHYLOGENY AND CONCLUSION

The first aspect of the bottom-up approach discussed in Chapter I is what has been primarily developed up to this point. This aspect defines biological systems as integrated subsystems and, ultimately, as quantifiable components. In this chapter, I will address the second aspect of this approach. This aspect attempts to define ancestral versions of extant biological systems in relationship to their functional limits in order to determine potential design points. This will consist of a research methodology I will call functional molecular phylogeny. This methodology primarily aims at determining potential points of design for genera. It does so by first examining the histories of extant alleles of proteins. It then utilizes the notions of protein functional limits and incomplete allelic purging from gene pools. This methodology follows three general steps:

1. An investigator begins by attempting to infer the alleles of the 'original' protein variants in a given species' gene pool. The first step in this is to examine the stepwise causal relationships between a gene's different alleles. These 'original' alleles would have existed in the archetype in the two Intelligent Design scenarios, were they the case. The approximate times these ancestral alleles came into existence could also be established. In addition to the species level, this procedure also applies to the population level and higher taxonomical levels such as the genus level.
2. After inferring a gene's original allele(s) in one species, the protein's functional limits of that original allele(s) are considered to determine if that allele was

compatible with the ancestral genetic background of other species that are presumably closely related. This procedure, then, also needs to be applied to the relevant genes in the genetic backgrounds of these presumably closely-related species. Such compatibility is necessary (if the relationship is essential for fitness) if reproductive compatibility between the two species existed at that time. Confirming such incompatibilities can suggest proposed design points (insofar as those genes are considered).

3. Finally, this procedure is repeated for multiple other protein-coding genes within the species (or other taxa being considered). The resulting pattern of proposed allelic design points can suggest whether that species exhibits a proposed design point for the genus.

As stated previously, the falsifiable prediction of Intelligent Design theories is the presence of a pattern of allelic design points at the origin of new genera. The inability to find any consistent pattern of proposed allelic design points at a proposed genus's design point is, then, one way to falsify the Intelligent Design scenarios. This fulfills the desirable quality of being empirically falsifiable raising the status of it as a scientific theory in a Popperian scheme. The methodology described in this chapter is a possible way to determine these results.

MOLECULAR CLOCKS, ANALOGY AND HOMOLOGY

A useful tool that has emerged in the field of molecular phylogeny has been studies of "molecular clocks". In these studies, DNA sequences perceived to be identical (the

same sequences/genes from the same species) or *homologous* (deriving from common ancestry) are compared either between different populations or species. These comparisons are used to determine approximate genealogies of these groups based on their degree of sequence homology. The maximum parsimony of this homologous relationship is used to configure their most likely lineage, given common descent. This is calculated by the combined rarity of the transition, transversion, insertion, and deletion events observed between these sequences. Approximate divergence times between the sequences can be determined by applying generally accepted probability rates of each mutation event.

Functional molecular phylogeny utilizes a modified version of this methodology. There are two main distinctions between molecular clock use in functional molecular phylogeny and in contemporary molecular phylogeny. One is the instances in which molecular clocks are employed. A second is certain aspects of their interpretation. These distinctions result from functional molecular phylogeny's primary end-goal – determining proposed design points. Contemporary molecular phylogeny proceeds from a position of descent by solely unguided natural selection uninterrupted by design points. This view suggests that analogous features from different species are derived from a feature in a common ancestor. They are analogous in that they code for nearly compositionally identical proteins that perform nearly identical particular functions. Likewise, certain similar, yet non-analogous proteins from different species are taken to be orthologously derived from common ancestral proteins. This ancestral protein's gene would have duplicated and since diverged sequentially and functionally. Thus, this method usually equates analogous DNA and protein sequences with homologous ones either across species

lines, or even higher taxonomical classifications as the case demands. There is typically little reluctance in classifying analogous sequences as homologous between different species where reasonable.

This is all quite legitimate because homology is the best explanation for analogies from a naturalistic evolutionary perspective, free of design points. For example, the human protein cytochrome c is analogous in function and sequence to a protein found throughout the animal kingdom, also named cytochrome c. Every time a sequentially and functionally similar protein is discovered in an animal, it is also called cytochrome c. These similarities are seen to exist because all analogous cytochrome c proteins among each species are seen to be homologous. The same holds for virtually all other instances where analogous proteins exist among different species. These cytochrome c proteins are paralogous - homologous proteins that each conserved their original particular functions. That is, they are not several different proteins that just happen to have these similarities (convergent evolution). Instead, they are truly the same protein that has been consistently functionally preserved throughout the common lineage of each species.

A reliable conclusion of functional molecular phylogeny can come only from a study of multiple genes in a species' genome, in addition to studies of the relevant genes in presumably closely-related species. Concluding a proposed design point for a species needs to be corroborated by several independent protein studies revealing a pattern of proposed allelic design points. Again, what the two Intelligent Design scenarios predict is a clear pattern of multiple proposed allelic design points converging at a specific historical point.

THE OVERALL APPROACH

Functional molecular phylogeny studies begin by attempting to reconstruct the histories of individual genes coding for proteins. These histories are developed from all individual alleles of that gene found throughout the given species. This is regardless of the populations in which they are found, and regardless of their frequency in those populations.

Codon Position -	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	
Allele 1	5'- AAT-	[CGC]	-GGT-TGT-CCA-	[TTA]	-CAT-GAA-TTC – 3'	
		N	R	G	C	P	L	H	E	F
Allele 2	5'- AAT-	[CGA]	-GGT-TGT-CCA-	[ATA]	-CAT-GAA-TTC – 3'	
		N	R	G	C	P	I	H	E	F
Allele 3	5'- AAT-	[CGA]	-GGT-TGT-CCA-	[ATT]	-CAT-GAA-TTC – 3'	
		N	R	G	C	P	I	H	E	F
Allele 4	5'- AAT-	[CGA]	-GGT-TGT-CCA-	[ACT]	-CAT-GAA-TTC – 3'	
		N	R	G	C	P	T	H	E	F

Figure 5.1. Set of Alleles Comprised of All Extant Variants in the Species.

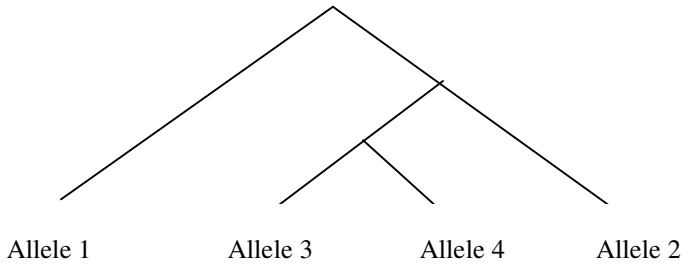

Figure 5.2. A Phylogenetic Tree of the Allele Set Based Upon a Stepwise Mutation Model.

After a gene is designated for study, samples of all extant alleles within the species are gathered for analysis (Figure 5.1). Sequence comparisons are then run within this set to determine the phylogenetic relationships between the different alleles based upon the maximum parsimony. This should reflect the stepwise mutational history of those alleles insofar as they reveal such a history. This comparison yields a phylogenetic tree of this allele set that reflects the stepwise mutational history of that gene in the species (Figure 5.2). It should be noted that my example used in Figure 5.1 and those in the remainder of this chapter are point mutations. However, functional molecular phylogeny applies not just to point mutations but to all types of heritable genetic mutations as mutation events. Other mutation events could involve large numbers of nucleotides, which can complicate assessments of phylogenetic trees. Point mutations are used to represent these mutation events in this chapter primarily for explanatory purposes.

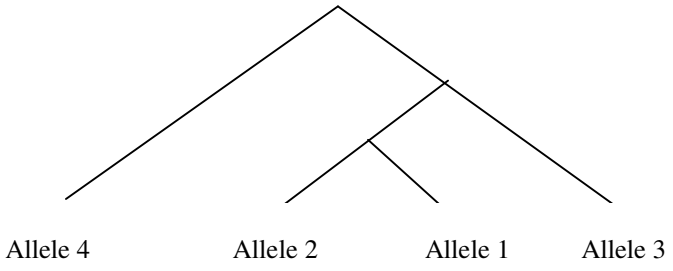

Figure 5.3. An Alternative Phylogenetic Tree of the Allele Set From Figure 5.1.

Allele 1 ⇔ Allele 2 ⇔ Allele 3 ⇔ Allele 4

Figure 5.4. The Symmetrical Stepwise Causal Relationships Between Alleles 1 through 4.

The phylogenetic tree in Figure 5.2 displays the alleles' causal relationships beginning at the divergence between Alleles 1 and 2. Allele 3 appears to have arisen as a permutation of Allele 2. Allele 4 apparently originated through a mutation in Allele 3. However, it quickly becomes apparent that such a tree is underdetermined in one sense. Furthermore, a tree such as that from Figure 5.3 could easily replace this one. Figure 5.3's tree reverses the structure of Figure 5.2's tree. This problem of polarity does not totally invalidate the causal relationships found between the alleles though. Rather it highlights the symmetrical nature of these causal relationships between the alleles, insofar as sequence alone is concerned (Figure 5.4). In these symmetrical relationships, Allele 3 could have yielded Allele 4 through stepwise mutations, or vice versa. However, Allele 1 did not directly yielded Allele 4 in a stepwise scheme. One could claim that it yielded it indirectly though. Here, it needs to be argued that Allele 4 arose from Allele 1 through extinct intermediates, and that compelling reasons exist for thinking so. There are as many possible phylogenetic trees for this gene's history as the symmetrical causal relationships of Figure 5.4 allow. The discovery of the true polarity of these relationships then needs to be further informed by some auxiliary considerations.

Codon Position -	1	2	3	4	5	6	7	8	9	
Allele 1	5’- AAT-	CGC	-GGT	-TGT-CCA-		TTA	-CAT-GAA-		TTC	– 3’
Allele 2	5’- AAT-	CGA	-GGT	-TGT-CCA-		ATA	-CAT-GAA-		TTC	– 3’
Allele 3	5’- AAT-	CGA	-GGT	-TGT-CCA-		ATT	-CAT-GAA-		TTC	– 3’
Allele 4	5’- AAT-	CGA	-GGT	-TGT-CCA-		ACT	-CAT-GAA-		TTC	– 3’
Allele 5	5’- AAT-	CAA	-AGT	-TGT-CCA-		ATA	-CAT-GAA-		GTG	– 3’
Allele 6	5’- AAT-	CAA	-AGA	-TGT-CCA-		GCA	-CAT-GAA-		GTG	– 3’
Allele 7	5’- AAT-	CAA	-AGA	-TGT-CCA-		GTA	-CAT-GAA-		GTG	– 3’
Allele 8	5’- AAT-	CAA	-ACT	-TGT-CCA-		ATA	-CAT-GAA-		GTG	– 3’

Figure 5.5. Dissimilar Allele Set Found Between Two Species.

Allele 1 ⇔ Allele 2 ⇔ Allele 3 ⇔ Allele 4

Allele 8 ⇔ Allele 5 ⇔ Allele 7 ⇔ Allele 6

Figure 5.6. Two Distinct Sets of Symmetrical Causal Relationships From the Alleles Found in Figure 5.5.

In some cases, a gene’s allele sequences within a species might exhibit a dissimilarity that cannot be easily reconciled with a stepwise mutation model (Figure 5.5). Before resorting to the proposal that these alleles arose via extinct intermediates, an alternative history can be considered. These alleles could exist in two or more distinct, non-overlapping phylogenetic lineages at the species level, yet each descend from two or more distinct ancestral alleles at a higher taxonomical level. Figure 5.6 demonstrates the two independent sets of symmetrical causal relationships that could yield these distinct phylogenies. Each allele’s sequence falls within one of these two distinct sets of symmetrical relationships. With this data, one could reasonable infer that the species originated with both ancestral alleles suggested by these phylogenies. These ancestral alleles’ ultimate origin could be considered by comparing the two resulting phylogenetic

trees with equivalent trees in other alleged closely-related species. This will be addressed in greater detail below. By this approach, one could investigate whether these distinct ancestral alleles within this species had a common origin further back in their history.

THE FUNDAMENTAL UNIT OF FUNCTIONAL MOLECULAR PHYLOGENY

Often in molecular phylogeny studies, relationships between individuals or between groups act as the primary basis for the investigation. These groups can be populations, species, or higher taxa. Individuals within these groups then often act as the fundamental units by which phylogenetic comparisons can be drawn. The fundamental unit in functional molecular phylogeny, however, is the alleles of the gene considered in the particular study. Which specific individuals carry those alleles is not a primary concern.

As discussed earlier, the possible phylogenetic histories are first uncovered by examining the basic symmetrical causal relationships of alleles themselves. This set of possible phylogenetic trees is then narrowed down by examining the general allelic distribution among the different groups being considered. This distribution is considered insofar as the alleles exist uniformly throughout that population. This is due to the possibility of allele sharing with other populations along the periphery of the given population. Alleles arising earlier tend to be distributed more uniformly among all the populations because these alleles arose before later populations broke off of the earlier population(s).

Population:	A	B	C	D	E	F
Alleles:	2	2,3	2,3,4	2	1	1,2

Figure 5.7. Distribution of Alleles Throughout Populations A-F in Species 1.

In contrast, many of the later alleles arose and remained within single populations after further population radiation and isolations. These later alleles are then more localized to both the populations where they arose, and populations descending from them. Figure 5.7 displays the alleles from Figure 5.1 distributed among the different populations of the species. This distribution suggests a phylogenetic tree that could account for the widespread depositing of Allele 2 among the different populations. It also needs to explain the smaller distribution of Alleles 1 and 3, and the single instance of Allele 4. This suggests a tree that is more in concert with that presented in Figure 5.2 than Figure 5.3.

Allele 1 – CG**C**… **T**TA

Allele 2 – CG**A**…**A**TA

Figure 5.8. Alleles 1 and 2 from Figure 5.1.

Instances could exist where, for example, an earlier allele was eliminated in many or all populations by any variety of means. This allele does thus not display a uniform distribution among the populations, despite being an earlier allele. Alleles 1 and 2 from Figure 5.1 exhibit a gap of two bases that suggests such a scenario (Figure 5.8). In this case, the general structure of the phylogenetic tree still holds because of Alleles 1 and 2.

This is due to the reliability of the stepwise mutation history of these alleles. Furthermore, the sequence of the extinct allele from this gap can be reconstructed using both Alleles 1 and Alleles 2. One of these alleles is the extinct allele's *parent* allele - the allele it was derived from; the other is its *daughter* allele - an allele that was derived from it. Which is which, however, is another issue that cannot be resolved simply by Alleles 1 and 2's causal relationships to the extinct allele. A tree displaying a reasonable historical relationship needs to be sought by considering allelic distributions. The case for such a tree will rarely be presented with one hundred percent certainty though (such is the problem of molecular clocks in general). However, a reasonable case could most often be made for the tree displaying maximum parsimony.

Alleles serve as the basis of a functional molecular phylogeny study instead of individual organisms for different reasons. One is that functional molecular phylogeny is not primarily concerned with particular relationships between, for example, different populations within a species. Rather, it is concerned with the ultimate origin of the species itself, including all the genetic diversity it exhibits. Allele frequencies within populations then do not add to the fact that those alleles simply exist within those populations. These alleles either exist in a population as its founding allele(s), or later permutations of the founder(s). Functional molecular phylogeny studies primarily deal with a gene's alleles solely in terms of their causal relationships to the original allele(s). They do not treat them as representatives of population fluctuations. As was discussed previously however, population origins in general play a major role in discovering a gene's correct causal history.

Additionally, each of a gene's alleles within a species exists as a link somewhere along the species' historical, stepwise mutation ladder. This ladder is the gene's history within that species. Elimination of dominant alleles can come about either by consistent selective pressure, or by random drift. Dominant alleles that are actively selected out of a gene pool will themselves fall into one of three categories:

1) Those whose proteins exist outside their global functional limits;

2) Those whose proteins exist inside their global functional limits, yet outside their local limits either when they arose, or at some time since;

3) Those whose proteins always existed within their global and local functional limits, but were eliminated due to quantitative inferiority.

Alleles from category 1) will likely not exist long enough to dissipate through the gene pool and form novel mutant alleles. However, alleles in categories 2) and 3) can possibly persist long enough to yield new alleles by mutation before becoming extinct. These new mutants could themselves be fit genes (falling outside the above three categories). As discussed in the previous chapter, recessive alleles are selected differently than their dominant counterparts. Because of this, recessive alleles could avoid total extinction by any of these three categories. However, they, along with their dominant counterparts, could still be eliminated through random drift after a sufficient amount of time.

In functional molecular phylogeny studies, alleles derived from extinct alleles can be compared to other daughter alleles to infer the sequence of the extinct allele. Likewise, these derived alleles can be compared to the eliminated parent allele's own parent allele if it had not been purged from the species gene pool. Both of these approaches can give

insight into the extinct alleles' sequences and causal relationships. This eliminated parent allele's own parent allele is the daughter alleles' 'grandparent' allele. From these comparisons the extinct parent allele's sequence, which differs by only one base pair, could be reconstructed. The relationship between Alleles 1 and 2 from Figure 5.1, which displays a two base pair difference, could then be reconciled in this way:

Allele 1 – CG**C**…**T**TA **Candidate Allele 1.5a** – CG**C**…**A**TA

Allele 2 – CG**A**…**A**TA **Candidate Allele 1.5b** – CG**A**…**T**TA

Reconciling small gaps in a functional phylogenetic history can be rather straightforward because of the unguided stepwise creation of allele permutations. In this case, Candidate Alleles 1.5a and 1.5b are the two candidates for the extinct allele. This extinct allele could have been a daughter allele of Allele 1 and the parent allele to Allele 2, or vice versa. On the other hand, this extinct allele could also have been the parent of both Alleles 1 and 2. It is unlikely that it is the daughter of both alleles, each having two independent origins. Let us assume for the moment, however, that this allele is an intermediate between Alleles 1 and 2. Rectifying the proper polarity of these alleles' phylogenetic tree still remains to be done. To address this problem, this species' allelic distribution could be compared with that of other species within its genus. This will be discussed more below.

One might suggest that a parental allele could be purged merely as a result of yielding a daughter allele. This daughter allele then somehow 'replaces' the parental allele by this mutation event. Mutation events that create new alleles, though, are incidental

events that occur only in individuals. They do not happen simultaneously throughout groups of individuals. Heritable genetic mutations occur as mistakes in DNA replication, recombination, etc. This is the reason why mutations occur only in individuals carrying that allele instead of in groups carrying it. That is, these mutation events happen independently of allelic fitness level. When new daughter alleles are formed, they are subject to two possible means of elimination. One way this elimination can come about is by selective pressures. This pressure can occur by competition from the parental allele or other alleles. Another mode of elimination is by random drift operating independently of any allelic competition. The parent alleles scattered throughout the remainder of the gene pool are not necessarily affected by this new daughter allele's creation either.[1] These remaining parental allele copies retain their standing in individuals carrying them. These copies then continue to be passed on to future generations unless they are themselves eliminated.

If the parent allele is selectively eliminated, its protein has fallen into either category 2) or 3) listed above. A category 2) scenario suggests that the species' perpetual co-adaptive drift altered the protein's local functional limits. This alteration was sufficient to displace the parent allele's protein outside these local limits. Other alleles' proteins, including those of the parent allele's own daughter allele(s), could exist within these new local limits. A category 3) scenario suggests that the parent allele's protein does exist within its local functional limits (whether new or old limits). However, it is eliminated due to a quantitative inferiority to a competing allele's protein, perhaps its daughter's. In either case, mutation events might create new competition for the parent allele's protein. It is

because of the fitness level of the parent allele's protein relative to that of new quantitatively superior daughter alleles that the parent allele will be eliminated though.

Thus, one should not expect to find historical mutation events in themselves somehow replacing parent alleles with daughter alleles. Perhaps, though, a strong difference in quantitative fitness level existed between the two alleles' proteins. It then could be possible for the daughter allele's protein to out-compete its parent allele's protein. It could even drive the parent allele to extinction in that population or in the species wholesale. Alternatively, many or all other alleles could have had an equally quantitatively inferior fitness level compared to the daughter allele. In this case, these other alleles likely have been also selected against, just as the parent allele was.

POPULATION AND GENUS STUDIES

In addition to its application to species, functional molecular phylogeny can also be applied at the population and genera levels to determine their original alleles. These studies normally proceed in much the same way as described for species studies.

Functional molecular phylogeny analyses of species begin by characterizing the symmetrical causal relationships between a gene's different alleles within a species. The proper allelic relationships are sought by appealing to general allelic distribution found among its sub-levels, the populations. Populations do not have sub-levels to appeal to in the same way that species do. Therefore, the true polarity of a population's allelic relationships is difficult to clarify while only considering the population level. To rectify this, other peripheral populations must be examined to reveal the relationship between

these populations' alleles and those of the population being studied. Otherwise, reconciling these symmetrical allelic relationships with the gene's actual history in the population does not seem very feasible. Simply considering one other population does not seem to suffice though. For if Population 1 displays Alleles 1, 2, and 3 while Population 2 displays Alleles 3, 4, and 5, this still falls prey to this problem of polarity. It is still unclear, then, whether the alleles of Population 1 descended from those of Population 2 or vice versa. The gene's true history becomes more apparent as more populations are considered. This culminates in the most accurate picture being presented by considering all populations in the species. It seems then that a functional molecular phylogeny population study is worthwhile in that it proceeds to higher level investigations. This includes parallel studies of other populations on up to species and genus level studies, and perhaps beyond.

A functional molecular phylogeny genus study does not fall prey to the type of polarity problem that a population study does. On the contrary, it has the potential to yield much more definitive results than even species' studies do. This is because the genus' sub-levels, its species, contain a property that the species' sub-levels, its populations, do not. That is, populations have the ability to interbreed, while species do so only in extreme situations (if at all). Thus one population's individuals could possibly interbreed with another population's individuals, disrupting the allelic distribution among populations. Species do not typically have this opportunity. A genus' species are then more likely to contain an allelic distribution that accurately reflects a gene's causal history within that genus.

Species 1:

Population:	A	B	C	D	E	F
Alleles:	2	2,3	2,3,4	2	1	1,2

Genus:

Species:	1	2	3	4	5
Alleles:	1,2,3,4	1,2,0	0,-1,-2	0,-1,-3,-4	0,-5,-6

Figure 5.9. Distribution of Alleles Throughout Populations A-F in Species 1 and Distribution of Alleles Throughout Species 1-5 in the Genus. The Numbering of the Alleles Reflects the Allele's Relative Time of Origin and Divergence From the Original Allele, Allele 0.

This allelic distribution among the species within a genus provides a preliminary means of determining relationships between these species. Ultimately it could aid in discovering whether different species belong to a common genus. If two species are descended from a common ancestor, they should carry analogous (indeed, homologous) alleles descended from a common ancestral allele(s). As exhibited previously in Figure 5.7, earlier alleles should be more uniformly distributed among a species' populations. Similarly, later alleles tend to be localized in populations where they arose. If two species descended from a common ancestor, the allelic distribution between them should reflect that causal history. The patterns this yields should be equivalent to those yielded between populations within a species, yet more defined. Figure 5.9 displays an extension of the allelic distribution of the populations considered earlier in Figure 5.7. It also examines that species' allelic distribution (Species 1) compared to other species within its genus. This figure illustrates how these causal history and allelic distribution patterns found among a

species' populations also apply to the genus level. The genus' allele distribution also helps clarify the polarity between Alleles 1 and 2 in Species 1 highlighted earlier. That is, whether Allele 1 or Allele 2 came first or neither. In examining only the allelic distribution among the genus' different species, it appears that Allele 2 descended from Allele 1. Allele 1 then appears to have descended from Allele 0, which has since most likely become extinct in Species 1. This distribution reveals a rough sketch of the species' phylogenetic relationships. However, further clarification will depend upon considering the causal relationships of the DNA sequences themselves.

The gene pools of a genus' different species could provide the allele set needed for a study of that genus. Perhaps, though, the common genus of target species was not known. In that case, analogous genes' alleles from gene pools of species suspected to exist within a common genus could be studied. To be considered for such a study, these candidate genes must first fulfill certain phylogenomic criteria. Such criteria dictate:

1) the genes must be analogous by sharing requisite similarities in sequence and in function;
2) the genes must reside in the same location in their respective genomes;
3) the genes must contain equivalent auxiliary considerations such as similarities in intron splicing sites, promoter sequences, etc.

Fulfilling these criteria addresses whether two or more genes are identical between two or more species. Homologous genes such as those yielded by gene duplication then do not adhere to these criteria. This only highlights, though, that these genes are not identical in the manner these criteria seek to demonstrate. It does not highlight whether or not they are

related at all. Such studies then suggest, in part, whether the examined species are capable of existing within a common genus. This extends insofar as the history of that gene is concerned.

ANCESTRAL LOCAL FUNCTIONAL LIMITS

After establishing a gene's history at a particular taxonomical level, the next step in functional molecular phylogeny is to investigate that gene's peripheral genes' histories. These are genes that code for the peripheral proteins in the organism's genetic background that the first gene's protein interacts with. Here, we can assume the taxonomical level in question is a species. Each essential protein variant must have been viable for a species' individuals to have originally bred successfully with each other. By being viable, these variants lay within their proteins' local functional limits. As stated numerously, these local functional limits are determined by a protein's peripheral protein counterparts. This applies both to an original breeding group and all subsequent breeding groups.

A gene's ancestral local functional limits can then be determined by comparing its ancestral sequences with those of its peripheral genes. This applies only insofar as all of these sequences were contemporaneous. This comparison does not establish the sequence(s) of the ancestral allele(s) of the gene, which has already been established by the gene's phylogenetic tree. Rather, this comparison should confirm the proposed ancestral sequence as a truly viable ancestral variant. This variant's sequence lies within its protein's local functional limits of that ancient period, as dictated by these peripheral proteins. Additionally, the polarity of a gene's phylogenetic tree can be further confirmed

by accounting for its ancestral local functional limits. This ancestral allele sequence(s) could also partly confirm the local functional limits of its peripheral proteins' ancestral alleles. Thus, each functional molecular phylogeny study of a gene contributes to verifying the entire genetic background of a species' original members.

Each protein's local functional limits provide the qualitative filter by which a functional molecular phylogeny investigation is confirmed. Indeed each allele within a gene's history must have met these qualitative standards when they existed. That is because at any given time these standards just are the protein's local functional limits. A quantitatively enhanced variant is then equivalent to its parent variant in qualitative fitness. In other words, both of these variants reside within their protein's local functional limits. In functional molecular phylogeny studies, quantitative enhancements are then treated equivalently to their quantitative inferiors. This is insofar as local functional limits are concerned.

This does not deny that quantitatively superior alleles could have a greater selective fitness than their quantitative inferiors. It rather affirms that both alleles' proteins' have the ability to successfully contribute to their system's particular functioning. Past this, one allele's protein could still be consistently preferred over the other due to competition, that is, because the other allele was being selected against. The preferred allele is then expected to become a prevalent allele for that gene. If a superior allele prevailed at an ancestral time-point and the inferior became extinct, this should be manifested in the subsequent species' progeny lines. Thus, this protein should display a candidate set of

ancestral alleles reflecting the superior allele's dominance within and between extant populations.

It is possible that some quantitatively superior ancestral alleles could have existed that offered only a slight selective advantage. These alleles, however, coexisted with a quantitatively inferior rival, yet never became prevalent in the ancestral gene pool. In such cases, it appears difficult to ascertain which allele was indeed quantitatively superior. The alleged quantitatively inferior allele in this case must not have had too poor a fitness level, otherwise it would have been out-competed by the superior allele and driven to extinction (if it was a dominant allele).

Ancestral local functional limits can also be utilized to address other issues. One issue is whether one species belonged to an ancient, more encompassing species or archetype at its time of origin. This analysis asks whether two species existed as a common species when their contemporaneous ancestral genetic backgrounds existed. In practice, an investigator could compare each protein's ancestral local functional limits from one species with ancestral sequences of the other and vice versa. Many of one species' sequences existing within the local functional limits of the other's is strong evidence for claiming a common species ancestry.

Arguing for a common ancestral species can be corroborated with multiple lines of evidence. These include both common allelic origins for each respective species' alleles and the genomic equivalence criteria considered above. One needs to make a cumulative case here for a functional phylogenetic inference of two species' common ancestry.

In this comparison of ancestral alleles and functional limits, one must keep in mind that both species' essential proteins must have been compatible if interbreeding was possible. A general theme in Chapter IV was that it is not the similarities, but the differences that matter in species' survival. If the ancestral alleles of all proteins except a handful of essential ones were compatible between two species, those species would not have been able to interbreed successfully. Their otherwise compatible genetic backgrounds would not be capable of overcoming this deficiency.

In another situation, two species that descended from a common ancestor prior to the emergence of the ancestral alleles are being studied. In this case, the alleles intermediate between them have all become extinct. This displays a gap between both sets of ancestral alleles larger than a single nucleotide separation. An expanded version of the earlier method used for determining Candidate Alleles 1.5a and 1.5b could be employed here. This method attempts to reconstruct these alleles' history by the most likely series of mutation events. Each reconstructed intermediate must conform to the protein's local functional limits at that ancestral time. A gap-reconstruction of peripheral proteins is needed to determine the first protein's local functional limits. A successful stepwise reconstruction shows that a proposed design point is not evident between these species for that specific protein.

CONCLUSION

As we have discussed in this chapter, the main benefit of functional molecular phylogeny comes in its cumulative power to distinguish proposed design points. This

applies specifically to the particular genes considered, and ultimately to the species carrying them. This methodology could uncover proposed design points in genera that are predicted by the two Intelligent Design scenarios. This mainly consists of discerning between gradualistic natural selection within genera and species, and proposed design points between genera, which the previous sections of this chapter outlined. It is patterns of such proposed design points between genera that are predictions of the two Intelligent Design scenarios described here. It is this potential falsifiability that elevates the status of these Intelligent Design scenarios as theories in the Popperian sense.

We saw several angles in this chapter by which functional molecular phylogeny can approach the issue of proposed allelic design points. We discussed how rough guidelines for functional molecular phylogeny studies can be drawn by examining allelic distributions throughout different taxonomical levels. Alleles arising prior to the divergence of two populations are normally expected to continue in those populations and possibly yield further alleles. The reality of this becomes increasingly evident as more populations within a species are compared further. This reveals the true history of the gene in that species (insofar as it can be known). This discerning of allelic history patterns also applies to different species within a genus. Recall that a genus within naturalistic evolutionary theory and a genus derived at a proposed design point (the two Intelligent Design scenarios) are not necessarily the same thing. Common descent by gradualistic natural selection still holds between the members within both genus definitions though. Within naturalistic evolutionary theory, however, common descent extends past the genus level and ultimately incorporates the entire tree of life. This historical relationship via

gradualistic natural selection does not extend past the genus level in the Intelligent Design paradigms (whatever contemporary taxonomical level they may extend to).[2] Rather, each genus in these paradigms encapsulates its own tree of life distinct from those of other genera.

We then examined how functional molecular phylogeny could address proposed allelic design points by applying local functional limits to the genus's ancestral alleles. These ancestral alleles could then be compared with the ancestral gene pools of potentially related genera. This could aid in determining whether or not such an interrelationship was reproductively possible. This conclusion applies to genera whose alleles were contemporaneous at their times of origin. The results of this contributes to addressing the prediction of proposed design points of genera by displaying potential allelic design points at each of these junctures, insofar as functional limits are concerned. The Intelligent Design paradigms both display points of design at the origin of their respective genera, exhibited, in part, by their functional limits.

There was then a final way by which this methodology could address the issue of proposed points of allelic design. Functional molecular phylogeny could attempt to reconstruct a gradualistic common descent between two species' analogous alleles that display an alleged gap of intermediate extinctions. It seems, though, that this application could be left open to too much interpretation. This might leave it susceptible to *ad hoc* modifications to either prove or disprove a particular reconstruction. This is a potential issue both for Darwinian evolutionists and design theorists seeking to redefine a specific genus from within these paradigms. A gap-reconstruction for these alleles is at least

possible as far as the global functional limits and potential local functional limits allow. However, this could only show that such a reconstruction was possible, not that it actually was the case. This problem also applies to the historical relationships of analogous sequences between higher taxa. Such a record should appear increasingly coarse-grained the higher up the investigation went on the taxonomical scale. This is due to group isolation over vast periods of time and to extinctions of many intermediate groups and alleles. This underscores the need to search for the predicted patterns of proposed allelic design points instead of singular instances.

An inherent confusion could then appear in many of these hard cases. To avoid this, investigators could first attempt to establish the underlying existence of proposed allelic design points by examining finer-grained groups that don't have these extinction gaps. A proposed design point for two analogous essential genes could, by itself, be evidence of a point of design between the two species carrying these genes respectively. However, if there was a design point at the genus-level, it is probably manifest in most genes within all participating species yielding a pattern of proposed allelic design points.

Establishing design points between different genera addresses whether or not design points, and thus the present Intelligent Design scenarios that predict them, could be manifested in biological life. It appears then that establishing the mere existence of patterns of allelic design points (leading to conclusions of design points for genera) in the clearest, most accessible biological areas is the first course of action. Addressing potential design points between species displaying an alleged gap of extinct intermediates and other hard cases then become easier to approach. In any case, the mere ability to falsify these

Intelligent Design scenarios by demonstrating the absence of such patterns (which they predict should be present), according to Popper, raises their statuses as scientific theories. Whether these theories are actually confirmed or denied is then left up to the practical investigator.

NOTES

1. There is an exception worth noting at this point. The creation of this new daughter allele replaces that particular copy of the parental allele that is normally passed on. Instead, this daughter allele and its allelic descendents are passed on in this line of the parent's lineage. Thus, the multitude of parental copies that could have been yielded in this line has been withheld from the gene pool.

2. Intervening-maintenance would produce genera that are at least partly related to other genera. They would not be entirely related like the genera produced within a naturalistic evolutionary scenario, perpetuated solely by unguided processes.

WORKS CITED

Andersson, Dan. "The Biological Cost of Antibiotic Resistance." Curr Opin Microbiol 2 (1999): 489-493.

Behe, Michael. Darwin's Black Box. New York: Free Press, 1996.

Dembski, William. "Evolution's Logic of Credulity: An Unfettered Response to Allen Orr" 2002. Online posting. Design Inference Website, 2002 Postings Section. http://www.designinference.com/. October 2003.

----------. No Free Lunch. Lanham: Rowman & Littlefield, 2002.

DeRosier, David. "Spinning Tails." Curr Opin Struct Biol. 5.2 (1995): 187-93.

Lewis, Edward. "A Gene Complex Controlling Segmentation in Drosophila." Nature 276 (1978): 565-570.

Pennock, Robert. "DNA by Design?." Debating Design. Ed. Michael Ruse. Cambridge: Cambridge UP, 2004. 130-148.

Popper, Karl. "Science: Conjectures and Refutations." Philosophy of Science. Ed. Martin Curd. New York: Norton & Company, 1998. 3-10.

----------. "The Problem of Induction." Philosophy of Science. Ed. Martin Curd. New York: Norton & Company, 1998. 426-432.

Ruse, Michael. "Creation-Science Is Not Science." Philosophy of Science. Ed. Martin Curd. New York: Norton & Company, 1998. 38-47.

Wallace, Bruce. "Coadaptation Revisited." Journal of Heredity 82.2 (1991): 89-95.

VITA

Name:	James William Schroeder
Address:	14819 River Vista N., San Antonio, TX. 78248
Email Address:	james_schroeder@hotmail.com
Education:	B.S., Genetics, Texas A&M University, 1998 M.A., Philosophy, Texas A&M University, 2005