Code, Constraint, and Contingency

A Biological Case for Intelligent Design: DNA, the Built-In Algorithms of Evolution, the Probabilities of Life's Origin, and a Structured Counter-Debate

Prepared for Bocharmer · Independent Scholarship · October 2026 · Companion to "Lawful Matter and the Signature of Mind"

Abstract

This paper argues that biology exhibits three features best explained by intelligent design (ID). First, DNA is a digital information system whose code is near-optimal among an astronomically large space of alternatives and displays deep structural symmetries. Second, the evolutionary process is not a blind lottery but runs on built-in, algorithm-like machinery: error correction, robustness buffers, regulated and targeted variation, and biases toward simple, symmetric forms. These features make evolution stable and productive, and they look like engineered search strategies. Third, the probability of a minimal living system arising from non-living chemistry by undirected processes, when estimated by mainstream researchers themselves, is extraordinarily small. The paper quantifies the relevant permutation spaces (DNA sequences, protein sequences, and possible genetic codes) and compares them with the probabilistic resources of Earth and the observable universe. A final section stages an eight-round debate in which the strongest objections from evolutionary biology and origin-of-life chemistry are presented in full and answered. The conclusion separates what the evidence establishes (the information-richness and robustness of biological systems; the unsolved state of abiogenesis) from what remains contested (whether those facts require design rather than as-yet-undiscovered natural processes).

Keywords: genetic code optimality; sequence space; evolvability; facilitated variation; error correction; origin of life; RNA world; error threshold; irreducible complexity; specified complexity; simplicity bias; biological symmetry.

1. Introduction and Scope

The companion paper argued that the laws and constants of physics are finely tuned so that chemistry rich enough for life is possible. This paper moves up one level, from the stage to the performance: given that chemistry is possible, what does the actual structure of living systems tell us?

Three clarifications frame the argument.

- **"From nothing."** In biology, the question is not the origin of matter (addressed in the companion paper) but *abiogenesis*: the emergence of the first self-replicating, information-bearing living system from non-living chemistry. Section 4 treats that question.
- **Common descent is not the target.** Several leading design theorists, including Michael Behe, accept common ancestry. The argument here concerns the *mechanism and information content* of life, not the genealogy of species. It is compatible with an Earth roughly 4.5 billion years old and a long history of life.
- **"Inherent algorithms."** By this we mean regulated, rule-governed processes encoded in organisms themselves that shape how genetic variation arises, is filtered, and is preserved. These are well documented in peer-reviewed biology; the dispute is over their explanation.

Technical passages are followed by short "In plain terms" summaries. Where a claim comes from design-advocacy literature that has not passed mainstream peer review, the paper says so.

2. DNA as an Information System

2.1 The digital architecture of the genome

DNA stores information as a linear sequence of four bases (A, C, G, T), read in triplets called codons. There are $4^3 = 64$ codons, which map onto 20 amino acids plus a stop signal. The mapping, the genetic code, is nearly universal across life (Crick 1968; Koonin and Novozhilov 2009). Information flows from DNA to RNA to protein, and proteins in turn replicate, repair, and regulate DNA. Formally, the genome is a discrete-symbol channel with an alphabet of size 4, a block code of length 3, and an error-correcting decoding apparatus (Yockey 2005).

The Watson–Crick double helix (Watson and Crick 1953) is itself a symmetry: each strand is the reverse complement of the other, so the molecule carries two copies of its information, which is the physical basis of both replication and repair. A subtler symmetry, Chargaff's second parity rule, holds that *within a single strand* the count of A approximately equals T and G approximately equals C, across very diverse genomes (Rudner, Karkas and Chargaff 1968; Albrecht-Buehler 2006). The reason for this intra-strand symmetry is still debated.

2.2 The permutation spaces: how large is the space of possibilities?

The number of possible sequences grows exponentially with length. For a DNA sequence of length n there are 4^n possibilities; for a protein of length L there are 20^L. The table below compares several biological spaces with the probabilistic resources available to search them.

Quantity	Expression	Approximate size
DNA sequences, 100 base pairs	4^{100}	10^{60}
DNA sequences, one 1,000-bp gene	4^{1000}	10^{602}
Protein sequences, 150 residues (a typical domain)	20^{150}	10^{195}
Possible genetic codes (64 codons → 21 meanings)	21^{64}	10^{84}
E. coli genome (≈4.6 million bp)	4^(4.6×10^6)	10^(2.8 million)
Human genome (≈3.1 billion bp)	4^(3.1×10^9)	10^(1.9 billion)
Atoms in the observable universe	—	$\approx 10^{80}$
Prokaryotic cells alive on Earth today	Whitman et al. 1998	$\approx 4–6 \times 10^{30}$
Cells that have ever lived on Earth	Crockford et al. 2023	order of $10^{39}–10^{40}$
Dembski's "universal probability bound"	10^{80} particles × 10^{45} s^{-1} × 10^{25} s	10^{150}

The arithmetic is not in dispute. Even a single average protein domain lives in a space (10^{195}) vastly larger than the number of cells that have ever existed ($\approx 10^{40}$) or atoms in the universe ($\approx 10^{80}$). Evolution has sampled only an infinitesimal fraction of sequence space. The live question, addressed in Sections 4.4 and 5 (Round 1), is what *fraction* of that space is functional, and how sequences reach the functional regions.

> In plain terms: there are more possible versions of a single average-sized protein than there are atoms in the universe, by a factor of more than a googol. Life has explored only a tiny sliver of these possibilities. Either the useful ones are common enough to be found by trial and error, or something guided the search.

2.3 The genetic code is one in a million

Of the roughly 10^{84} conceivable genetic codes, the one life uses is remarkably good at minimizing the damage done by errors. Codons that differ by a single base usually code for the same amino acid or a chemically similar one, so most point mutations and mistranslations are harmless or mild. Freeland and Hurst (1998), comparing the natural code with a million random alternatives and accounting for the known biases in mutation and mistranslation, found it better than all but about one in a million. Itzkovitz and Alon (2007) found that the code is also nearly optimal for a second, independent purpose: allowing protein-coding sequences to carry additional overlapping signals (such as splicing and regulatory sites) without disrupting the protein. A code that is simultaneously near-optimal for error tolerance and for carrying parallel information is, on the design view, the signature of a multi-objective engineering solution.

The code's table also shows internal symmetries: the second codon position largely determines the chemical class of the amino acid (hydrophobic versus hydrophilic), and the third position is largely redundant ("wobble"), which functions as an error buffer.

3. The Built-In Algorithms of Evolution

Neo-Darwinism in its classical form treated mutation as random noise and natural selection as the sole creative filter. Twenty-first-century molecular biology has revealed a far richer picture: organisms contain elaborate, rule-governed systems that control the rate, location, and kind of genetic change. The design argument interprets these as *inherent algorithms* that keep evolution stable and productive.

3.1 Error correction: keeping the message intact

DNA is chemically fragile. Tomas Lindahl, who later shared the 2015 Nobel Prize in Chemistry, showed that spontaneous hydrolysis, oxidation, and deamination damage DNA continually, on the order of thousands to tens of thousands of lesions per cell per day (Lindahl 1993). Life survives because of a layered error-correction architecture:

- **Base selection** by DNA polymerase (error rate roughly 10^{-4} to 10^{-5}).
- **Proofreading** by a built-in 3′→5′ exonuclease (improves fidelity about 100-fold).
- **Mismatch repair** after replication (another 100- to 1,000-fold).
- **Dedicated repair pathways** for specific damage types: base excision, nucleotide excision, double-strand-break repair by homologous recombination and end joining.

The combined result is a germline mutation rate in humans of about 1.2×10^{-8} per base pair per generation (Kong et al. 2012), an overall fidelity improvement of roughly a million-fold over the raw chemistry. This is precisely what information theory requires: Eigen (1971) showed that a replicator can maintain a genome of length L only if its per-site error rate is below roughly 1/L. Large genomes are possible *only* because of this multi-layer error-correcting system.

> In plain terms: copying DNA is like copying a library by hand. Cells use spell-check, then a proofreader, then an editor, then specialist repair crews. Without all of them, long genomes like ours would decay into nonsense within a few generations.

3.2 Robustness and evolvability: stable yet adaptable

A system that never changes cannot adapt; a system that changes too freely falls apart. Living systems resolve this tension in a way that biologists describe as surprising. Andreas Wagner (2008) showed that robustness (tolerance of mutations) actually *promotes* evolvability: because many genotypes produce the same phenotype, populations can drift across large "neutral networks" in sequence

space without loss of function, and from different points on those networks they can access a wide variety of new phenotypes. Kirschner and Gerhart (1998; Gerhart and Kirschner 2007) proposed the theory of **facilitated variation**: organisms are built from conserved core processes (such as signaling pathways and cytoskeletal systems) linked by weak, easily rewired regulatory connections, with exploratory behavior and modular compartments. This architecture channels random mutation into a limited range of viable, potentially useful variants.

On the design interpretation, this is what an engineer would build: a modular, fault-tolerant platform with standardized interfaces, so that small changes in configuration produce functional new outputs rather than system crashes.

3.3 Regulated and targeted variation

Several mechanisms generate genetic change in a controlled way:

- **Stress-induced mutagenesis.** In bacteria under stress, the SOS and general stress responses switch on error-prone polymerases, transiently raising mutation rates when the current genome is failing (Galhardo, Hastings and Rosenberg 2007).
- **Somatic hypermutation.** In antibody-producing B cells, the enzyme AID deliberately introduces mutations at rates about a million times the background, *targeted* to the antibody variable regions, followed by selection for better binding (Di Noia and Neuberger 2007). This is a literal, regulated mutation-and-selection algorithm operating inside our bodies.
- **CRISPR-Cas immunity.** Bacteria and archaea insert fragments of viral DNA into their own genomes as a heritable memory of past infections. Koonin and Wolf (2009) describe this as a genuinely Lamarckian mechanism: acquired information is inherited.
- **Natural genetic engineering.** James Shapiro (2005; 2011) catalogs how cells cut, splice, and rearrange their own DNA using transposons, recombinases, and reverse transcriptases, often in response to environmental signals, and argues that genome change is better described as cellular "engineering" than as accident.

A further, more contested line of evidence concerns mutation bias. Monroe et al. (2022) reported in *Nature* that in the plant *Arabidopsis thaliana*, mutation rates are lower in essential genes and functionally constrained regions, suggesting that epigenetic features direct mutations away from where they would do most harm. That result has been forcefully challenged (Wang et al. 2023; see Round 4) and should be regarded as unsettled.

3.4 Evolution as a learning algorithm, and a bias toward symmetry

Computer scientists have formalized evolution as a kind of search or learning algorithm. Leslie Valiant (2009), a Turing Award laureate, modeled evolvability within computational learning theory and showed that only certain classes of functions can be efficiently "learned" by mutation and selection. Watson and Szathmáry (2016) argued that evolution, through the structure of development and gene-regulatory networks, can generalize from past selection in ways formally analogous to machine learning.

Most striking for the theme of symmetry, Johnston et al. (2022) showed in the *Proceedings of the National Academy of Sciences* that the genotype-to-phenotype maps of biology are strongly biased toward outputs with short algorithmic descriptions, a property called **simplicity bias** (see also Dingle, Camargo and Louis 2018). Because symmetric structures can be described more briefly, they are exponentially more likely to arise as variants. The authors use this to explain why protein complexes and RNA structures are far more symmetric than random expectation would suggest.

The design proponent draws two lessons. First, an algorithm that reliably finds functional, symmetric structures in an otherwise hopeless search space is doing work that requires information about the space. William Dembski and Robert Marks (2009), writing in *IEEE Transactions on Systems, Man, and Cybernetics*, call this **active information**: the measurable degree to which a search outperforms blind sampling, which must come from somewhere. Second, simplicity bias arises from the mathematical and physical structure of the genotype-to-phenotype map, which is ultimately grounded in the laws of chemistry discussed in the companion paper. The orderliness of biology may therefore be "front-loaded" into the physics.

> In plain terms: evolution does not search blindly. Living things carry systems that control when and where DNA changes, protect what matters, and steer variation toward workable, often symmetric, designs. Whether those systems were themselves produced by evolution or reflect built-in design is the heart of the debate.

4. The Probability of Life Arising from Non-Life

4.1 What must be explained

Even the simplest free-living cells are complex. The synthetic minimal bacterium JCVI-syn3.0, produced by the J. Craig Venter Institute, required 473 genes, and the functions of about a third of them were unknown at the time (Hutchison et al. 2016). Any first living system, however simpler, needed at minimum: (i) an information-bearing polymer, (ii) a means of copying it with adequate fidelity, (iii) a boundary (membrane), and (iv) an energy-harvesting metabolism, all coupled together.

4.2 Probability estimates from the literature

Hubert Yockey (1977; 2005), an information theorist, applied Shannon theory to the information content of proteins and concluded that spontaneous biogenesis was effectively impossible by chance. Fred Hoyle's well-known estimate of 1 in 10^40,000 for producing a cell's complement of enzymes appeared in popular books rather than peer-reviewed journals and rests on simplified assumptions.

The most notable estimate comes from a leading evolutionary biologist with no sympathy for design. Eugene Koonin of the U.S. National Center for Biotechnology Information, in a peer-reviewed paper in *Biology Direct* (2007; expanded in Koonin 2011), estimated that the chance emergence of a minimal system coupling replication and translation within our observable universe has a probability so small (his appendix gives a figure on the order of 10^{-1018}) that it would effectively never happen. Koonin's own solution was not design but an infinite multiverse under eternal inflation, in which even fantastically improbable events occur somewhere. Note that a figure of 10^{-1018} exceeds Dembski's universal bound of 10^{-150} by more than 860 orders of magnitude.

4.3 Specific chemical hurdles

- **Eigen's paradox.** Without accurate copying, a replicator cannot maintain a genome long enough to encode an accurate copying enzyme; without such an enzyme, accurate copying is impossible (Eigen 1971).
- **Homochirality.** Life uses only left-handed amino acids and right-handed sugars, but ordinary chemistry produces equal mixtures. How a single handedness arose is still debated (Blackmond 2010).
- **Nucleotide synthesis.** Ribose and nucleotides are unstable and their prebiotic syntheses tend to produce complex tars, which Steven Benner has called the "asphalt problem" (Benner, Kim and Carrigan 2012).
- **Metabolism-first difficulties.** Leslie Orgel (2008), a founder of origin-of-life research, argued that self-organizing metabolic cycles on the prebiotic Earth were implausible without enzymes. Robert Shapiro (2006) argued the reverse, that replicator-first scenarios are implausible. Each camp's critique of the other remains largely unanswered.

4.4 The rarity of function in sequence space

How rare are functional proteins? Douglas Axe (2004), in the *Journal of Molecular Biology*, used mutagenesis experiments on a 150-residue β-lactamase domain to estimate that only about 1 in 10^{77} sequences of that length would adopt a working enzyme fold. If accurate, functional folds would be essentially unreachable by unguided search starting from random sequences. (Axe is a design advocate; his estimate is contested; see Round 1.)

4.5 The time window

Liquid water may have been present on Earth by about 4.4 billion years ago (Wilde et al. 2001). Carbon of possibly biological origin occurs in a zircon dated to 4.1 billion years ago (Bell et al. 2015), and a recent phylogenetic analysis places the last universal common ancestor (LUCA) at about 4.2 billion years ago, already a complex prokaryote-grade organism (Moody et al. 2024). If correct, life appeared geologically quickly, leaving at most a few hundred million years for abiogenesis, and LUCA was already sophisticated.

> In plain terms: even leading mainstream scientists estimate that the chance of a minimal living system forming by random chemistry in our universe is effectively zero. Life also appeared on Earth very early. Researchers who reject design typically turn to the idea of an infinite number of universes, or to unknown chemical shortcuts, to bridge the gap.

5. Synergy and Symmetry in Living Systems

A central question is how life achieves such synergy and symmetry on Earth. Both are pervasive.

- **Interdependence at the core.** DNA needs proteins to be copied; proteins need DNA (via RNA) to be specified; the ribosome, which makes all proteins, is itself built of RNA and proteins. Translation requires about 20 aminoacyl-tRNA synthetases, each recognizing both its amino acid and its tRNA. Behe (1996) calls systems in which several parts must be present simultaneously for any function "irreducibly complex."
- **Endosymbiosis.** Mitochondria and chloroplasts descend from free-living bacteria engulfed by host cells (Sagan 1967). The eukaryotic cell is a cooperative partnership whose integration required extensive coordinated gene transfer.
- **Ecological synergy.** Photosynthesis and respiration form a planetary oxygen–carbon cycle; nitrogen fixers, decomposers and producers sustain each other in interlocking cycles.
- **Symmetry.** Symmetry appears at every scale: the double helix, palindromic regulatory sites, the dimeric and higher symmetric assemblies of most protein complexes, radial and bilateral body plans, and Fibonacci phyllotaxis in plants.

The design proponent interprets this layered integration as purposive engineering: an ecosystem of mutually dependent, mutually supporting subsystems, of the kind that in human experience arises only from planning.

6. Summary of the Positive Case

1. DNA is a digital code; the space of possible sequences and codes is astronomically large (Section 2.2).
2. The actual genetic code is near-optimal for multiple independent objectives (Section 2.3).

3. Organisms contain built-in algorithms (error correction, robustness, regulated and targeted variation, simplicity bias) that make evolution stable and productive (Section 3).
4. Mainstream estimates of unguided abiogenesis yield probabilities far below any reasonable probabilistic resource (Section 4).
5. Living systems display deep interdependence and symmetry at every scale (Section 5).
6. In human experience, systems with these properties (codes, error correction, efficient search, integrated modular design) arise from intelligence.
7. Therefore, intelligent design is a reasonable, and arguably the best, explanation of the information and organization of life.

7. The Counter-Case: A Structured Debate

The following exchange presents the strongest objections from the peer-reviewed literature in the voice of a **Critic**, with replies from a **Proponent**.

Round 1: "Probability calculations assume one-shot random assembly"

Critic: Big-number arguments compute the odds of assembling a modern protein or cell all at once, by pure chance. No biologist believes that happened. Evolution is *cumulative*: selection preserves partial improvements, so the search is incremental, not a lottery. And functional sequences are far more common than Axe claims. Keefe and Szostak (2001), in *Nature*, screened a library of random 80-residue proteins and found ATP-binding proteins at a frequency of about 1 in 10^{11}. Neme et al. (2017) found that random sequences expressed in *E. coli* frequently have biological effects. Axe's 10^{-77} figure measures how many sequences can perform one *particular* modern function in one *particular* fold, not how many can perform *some* useful function. Protein space is also traversed by duplication and divergence of existing folds (Ohno 1970), not by fresh random draws.

Proponent: Cumulative selection requires a replicator, so it cannot explain the origin of the first replicator; that is why the abiogenesis estimates in Section 4 matter, and why Koonin, not a design theorist, calculated such a small number. On functional rarity, Keefe and Szostak's ATP-binders were weak binders selected by an intelligent experimental protocol, not working enzymes, and the gap between 10^{-11} and 10^{-77} shows that the frequency of function depends steeply on how much function is required. Real cells need many coordinated, high-performance enzymes.

Round 2: "Origin-of-life chemistry is steadily closing the gap"

Critic: Since Miller (1953) produced amino acids from simple gases, prebiotic chemistry has advanced dramatically. Powner, Gerland and Sutherland (2009) found a plausible route to activated pyrimidine nucleotides that bypasses the free-ribose problem, and Patel et al. (2015) showed that a single "cyanosulfidic" network can produce precursors of nucleic acids, amino acids, and lipids together. Ribozymes have been evolved that replicate RNA (Lincoln and Joyce 2009) and copy long templates (Wochner et al. 2011; Attwater et al. 2018). Homochirality can be amplified by autocatalysis (Soai et al. 1995) and by crystallization on magnetic surfaces (Ozturk et al. 2023). Koonin's number assumes no such stepwise chemistry; it is a calculation of what *does not* happen, not a measure of what does. "We don't know yet" is not evidence of design.

Proponent: These are genuine achievements, and they are carried out by highly skilled chemists controlling purified reagents, sequence of steps, and conditions, which itself illustrates the role of intelligence. No laboratory has yet produced a self-sustaining, evolving replicator from plausible prebiotic starting materials. Senior figures across the field, including Orgel (2008), Shapiro (2006), and Benner, have each argued that the leading rival scenarios face fundamental problems. The design argument does not rest on ignorance; it rests on the positive fact that codes and coupled information systems are, in all known cases, products of mind.

Critic: That last claim is circular: the only "codes" we know besides biology are human ones, so of course they are made by minds. Using that to infer that the biological code was also designed assumes the conclusion.

Proponent: It is an inductive inference from all known *causally established* cases, the same form of reasoning used in archaeology and in the search for extraterrestrial intelligence. It is defeasible, and a demonstrated natural origin of a code would defeat it.

Round 3: "The built-in algorithms evolved too"

Critic: The proponent treats evolvability machinery as evidence of foresight, but these systems are well explained by selection. Error correction is favored because organisms with higher fidelity leave more descendants. Stress-induced mutagenesis is favored because it raises variation only when the current genome is failing. Robustness and neutral networks arise naturally from the many-to-one structure of genotype-to-phenotype maps (Wagner 2008), and Watson and Szathmáry (2016) explain how evolution can come to "learn" without foresight. Above all, Johnston et al. (2022), the very paper cited for symmetry, explicitly argue that symmetry *emerges spontaneously* from the algorithmic nature of evolution, with no designer required. Dembski and Marks's "active information" has been criticized for assuming a uniform baseline search and for relabeling the structure of the fitness landscape as information (Häggström 2007).

Proponent: The critic is right that each mechanism can be given an adaptive story once a sophisticated replicating cell exists. But every such story presupposes the very system it seeks to explain: selection for better repair requires a cell that already repairs. The deeper point concerns the *landscape*. Johnston et al. show that simplicity and symmetry follow from the mathematics of genotype-to-phenotype mapping, which is set by chemistry and physics. That pushes the explanatory question back to why the physical and chemical landscape is so structured that efficient search succeeds, which is exactly where the companion paper's fine-tuning argument begins. Häggström's critique shows that landscape structure does the work; the design argument asks where that structure came from.

Round 4: "Mutations are random with respect to fitness"

Critic: Luria and Delbrück (1943) showed that bacterial mutations conferring resistance arise *before* exposure to the selective agent, not in response to it. Seventy years of experiments have confirmed that mutation is random with respect to its effect on fitness. Monroe et al.'s claim of adaptive mutation bias in *Arabidopsis* was reanalyzed by Wang, Ho, Hurst and Yang (2023), also in *Nature*, who concluded that the pattern was largely an artifact of sequencing and mapping errors. Somatic hypermutation and CRISPR are themselves products of evolution and operate within narrow domains.

Proponent: Granted on Monroe et al.; the paper does not rely on it. But "random with respect to fitness" at the level of individual mutations is fully compatible with highly structured variation at the level of the system: where mutations are permitted (targeted hypermutation), when they occur (stress responses), how they are buffered (the genetic code), and which phenotypes they can reach (facilitated variation). The point is that variation is *regulated*, not that individual mutations are foresighted.

Round 5: "Irreducible complexity has been answered"

Critic: Behe's showcase examples have received detailed evolutionary explanations. Pallen and Matzke (2006) showed that the bacterial flagellum is a family of diverse systems built from parts homologous to secretion systems and other machinery, a mosaic of co-opted components. Bridgham, Carroll and Thornton (2006) experimentally reconstructed how a hormone-receptor pair evolved by "molecular exploitation," in which a receptor existed before its partner hormone. Lenski's long-term experiment documented *E. coli* evolving the ability to use citrate aerobically, a new function arising through a multi-step historical sequence (Blount et al. 2008; 2012). Digital-organism experiments show complex features evolving through intermediate steps (Lenski et al. 2003). Gray et al. (2010) describe "constructive neutral evolution," a route by which irreducible

interdependence can arise without selection, as parts that were once optional become indispensable.

Proponent: These studies are valuable, but the gains they document are modest relative to the systems in question. The citrate innovation largely rewired an existing transporter gene to a new promoter (Blount et al. 2012) and took about 30,000 generations even in a large population. Durrett and Schmidt (2008), writing critically of Behe, nonetheless calculated that a coordinated pair of specific mutations would take on the order of 200 million years to appear in the human lineage. Weinreich et al. (2006) found that of 120 possible mutational paths to a five-mutation antibiotic-resistance enzyme, only a small minority were accessible to selection. Homology of flagellar parts establishes common origin, not a step-by-step pathway with selective advantage at each step.

Critic: Durrett and Schmidt's point was that Behe's required waiting times were vastly *overstated*, and Lynch (2005) showed that under realistic population-genetic parameters, Behe and Snoke's (2004) estimates shrink by orders of magnitude. Weinreich's result shows that evolution is constrained, not that it is impossible; some paths remain open, and evolution found them.

Round 6: "The genetic code's optimality has natural explanations"

Critic: Freeland and Hurst's "one in a million" also means that better codes exist; the code is good, not optimal. Its structure has several candidate natural explanations: chemical affinities between amino acids and codons (Yarus, Widmann and Knight 2009), co-evolution with the biosynthetic pathways of amino acids (Wong 1975), and Crick's (1968) "frozen accident," locked in because changing a code in use is lethal. Massey (2008) showed that substantial error-minimization can emerge as a neutral by-product of code expansion, without selection for it.

Proponent: Koonin and Novozhilov (2009) concluded that none of these theories is individually adequate and called the code's origin a "universal enigma." A code that is near-optimal for error minimization *and*, independently, for carrying overlapping information (Itzkovitz and Alon 2007) is harder to explain as a by-product of a single process.

Round 7: "Bad design and junk DNA"

Critic: If life were designed, why so much apparent inefficiency? Michael Lynch (2007) argued that much genomic complexity, including introns and mobile elements, reflects the weakness of selection in small populations rather than adaptive design. Comparative genomics suggests only about 8 percent of the human genome is under selective constraint (Rands et al. 2014), and Graur et al. (2013) sharply criticized ENCODE's claim that 80 percent of the genome is "functional." Pseudogenes, endogenous retroviruses shared across primates, and anatomical oddities such as

the detour of the recurrent laryngeal nerve are expected on common descent with modification, not on fresh design.

Proponent: ENCODE (2012) found biochemical activity across most of the genome, and while "activity" is not the same as function, the list of functional non-coding elements continues to grow (Kellis et al. 2014 discuss the competing definitions). Some cases once cited as bad design have turned out to be functional: Labin and Ribak (2010) found that the "backward" vertebrate retina has glial cells that act as optical fibers improving acuity. In any case, the design argument concerns the existence of sophisticated information systems, not that every feature is optimal; design does not require perfection, and a designed system can also degrade.

Round 8: "Design is not a testable scientific hypothesis"

Critic: Elliott Sober (2007), in the *Quarterly Review of Biology*, argued that ID makes no testable predictions because, unless something is specified about the designer's goals and abilities, a designer is compatible with any observation. Every gap filled by natural science shrinks ID's domain, a "god of the gaps" pattern. The scientific approach is to keep investigating mechanisms.

Proponent: Design theorists have offered predictions: that apparently "junk" DNA would prove largely functional, that functional protein folds would be rare, and that origin-of-life research would not find a simple undirected path to a coded replicator. These can turn out false. The design inference is also not a substitute for mechanistic research but a claim about what kind of cause best explains information-rich systems.

Critic: Those "predictions" are vague enough to survive almost any outcome, and the junk-DNA prediction appears to be faring poorly given the constraint data. Common descent, by contrast, made precise, risky predictions, such as the nested pattern of shared pseudogenes and a formally testable universal common ancestry (Theobald 2010), that were confirmed.

Debate scorecard

Issue	Strongest critic source	Strongest proponent source	Assessment
Probability of proteins	Keefe & Szostak 2001; Neme et al. 2017	Axe 2004	Depends on function required; estimates span 60+ orders of magnitude
Origin of life	Powner et al. 2009; Patel et al. 2015	Koonin 2007; Orgel 2008; Shapiro 2006	Genuine progress; no full pathway yet; strongest point for design
Evolvability machinery	Wagner 2008; Watson & Szathmáry 2016	Kirschner & Gerhart; Dembski & Marks 2009	Mechanisms real; origin explanation contested
Mutation randomness	Luria & Delbrück 1943; Wang et al. 2023	Di Noia & Neuberger 2007; Shapiro 2011	Random per mutation; regulated as a system
Irreducible complexity	Pallen & Matzke 2006;	Behe & Snoke 2004;	Specific cases partly

Issue	Strongest critic source	Strongest proponent source	Assessment
	Bridgham et al. 2006; Gray et al. 2010	Weinreich et al. 2006	answered; general rate questions open
Genetic code	Massey 2008; Yarus et al. 2009	Freeland & Hurst 1998; Itzkovitz & Alon 2007	Near-optimal; origin unresolved
Junk DNA / bad design	Lynch 2007; Rands et al. 2014; Graur et al. 2013	ENCODE 2012; Labin & Ribak 2010	Weight of current evidence favors critic
Testability	Sober 2007; Theobald 2010	Design-theorist predictions	Favors critic on predictive precision

8. Assessment and Conclusion

The evidence supports conclusions of different strength.

First, life is an information system of extraordinary sophistication. The genetic code is near-optimal for multiple objectives; cells run layered error correction; and evolution operates through regulated, buffered, and biased variation that makes it far more productive than blind search. These facts are established in mainstream peer-reviewed biology and are not in dispute.

Second, the origin of life remains unexplained. This is the design argument's strongest ground. Prebiotic chemistry has made real progress, but no complete undirected pathway to a coded, self-replicating system exists, and some of the most respected researchers in the field, Koonin among them, have judged chance origin within our universe to be effectively impossible. The live alternatives are design, a multiverse, or as-yet-unknown self-organizing chemistry.

Third, once replicating life exists, evolutionary biology has strong, detailed answers to many specific design arguments. The flagellum, hormone receptors, citrate use, and the genetic code all have published evolutionary accounts of varying completeness. The design reading of the built-in algorithms is coherent but does not, by itself, rule out their having evolved; its force depends on the deeper question of why the chemical and physical landscape is structured so that such search succeeds.

The strongest form of the biological design argument is therefore not that evolution cannot work, but that evolution works *remarkably well* because of information, error correction, and landscape structure that are themselves in need of explanation, and that the origin of the first information-bearing replicator lies, on present evidence, beyond the reach of known undirected processes. That argument connects naturally with the fine-tuning argument of the companion paper. Whether it ultimately succeeds depends on future origin-of-life research and on philosophical judgments about inference to design, on which informed scholars continue to disagree.

References

Albrecht-Buehler, G. (2006). Asymptotically increasing compliance of genomes with Chargaff's second parity rules through inversions and inverted transpositions. *PNAS* 103(47), 17828–17833. https://doi.org/10.1073/pnas.0605553103

Attwater, J., Raguram, A., Morgunov, A. S., Gianni, E., & Holliger, P. (2018). Ribozyme-catalysed RNA synthesis using triplet building blocks. *eLife* 7, e35255. https://doi.org/10.7554/eLife.35255

Axe, D. D. (2004). Estimating the prevalence of protein sequences adopting functional enzyme folds. *Journal of Molecular Biology* 341(5), 1295–1315. https://doi.org/10.1016/j.jmb.2004.06.058

Behe, M. J. (1996). *Darwin's Black Box: The Biochemical Challenge to Evolution.* New York: Free Press. (Not peer-reviewed.) https://archive.org/search?query=Darwin%27s+Black+Box+Behe

Behe, M. J., & Snoke, D. W. (2004). Simulating evolution by gene duplication of protein features that require multiple amino acid residues. *Protein Science* 13(10), 2651–2664. https://doi.org/10.1110/ps.04802904

Bell, E. A., Boehnke, P., Harrison, T. M., & Mao, W. L. (2015). Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. *PNAS* 112(47), 14518–14521. https://doi.org/10.1073/pnas.1517557112

Benner, S. A., Kim, H.-J., & Carrigan, M. A. (2012). Asphalt, water, and the prebiotic synthesis of ribose, ribonucleosides, and RNA. *Accounts of Chemical Research* 45(12), 2025–2034. https://doi.org/10.1021/ar200332w

Blackmond, D. G. (2010). The origin of biological homochirality. *Cold Spring Harbor Perspectives in Biology* 2(5), a002147. https://doi.org/10.1101/cshperspect.a002147

Blount, Z. D., Borland, C. Z., & Lenski, R. E. (2008). Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. *PNAS* 105(23), 7899–7906. https://doi.org/10.1073/pnas.0803151105

Blount, Z. D., Barrick, J. E., Davidson, C. J., & Lenski, R. E. (2012). Genomic analysis of a key innovation in an experimental Escherichia coli population. *Nature* 489, 513–518. https://doi.org/10.1038/nature11514

Bridgham, J. T., Carroll, S. M., & Thornton, J. W. (2006). Evolution of hormone-receptor complexity by molecular exploitation. *Science* 312(5770), 97–101. https://doi.org/10.1126/science.1123348

Crick, F. H. C. (1968). The origin of the genetic code. *Journal of Molecular Biology* 38(3), 367–379. https://doi.org/10.1016/0022-2836(68)90392-6

Crockford, P. W., Bar-On, Y. M., Ward, L. M., Milo, R., & Halevy, I. (2023). The geologic history of primary productivity. *Current Biology* 33(21), 4741–4750. https://doi.org/10.1016/j.cub.2023.09.040

Dembski, W. A., & Marks, R. J. (2009). Conservation of information in search: measuring the cost of success. *IEEE Transactions on Systems, Man, and Cybernetics A* 39(5), 1051–1061. https://doi.org/10.1109/TSMCA.2009.2025027

Di Noia, J. M., & Neuberger, M. S. (2007). Molecular mechanisms of antibody somatic hypermutation. *Annual Review of Biochemistry* 76, 1–22. https://doi.org/10.1146/annurev.biochem.76.061705.090740

Dingle, K., Camargo, C. Q., & Louis, A. A. (2018). Input–output maps are strongly biased towards simple outputs. *Nature Communications* 9, 761. https://doi.org/10.1038/s41467-018-03101-6

Durrett, R., & Schmidt, D. (2008). Waiting for two mutations: with applications to regulatory sequence evolution and the limits of Darwinian evolution. *Genetics* 180(3), 1501–1509. https://doi.org/10.1534/genetics.107.082610

Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58, 465–523. https://doi.org/10.1007/BF00623322

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74. https://doi.org/10.1038/nature11247

Freeland, S. J., & Hurst, L. D. (1998). The genetic code is one in a million. *Journal of Molecular Evolution* 47(3), 238–248. https://doi.org/10.1007/PL00006381

Galhardo, R. S., Hastings, P. J., & Rosenberg, S. M. (2007). Mutation as a stress response and the regulation of evolvability. *Critical Reviews in Biochemistry and Molecular Biology* 42(5), 399–435. https://doi.org/10.1080/10409230701648502

Gerhart, J., & Kirschner, M. (2007). The theory of facilitated variation. *PNAS* 104(suppl. 1), 8582–8589. https://doi.org/10.1073/pnas.0701035104

Graur, D., Zheng, Y., Price, N., Azevedo, R. B. R., Zufall, R. A., & Elhaik, E. (2013). On the immortality of television sets: "function" in the human genome according to the evolution-free gospel of ENCODE. *Genome Biology and Evolution* 5(3), 578–590. https://doi.org/10.1093/gbe/evt028

Gray, M. W., Lukeš, J., Archibald, J. M., Keeling, P. J., & Doolittle, W. F. (2010). Irremediable complexity? *Science* 330(6006), 920–921. https://doi.org/10.1126/science.1198594

Häggström, O. (2007). Intelligent design and the NFL theorems. *Synthese* 154(2), 217–230. https://philpapers.org/s/H%C3%A4ggstr%C3%B6m%20Intelligent%20design%20and%20the%20NFL%20theorems

Hutchison, C. A., et al. (2016). Design and synthesis of a minimal bacterial genome. *Science* 351(6280), aad6253. https://doi.org/10.1126/science.aad6253

Itzkovitz, S., & Alon, U. (2007). The genetic code is nearly optimal for allowing additional information within protein-coding sequences. *Genome Research* 17(4), 405–412. https://doi.org/10.1101/gr.5987307

Johnston, I. G., Dingle, K., Greenbury, S. F., Camargo, C. Q., Doye, J. P. K., Ahnert, S. E., & Louis, A. A. (2022). Symmetry and simplicity spontaneously emerge from the algorithmic nature of evolution. *PNAS* 119(11), e2113883119. https://doi.org/10.1073/pnas.2113883119

Keefe, A. D., & Szostak, J. W. (2001). Functional proteins from a random-sequence library. *Nature* 410, 715–718. https://doi.org/10.1038/35070613

Kellis, M., et al. (2014). Defining functional DNA elements in the human genome. *PNAS* 111(17), 6131–6138. https://doi.org/10.1073/pnas.1318948111

Kirschner, M., & Gerhart, J. (1998). Evolvability. *PNAS* 95(15), 8420–8427. https://doi.org/10.1073/pnas.95.15.8420

Kong, A., et al. (2012). Rate of de novo mutations and the importance of father's age to disease risk. *Nature* 488, 471–475. https://doi.org/10.1038/nature11396

Koonin, E. V. (2007). The cosmological model of eternal inflation and the transition from chance to biological evolution in the history of life. *Biology Direct* 2, 15. https://doi.org/10.1186/1745-6150-2-15

Koonin, E. V. (2011). *The Logic of Chance: The Nature and Origin of Biological Evolution*. Upper Saddle River, NJ: FT Press. https://archive.org/search?query=Logic+of+Chance+Koonin

Koonin, E. V., & Novozhilov, A. S. (2009). Origin and evolution of the genetic code: the universal enigma. *IUBMB Life* 61(2), 99–111. https://doi.org/10.1002/iub.146

Koonin, E. V., & Wolf, Y. I. (2009). Is evolution Darwinian or/and Lamarckian? *Biology Direct* 4, 42. https://doi.org/10.1186/1745-6150-4-42

Labin, A. M., & Ribak, E. N. (2010). Retinal glial cells enhance human vision acuity. *Physical Review Letters* 104, 158102. https://doi.org/10.1103/PhysRevLett.104.158102

Lenski, R. E., Ofria, C., Pennock, R. T., & Adami, C. (2003). The evolutionary origin of complex features. *Nature* 423, 139–144. https://doi.org/10.1038/nature01568

Lincoln, T. A., & Joyce, G. F. (2009). Self-sustained replication of an RNA enzyme. *Science* 323(5918), 1229–1232. https://doi.org/10.1126/science.1167856

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature* 362, 709–715. https://doi.org/10.1038/362709a0

Luria, S. E., & Delbrück, M. (1943). Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28(6), 491–511. https://doi.org/10.1093/genetics/28.6.491

Lynch, M. (2005). Simple evolutionary pathways to complex proteins. *Protein Science* 14(9), 2217–2225. https://doi.org/10.1110/ps.041171805

Lynch, M. (2007). The frailty of adaptive hypotheses for the origins of organismal complexity. *PNAS* 104(suppl. 1), 8597–8604. https://doi.org/10.1073/pnas.0702207104

Massey, S. E. (2008). A neutral origin for error minimization in the genetic code. *Journal of Molecular Evolution* 67(5), 510–516. https://doi.org/10.1007/s00239-008-9167-4

Miller, S. L. (1953). A production of amino acids under possible primitive Earth conditions. *Science* 117(3046), 528–529. https://doi.org/10.1126/science.117.3046.528

Monroe, J. G., et al. (2022). Mutation bias reflects natural selection in Arabidopsis thaliana. *Nature* 602, 101–105. https://doi.org/10.1038/s41586-021-04269-6

Moody, E. R. R., et al. (2024). The nature of the last universal common ancestor and its impact on the early Earth system. *Nature Ecology & Evolution* 8, 1654–1666. https://doi.org/10.1038/s41559-024-02461-1

Neme, R., Amador, C., Yildirim, B., McConnell, E., & Tautz, D. (2017). Random sequences are an abundant source of bioactive RNAs or peptides. *Nature Ecology & Evolution* 1, 0217. https://doi.org/10.1038/s41559-017-0127

Ohno, S. (1970). *Evolution by Gene Duplication*. Berlin: Springer. https://doi.org/10.1007/978-3-642-86659-3

Orgel, L. E. (2008). The implausibility of metabolic cycles on the prebiotic Earth. *PLoS Biology* 6(1), e18. https://doi.org/10.1371/journal.pbio.0060018

Ozturk, S. F., et al. (2023). Origin of biological homochirality by crystallization of an RNA precursor on a magnetic surface. *Science Advances* 9(23), eadg8274. https://doi.org/10.1126/sciadv.adg8274

Pallen, M. J., & Matzke, N. J. (2006). From The Origin of Species to the origin of bacterial flagella. *Nature Reviews Microbiology* 4, 784–790. https://doi.org/10.1038/nrmicro1493

Patel, B. H., Percivalle, C., Ritson, D. J., Duffy, C. D., & Sutherland, J. D. (2015). Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nature Chemistry* 7, 301–307. https://doi.org/10.1038/nchem.2202

Powner, M. W., Gerland, B., & Sutherland, J. D. (2009). Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* 459, 239–242. https://doi.org/10.1038/nature08013

Rands, C. M., Meader, S., Ponting, C. P., & Lunter, G. (2014). 8.2% of the human genome is constrained: variation in rates of turnover across functional element classes in the human lineage. *PLoS Genetics* 10(7), e1004525. https://doi.org/10.1371/journal.pgen.1004525

Rudner, R., Karkas, J. D., & Chargaff, E. (1968). Separation of B. subtilis DNA into complementary strands. III. Direct analysis. *PNAS* 60(3), 921–922. https://doi.org/10.1073/pnas.60.3.921

Sagan, L. (1967). On the origin of mitosing cells. *Journal of Theoretical Biology* 14(3), 255–274. https://doi.org/10.1016/0022-5193(67)90079-3

Shapiro, J. A. (2005). A 21st century view of evolution: genome system architecture, repetitive DNA, and natural genetic engineering. *Gene* 345(1), 91–100. https://doi.org/10.1016/j.gene.2004.11.020

Shapiro, J. A. (2011). *Evolution: A View from the 21st Century*. Upper Saddle River, NJ: FT Press. https://archive.org/search?query=Evolution+A+View+from+the+21st+Century+Shapiro

Shapiro, R. (2006). Small molecule interactions were central to the origin of life. *Quarterly Review of Biology* 81(2), 105–126. https://doi.org/10.1086/506024

Soai, K., Shibata, T., Morioka, H., & Choji, K. (1995). Asymmetric autocatalysis and amplification of enantiomeric excess of a chiral molecule. *Nature* 378, 767–768. https://doi.org/10.1038/378767a0

Sober, E. (2007). What is wrong with intelligent design? *Quarterly Review of Biology* 82(1), 3–8. https://doi.org/10.1086/511656

Theobald, D. L. (2010). A formal test of the theory of universal common ancestry. *Nature* 465, 219–222. https://doi.org/10.1038/nature09014

Valiant, L. G. (2009). Evolvability. *Journal of the ACM* 56(1), 3. https://doi.org/10.1145/1462153.1462156

Wagner, A. (2008). Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society B* 275(1630), 91–100. https://doi.org/10.1098/rspb.2007.1137

Wang, L., Ho, A. T., Hurst, L. D., & Yang, S. (2023). Re-evaluating evidence for adaptive mutation rate variation. *Nature* 619, E52–E56. https://doi.org/10.1038/s41586-023-06314-y

Watson, J. D., & Crick, F. H. C. (1953). Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 171, 737–738. https://doi.org/10.1038/171737a0

Watson, R. A., & Szathmáry, E. (2016). How can evolution learn? *Trends in Ecology & Evolution* 31(2), 147–157. https://doi.org/10.1016/j.tree.2015.11.009

Weinreich, D. M., Delaney, N. F., DePristo, M. A., & Hartl, D. L. (2006). Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312(5770), 111–114. https://doi.org/10.1126/science.1123539

Whitman, W. B., Coleman, D. C., & Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *PNAS* 95(12), 6578–6583. https://doi.org/10.1073/pnas.95.12.6578

Wilde, S. A., Valley, J. W., Peck, W. H., & Graham, C. M. (2001). Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. *Nature* 409, 175–178. https://doi.org/10.1038/35051550

Wochner, A., Attwater, J., Coulson, A., & Holliger, P. (2011). Ribozyme-catalyzed transcription of an active ribozyme. *Science* 332(6026), 209–212. https://doi.org/10.1126/science.1200752

Wong, J. T.-F. (1975). A co-evolution theory of the genetic code. *PNAS* 72(5), 1909–1912. https://doi.org/10.1073/pnas.72.5.1909

Yarus, M., Widmann, J. J., & Knight, R. (2009). RNA–amino acid binding: a stereochemical era for the genetic code. *Journal of Molecular Evolution* 69(5), 406–429. https://doi.org/10.1007/s00239-009-9270-1

Yockey, H. P. (1977). A calculation of the probability of spontaneous biogenesis by information theory. *Journal of Theoretical Biology* 67(3), 377–398. https://doi.org/10.1016/0022-5193(77)90044-3

Yockey, H. P. (2005). *Information Theory, Evolution, and the Origin of Life*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511546433