Multiple Overlapping Genetic Codes Profoundly Reduce the Probability of Beneficial Mutation

George Montañez[1], Robert J. Marks II[2], Jorge Fernandez[3] and John C. Sanford[4§]

[1]*Computer Science Department, One Bear Place #97356, Baylor University, Waco, TX 76798–7356.* [2]*Electrical & Computer Engineering, One Bear Place #97356, Baylor University, Waco, TX 76798–7356.* [3]*FMS Foundation, 7160 Stone Hill Rd. Livonia, NY 14487.* [4]*Department of Hort Sci, Hedrick Hall, NYSAES, Cornell University, Geneva, NY 14456, USA.* ([§]*Corresponding author: jcs21@cornell.edu*)

Abstract

There is growing evidence that much of the DNA in higher genomes is poly-functional, with the same nucleotide contributing to more than one type of code. Such poly-functional DNA should logically be multiply-constrained in terms of the probability of sequence improvement via random mutation. We describe a model of this relationship, which relates the degree of poly-functionality and the degree of constraint on mutational improvement. We show that: a) the probability of beneficial mutation is inversely related to the degree that a sequence is already optimized for a given code; b) the probability of beneficial mutation drastically diminishes as the number of overlapping codes increases. The growing evidence for a high degree of optimization in biological systems, and the growing evidence for multiple levels of poly-functionality within DNA, both suggest that mutations that are unambiguously beneficial must be especially rare. The theoretical scarcity of beneficial mutations is compounded by the fact that most of the beneficial mutations that do arise should confer extremely small increments of improvement in terms of total biological function. This makes such mutations invisible to natural selection. Beneficial mutations that are below a population's *selection threshold* are effectively neutral in terms of selection, and so should be entirely unproductive from an evolutionary perspective. We conclude that beneficial mutations that are unambiguous (not deleterious at any level), and useful (subject to natural selection), should be extremely rare.

Key words: beneficial mutation, probability, multiple codes, overlapping codes, ENCODE, poly-functional DNA, selection threshold

1. Introduction

It is almost universally acknowledged that beneficial mutations are rare compared to deleterious mutations [1–10]. However, there is controversy regarding just how rare beneficial mutations actually are. It appears that beneficial mutations may be too rare to actually allow the accurate measurement of how rare they are [11]. For

decades it has been widely thought that beneficial mutations might be as rare as one in a million [12, 13]. However, more recently some have argued that beneficial mutations might be much more common [14, 15].

The actual rate of beneficial mutation is a crucial question, because if determines both the speed and the direction of genetic change. If beneficial mutations are extremely rare, this profoundly limits the *rate and range* of all forward genetic change. Furthermore, to the extent that beneficial mutations may be extremely rare, the question arises — "how can there be any net gain in total biological fitness?" This question arises because it is widely recognized that in large genomes most mutations should have very small effects, and so large numbers of low-impact deleterious mutations should not be subject to purifying selection [16-33]. This means that over time large numbers of such deleterious mutations should accumulate continuously, leading to ever-increasing genetic load [29-33]. In order to halt such genetic deterioration, one must invoke the continuous amplification of a large number of beneficial mutations to fully compensate for all the accumulating deleterious mutations [34–36].

Fisher addressed the problem of the rarity of beneficial mutations as long ago as 1930 [37]. He argued that beneficial mutations might be quite common. He used the illustration of focusing a microscope. A random change in focal length has a nearly equal chance of either improving or diminishing the focus, assuming three things: a) the microscope is significantly out of focus, b) the change in focus is very small, and c) focus is just a one dimensional trait (a single knob — turned either up or down). We now know that Fisher's three necessary conditions do not apply to the real biological world. Biological systems are highly optimized (the microscope is not substantially out of focus), a beneficial mutation must be subject to selection, so its biological effect must not be too small (so very tiny changes in focus are not feasible), and fitness is extremely multi-dimensional (there is much more to biological functionality than optimizing a single parameter such as focal length).

Fisher acknowledged that focusing a microscope just involves optimization in a single dimension, and conceded that to the extent that fitness is not a simple one-dimensional trait, his analogy would break down. He went on to show that as the number of "dimensions" of fitness increased, the probability of beneficial mutation should rapidly decrease. This insight was profound, yet in his day he could not have realized how extremely multi-dimensional biological fitness really is. Fisher lived before the revolution in biology — he knew nothing of cell biology, molecular biology, or molecular genetics. We now know that total biological fitness is multi-dimensional in the extreme. In a sense, every functional nucleotide within a genome adds another dimension to the fitness equation. So in a sense Fisher's allegorical "microscope" has millions of knobs that must be focused simultaneously and interactively.

In the last decade, we have discovered still another aspect of the multi-dimensional genome. We now know that DNA sequences are typically "poly-functional" [38]. Trifanov previously had described at least 12 genetic codes that any given nucleotide can contribute to [39,40], and showed that a given base-pair can contribute to multiple overlapping codes simultaneously. The first evidence of overlapping protein-coding sequences in viruses caused quite a stir, but since then it has become recognized as typical. According to Kapronov et al., "it is not unusual that a single base-pair can be part of an intricate network of multiple isoforms of overlapping sense and antisense transcripts, the majority of which are unannotated" [41]. The ENCODE project [42] has confirmed that this phenomenon is ubiquitous in higher genomes, wherein a given DNA sequence routinely encodes multiple overlapping messages, meaning that a single nucleotide can contribute to two or more genetic codes. Most recently, Itzkovitz et al. analyzed protein coding regions of 700 species, and showed that virtually all forms of life have extensive overlapping information in their genomes [43]. So not only are there many "knobs" in Fisher's microscope analogy, each one can affect multiple traits simultaneously and interactively.

In light of these new developments, it is timely to reexamine the question of the probability of beneficial mutation, the utility of Fisher's model, Fisher's Theorem, and Fisher's insight about multiple fitness dimensions. This paper examines the probability of a selectable beneficial mutation arising within a DNA sequence that is functional (hence must be significantly optimized), and contains multiple overlapping codes.

2. Method and Results

2.1 *The Model*

For illustration, in Figure 1 we show a hypothetical 100 base pair sequence, which participates in 12 partially overlapping codes.

Starting Assumptions:

1. We only consider here a "functional sequence". We assume this sequence is not primarily "junk DNA", but that for the most part it encodes information, yet we allow for rare nucleotide sites within the functional sequence that are perfectly neutral.
2. Each nucleotide within the functional genome is classified by level (L_1–L_{12}), depending on how many codes it contributes to. A nucleotide that does not contribute to a given code is considered neutral relative to that code. A nucleotide which does not contribute to any of the codes is considered perfectly neutral and will be designated L_0.

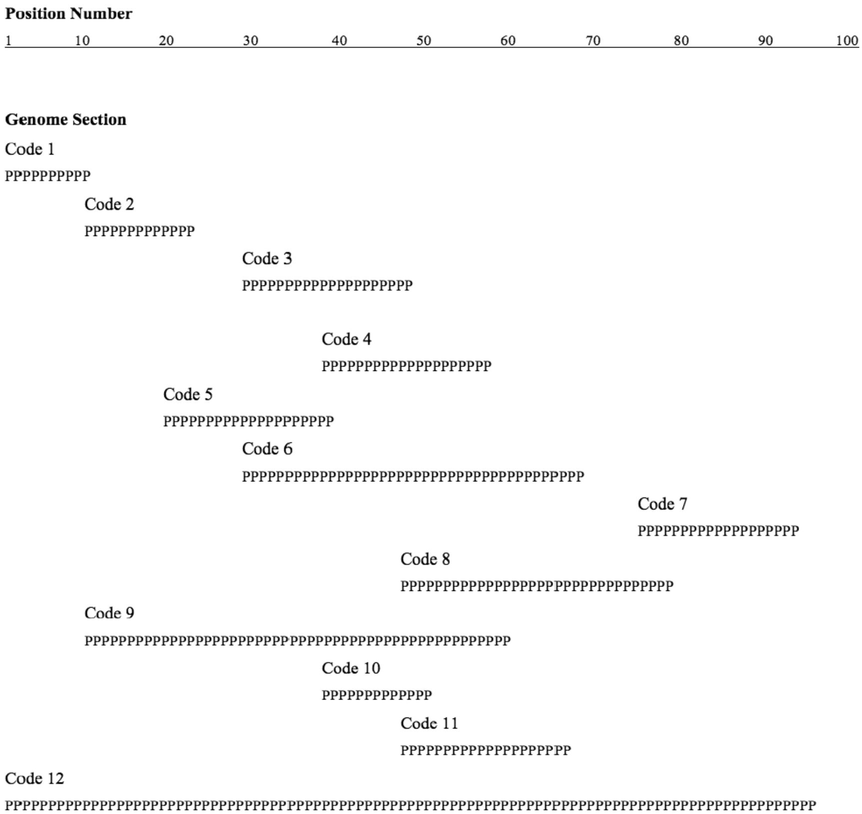

Fig. 1. A model nucleotide sequence of 100 bases that encodes 12 partially overlapping codes. Each sub-section represents the positions of the Genome Section that participate in that particular code. For example, only the first 10 positions of the Genome Section participate in Code 1 whereas all except the last 5 positions of the Genome Section participate in Code 12. Nucleotide positions of the Genome Section that do not fall into any code are considered entirely neutral with respect to those codes, since they play no part in what the function of those codes may be. In that regard, these neutral positions are not part of the functional genome (at least with respect to those specific codes).

3. Consistent with commonly used evolution models [41, 44–46], we assume the optimization of a composite organism is determined by a single fitness function. The contribution of each code to fitness is assumed to arise by aggregation of constraint commonly found in multi-objective optimization [47-50].
4. We assume a high degree of optimization within each code, although this assumption can be relaxed, and is a tunable parameter within the model. For the analysis and discussion we assume 99.9% of the nucleotide positions defining a code are already an optimal nucleotide.

5. For those nucleotides that are part of a given code but are not yet the optimal nucleotide for that code, we assume that only one of the three alternative nucleotides will be an improvement relative to the existing sub-optimal nucleotide. Mutations at such sites will therefore have one-third chance of being beneficial, but will still have a two-thirds probability of being deleterious. Another way to say this is that for a given site, relative to a given code, there is a hierarchy of most desirable nucleotide (ranked first) to least desirable nucleotide (ranked fourth), and as a rule the non-optimal nucleotide is ranked second, rather than third or fourth. This reflects the idea that even if non-optimal, the existing nucleotide is not truly random.
6. We assume independence in the designation of different optimal codes, such that the nucleotide deemed optimal for a given code (at a position) is chosen independently of the other nucleotides that are optimal for codes that may overlap at that position. In other words, for position 1, the first code may view G as the optimal nucleotide, whereas the second code may consider T the optimal nucleotide, or both may consider C optimal, etc. Although the nucleotide at a position may be shared by several codes (in the case of overlap), we assume that a nucleotide for an optimal code sequence is chosen only with respect to other nucleotides within that same code, and not with respect to other codes which may or may not overlap with it on the genome section currently or in the future. Modeling these optimal code sequence decisions as independent gives rise to the Bernoulli model presented here.
7. Lastly, we make the simplifying assumption that beneficial and deleterious mutations have "unit magnitude" effects, such that if one of each is present, their combined effects effectively cancel out (See Discussion).

2.2 *Analyses*

We analyzed how overlapping codes affect the probability of beneficial mutation in three ways. The first analysis involved a very simple calculation of how multiple overlapping codes affect the theoretical probability of an "unambiguously beneficial mutation". We define an unambiguously beneficial mutation as a mutation that causes a benefit in at least one code, without causing any deleterious effect in any other code. The second analysis is more involved, and examines the probability of a "net-effect beneficial mutation". A net-effect beneficial mutation is a mutation that improves more codes than it disrupts. The last analysis involves an empirical analysis of how overlapping English words (i.e. as in a crossword puzzle), affect the probability of creating a new valid word.

2.2.1 *First Level of Analysis:*

When we consider poly-functional nucleotide sites, it is relatively simple to calculate the probability of mutations which are "unambiguously beneficial" (i.e., beneficial in one code, and not deleterious in any other code). For example, let us assume all codes are 99.9% optimized, (such that 99.9% of all mutations will be deleterious for any given code). Even for that one-in-a-thousand site which is sub-optimal, on average only about half of the nucleotide substitutions at such a site will be an improvement (which in this simple analysis we can ignore). For L_1 nucleotides, the rate of unambiguously beneficial mutations will be at best one in 10^3, for L_2 nucleotides this rate will be at best one in 10^6, and for L_3 nucleotides this rate will be at best one in 10^9. Generalized, for a L_n nucleotide, the rate will be at best one in 10^{3n}. Overlapping codes, by their very nature, make unambiguous mutations vanishingly rare. This means that within all poly-functional nucleotide sites, essentially all "beneficial mutations" will at best be *ambiguously* beneficial, being beneficial at just one level, but simultaneously being deleterious at one or more levels. Therefore at any poly-functional nucleotide site, a "beneficial" mutation will almost always still consistently have deleterious effects, systematically eroding the total amount of information in the entire information system.

2.2.2 *Second Level of Analysis:*

We can calculate the probability of a net-effect beneficial mutation for each nucleotide level (L_1–L_{12}) as described below.

Within a given code, assume that sequences are highly optimized. We use p(optimal) = 99.9% = 0.999 of all nucleotides being optimal in our recurring example. In the case of optimal nucleotide bases, any change is deleterious, assuming no neutral changes. Therefore, only $r = 1$ — p(optimal) = 0.1% = 0.001 are subject to beneficial mutation. There are no absolutely neutral positions in any given code, because by definition such a position is not part of that code. The conditions for net beneficial or net deleterious changes, therefore, are as follows:

To be a net-beneficial mutation:

- The current nucleotide in that position must be non-optimal AND
- The change must be to a beneficial nucleotide, which occurs with a 1/3 probability, denoted as

$$p(\text{beneficial} \mid \text{non-optimal})$$

To be a net-deleterious mutation:

- The current nucleotide in that position can be optimal OR
- The change must be to a deleterious nucleotide, which occurs with a 2/3 probability, denoted as

$$(1 - p(\text{beneficial} \mid \text{non-optimal}))$$

Given these assumptions, we calculate the probability that for a single code a mutation at a uniformly random chosen position is beneficial as follows, according to the law of total probability:

$$\begin{aligned} p(B) &= \\ &p(\text{non-optimal}) \times p(\text{benefical} \mid \text{non-optimal}) + p(\text{optima1}) \times p(\text{benefical} \mid \text{optimal}) \\ &= p(\text{non-optimal}) \times p(\text{benefical} \mid \text{non-optimal}) + 0 \\ &= (1 - p(\text{optimal})) \times p(\text{benefical} \mid \text{non-optimal}) \\ &= r \times p(\text{benefical} \mid \text{non-optimal}) \end{aligned}$$

Given the stated assumptions, for any single code, a mutation at a position chosen at random that mutates has a probability of being a *beneficial* (**B**) mutation equal to $p(B) = (1/3)r = 0.00033$. This, in turn, means that a random position that mutates has a probability of being a *deleterious* (**D**) mutation equal to $1 - p(B) = 0.99967$.

A mutation occurring to a single nucleotide may be beneficial or deleterious for any given code (as per previous discussion, neutral cases are excluded). Let's consider a few specific cases before generalizing:

(1) If the nucleotide is a L_1 nucleotide then there is only one possibility: a mutation will be either beneficial (**B**) or deleterious (**D**) with $p(B) = 0.00033$ and $1 - p(B) = 0.99967$.

(2) If the nucleotide is a L_2 nucleotide then there will be four possibilities: 1) a mutation may be beneficial for both codes (**B,B**); 2) a mutation may be beneficial to the first code and deleterious to the second code (**B,D**); 3) a mutation may be deleterious to the first code and beneficial to the second code (**D,B**) or, 4) a mutation may be deleterious to both codes (**D,D**). For such nucleotide positions, there is a value for each code, each of which is either beneficial or deleterious. We will make the simplifying assumption that where there is a beneficial effect in one code and a deleterious effect in another code, these effects will essentially cancel, leaving a neutral effect. Therefore (**B,B**) will be beneficial, (**D,D**) will be deleterious, while (**D,B**) and (**B,D**) will be neutral. In this case,

$p(\text{beneficial}) = p(\text{B,B}) = p(\text{B})^2 = 1.11 \times 10^{-7}$
$p(\text{neutral}) = p(\text{B,D}) + p(\text{D,B}) = 2 \times p(\text{B}) \times (1-p(\text{B})) = 6.66 \times 10^{-4}$
$p(\text{deleterious}) = p(\text{D,D}) = (1-p(\text{B}))^2 = 0.99933$

(3) In all other cases, where more than two codes are involved, there can be more than two factors to consider. For example, for L_3 positions, there are three levels of mutational effect.
(4). If the nucleotide is a L_N nucleotide, there will be 2^N possibilities. To generalize:

Let L_i be the level of a particular nucleotide. Combining all of the above, and formulating the binomial within our model parameters, if there are N codes and an L-level nucleotide, then the probability of a beneficial mutation for this L-level nucleotide, $p(B)_L$, is obtained with the binomial distribution [42]

$$p(B)_L = \sum_{k=\left\lceil \frac{L+1}{2} \right\rceil}^{L} \binom{L}{k} p(B)^k (1-p(B))^{L-k} \tag{1}$$

where L is the number of codes, $\left\lceil \frac{L+1}{2} \right\rceil$ is the minimum number of codes that constitute a majority (with the brackets denoting the ceiling function), and

$$p(B) = (1-p(\text{optimal})) \times p(\text{beneficial} \mid \text{non-optimal})$$

with p(optimal) denoting the probability that a nucleotide is already optimal.

In similar fashion, the probability of a deleterious mutation for this L-level nucleotide, $p(D)_L$, is obtained with:

$$p(D)_L = \sum_{k=\left\lceil \frac{L+1}{2} \right\rceil}^{L} \binom{L}{k} p(B)^{L-k} (1-p(B))^k \tag{2}$$

In general, the probability of a neutral mutation is

$$\begin{aligned} p(neutral)_L &= 1-(p(B)_L + p(D)_L) \\ &= \binom{L}{\frac{L}{2}} (p(B)(1-p(B))^{\frac{L}{2}} \times \delta(L \text{ is even}) \end{aligned} \tag{3}$$

where δ (L is even) is one when L is even and is zero otherwise. When L is even, $p(B)_L = 1-p(D)_L$. For $p(B) << 1$ (in other words, when $p(B)$ is near zero), this becomes approximately true for large odd L.

The value of $p(B)_L$ (the probability of a beneficial mutation) in equation (1) rapidly goes to zero for increasing L when $p(B) << 1$. Because differentiating

between of probabilities like 10^{-11} and 10^{-22} is intuitively challenging, we propose use of the information measure [51, 52, 53]

$$I\,[L_+] = -\log_2(p(B)_L)$$

$I[L_+]$ measures the probability in terms of flips of a fair coin. If $I[L_+] = 3$ bits, for example, the corresponding probability is the same as forecasting the result of three flips of a fair coin, i.e. $p = (½)^3 = 0.125$. $I[L_+] = N$ bits corresponds to a probability of $p = (½)^N$. To place this measure in perspective, there are 10^{15} square millimeters in an area of 1000 square kilometers. The probability of two people choosing the same square millimeter is thus 10^{-15}. Since $-\log_2(10^{-15}) = 50$ bits, the success probability is the same as the probability of predicting 50 sequential outcomes of the flipping of a fair coin.

A plot of $I[L_+]$ is shown in Figure 2 as a function of L for various values of r, where $r = 1–p$(optimal). The plots rapidly approach improbable values. For $r = 0.001$, a value of $L = 12$, $p(B)_L = 4.15 \times 10^{-22}$ or $I[L_+] = 71$ bits. The chance of choosing the same millimeter twice in a distance of 100 light years (10^{-21}) is more probable.

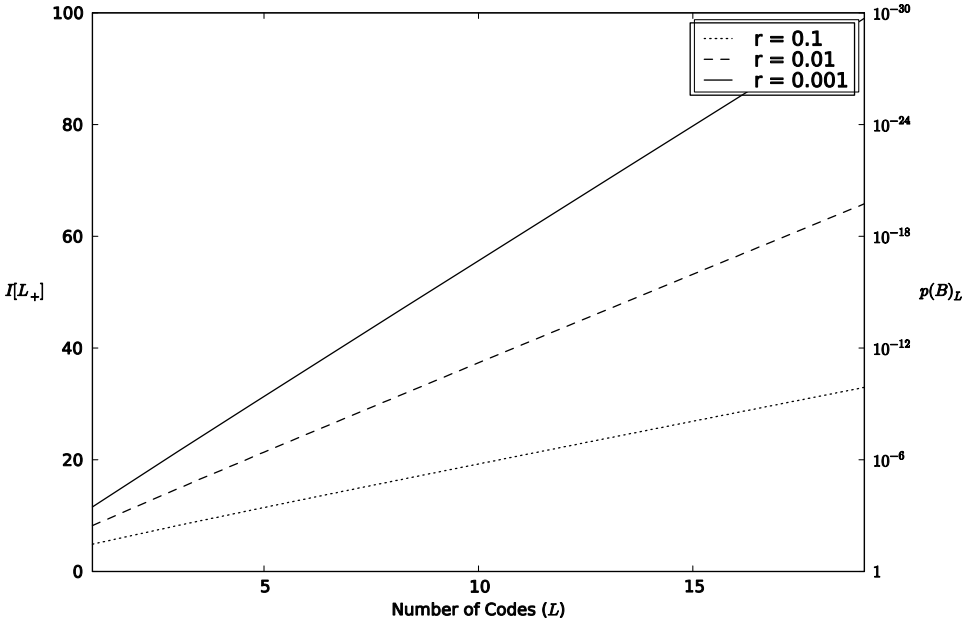

Fig. 2. Plot of $I[L_+]$ (information, in bits) versus L for various values of $r = 1$ — p(optimal). Even numbered codes are omitted for clarity. Since the probability of a beneficial mutation, $p(B)_L$, decreases exponentially with increasing L, the logarithmic information measure $I[L_+]$ increases linearly with increasing L. The right-hand scale indicates the probability of net beneficial mutation, using standard scientific notation. The three lines represent the cases where the overlapping codes are weakly optimized (10% of nucleotides are sub-optimal), moderately optimized (1% of nucleotides are sub-optimal), and highly optimized (0.1% of nucleotides are sub-optimal).

Our analysis suggests that increasing either the number of overlapping codes or the degree of optimization has negative effects on the probability of producing a beneficial mutation. A high degree of optimization makes beneficial mutations unlikely — even when considering just one code. As more codes are considered, the probability of beneficial mutation diminishes rapidly, as is shown in Figures 3, 4 and 5. The ratio of beneficial to deleterious mutations decreases so rapidly that for L_3 nucleotides in highly optimized sequences, the number of deleterious mutations expected before the first beneficial arose would be greater than the genome size of a typical bacterium. For L_5 nucleotides, the number of deleterious mutations expected before the first beneficial arose would be greater than the genome size of a typical mammal. While relaxing the optimization assumption reduces the severity of the problem (as can be seen in Figure 4), increasing the number of overlapping codes diminishes the likelihood of attaining a net beneficial mutation even for weakly optimized systems. If we allow, within a functional sequence, for overall optimization values as low as 50%, deleterious mutations remain roughly a thousand times more likely than beneficial mutations in the presence of twelve overlapping codes. As the organism becomes more optimized, the probability of receiving an overall beneficial mutation falls rapidly.

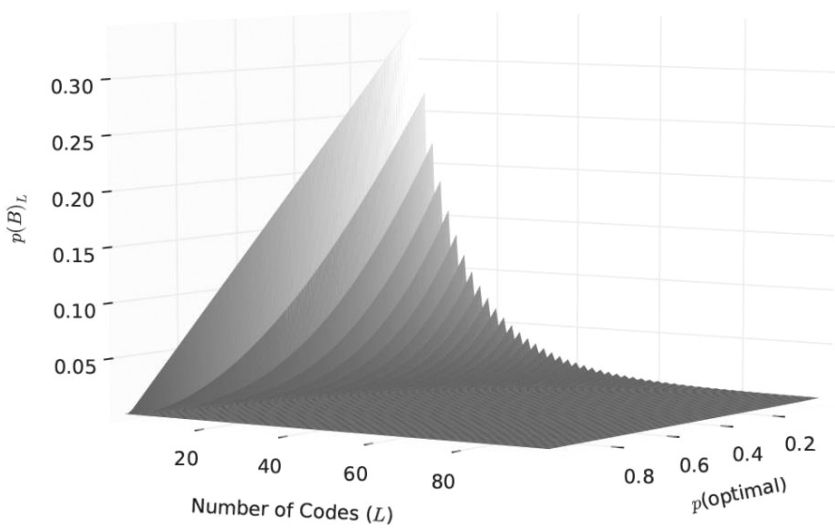

Fig. 3. Number of Codes (L) and p(optimal), plotted against $p(B)_L$, for one to one-hundred codes, showing the general behavior of the model as L increases. The probability of an overall beneficial mutation, $p(B)_L$, decreases exponentially with increasing L.

(Note: The spikes on the surface of the plot, visible near the rear plane of the figure, result from the difference between the majority of an even number of codes and the majority of an odd number of codes. For example, six is the majority for ten codes (60% of total); whereas six is also the majority for eleven codes (only 54% of total). The disparity declines with increasing L.)

We are forced to conclude that the poly-functionality of DNA profoundly affects the expected rate of beneficial mutations. The growing evidence for poly-functional DNA therefore suggests that unambiguously beneficial mutations should be vanishingly rare.

2.2.3 *Third Level of Analysis:*

To further test the effect of multiple constraints on the appearance of beneficial mutations, we constructed a simple poly-constrained artificial system based on English crossword puzzles. Crossword puzzles, for our purpose, are simply collections of words with overlapping, shared letters among some of the words. Figure 6 contains an illustration of such puzzles. We are most familiar with two-dimensional crossword puzzles, where up to two words may share a single letter, but crossword puzzles can be extended to many dimensions. An L-dimensional crossword puzzle is here defined as a collection of words, such that up to L words may share a single, overlapping letter, for one or more letters in the puzzle. Each

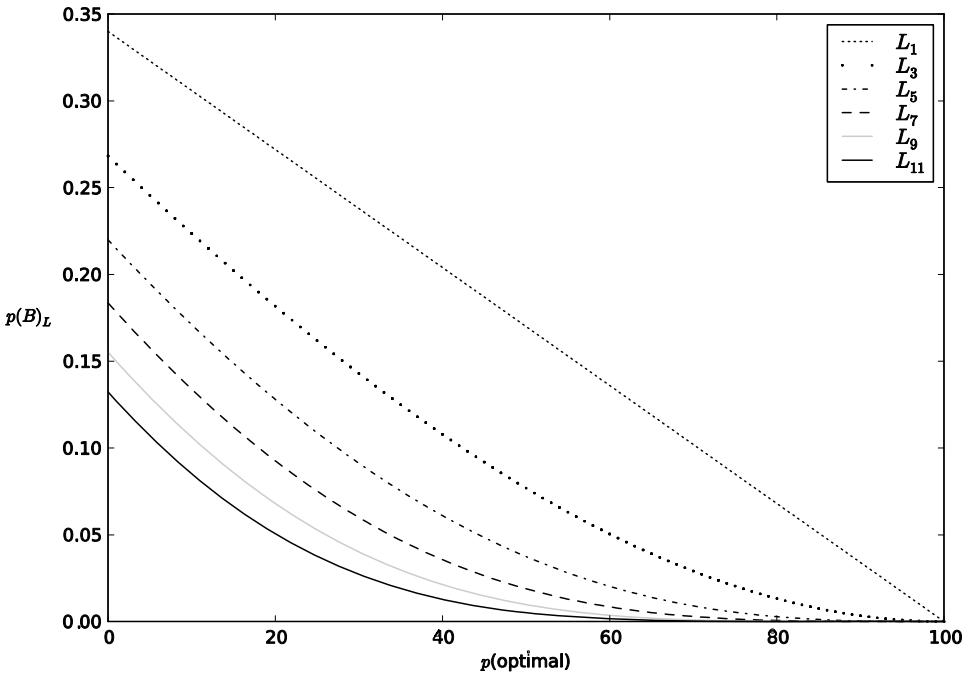

Fig. 4. $p(B)_L$ for different p(optimal) using a fixed p(beneficial | non-optimal) value of 0.34. Even numbered codes are omitted for clarity. If more than 80% of nucleotides are optimized, the probability of a beneficial mutation is near zero for $L \geq 5$.

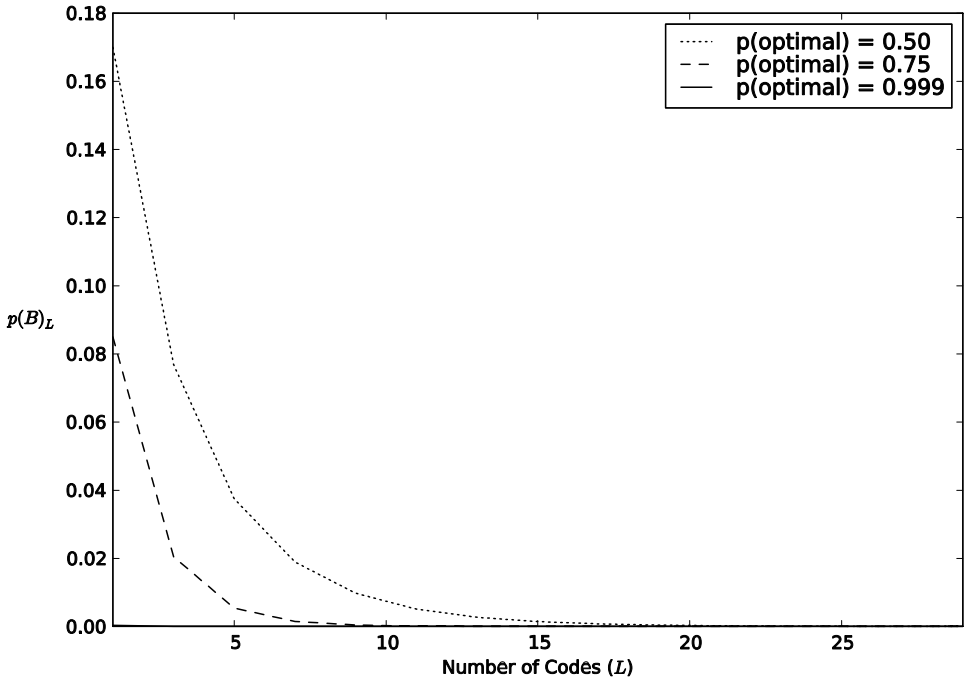

Fig. 5. Exponential decay of $p(B)_L$ as the number of codes (L) increases. Even numbered codes are omitted for clarity. The line for p(optimal) = 0.999 is indistinguishable from the horizontal axis.

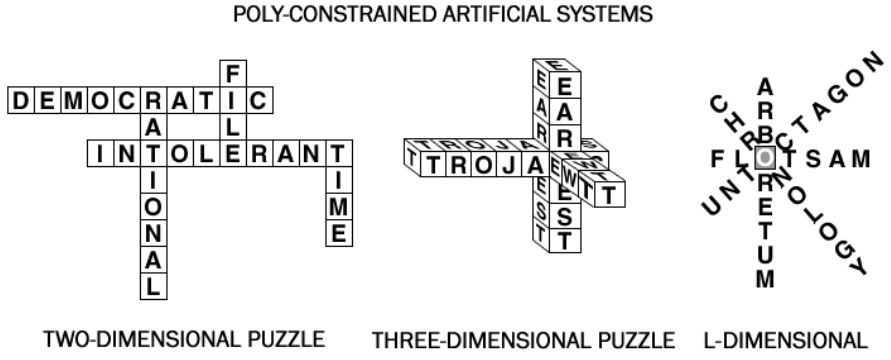

Fig. 6. Crossword puzzles are familiar poly-constrained systems. Intersecting words create constraints on overlapping letters, such as the **E** of **FILE** in the first puzzle. Although a viable, functional mutation may change **FILE** to **FILL**, this would simultaneously change **INTOLERANT** to the non-functional **INTOLLRANT**, a non-word. As we increase the number of dimensions, the number of overlapping words can increase as well, further preventing beneficial changes.

overlap forms a constraint on our puzzle, which limits the possible letters that are allowed in a given position. Increasing the number of words that share a single letter increases the number of constraints on that particular letter, and limits the number of values that letter position may take.

It is known that English words can be transformed into other English words via substitutions of single letters, such as changing the T in RAT to a P, forming RAP. When a letter is constrained within a puzzle, however, changes can affect more than one word simultaneously. A change to a letter may result in a new English word at one level, but render a second word that shares the letter non-functional (non-English). For example, if we have both DOG and GRATEFUL overlapping in a puzzle and sharing a common G, then changing DOG to DOT would change GRATEFUL to the non-English TRATEFUL, which is a deleterious change. However, in some cases we can make an overall beneficial substitution, such as when DOG and GO overlap on the G, and we change GO to the word TO. If our model is correct, then increasing the number of words that overlap should negatively affect the probability of overall beneficial mutations occurring. Therefore, using our simple artificial system, we examine the degree to which overlapping constraints prevent net beneficial mutations from occurring in *L*-dimensional crossword puzzles. In this section, we define a *beneficial mutation* as any change in a word that results in another English word, for a non-optimal position. Mutated words were checked against a text file containing 113,809 official Scrabble® words to confirm whether or not they were functional English words, and if they were found in the file, the change was counted as beneficial *for that word*, as long as the word was not already optimal. If multiple words were changed by a single mutation, we compared how many of the changes were beneficial to how many were deleterious. When the majority of the changes were beneficial, the mutation was counted as beneficial.

We tested groups of 1, 3, 5, 7, 9 and 11 words that contained an overlapping, shared letter. To construct the groups of words, we randomly selected a single letter from the alphabet with uniform probability, and randomly selected a sample of *L* words containing the letter uniformly from our list of possible words. We assumed the overlap occurred at the first instance of the chosen letter within each word. This resulted in an *L*-dimensional puzzle, with the shared letter being the single point of overlap among all words.

Next, we selected a new letter at random from the alphabet (excluding the current letter) with uniform probability, and changed the letter in each of the words. If the change resulted in other English words for the majority of the words in our group, we counted the mutation as beneficial overall. We also introduced a notion of optimization, so that the overlapping letter had a probability, p(optimal), of already being the 'optimal' letter at that position, meaning that for all the possible words that could occur by varying that letter, the current one was already the best. If a word was already optimal, then any mutation at the shared letter counted as deleterious, regardless of whether or not it resulted in another English word.

Figure 7 shows the results of our tests, based on ten-million empirical trials. We found that the estimated probability that a uniform random letter change to a

randomly selected English word would result in another word was roughly 1.65% (using a p(optimal) value of 0.0). As we increase either the level of optimization or the number of overlapping words (L), this probability drops as expected. If more than five overlapping words are present, then the probability of making a change that is beneficial for the majority of words on the shared letter is empirically less than one in 10^7. For eleven overlapping words (similar to eleven overlapping codes in our biological model), we were unable to find a single example of an overall beneficial change during our tests. Therefore, we find the same dearth of unambiguously beneficial mutations in simple poly-constrained systems such as crossword puzzles, due to constraints imposed by the presence of interlocking, mutually dependent systems.

2.2.4 *Summary of Results:*

Having overlapping genetic codes profoundly reduces the probability of beneficial mutation. This is most dramatically seen when we consider unambiguous

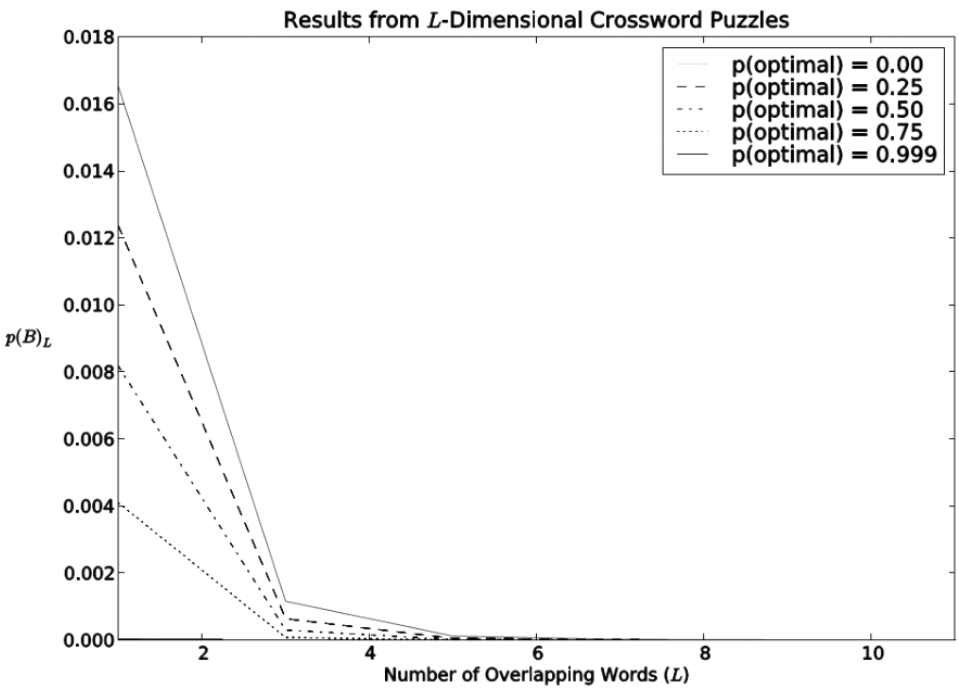

Fig. 7. Empirical results from ten-million trials, plotting the probability of achieving an overall beneficial mutation, $p(B)_L$, when mutating a shared letter among L words. Beneficial mutations were defined as changing a non-optimal word (with probability determined by p(optimal)) to another English word. Graph contains data points for odd numbered L only. The line for p(optimal) = 0.999 is indistinguishable from the horizontal axis.

beneficial mutations — which are not deleterious for any one of the overlapping codes. For example, for those nucleotides that contribute to just three different overlapping codes, assuming each code is 99.9% optimized, less than one in a billion mutations will be unambiguously beneficial. For net-effect beneficial mutations, having three overlapping codes still reduces the probability of beneficial mutation down to less than one per 10^6. When we experimentally test our basic model using a real information system (overlapping English words in the context of a crossword puzzle), we see empirical confirmation of our genetic analysis (even though our only requirement is that a letter substitution creates a new valid English word). Assuming no optimization (namely $p(\text{optimal}) = 0.0$), the probability of having a productive letter substitution within a single word is 1.65%, but when a letter substitution occurs where just three words overlap and $p(\text{optimal}) \geq 0.75$, the probability drops to 7.64×10^{-5}. For nine overlapping words and $p(\text{optimal}) \geq 0.75$ it drops to less than 10^{-7}. Our results clearly show that overlapping codes reduce the potential for beneficial mutation in a most profound way, even for moderately optimized systems.

3. Discussion

Beneficial mutations in nature appear to be so rare that after decades of research we still cannot empirically determine just how rare they are [11]. This suggests they are very rare indeed. There are many reasons to believe that beneficial mutations must be very rare. A mutation is a component of an organism's genetic specifications. Specifications are, by definition, specific. For life to be life requires an exquisite degree of specification — optimization that is hard for us to understand, involving global integration of thousands of systems which have hundreds of thousands of interactions [54]. What is being specified are all the instructions for the establishment, maintenance, and operation of a network of countless biological functions. These functions are integrated into a single elaborate system that is more complex than anything man has ever designed. Each biological specification is encoded by strings of characters (nucleotides or amino acids) that are very specific (and hence very unlikely), with each character having meaning only in the *context* of many other characters — like letters in a book or like the binary bits comprising a computer code. Any random change in such a set of specifications causes some loss of useful information — with a very high degree of probability. The more that each character is contextually interactive with other characters, the less feasible it becomes to improve a set of specifications via random character changes, because each character is multiply *constrained* by its many contextual relationships.

It has often been argued that life's specifications must be very *unconstrained*, citing "junk DNA", synonymous sites in protein coding regions, and the general concept of "bad design". However since the ENCODE project the term "junk DNA" has been largely abandoned [42,55]. "Synonymous mutations" have been shown to be biologically very important [56]. Arguments of bad design have assumed we understand every possible design constraint for a given biological component — which seems unreasonable in light of evidence for poly-functionality of most biological components.

It is now clear that biological systems are very robust and can tolerate much genetic damage. While many in the past have argued that this is due to a general lack of specificity (many sequences will do), this no longer seems reasonable. It now seems more likely that biological systems are robust because of many levels of auto-regulation, self-correction, and countless back-up systems. The new field of systems biology informs us of near-optimality in biological systems, and this appears to be ubiquitous. Such ubiquitous optimality is only conceivable given extremely specific (hence extremely constrained) genetic specifications. Such nearly-optimal genetic specifications should inherently be very difficult to improve, especially when limited to changes which only arise as rare, random, and isolated events.

The discovery of ubiquitous poly-functional DNA is profound, and forces us to reassess our understanding of the degree of genetic specificity and the probability of beneficial mutation. Trifanov pioneered the concept that genomes have a multiplicity of codes and such codes can overlap [40,41]. He showed that a given nucleotide site can participate in multiple genetic codes (with the standard protein code being just being one such code). This is the basic meaning of "poly-functional DNA" [38]. Regrettably, Trifanov's profound discovery generated limited interest. However the ENCODE project has validated the importance of his ideas, and has shown that poly-functional DNA appears to be ubiquitous in higher genomes.

To illustrate how a single nucleotide pair can participate in many different codes, let us consider some of the multiple functions a given nucleotide can participate in (each of these modes of functionality has its own code). A given nucleotide could be: 1) part of an isochore structure; 2) part of a nucleosome binding site; 3) part of a cohesion binding site; 4) part of a transcriptional promoter or enhancer; 5) part of numerous forward-strand RNA transcripts, each with its own transcriptional start and stop points; 6) part of numerous reverse-strand RNA transcripts, each within its own transcriptional start and stop points; 7) part of an mRNA splice site; 8) part of an antisense RNA; 9) part of a nucleo-protein complex; 10) part of several alternately-spliced proteins within the source genic region; 11) part of several alternately-spliced proteins between different genic regions; 12) part of the genome which regulates alternative splicing of proteins; 13)

part of the 3-dimensional organization of the chromosome; 14) part of the 3-dimensional organization of the entire genome; 15) part of the machinery which transports genic regions to active regions of transcription within the nucleus; 16) part of a site for attachment to the nuclear membrane; and 17) part of other undiscovered coding structures.

Given that a single nucleotide pair can potentially participate in so many different codes simultaneously, it should be obvious that this allows data amplification without increasing genome size, and so reflects a very sophisticated form of data compression. One interesting requirement of overlapping codes is that each code must be partially "degenerate" (imperfect) to create the "flexibility" required to allow other overlapping codes. Such degeneracy might appear to the casual observer as an example of bad design, but would actually reflect extreme optimization.

Poly-functional DNA has several implications. Firstly, it is difficult to understand how poly-functional DNA could arise through random isolated mutations. In illustration, when we write, it is difficult to compose a good paragraph (although with training our minds accomplish this with apparent ease). It involves a great deal of optimization because the letters interact, the words interact, the sentences interact, and the ideas interact. But imagine if it was required that such a paragraph had to also have several other messages, using different languages, embedded within it (i.e., using every-other-letter codes, or by reading parts of the message backwards). It would obviously be vastly more difficult to compose a coherent paragraph. The chance of random letter changes creating these types of overlapping messages (in multiple languages) seems incredible, and the chance that natural selection could sort out all the possible interactions also seems incredible.

Given an existing poly-functional DNA sequence, it would seem inordinately difficult to improve it via random mutation. This is at the heart of this paper's analysis. Poly-functional DNA by its very nature is ultra-specific, highly-optimized, and hence highly-constrained. This paper shows that when a nucleotide participates in more than one code, a mutation at that site is going to almost certainly be deleterious relative to the first code, and even when a mutation is beneficial in the first code, it will still almost certainly be deleterious in one or more of the other codes. Hence a mutation at a poly-functional site will at best be only "ambiguously beneficial" — still being deleterious at one or more other levels. The exact degree to which nucleotides participate in two or more codes is still unknown, but if it is at all common, it should profoundly reduce the probability of mutations which are unambiguously beneficial.

Mutations that affect more than one code are pleiotropic, in that they have multiple biological effects. This is consistent with what geneticists have known for

many decades — most known mutations are pleiotropic at some level — affecting more than one biological trait. In the case of most human genetic pathologies, the multiple effects of a mutation are usually all negative. In the rare case of an ambiguously beneficial mutation, a certain beneficial effect will be combined with one or more deleterious effects (for example, carriers of the mutation for sickle cell anemia are more resistant to malaria — but suffer from impaired hemoglobin function and reduced red blood cell counts).

In our analysis we have for simplicity assumed that if a mutation has a single beneficial effect and a single deleterious effect, it is counted as neutral. However this is not realistic because we can logically expect most such ambiguous mutations to have a net deleterious effect. This is because, not only is it more likely for a random change to damage an optimized system than improve it, the nature of that damage will tend to be more pronounced than any potential improvement. Within a highly optimized genetic system, mutational damage can range from very slight to lethal — but improvements will consistently be only very slight. For example, certain spelling errors in a plane's assembly manual could cause the plane to fly twice as slow, but no spelling error can be expected to cause the plane to fly twice as fast. Therefore selection for the ambiguous beneficial mutation is especially problematic — the positive and negative effects will tend to cancel out, but the deleterious effect will tend to overshadow the beneficial effect.

The analysis in this paper provides strong evidence that the discovery of multiple overlapping codes requires us to re-adjust downward our estimates of the rate of beneficial mutation. At the same time, the newly emerging field of systems biology strongly points to a very high degree of optimization in all biological systems, and this also requires us to adjust downward our estimate of the rate of beneficial mutation. Lastly, there is clearly a selection threshold [57], wherein below a certain limit, all low impact beneficial mutations must become invisible to natural selection. Using realistic biological conditions, it appears that in a large genome, at least 99% of all beneficial mutations should be so subtle as to be un-selectable [57]. So the rate of *useful* beneficial mutations should be at least two orders of magnitude less that the rate of actual beneficial mutations. Taking this into consideration, this suggests we should reduce the probabilities reported in this analysis by another two orders of magnitude. Although we do not quantitatively analyze the problem of drift in this paper, it is important to note that the vast majority of beneficial mutations that do arise, and are above the selection threshold of the population, are still lost due to genetic drift.

Logic and mathematical analysis persuade us that unambiguous beneficial mutations should be extremely rare. This is consistent with the apparent absence of documented mutations that are unambiguously beneficial (i.e., beneficial at one or more levels, while not deleterious on any level). To our knowledge there is no case of a

mutation which is unambiguously beneficial and which has been shown to distinctly improve the inner workings or an organism. Certainly there are numerous documented cases of simple adaptations to an external environment factor, but these special cases have little bearing on how most of the information within a genome arose — because most of a genome's information specifies life's internal workings.

The long-term *E. coli* experiments of Lenski *et al.* [58] have been widely acclaimed as "proof of evolution before our very eyes". Such evolution would suggest that numerous beneficial mutations were arising. It is useful to examine these claims more carefully. The *E. coli* in these long-term experiments (which involved vast numbers of cells over vast numbers of generations), did not appear to evolve any new functions. The only changes that were observed involved adaptations to the specific artificial growth medium. This type of adaptive change to an external factor is only a superficial improvement — it does not explain how the *E. coli* genome arose, nor how the information specifying the bacteria's internal workings arose. Moreover, those studies failed to show any specific mutation which was unambiguously beneficial. In fact, it is clear that most of the adaptive mutations involved loss of function mutations — including deletions of genetic material [59]. It should be obvious that genetic material not essential for a given environment, if inactivated or deleted, can decrease metabolic load, and so can allow more total growth in that given medium. But all such broken genes and deletions clearly involve a net loss of information, and there is no question that the resulting bacteria became less "fit" in the broader and truer sense. Such strains of bacteria would immediately go extinct in virtually any natural environment.

In that enormous evolutionary experiment, the closest instance to an unambiguously beneficial mutation was a mutation that allowed the bacteria to utilize citrate from the artificial medium [60]. However, this did not actually involve evolution of a new function — the *E. coli* already had all the machinery needed for metabolizing citrate, but the citrate could not normally pass through the bacteria's external membrane. In light of the work of Behe [61], in such a case the most likely explanation for this mutant strain would be a loss-of-function mutation that would result in a leaky membrane. Certainly no exhaustive research was done to prove that the mutation in question had zero deleterious effect.

3.1 *Possible Objections*

Contrary to the thesis of this paper, some scientists have argued that beneficial mutations might be extremely common — even approaching 50% of all non-neutral mutations [14,15,37]. The concept that beneficial mutations might be extremely common traces back to some simple mental constructs suggested by

Fisher [37]. Fisher's most famous illustration was the example of focusing a microscope. If the microscope is significantly out of focus and one makes a small random adjustment, there is roughly a 50% chance of improving the focus (this would only be true for extremely small adjustments). Fisher argued that in the same way, a very low impact mutation might have roughly a 50% chance of improving fitness (in his day the near-neutral mutation problem had not yet been identified, and he apparently did not consider that such a low-impact mutation might be inherently un-selectable). When Fisher developed this illustration, DNA had not yet been discovered, genes seemed to be very simple (beads on a string), and the nature of mutation was unknown. With the advent of molecular genetics it is now evident why this analogy simply is not applicable.

Fisher knew mutations happened, but he did not know what they really were. We now know mutations are essentially spelling errors in the assembly manual of the cell. There are some small isolated parts of the genome (such as gene promoters), which can act like an electric rheostat or like a microscope's focusing knob. Mutations within these special regions can raise of lower a gene's expression level — and in this special case mutations that can increase expression can conceivably be almost as common as those that decrease expression. For example, mutations in the promoter region of the growth hormone gene might cause either giants or dwarfs. These special variable switches within DNA appear to function for the purpose of fine-tuning a trait such as height. But these special cases do not reflect the true nature of total fitness (total biological functionality), and do not reflect the way most of the genome functions. A change in height can only result in two possibilities — taller or shorter. But overall biological fitness is inherently multi-dimensional, it involves a multitude of separate traits and is contingent upon millions of nucleotides, and requires very precise genetic specifications. When a single trait is defined by just 100 functional nucleotides, that trait's genetic optimum is an extremely specific set of 100 base pairs (one specific set of 4^{100} sets, or one in 10^{60}). If that trait is anywhere near its optimum, then there are a multitude of mutations which can make the trait worse, but there are very few opportunities to make the trait better. This is analogous to a random letter change in a text that results in a superior text. As a message becomes more and more complex and refined, a text change must be more and more specific in order to enhance that message, and hence the greater the constraint for achieving improvement via any random change. As this paper shows, the recent discovery of poly-functional DNA vastly compounds this problem. To his credit, Fisher acknowledged that the chance of improvement via a random change must approach zero — either when the focus is already nearly optimized, as the size of the change in focus grows larger, or as the number of dimensions defining the trait (i.e., overall fitness) becomes larger [37].

There is another aspect of Fisher's theoretical work, which arose because he did not understand that genes specify information and that mutations are just errors in genetic specifications. Fisher imagined that all biological variation arose symmetrically. In the case of the focusing knob on a microscope, the knob turns equally well both ways, and Fisher imagined this would be equally true for mutations affecting any biological trait — such as height or vigor. There would be just as many mutations that increased performance as diminished it. This is the error underlying Fisher's famous "Fundamental Theorem of Natural Selection" [37]. Given a population with performance levels following a bell-shaped curve, he reasoned that any level of selection will always remove at least some of the underperformers and will favor at least some of the higher performing individuals. This would consistently yield higher mean performance in the next generation. He then assumed new mutations would arise creating new variation *symmetrically* around the new mean. This is what led Fisher to believe he had a mathematical proof that continuous evolutionary improvement was unavoidable. But we now know that mutations are essentially word-processing errors in the DNA, so new variation will be *extremely asymmetrical* and will be almost exclusively deleterious. So, for example, apart from a small set of mutations within its promoter region, mutations deleterious for a gene's function will be much more common than mutations for enhanced function — invalidating Fisher's Theorem, and negating his simple microscope analogy.

When we consider the organism as an integrated whole, we conclude beneficial mutations should generally be very rare for the reasons discussed above. We can only rationalize that beneficials might be common when considering one tiny component of fitness at a time, such as height. When we do this we artificially make fitness seem one-dimensional — analogous to Fisher's example of focusing his microscope. Within this very limited context, most of the constraints on what constitute a "beneficial" mutation disappear. For example, in terms of malaria resistance, a deleterious mutation in the hemoglobin gene can be defined as "beneficial", even though it is actually a semi-lethal mutation. Under this type of very limited one-dimensional analysis, the rate of beneficial mutation can appear much higher than it really is. This is especially true in the case of those rare mutations that strongly interact with major environmental factors that are external to the organism (i.e., antibiotic resistance). Relative to just that single component of the entire biological system, one can expect a reasonable probability of beneficial mutation. This is because any genetic change that interacts with that specific external factor has a nearly equal probability of making that factor's impact either better or worse. This allows biological fine-tuning for a single isolated trait, relative to a single external factor. In these special cases Fisher's microscope analogy has some validity, so that relative to that single trait (or within a single code),

random mutations can have a reasonable probability of being beneficial. This may explain why most examples of beneficial mutation involve a form of adaptation to a local condition. However, most genomic information does not involve adaptation to specific high-impact external factors, but rather specifies a labyrinth of complex, integrated, and optimized biological functions internal to a living system. The important distinction between adaptation to some local external condition versus maintenance of total genomic integrity is illustrated by a recent study. That study showed that specific adaptive mutations within a mutagenized population, when tested in a particular environment, obscured, but did not halt genetic degeneration [62].

A few recent studies have inferred extremely high beneficial mutations rates, based on data from mutation accumulation (MA) experiments [14,15]. These MA experiments have significant problems. No actual mutations were actually seen, the beneficial and deleterious mutation rates were only inferred based upon the differential growth rates of a limited number of isolated strains. These experiments were not capable of identifying the vast majority of subtle mutations that arose in the populations. They could only detect those few mutations that had large effects and affected a single trait (growth rate on a given medium) making inferences about total mutation rates entirely unwarranted. The observed effects in these two studies could be attributed to a specific one-dimensional adaptation, which could arise due to a specific mutational hotspot, or could even be due to an epigenetic effect. Lastly, unintentional selection could not be rigorously precluded.

Given the one-dimensional nature of these MA experiments, a relatively high rate of beneficial mutation is not unexpected because only one trait was measured, making fitness appear one-dimensional (like Fisher's microscope), or like a simple one-dimensional trait such as height. In both of these studies, fitness was measured only in a very narrow sense and in a very specific and unnatural environment. Instead of total fitness, what was being measured was the degree of biological fine-tuning to a very specific and very artificial circumstance. In one case [14], the researchers tested the ability of yeast strains that were initially grown under minimal selection conditions (to allow mutations to accumulate), to then grow slightly faster than the source strain in the same artificial medium where the mutations had been accumulating. In that study 5.75% of the derived lines grew faster than the parental strain, under those specific conditions. In a very similar yeast experiment [15], the researchers again minimized selection to allow mutation accumulation, and then tested derived strains for ability to compete with the parental genotype in artificial medium. In the second study 25% of the derived lines out-grew the parental strain. In both cases the researchers used extremely narrow and unnatural criteria for measuring "fitness", and the singular traits they focused on might easily have been affected (for better or worse) by very simple genetic or

epigenetic variations. However, natural selection, as it occurs in the natural world, must act on "fitness" in a much fuller sense — it must involve all heritable traits, all functional nucleotides, all codes, all relevant environments, and all phases of the life cycle. The authors of one of these two studies freely acknowledge these types of limitations on the interpretation of their study (including the possibility of unintentional selection) and state: "the large proportion of beneficial mutations observed in our experiment may in part reflect a combination of factors: the ancestor's distance from the fitness optimum, yeast's recent genome duplication, our examination of only a single environment and life-history stage, and the recessive nature of deleterious mutations" [14].

The two isolated reports mentioned above, which claim very high rates of beneficial mutation, are inconsistent with a much broader range of observations. For example, the *net effect* of currently observed human mutation is universally recognized as being distinctly deleterious, and hence clearly represents a serious problem in terms of public health. This is made obvious by the fact that there are thousands of Mendelian pathologies documented in man, in spite of the tendency for natural selection to eliminate such mutations from the population. Conversely, there are only a handful of putative beneficial mutations commonly cited for man, despite the tendency for natural selection to amplify such mutations. Moreover, the "benefit" of most such mutations is typically equivocal, usually being defined as beneficial in only a very narrow sense (as in the case of sickle cell anemia).

Another possible argument against the thesis of this paper might be that it is contradicted by a substantial volume of scientific literature that uses DNA sequence comparisons to infer historical positive selection events for great numbers of putative beneficial mutations. It is important to realize that the vast majority of the putative beneficial mutations claimed in these papers are just observed alternative nucleotides — with no known biological function (the presumed benefits being inferred, not being in any way understood or observed). We naturally acknowledge the operation of selection for beneficials in the past, but argue that such selection is severely constrained by the reality of very low rates of beneficial mutations, as this study and common sense both demand. It is noteworthy that a significant part of this body of literature that claims proof of so much positive selection in the past (based upon observed sequence variability in the present), may suffer from systematic error and is now being challenged [43,54,55]. Inferences of specific positive selection events in the past, based solely upon sequence data and allele frequencies, are mere historical inferences. The observed sequence variations might be explained using alternative mechanisms such as differential mutation rates or ordinary statistical fluctuations.

A final possible argument against the thesis of this paper might be that our analysis involved point mutations, but did not consider duplications. Some might

argue that genetic duplications are especially likely to be beneficial. However in terms of immediate effects, duplications are more likely than other mutations to cause harm. Duplications are more likely to be immediately deleterious because unlike point substitutions, they scramble the genome — causing frame shifts and generally disrupting genomic context and architecture. Like the duplication of letters, words, or paragraphs in a regular text — genomic duplications add nothing, but systematically disrupt context. Furthermore, unlike other types of mutations, duplications increase metabolic load for the host cell in terms of DNA replication, repair, transcription, and translation. So if a duplication is neutral in terms of information, it is then by definition a deleterious mutation due to increased metabolic cost.

Can it be argued that even if duplications are not immediately beneficial, they might still be beneficial in the long run, producing large reservoirs of "junk DNA", which could then serve as a breeding ground for future evolutionary "experimentation and innovation"? The concept of building up a large amount of "junk DNA" in the genome for possible long-term evolutionary benefit has several flaws. Firstly, the most recent evidence [35,54] suggests that the genome is mostly functional and that so there is little junk DNA. Secondly, the huge metabolic cost of junk DNA would be immediately deleterious. Thirdly, long-term benefits would be remote and hypothetical, while selection only operates in the present and cannot anticipate future benefits. Fourthly, even within junk DNA, mutations can still be deleterious due to negative interactions with the functional genome. Lastly, the prospects for beneficial mutations arising within junk DNA is very problematic, because like a letter within a text, no nucleotide is good or bad in itself, but only in the context of many other nucleotides. Within the context of a non-functional array of letters, it is not reasonable to expect a spelling error to ever create useful information. Single letters outside of a functional context cannot take on a function of their own. In the same way, within any DNA sequence that is truly neutral "junk", there is no frame of reference for defining a point substitution as being either beneficial or deleterious in terms of useful information. There is no functional context within which beneficial mutations could arise — with one major exception. Ironically, there is one type of beneficial mutation that should arise systematically within junk DNA — *deletions*. Essentially all deletions within junk DNA should be beneficial, due to improved metabolic efficiency. The larger the deletion — the more the benefit, and so the stronger the selective advantage. So to the extent that selection is actually operational, all junk DNA should be systematically deleted. This should happen long before enough beneficial mutations might accumulate within the junk DNA to give it a new and meaningful biological function.

4. Conclusions

Our analysis confirms mathematically what would seem intuitively obvious — multiple overlapping codes within the genome must radically change our expectations regarding the rate of beneficial mutations. As the number of overlapping codes increases, the rate of potential beneficial mutation decreases exponentially, quickly approaching zero. Therefore the new evidence for ubiquitous overlapping codes in higher genomes strongly indicates that beneficial mutations should be extremely rare. This evidence combined with increasing evidence that biological systems are highly optimized, and evidence that only relatively high-impact beneficial mutations can be effectively amplified by natural selection, lead us to conclude that mutations which are both selectable and unambiguously beneficial must be vanishingly rare. This conclusion raises serious questions. How might such vanishingly rare beneficial mutations ever be sufficient for genome building? How might genetic degeneration ever be averted, given the continuous accumulation of low impact deleterious mutations?

***Addendum**: We append the following reference which appeared following the finalization of this chapter, which shows evidence that mammalian genes have extensive overlapping functions ("Locating protein-coding sequences under selection for additional, overlapping functions in 29 mammalian genomes." Lin MF, Kheradpour P, Washietl S, Parker BJ, Pedersen JS, Kellis M. Genome Res. 2011 Nov;21(11):1916–28. Epub 2011 Oct 12). We also append another significant paper ("The genetic code is nearly optimal for allowing additional information within protein-coding sequences", Itzkovitz S., Alon U., Genome Res. 2007 Apr; 17(4):405-12. Epub 2007 Feb 9).*

References

1. Kibota T, Lynch M (1996) Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*. Nature 381:694–696.
2. Charlesworth B, Charlesworth D (1998) Some evolutionary consequences of deleterious mutations. Genetica 103: 3–19.
3. Elena S, et al (1998) Distribution of fitness effects caused by random insertion mutations in Escherichia coli. Genetica 102/103: 349–358.
4. Gerrish P, Lenski R N (1998) The fate of competing beneficial mutations in an asexual population. Genetica 102/103:127–144.
5. Crow J (2000) The origins, patterns, and implications of human spontaneous mutation. Nature Reviews 1:40–47.

6. Bataillon T (2000) Estimation of spontaneous genome-wide mutation rate parameters: whither beneficial mutations? Heredity 84:497–501.
7. Imhof M, Schlotterer C (2001) Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. Proc Natl Acad Sci USA 98:1113–1117.
8. Orr H (2003) The distribution of fitness effects among beneficial mutations. Genetics 163: 1519–1526.
9. Keightley P, Lynch M (2003) Toward a realistic model of mutations affecting fitness. Evolution 57:683–685.
10. Barrett R, et al (2006) The distribution of beneficial mutation effects under strong selection. Genetics 174:2071–2079.
11. Bataillon T (2000) Estimation of spontaneous genome-wide mutation rate parameters: whither beneficial mutations? Heredity 84:497–501.
12. Elena S, et al (1998) Distribution of fitness effects caused by random insertion mutations in *Escherichia coli*. Genetica 102/103: 349–358.
13. Gerrish P, Lenski R N (1998) The fate of competing beneficial mutations in an asexual population. Genetica 102/103:127–144.
14. Joseph S, Hall D (2004) Spontaneous mutations in diploid *Saccharomyces cerevisiae*: More beneficial than expected. Genetics 168:1817–1825.
15. Dickinson W (2008) Synergistic fitness interactions and a high frequency of beneficial changes among mutations accumulated under relaxed selection in *Saccharomyces cerevisiae.* Genetics 178:1572–1578.
16. Muller HJ (1964) The relation of recombination to mutational advance. Mutat Res 1:2–9.
17. Kimura M (1968) Evolutionary rate at the molecular level. Nature 217:624-626.
18. Ohta T (1973) Slightly deleterious mutant substitutions in evolution. Nature 246: 96–98.
19. Ohta T (1974) Mutational pressure as the main cause of molecular evolution and polymorphism. Nature 252:351–354.
20. Ohta T (1992) The nearly neutral theory of molecular evolution. Ann Rev Ecol Syst 23:263–286.
21. Ohta T (2002) Near-neutrality in evolution of genes and gene regulation. Proc Natl Acad Sci USA 99:16134–16137.
22. Kimura M (1979) Model of effectively neutral mutations in which selective constraint is incorporated. Proc Natl Acad Sci U S A 76:3440–3444.
23. Kimura M (1983) Neutral Theory of Molecular Evolution. Cambridge Univ. Press, New York.
24. Kondrashov AS (1995) Contamination of the genome by very slightly deleterious mutations: why have we not died 100 times over? J Theor Biol 175:583–594.
25. Lynch M, Conery J, Burger R (1995) Mutation accumulation and the extinction of small populations. Am Nat 146:489-518.

26. Lynch M, Conery J, Burger R (1995) Mutational meltdown in sexual populations. Evolution 49(6):1067–1080.
27. Higgins K, Lynch M (2001) Metapopulation extinction caused by mutation accumulation. Proc Natl Acad Sci USA 98:2928–2933.
28. Loewe L (2006) Quantifying the genomic decay paradox due to Muller's ratchet in human mitochondrial DNA. Genet Res 87:133–159.
29. Sanford J, Baumgardner J, Gibson P, Brewer W, ReMine W (2007) Using computer simulation to understand mutation accumulation dynamics and genetic load. In 7th International Conference on Computational Science, Beijing, China, May 27–30, 2007 Proceedings, Part II, LNCS 4488. Edited by Shi Y, van Albada GD, Dongarra J, Sloot PMA, Berlin/Heidelberg: Springer-Verlag: 386–392.
30. Gibson P, Baumgardner JR, Brewer WH, Sanford JC (2013) Can purifying natural selection preserve biological information? In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 232–263.
31. Brewer WH, et al (2013) Using numerical simulation to test the "mutation-count" hypothesis. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 298–311.
32. Baumgardner JR, et al (2013) Can synergistic epistasis halt mutation accumulation? Results from numerical simulation. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 312–337.
33. Nelson CW, et al (2013) Computational evolution experiments reveal a net loss of genetic information despite selection. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 338–368.
34. Ohta T (1973) Slightly deleterious mutant substitutions in evolution. Nature 246: 96–98.
35. Hirsch A, Fraser H (2001) Protein dispensability and rate of evolution. Nature 411:1046–1049.
36. Poon A, Otto S (2000) Compensating for our load of mutations: freezing the meltdown of small populations. Evolution 54: 1467–1479.
37. Fisher R (1930) The genetical theory of natural selection. Clarendon Press, Oxford. Pages 38–41.
38. Sanford J (2008) Genetic Entropy and the Mystery of the Genome. FMS Publications, NY. Pages 131–142.
39. Trifonov EN (1989) Multiple codes of nucleotide sequences. Bull of Mathematical Biology 51:417–432.
40. Trifanov EN (1997) Genetic sequences as products of compression by inclusive superposition of many codes. Mol Biol 31:647–654.

41. Kapranov P, et al (2005) Examples of complex architecture of the human transcriptome revealed by RACE and high density tiling arrays. Genome Res 15:987–997.
42. Birney E, et al (2007) Encode Project Consortium: Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447:799–816.
43. Itzkovitz S, Hodis E, Sega E (2010) Overlapping codes within protein-coding sequences. Genome Res. 20:1582–1589.
44. Schneider, TD (2000) Evolution of biological information. Nucleic Acids Res 28 (14): 2794–2799.
45. Montañez G, Ewert W, Dembski WA, Marks II RJ (2010). Vivisection of the ev Computer Organism: Identifying Sources of Active Information. BIO-Complexity, 2010 (3):1–6
46. Lenski RE, Ofria C, Pennock RT, Adami C (2003) The evolutionary origin of complex features. Nature 423:139–144
47. Ewert W, Dembski WA, Marks II RJ (2009) Evolutionary Synthesis of Nand Logic: Dissecting a Digital Organism,. Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics. San Antonio, TX, USA: 3047–3053.
48. Goh CK, Tan KC (2009) A Competitive-Cooperative Coevolutionary Paradigm for Dynamic Multiobjective Optimization. IEEE Transactions on Evolutionary Computation, 13 (1): 103–127
49. Rachmawati L, Srinivasan, D (2009) Multiobjective Evolutionary Algorithm With Controllable Focus on the Knees of the Pareto Front. IEEE Transactions on Evolutionary Volume: 13 , Issue 4: 810–824.
50. Farina, M., Deb, K., Amato, P. (2004) Dynamic multiobjective optimization problems: test cases, approximations, and applications. IEEE Transactions on Evolutionary Computation, Vol 8 , Issue 5: 425–442
51. Marks II RJ (2009) Handbook of Fourier Analysis and Its Applications. Oxford University Press.
52. Cover TM. Thomas JA (2006) Elements of Information Theory, 2nd edn. Wiley-Interscience, New York.
53. Dembski WA, Marks II RJ (2009) Conservation of Information in Search: Measuring the Cost of Success. IEEE Transactions on Systems, Man and Cybernetics A, Systems and Humans: 39(5): 1051–1061.
54. Stumpf M, et al (2008) Estimating the size of the human interactome. Proc Natl Acad Sci USA 105:6959–6964.
55. Wells J (2013) Not junk after all: non-protein-coding DNA carries extensive biological information. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 210–231.

56. Kudla G., Murray AW, Tollervey D, Plotkin JB (2009) Coding-sequence determinants of gene expression in *Escherichia coli*. Science 324:5924.
57. Sanford JC, Baumgardner JR, Brewer WH (2013) Selection threshold severely constrains capture of beneficial mutations. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 264–297.
58. Lenski RE, Travisano M (1994) Dynamics of adaptation and diversification: A 10,000-generation experiment with bacterial populations. Proc Natl Acad Sci USA 91:6808-6814.
59. Cooper VS, Lenski RE (2000) The population genetics of ecological specialization in evolving *Escherichia coli* populations. Nature 407:736–739.
60. Blount ZD, Borland CZ, Lenski RE (2008) Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli.* Proc Natl Acad Sci USA 105:7899–7906.
61. Behe MJ (2013) Getting there first: an evolutionary rate advantage for adaptive loss-of-function mutations. In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological Information — New Perspectives. World Scientific, Singapore, pp. 450–473.
62. Springman R, Keller T, Molineux I, Bull J (2009) Evolution at a high imposed mutation rate: adaptation obscures the load in phage T7. Genetics (26 October 2009), genetics.109.108803.
63. Hughes A (2007) Looking for Darwin in all the wrong places: the misguided quest for positive selection at the nucleotide level. Heredity 99:364–373.
64. Hughes A (2008) The origin of adaptive phenotypes. Proc Natl Acad Sci USA 105: 13193–13194.
65. Kryukov G, Schmidt S, Sunyaev S (2005) Small fitness effect of mutations in highly conserved non-coding regions. Human Molecular Genetics 14: 2221–2229.