

Mathematical Challenges to Macroevolution

John B. Andelin

Williston, North Dakota, USA

Email: andelin2022@gmail.com

How to cite this paper: Andelin, J.B. (2022) Mathematical Challenges to Macroevolution. *Journal of Applied Mathematics and Physics*, 10, 3385-3391.

<https://doi.org/10.4236/jamp.2022.1011224>

Received: October 6, 2022

Accepted: November 26, 2022

Published: November 29, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The theory of evolution was advanced by Darwin in 1859, prior to Mendel's experiments demonstrating the particulate nature of inheritance. The modern synthesis was formulated in the early 1940s, well before the concept of coded information was understood. This paper outlines four mathematical challenges to the modern synthesis, which are based on current understanding of the proposed mechanisms of evolutionary change within the constraints of experimental molecular biology.

Keywords

Evolution, Convergence, Probability, Binomial Distribution Formula, Ex Post Facto Probability, Post-Hoc Probability, Modern Synthesis

1. Introduction

The modern theory of evolution is founded on the belief that mutations occur randomly across the genome, and that mutation rates are constant. This view has been challenged by numerous studies, which introduce uncertainties in calculating the probabilities that a given mutation will appear within the genome to allow for natural selection [1]. Nevertheless, models of evolution are generally based on the assumption that mutations do not affect certain genes more than others, and that their appearance is not influenced by the environment or other factors.

In this paper, four mathematical challenges to neo-Darwinism are presented. The calculations are based on the unpredictability of mutations. In view of mutation bias which preferentially effects non-coding DNA sequences, these calculations understate the improbabilities involved in many evolutionary proposals.

2. Ex Post Facto Probability

The theory of evolution attempts to explain historical events. It is often con-

tended that probability challenges cannot be applied *ex post facto*, or “after the fact”. In other words, it is argued that you cannot look at an existing reality and determine the probability that it arose through a random process. In attempting to refute probability challenges to evolution, David H. Bailey argued that evolutionary mechanisms cannot be mathematically challenged, stating that “*Any argument based on post-hoc probability reckoning is effectively nullified*” [2]. As an example, he noted that any random shuffle of a deck of 52 cards results in a sequence that has an *ex post facto* probability of 1.25×10^{-68} . Hence, once a deck is shuffled, you cannot look at the sequence *ex post facto* and calculate the probability that such a sequence would have been created. It is implied that just as every shuffle creates an improbable sequence, random mutations can explain the existence of improbable sequences of DNA. This is a very common error of applied mathematical logic. With mathematical certainty, a random shuffle of a deck will result in a random sequence. If someone shuffled a deck of cards and the sequence was perfectly ordered, with all hearts, spades, diamonds and clubs grouped together, one could conclude (*ex post facto*) that the deck was not adequately shuffled. Just as a shuffled deck of cards will always produce a random sequence, random mutations always result in unpredictable nucleotide substitutions distributed over the genome.

Ex post facto probability arguments are routinely employed in criminal trials, and life-and-death decisions are made based on those improbabilities. For example, consider a defendant, Jack Whiting, who is accused of pushing his wife off a cliff to her death. The prosecution argues that only 100 women die each year from falling off cliffs. Since there are 68 million married women in the country, the chance that Jack’s wife would accidentally fall off a cliff was only one in 680,000. Thus, Jack must be guilty. If this was the only evidence presented, a jury would not be convinced that Jack is guilty. Jack might have been just as unlucky as the other 99 husbands in the country who lost their wives to falling off cliffs. In this case, an *ex post facto* probability calculation is a fallacious argument because, like a shuffled deck of cards, only one result is considered.

Now, imagine that the prosecution produced documentation that Jack’s two former wives during the past ten years also died from accidentally falling off cliffs. Any jury member would be highly suspicious, because no husband could be that unlucky. It is self-evident that random chance cannot result in a pattern of outcomes that is not random. The probability of three events of this type occurring within ten years in one family is about one in 2.6 quadrillion. This probability is calculated *ex post facto*. A shuffled deck of cards can produce any of trillions of trillions of possible sequences. Three separate shuffles cannot result in the same sequence with each shuffle.

Consider another example. Suppose someone flipped a coin 100 times and achieved 60% heads. Suppose he claimed these were fair, random tosses. The *ex post facto* probability of those being random, fair tosses is one chance in 37. So, it is possible that the coin was fair, but unlikely. Now suppose he claimed to get

6000 heads out of 10,000 tosses. That is impossible with a fair coin, because the ex post facto probability is 8.7×10^{-90} . These calculations illustrate an important probability principle. The greater the number of trials is, the more certain the outcome will be constrained to a random result. This principle is mathematically described by the central limits theorem.

3. Challenge #1: Proposed Evolutionary Pathways Require a Directional Pattern of Mutations

The genomes of higher organisms consist of hundreds of millions to billions of nucleotide base pairs. An average of around 100 - 300 point mutations occur per generation in higher organisms [3]. Thus, the probability that a birth will contain a specific nucleotide substitution (A, T, C or G) is one in billions. For random mutations to result in the incremental improvement of a structure, those mutations would need to repeatedly affect specific genes. Just as millions of random coin tosses cannot result in 60% heads, random mutations cannot result in larger numbers of mutations that preferentially affect certain genes over other genes. With this in mind, proposed pathways of evolution can be evaluated (ex post facto), and it can be determined whether or not the mutations required to effectuate such changes could have been random.

It is commonly believed that if a pathway of functional continuity can be envisioned, a plausible evolutionary pathway has been proposed. For example, the evolution of the eye is often explained by a pathway of gradually increasing curvature of the retina, with each step resulting in slightly increasing visual acuity. Generally ignored is the calculation of the probability of the mutations that create a curved retina to appear in a given population. For example, Nilsson and Pelger proposed a model of eye evolution in fish, beginning with a flat photo-sensitive spot [4]. They proposed 1829 steps of incrementally increasing curvature of the retina and narrowing of the pupil, and demonstrated that each step would have been functionally superior to the preceding step. In the article, they conclude that a camera eye could have easily evolved in 364,000 years. However, the probability of those 1829 mutations appearing in a given population of fish over a specified number of generations is not calculated.

The following is a calculation of the probability of these changes to appear in the genome. Assume a genome size of 1.5 billion base pairs. Assume that one point mutation results in one step of an increasingly curved retina (a very generous assumption). Assume a population size of 20,000. Assume 1000 viable offspring per generation. Assume a mutation rate of 150 point mutations per generation. Assume a fixation rate of 0.002 for each favorable mutation. Using the binomial distribution formula, the probability of such a pattern of mutations over 364,000 years is 1.5×10^{-1423} (see calculation below*).

*Binomial Distribution:

Calculation of probability of the evolution of a globe-shaped eye (Nilsson and Pelger) in a population of fish:

$$\text{Probability mass: } f(x, n, p) = {}_n C_x p^x (1-p)^{n-x}$$

Probability of success: inverse of genome size ($1/1.5 \times 10^9$) \times 0.33 (3 possible nucleotide substitutions for each position) \times fixation rate (0.002) = 4.4×10^{-13} .

Number of trials (n): number of reproducing pairs in population ($20,000/2$) = $10,000 \times$ number of mutations per generation, per germ cell ($150/2 = 75$) \times number of offspring per generation (1000) \times number of generations (364,000) = 2.73×10^{14} success number: 1829 (incremental steps of eye evolution).

Results

Probability mass (f): 1.43×10^{-1423} .

Lower cumulative distribution (P): 1.

Upper cumulative distribution (Q): 1.53×10^{-1423} .

Although the numbers used in this calculation, may be disputed, it is apparent that the probability of such a pattern of mutations appearing in a finite population over a limited number of generations can be rounded off to zero. Think about achieving 60% heads with 10,000 coin tosses. With mathematical certainty, random mutations will not deviate from predictable random outcomes to preferentially effect specific genes.

It has been argued that these calculations are flawed, because DNA is a complex code, and incremental improvement does not necessarily involve similar mutations in the same genes. The lack of a one-to-one correspondence between nucleotides and phenotype is irrelevant in applied mathematical probability. What is critical is to examine whether or not the effect is random or specific. For example, if someone bet number 7 on a roulette wheel with 37 slots, the chance of winning on a single spin is one in 37. If someone bet 10 times on a roulette wheel, the likelihood of winning 8 out of 10 times is about one chance in 82 billion. Whether he places all 10 bets on number 7 or uses a different number with each spin...the odds are the same. The parameter to measure is his success (winning). By the same token, if mutations sequentially create similar outcomes (such as an increasingly curved retina), the probability of those mutations is measured by effect. It is irrelevant what mutations allegedly produce those changes.

4. Challenge #2: Molecular Convergences

Several studies have been published which document the existence of molecular convergences in multiple species. For example, the genes that encode for the protein prestin (essential in echolocation) are similar in echolocating bats and dolphins. The three auditory genes of echolocation are composed of about 20,000 specific nucleotide bases [5] [6]. These genes are believed to have evolved separately at least three independent times, to result in nearly identical nucleotide sequences in isolated species.

The existence of these convergences cannot be attributed to selection of random mutations. It is mathematically certain that a random process cannot result in identical outcomes if the number of possible outcomes is too great. This forms the basis of cybersecurity and protects billions of dollars from theft. This

identical principle is also used in courts of law to determine if a suspect's DNA links him or her to a crime scene. The only biologic reason why a person's DNA cannot match the DNA of someone else is the mathematical improbability that it will match. That calculated improbability is used as a gold standard to prove guilt or innocence. In considering the origin of molecular convergences, the creation of just 100 identical nucleotide sequences cannot occur in specific genes in different lineages by chance. This is because the number of combinatorial possibilities of 100 nucleotide sequences is 1.61×10^{60} .

A similar comparison could be made in determining if a student has plagiarized an essay written by someone else. If the text is sufficiently similar, he is guilty. This judgement is determined by the extreme *ex post facto* improbability that his words will coincidentally match the words of someone else. This principle is universally understood and accepted without doing any mathematical calculations.

5. Challenge #3: The Co-Evolution of Complex Integrated Parts

Is it possible for a proverbial monkey to create meaningful text by random keystrokes? Consider the probabilities involved. There are about 10^{25} possible 100-character sentences in English. That is about one million times the number of grains of sand on earth. With that many possibilities, could random input of 100 keystrokes find one of those sentences? Of course, everyone knows it's impossible, but why? Because the number of possible random sequences of 100 letters is about 10^{142} . The reason a monkey cannot type any meaningful sentence in any language is that a random process cannot generate a functionally integrated outcome in another system. This is because the search space is always so vast. This is a self-evident mathematical fact that applies to all informational systems and is universally understood.

Consider Jack Whiting who was accused of murder in the previous example. Suppose that, in addition to discovering that his previous two wives had also fallen off cliffs, the prosecution produced evidence that he took out a large life insurance policy on the wife that just died. That would have functionally integrated meaning to the case, and would be used as further evidence that Jack was guilty. With each line of corroborating circumstantial evidence presented, the probability that all of these integrated factors are coincidental diminishes. In this manner, circumstantial evidence is used to convict or exonerate suspects by *ex post facto* improbability arguments.

Complex integrated parts are ubiquitous in biologic systems. The evolution of the eye includes much more than a curved retina. A system of integrated parts includes an iris that reacts to light, complex eye fluids, biochemical systems of vision, an optic nerve that needs to incrementally increase in size, a transparent lens with a specific microanatomy, occipital lobes of the brain to translate binary code into 3D images, a transparent cornea, tear glands and ducts, muscles to move the eye, and other anatomic endowments. This obstacle is often dismissed,

because it is argued that an eye could function with a very few of these integrated parts. This begs the question. How could random mutations affect specific genes to result in a functionally integrated outcome? In the previous calculation of eye evolution, the probability would need to include not just the likelihood of the origin of an increasingly curved retina, but of all functionally integrated parts. If just ten functionally integrated components of vision are considered, this could result in a calculated probability on the order of $10^{-14,000}$.

6. Challenge #4: The Nature of Coded Information

In recent years, the characteristics of DNA are being increasingly compared to the informational systems of binary code of computer software. The nucleotide sequences of DNA are characterized by overlapping messages, data compression and intricate algorithms, with long strings of nucleotides that control other sequences. Many sequences are polyfunctional in nature. Bill Gates has stated, “*DNA is like a computer program but far, far more advanced than any software ever created.*” [7]

The neo-Darwinian synthesis proposes that DNA can be incrementally improved. This presumes that single nucleotide substitutions can impart gradually improving functionality [8] [9]. With an understanding of the nature of coded information, this is clearly impossible. The meaning of a sentence cannot be changed gradually by random substitutions of individual letters. Since the large majority of mutations are considered neutral, multiple nucleotides must be changed before natural selection can act (cumulative selection). Some have suggested that multi-base pair mutations (deletions, inversions, frameshifts, etc.) could have occurred. This proposal only magnifies the improbability barriers. Given the size of the human genome and the probability of specific mutations, the changing of just 10 nucleotides of human DNA to a specific result (before natural selection can act) in a single birth is around one chance in 10^{99} . In perspective, this is about as likely as selecting a specific atom in the known universe, divided by a factor of about ten quintillion.

7. Conclusion

Many proposed evolutionary pathways focus exclusively on patterns of phenotypic change in successive generations. These arguments are identical to those used in the nineteenth century, and ignore the molecular basis for such changes. With our current understanding of DNA and the particulate nature of inheritance, proposed models of evolution should consider the improbability that specific mutations would need to occur to allow natural selection to act. With the understanding of ex post facto probability, most proposals of complex macroevolutionary change decisively fail under the presumption that mutations are random.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Bundell, S. and Thompson, B. (2022, January 19) Why Mutation Is Not as Random as We Thought. Nature Podcast.
- [2] Bailey, D.H. (2020, January 3) Do Probability Arguments Refute Evolution? Math Scholar.
- [3] Sanford, J.C. (2008) Genetic Entropy and the Mystery of the Genome. Third Edition, FMS Publications, Waterloo, 4.
- [4] Nilsson, D.-E. and Pelger, S. (1994) A Pessimistic Estimate of the Time Required for an Eye to Evolve. *Biological Sciences*, **256**, 53-58. <http://www.jstor.org/stable/49593>
<https://doi.org/10.1098/rspb.1994.0048>
- [5] Parker, J., Tsagkogeorga, G., Cotton, J., *et al.* (2013) Genome-Wide Signatures of Convergent Evolution in Echolocating Mammals. *Nature*, **502**, 228-231.
<https://doi.org/10.1038/nature12511>
- [6] Shen, Y.-Y., Liang, L., Li, G.-S., Murphy, R.W. and Zhang, Y.-P. (2012) Parallel Evolution of Auditory Genes for Echolocation in Bats and Toothed Whales. *PLOS Genetics*, **8**, e1002788. <https://doi.org/10.1371/journal.pgen.1002788>
- [7] Gates, B. (1995) The Road Ahead. Viking Penguin, New York.
- [8] Lee, A. (2011, October 5) 4 Beneficial Evolutionary Mutations that Humans Are Undergoing Right Now.
- [9] Scacheri, C.A. and Scacheri, P.C. (2015) Mutations in the Noncoding Genome. *Current Opinion in Pediatrics*, **27**, 659-664.
<https://doi.org/10.1097/MOP.0000000000000283>

Appendix

For more information on this subject, including a free downloadable copy of the book *Evolution: Mask of Science*, visit <https://www.maskofscience.com/>.