Towards a Creationary Classification of Mutations

Jonathan L. Bartlett, 4208 W San Antonio, Broken Arrow, OK 74012

Abstract
Mutations are normally classified according to their proximal effect on an organism’s fitness, whether beneficial, deleterious, or neutral. While this is a very useful first-pass categorization of mutations, the realization that mutations are not always haphazard, but in fact may be part of a regulated design, means that creationists should be looking for a deeper classification of mutations based on whether or not they conform to their organism’s design. Design-consistent mutations are those which occur within the pattern expected by the genome’s architecture, and design-inconsistent mutations are those which occur outside of the genome’s architecture. Features such as metabolic consistency, mutational mechanism, mutation rate, reversibility, and preservation of genome semantics can be used by biologists to assess whether or not a mutation is design-consistent or design-inconsistent.

Keywords: mutation, design, molecular biology, genome, genome architecture, simple sequence repeats, recombination, transposable elements, semantics, information theory

Introduction
Modern biochemistry has shown that the cell is a much more fascinating piece of machinery than ever would have been expected a century earlier. It has moved in our understanding from being a blob of protoplasm to an intricate wonder of nanotechnology. Likewise, our understanding of the genome and its intricacies has increased by leaps and bounds over the last few decades. While it was previously thought that only protein-coding genes would be relevant, it is now known that the regulation of protein-coding genes is just as important, if not more so. While it was previously thought to be intellectually irresponsible to consider a biological function for transposable elements, we are now understanding their ubiquity and importance in shaping the genome (Sternberg 2002).

Our understanding of the process of mutation is undergoing a similar revolution. Historically, creationists and evolutionists have been in agreement that mutations which occur in organisms are haphazard—that is, there is no designed purpose for them. However, this understanding is beginning to change. It is becoming increasingly apparent that the cell itself can induce mutational processes in the right genes to produce beneficial changes at appropriate times. The most well-studied of these systems is the somatic hypermutation (SMH) system in the vertebrate immune system. In order to increase the binding potential of immunoglobulins to antigens, the vertebrate immune system introduces mutations to a specific region of a specific gene which attaches to the B-cell entirely, and focuses only on the region of the gene which binds to the antigen. It is not deterministic—that is, the specific changes which are made seem to be stochastic—but the changes are focused to the right gene in the right situation, bypassing well over 99.99% of the genome and focusing on the correct few hundred base pairs which would matter.

More and more examples of focused mutations have been explored. Some transposable elements are triggered in direct response to specific cell stressors. For instance, Hall (1999) showed that E. coli can use insertion sequences to activate the gene required to metabolize beta-glucoside sugars.

Simple sequence repeats (SSRs) have shown to mutate primarily in copy-number, acting as a genomic tuning knob or state switch. King, Trifonov, and Kashi (2006) enumerate several, including an AC repeat in the lic2A gene of H. influenzae to switch between three states—off, low expression, and high expression—as it alters the reading frame on which the ATG start codon is found.

Semi-palindromic DNA often points to potentially beneficial areas where mutations might take place. For instance, one study of the genetic adaptation of E. coli to low glucose concentrations found identical modifications of the mgl operator sequence (mglO) in multiple populations (Notley-McRobb and Ferenci 1999). These were later found to be in loops that were near stem-loop structures in DNA, formed by semi-palindromic sequences (Wright 2004).

Caporale has synthesized this new research into what she terms an “implicit genome”. That is, a
The Need for a Second-Order Classification of Mutations

In the current literature of both creationists and evolutionists, mutations are often classified according to their effect on an organism's survival within a specific environment (Baumgardner et al. 2008; Eyre-Walker and Keightley 2007). A mutation is considered "beneficial" if it helps the organism survive, "deleterious" if it hinders the organism, and "neutral" or "nearly-neutral" if there is little observed effect. This convention of using beneficial/neutral/deleterious for categorization is useful because it can often be measured directly and quantitatively.

Unfortunately, many people take these categorizations to signify more than can be inferred from them. For example, those who view all mutations as haphazard often view beneficial mutations as evidence of evolution. Likewise, those who view mutations as being possibly internally generated by a guided process within an organism may view the fact that a mutation is beneficial as being prima facie evidence that a mutation was haphazard.

Upon closer examination, however, knowing whether a mutation is "beneficial", "deleterious", or "neutral" does not by itself tell us whether or not a given mutation occurred according to design or not.

For instance, in order to survive large-scale environmental changes, populations of organisms may keep a supply of organisms with alternative biochemical configurations through a mutational process (Bartlett 2008). Compared to the overall population of organisms, these mutations would actually be phenotypically deleterious, even though they are part of the overall biological design for hedging against possible environmental changes.

Likewise, a phenotypically beneficial mutation is not necessarily part of an overall design. Behe has termed this sort of event "trench warfare"—the mutation may give a phenotypic advantage within a competitive environment, but at the cost of debilitating some important function of the organism. An easy example of this would be sickle-cell anemia—while it may be beneficial in some circumstances because it prevents malaria, the way it debilitates the person who has it overall leads to the conclusion that this was not a designed feature (Behe 2007). Many other such locally beneficial mutations which have an overall deleterious effect on the complexity of a cell's biochemistry have been documented (Anderson 2005).

Therefore, while the scale of beneficial/neutral/deleterious works well for a first-order classification of mutations, creation biologists should be looking deeper into a second-order classification based on its consistency with the design of the organism. This classification separates mutations into one of two possibilities—"design-consistent" mutations, and "design-inconsistent" mutations. A "design-consistent" mutation is one which appears to have occurred within the genome's implicit range, and a "design-inconsistent" mutation is one which appears to be haphazard (that is, philosophically random as described by Bartlett (2008)).

Guidelines for Determining Second-Order Classification

The problem with a second-order classification system is that since we do not have total knowledge of the original plan, it makes it difficult to determine whether a mutation was consistent with that design or not. However, creationists can apply theological concepts to achieve a basic understanding of the plan, which can then illuminate our investigations, even in absence of full prior knowledge of the full plan. The notion of a Genesis "created kind" (called baramins in creation biology) is a key theological notion which will aid our investigation. Because God created the animals according to their kinds (Genesis 1:11; 1:21; 1:24; 1:25), it can be presumed that the biological plan is a sort of dynamic stasis, where basic patterns are preserved, yet variance is allowed to aid in both survival of the baramin and the fulfillment of the baramin's role. The extent of the intended dynamic and static elements of the baramin are not known a priori. The criteria presented here should help creation biologists take the basic theological concepts provided by scripture, and combine these with the data of biology to achieve a fuller understanding of the pattern of life that God created.

The following are several parameters which can be considered which will help make the second-order classification of mutations based on the assumed dynamic stasis of the original kinds.
Metabolic consistency

One of the major ways of determining whether or not a mutation is design-consistent is to look at the internal operation of the organism, examining the consistency of function of organisms with and without the mutation. Design-consistent mutations should maintain internal consistency whether or not they are beneficial within the current environment. It is hard to separate problems which occur from internal metabolic problems as opposed to those caused from the outside. However, a decent test of metabolic consistency in single-celled organisms would be whether or not a mutation caused metabolic problems for an organism which was growing in a nutrient-rich environment free of competition, predation, and toxic compounds. Similar stress-free environments could be constructed for testing the metabolic consistency of mutations in multicellular organisms.

Sickle-cell anemia, for instance, while it is beneficial as far as preventing malaria, causes large-scale functional problems for the organism even in the best environments. Therefore, the mutation causing sickle-cell anemia is metabolically inconsistent, even though it may provide benefit in certain circumstances.

Anderson (2005) pointed out that most drug-resistant mutations of bacteria involve a fitness cost in most normal environments. However, for some of the mutations cited, the fitness cost was not severe, and therefore, by this criteria, would be considered metabolically consistent. In Mycobacterium tuberculosis, for example, some mutations which confer rifampin resistance also do well in normal cultures. In the three mutations isolated by Billington, McHugh, and Gillespie (1999), one mutation type had no relative fitness decreases, and another one had only moderate relative fitness decreases, thus indicating that they are metabolically consistent. One of the mutations had a drastic reduction in relative fitness, indicating that this mutation was probably not metabolically consistent.

Mutational mechanism

A mutation which is in response to a specific stress or group of stresses, or is timed to occur with a particular stage of life for an organism, or for which there is an enzyme whose core function is to produce such a mutation, is likely to be design-consistent. Williams uses mutational mechanism as the primary differentiator between design-consistent mutations, which he terms as “recombination”, and design-inconsistent mutations, which he terms as simply “mutations” (Williams 2005). Other creationists have followed this approach as well (Ashcraft 2003). This criteria by itself is not sufficient, especially considering that the physical cause of many mutations is unknown. In addition, some design-consistent mutational events may be the result of DNA sequence alone, and may occur without specialized enzyme assistance. In addition, some mutational mechanisms may become mistargeted due to either a problem in the mutational mechanism or the sites which are targeted.

An obvious example of a mutational mechanism would be the V(D)J recombination system used in the production of immunoglobulins. This system recombines three types of gene fragments (variable, diversity, and joining) into a nearly-limitless array of immunoglobulins. This system uses a combination of a pair of enzymes (RAG1 and RAG2) and a recombination signal sequence (RSS), to cut and splice gene fragments at the appropriate locations (Market and Papavasilious 2003). However, this targeting mechanism can also lead to cancers when cryptic RSS sequences (sequences similar to the RSS used by V(D)J reactions, but in other parts of the genome) become accessible to the recombination enzymes (Schlissel, Kaffer, and Curry 2006).

As mentioned earlier in the paper, Hall (1999) showed that E. coli can use insertion sequences to activate the gene required to metabolize beta-glucoside sugars. Because this mutation only occurs under the conditions where the mutation is needed, it can be inferred that this is the result of a cellular mechanism.

Mutation rate

A mutation which occurs at a significantly higher rate than the average mutation rate for the organism is likely to be design-consistent. This points to the mutation being part of the phenotype of the organism, rather than the mutation occurring arbitrarily. Even though we are living in a post-fall world, we assume that most of our biological systems function properly according to their design on a daily basis. Thus, a high mutation rate, especially across an entire population, is suggestive that a mutation is design-consistent.

As with other considerations, this one must be used with care. Problematic mutations can cause a site to become mutationally active when it should not be. One test suggested by Bartlett (2008) in differentiating between a design-consistent and a design-inconsistent hotspot would be to compare the average fitness effects of mutations at that hotspot with the average fitness effects of induced arbitrary mutations. If the mutations in the hotspot tend to be metabolically consistent, and their relative fitness is greater than the relative fitness of organisms with arbitrary mutations, then this is evidence of a mutational hotspot being design-consistent.

In the human genome, this appears to hold as a general pattern. Chuang and Li (2004) have noted that mutational hotspots tend to occur in genomic regions involved in extracellular communication,
while mutational coldspots tend to occur in cellular housekeeping functions. Thus, the mutations seem to be targeted at areas which would provide adaptation to new environments, and silenced at areas which would more likely cause metabolic inconsistencies.

**Reversibility**

If a mutation is easily reversible (that is, the frequency of reversion is significantly greater than what would be expected from the overall mutation rate), this is good evidence that the mutation is design-consistent. If one of the purposes of mutations is to provide a hedge against environmental changes, then it would be reasonable to think that if the hedge is successful, the organism needs to be able to make a future hedge of going back to the original configuration. Therefore, reversibility is a key indicator of design-consistent mutations.

SSRs are quite interesting subjects because, in general, they are both highly mutable as well as being easily reversible. Historically, they have been viewed as evolutionary byproducts, or junk DNA. However, current research is continually finding new ways in which SSRs allow for the genome to adapt to changing circumstances (Kashi and King 2006).

**Preservation of genomic semantics**

Every day we are learning more and more about the semantics of the genome. The genome’s semantics can be considered its internal logic—how different sequences serve to regulate and format the genome’s function and adaptation (Sternberg 2002). In computer programming, a program’s semantics is the set of rules, conventions, and axioms which are assumed to hold true as the program progresses.

Many mutations that lead to sickness are those which cause certain regions of the genome to operate in a semantically different way than before. For instance, some mutations have transformed a non-SSR sequence into an SSR. This causes the cell to modify the SSR’s copy number where it would not have before (King, Trifonov, and Kashi 2006). One example of this is in a heritable form of colorectal cancer. In this cancer, a T to A mutation creates a novel SSR. In later generations, the expansion of this SSR leads to colorectal cancer (Laken et al. 1997). Other types of signals may also be created improperly due to mutation, which could cause other systems to act on the wrong site (see for instance the previous discussion about V(D)J recombination).

Therefore, in many cases, the altering of the semantics of the genome often points to a design-inconsistent mutation. However, this criteria should be used with care. DNA might contain a large quantity of meta-information. Meta-information is “information about information”. If, for instance, the cell has sufficient meta-information about the roles of different DNA sequences (as opposed to simply the sequences themselves), there is no reason to think that it could not add or remove semantic elements as needed, using the meta-information as a guide. Dipterans, for example, have a class of transposable elements known as mini-me elements (microsatellite initiating mobile elements), which are retrotransposons that contain primers for SSRs (Wilder and Hollocher 2001). The functions of these are not well-characterized, but their abundance within Diptera (comprising, for instance, 1.2% of the *Drosophila melanogaster* genome), indicates that they are probably a part of the organism’s design.

Since scientific knowledge of genome semantics is still in its infancy, other factors such as metabolic consistency and precision should be considered in determining whether or not a violation of genomic semantics has taken place, or if the mutation is simply a part of a higher-level or undiscovered semantic within the cell.

**Other Considerations**

These heuristic guidelines are certainly not complete, nor do any of them stand on their own. Nonetheless, when used in combination, they can open up a new way for creationists to look at mutational processes within cells. Hopefully as this categorization is applied, additional considerations will be supplied to this list.

Also note that the weighting of the various criteria are dependent on our understanding of God’s general plan for organisms. If the goal is a dynamic stasis, then reversibility should be a heavily weighted factor. If the goal is for life to proceed according to a specific direction (for instance, see Gene 2009), then mutation rate should be weighted higher.

**Conclusions**

While the first-level classifications of mutations as beneficial, neutral, or deleterious is certainly useful, the perspective of creation biology can add additional depth by looking at whether or not a mutation is consistent with the cell’s overall architecture, or if the mutation goes against that architecture. Only distinguishing between mutations on the basis of immediate fitness gains or losses can blind biologists to the architecture in which those mutations are made. By differentiating between design-consistent and design-inconsistent mutations, creation biologists can better understand the overall architecture of each baramin, and ultimately understand both God’s

---

1 Gene (2009) argues from a vaguely-theistic evolutionary perspective. However, the mechanism he proposes is quite intriguing. It is not specific to theistic evolution and may prove useful to creation biology.
purposes for these organisms as well as how humans can best cooperate with their designs in ecological and biotechnical endeavors.

Appendix—Experimental Methodology
A basic outline of how an experimenter could test each of these criteria is given below. Because experimental mutational studies are generally easier on single-celled organisms which can be grown in a laboratory setting, these methodologies are tailored to that environment. It is the hope of this author that further work can expand upon these methodologies and demonstrate the most effective ways of experimentally determining whether mutations are design-consistent or design-inconsistent.

- **Metabolic Consistency**—to determine metabolic consistency, plate both the wild-type and the mutant-type on several different media, and examine the relative fitness of the mutant-type on each. Compare the best fitness scores of both the wild-type and mutant-type. If the relative fitness of the mutant-type on its best media is significantly lower than the fitness of the wild-type on its best media, then the mutation is probably metabolically inconsistent. Further experimental work comparing relative fitness values of mutations in the context of the other criteria is required for determining what the relative fitness value ranges should be for a mutation to be considered metabolically consistent.

- **Mutational Mechanism**—this criteria is easy to rule in (by knowing a mutational mechanism capable of producing the mutation) but hard to rule out. If no known mutational mechanism causes the mutation, but a specific environmental inducer causes a specific or semi-specific mutational response, then it is reasonable to infer a mutational mechanism even without knowing what it is.

- **Mutation Rate**—the rate of mutation should be compared to other sites within this organism’s genome. If the production of this mutation is significantly higher than the average site for the organism, then a high mutation rate is established.

- **Reversibility**—the reversion rate for a given mutation should be compared to the mutation rate for the mutant’s original production. It is best if the specific gene sequences can be obtained for the wild-type, the mutant, and the reversion, in order to determine if the reversion is a true genetic reversion, or if a different mutation supplied the missing functionality. If the original mutation was a spontaneous mutation by the fluctuation test (Luria and Delbrück 1943) or the Lederberg test (Lederberg and Lederberg 1952), then it is reasonable to think that the reversions should be spontaneous, too. If the original mutation is adaptive, however, then the reversion mutation might require a specific environmental signal to occur, and it might take some effort for the experimenter to determine what this signal is.

- **Preservation of Genomic Semantics**—the interplay of the mutant sequence and the surrounding genetic context should be examined. Known genetic sequence motifs should be identified, as well as the mutation’s impact on those motifs. If the mutation corrupts an existing motif, then it is likely that the mutation is violating genomic semantics. However, if a set of mutations are consistently altering motifs in narrow ways, then it is possible that the mutations are following an as-yet-unknown semantic rule within the cell.

References


