Unmasking the Genome: The Collapse of the Junk-DNA Paradigm

A Creation–Design Review of Failed Expectations, Predictive Priority, and the Functional Noncoding Genome

D. Budinsky [1], *M. Nailor* [1]
Independent Researchers

Truth in Research • Volume 1 • Issue 1 • 2026 • ISSN 3155-6302

ABSTRACT

For much of the late twentieth century, a powerful version of the junk-DNA thesis portrayed most non-protein-coding DNA as phenotypically inert residue: neutral sequence, selfish sequence, broken genes, and fossil viral insertions carried by genomes without organism-level purpose. That thesis was not merely a nickname. It was a research expectation generated and reinforced by influential evolutionary arguments. This review tests that expectation against the subsequent record. We distinguish biochemical activity, causal biological function, and selected-effect function; therefore, we do not equate ENCODE's 80.4% biochemical coverage with proof that 80.4% of nucleotides are indispensable. Even under that stricter standard, however, the strong junk paradigm has failed. Direct perturbation and mechanistic studies now assign regulatory, architectural, developmental, immune, reproductive, and disease-related roles to noncoding RNAs, pseudogenes, satellite repeats, transposable elements, and endogenous retroviral sequences. The failure had a measurable scientific cost: repetitive regions were routinely masked from genome-wide analyses, pseudogenes were excluded from functional screens, and clinical genetics concentrated overwhelmingly on protein-coding sequence. The historical prediction record is equally important. Long before large-scale genome annotation, intelligent-design and creation advocates placed an unusually broad, explicit, theory-derived prediction on record: DNA labeled "junk" should commonly prove functional, and investigators should search it for purpose. Forrest Mims made the point in 1994, William Dembski stated the design versus evolution contrast in 1998, and Linda Walkup and Todd Wood extended it in 1999–2000 to mobile and repetitive elements, including endogenous retroviral sequences. The modern evidence establishes that the confident strong-junk expectation was empirically unproductive, that design advocates won a genuine directional prediction, and that future research should begin with disciplined functional agnosticism rather than a presumption of genomic waste.

Keywords: *junk DNA, noncoding DNA, ENCODE, endogenous retroviruses, pseudogenes, long noncoding RNA, genome function, intelligent design*

1. Introduction

The phrase "junk DNA" compressed a scientific judgment into two words. Protein-coding exons occupy only a small fraction of the human genome, and the remainder was widely described as disposable debris, selfish sequence, or the fossil record of unguided genomic history. In its strongest form, the claim was not simply that some DNA lacks function. Every realistic model creationist, design-based, or evolutionary, can accommodate damaged, parasitic, redundant, or presently dispensable sequences. The strong claim was that most noncoding DNA should be expected to lack organism-level function because undirected evolutionary processes readily accumulate it.

That expectation failed.

The modern genome is not a protein-coding archipelago surrounded by an inert sea. It is a cell-type-specific, developmentally staged, three-dimensional regulatory system. Noncoding sequences help determine when genes are used, where chromosomes fold, how centromeres assemble, how stem-cell states are maintained, how placentas fuse, how innate immunity is triggered, and where disease-causing variation resides. Many of the most striking examples occur in the very categories once offered as paradigmatic junk: repetitive DNA, transposable elements, pseudogenes, and endogenous retroviruses (ERVs).

This paper advances four claims.

First, the strong junk-DNA thesis was an influential evolutionary expectation, not a caricature invented by critics. King and Jukes (1969), Ohno (1972), Doolittle and Sapienza (1980), and Orgel and Crick (1980) supplied explicit statements of phenotypically inert, selfish, or parasitic DNA.

Second, its empirical collapse is not rescued by pointing out that biochemical activity is not identical to organismal function. That distinction is correct and essential. Yet direct genetic perturbation, developmental experiments, disease association, biochemical mechanism, and chromosome architecture studies now demonstrate causal functions across categories once dismissed wholesale.

Third, the paradigm imposed a cost on science. The label licensed gene-centric methods that masked repeats, omitted pseudogenes, and underanalyzed noncoding clinical variants. Technology also contributed to these blind spots, especially the difficulty of assembling and mapping repetitive sequences. But a technical obstacle becomes a scientific setback when a theory supplies permission to ignore the obstacle instead of treating it as a frontier.

Fourth, creation and intelligent-design advocates made a real prediction. They did not discover every relevant mechanism first, and they were not the first scientists ever to suspect regulatory roles in repetitive DNA. Their priority is more precise, and more defensible: within the modern contest of origins frameworks, design advocates publicly predicted that broad function would be found in DNA then labeled junk and explicitly recommended searching for it. That is predictive priority, not discovery priority. The distinction strengthens rather than weakens the case.

2. Method and Scope

This article is a structured historical and predictive review, not a formal systematic meta-analysis. It compares: (1) primary historical statements about junk, selfish, and repetitive DNA; (2) publicly dated creation and design predictions; (3) genome-wide biochemical maps; (4) direct perturbation and mechanistic studies; and (5) reviews documenting methodological blind spots. Sources were selected to represent both sides of the functional debate. Accordingly, ENCODE's expansive biochemical results are presented alongside evolutionary-constraint estimates and published criticism of ENCODE's language (ENCODE Project Consortium, 2012; Graur et al., 2013; Kellis et al., 2014; Rands et al., 2014).

The paper addresses the human genome most often but uses mammalian and model-organism evidence where a mechanism is experimentally clearer. "Evolutionary community" refers to historically influential programs and expectations, not to every evolutionary biologist. Evolutionary biology has never been monolithic. The existence of early dissenters is part of the record and is discussed below.

3. Definitions: Noncoding Is Not Junk, and Activity Is Not Enough

Three distinctions prevent confusion.

3.1 Noncoding DNA

Noncoding DNA is DNA that does not encode the amino-acid sequence of a protein. It includes introns, untranslated regions, enhancers, silencers, promoters, origins of replication, structural repeats, noncoding-RNA genes, transposable elements, and many sequences whose roles remain unknown. “Noncoding” is a biochemical description. It is not a verdict of uselessness.

3.2 Junk DNA

In this paper, “junk DNA” means sequence asserted to lack a selected or organism-level role, not merely a sequence whose function has not yet been demonstrated. This definition separates legitimate uncertainty from a negative ontological claim. Unknown is not the same category as useless.

3.3 Three senses of function

Genome research commonly uses at least three senses of function (Kellis et al., 2014):

- Biochemical activity: a sequence is transcribed, bound by a protein, chemically modified, or located in accessible chromatin.
- Causal-role function: changing the sequence or its activity changes a cellular or organismal outcome.
- Selected-effect function: a sequence is maintained because natural selection preserves its contribution to fitness.

These measures overlap, but they are not interchangeable. Transcription can be noisy. A binding event can be incidental. A knockout can miss a conditional, redundant, developmental, silent, epigenetic or stress-specific role. Conservation can miss lineage-specific functions or functions maintained at the level of spacing, composition, or structure rather than exact bases. A serious assessment therefore integrates evidence rather than treating any single assay as an infallible oracle.

This review rejects two opposite errors. It rejects the claim that every transcribed nucleotide is thereby proven essential. It also rejects the claim that lack of present evidence, lack of deep conservation, or lack of phenotype in one laboratory condition proves nonfunction.

4. The Historical Claim at Full Strength

The strong junk thesis arose from real theoretical arguments. King and Jukes (1969) reasoned that either approximately 99% of mammalian DNA was not *“true genetic material”* in the sense of transmitting phenotype-affecting mutational change, or estimates of gene number were grossly wrong. Ohno (1972) popularized “junk DNA” while arguing from a mutational load that only a limited number of loci could carry indispensable information. These were not merely observations that most DNA does not encode protein. They were inferences that most of the genome could not carry phenotype-relevant function.

The argument sharpened in 1980. Doolittle and Sapienza (1980) proposed that selection within genomes would inevitably generate DNA whose only function was its own survival and that no

phenotypic or evolutionary function need be assigned to many transposable and repetitive sequences. Orgel and Crick (1980) described “selfish DNA” as an ultimate genomic parasite. These papers were scientifically serious and historically influential. They supplied a positive evolutionary rationale for expecting large quantities of organismally useless sequences.

The historical point is that evolutionary neutralist and selfish-DNA frameworks made extensive nonfunction unsurprising and, for many categories, expected. When functions later appeared, evolutionary explanations were expanded and retrofitted to accommodate them. Accommodation, however, is not the same as having made a better prior prediction.

Table 1. Competing expectations in the historical record

Date and source	Public expectation	Predictive significance
King and Jukes (1969)	Approximately 99% of mammalian DNA may lack phenotype-affecting genetic function	Strong nonfunction expectation
Ohno (1972)	Mutational-load reasoning supports a small functional fraction and abundant “junk”	Popularized the genome-as-junk framework
Doolittle & Sapienza (1980); Orgel & Crick (1980)	Repetitive/mobile DNA may persist for its own replication without organism-level function	Explicit selfish/parasitic expectation
Mims (1994)	Apparent filler may perform timing or control functions analogous to NOPs in software	Archival design analogy; rejected letter, not peer-reviewed
Dembski (1998)	Evolution expects much useless DNA; design expects DNA, as much as possible, to exhibit function	Broad, explicit design-versus-evolution prediction
Wood (1999); Walkup (2000)	Mobile and repetitive elements were designed for regulated variation and should show site bias and trait association	Specific creation-model predictions before modern functional genomics

5. Predictive Priority: What Intelligent Design Said Before the Genomic Revolution

5.1 The historical claim

Within the modern origins debate, intelligent-design and creation advocates were first to place a broad, explicit, theory-derived prediction of widespread function in so-called junk DNA on public record while the dominant junk and selfish-DNA narratives still treated extensive nonfunction as expected.

This is not retroactive wordplay. The record is dated.

5.2 Mims and Dembski: search where “junk” discourages inquiry

In a December 1994 letter rejected by Science, christian creationist and electronics writer and designer Forrest Mims compared supposed filler DNA with NOP instructions in digital controllers. A NOP does not calculate a result, yet a sequence of NOPs can regulate timing. Mims suggested that the junk-DNA puzzle might be solved more rapidly if investigators with design experience entered molecular biology (Mims, 1994). Because the letter was rejected but later archived on the author’s website, it is evidence of closed minded prior reasoning from the evolutionary community who already had made up their mind about the topic.

Ph.D. creationist William Dembski made the comparison explicit in a published 1998 essay. He argued that an evolutionary patchwork view expects much useless DNA, whereas design expects

DNA, *"as much as possible,"* to exhibit function. His operational conclusion was direct: design encourages scientists to look for function where evolutionary expectations discourage it (Dembski, 1998). Whatever one thinks of Dembski's larger program, this is an unambiguous, risky, and temporally prior prediction.

5.3 Walkup and Wood: mobile elements as designed systems of variation

The most specific creationist prediction came from Linda Walkup and Todd Wood. At Baraminology '99, held at Liberty University on 5–7 August 1999, Wood presented an unpublished proposal titled "The AGEing Process: Post-Flood Intrabaraminic Diversification Caused by Altruistic Genetic Elements (AGEs)" (Wood, 1999). Walkup's 2000 Journal of Creation review described and extended the hypothesis. She argued that transposable and repetitive elements could have been designed to move or recombine in order to generate rapid variation within created kinds, with some copies later damaged or silenced (Walkup, 2000).

Walkup did not stop at a general expectation of function. She listed empirical requirements. Candidate AGEs should show insertion-site specificity or evidence of former specificity; horizontal transfer into germline DNA should occur; and mobile or repetitive elements should be associated preferentially with genes affecting size, morphology, coloration, and specialized environmental adaptation rather than housekeeping genes. She also predicted that the distribution of such elements could conflict with conventional phylogenetic expectations. Wood later published a fuller model in Origins (Wood, 2002).

These claims include speculative biblical young earth historical components that have not been independently established. Yet the functional direction of the prediction is unmistakable: mobile and repetitive DNA should be investigated as regulated, trait-linked, potentially adaptive genomic machinery, not presumed parasitic litter. Modern research has confirmed site preferences, horizontal transfer, stress-responsive activation, regulatory network wiring, and trait-associated effects across many transposable-element systems (Chuong et al., 2017; Rebollo et al., 2012). Taxon-restricted ERV families and lineage-specific regulatory deployments also broadly match the expectation of group-specific use, though they do not by themselves identify a biblical "kind."

5.4 A prediction won is still a prediction won

Evolutionary biologists can and have retrofitted many functional discoveries through exaptation or co-option. But a later mechanism does not erase the earlier failed prediction record. The relevant question is not *"Can a flexible theory that can accommodate anything tell a story after the result?*" It is *"Which framework told investigators, before the result, where to look and what general class of outcome to expect?*" On that test, design advocates earned genuine success down the line.

6. ENCODE: A Real Discovery, a Real Controversy, and the Wrong Escape Route

The ENCODE Project Consortium (2012) reported that 80.4% of the human genome participates in at least one measured RNA or chromatin associated biochemical event in at least one tested cell type. This finding demolished the picture of noncoding DNA as uniformly silent. It did not prove that 80.4% of bases are indispensable, selected, or organismally functional.

Critics correctly objected when biochemical activity was rhetorically converted into essential function (Graur et al., 2013). Comparative-genomic estimates have been much smaller. Rands et al. (2014), for example, estimated that 8.2% of the human genome 7.1% to 9.2% is presently constrained by

purifying selection, while only 2.2% retained constraint in both human and mouse. Kellis et al. (2014) explained why biochemical, evolutionary, and genetic approaches measure different dimensions of function.

These numbers are not a referendum with one winner. They answer different questions.

- ENCODE asks where reproducible biochemical events occur across sampled cells and conditions.
- Constraint analyses ask which bases bear a detectable history of selection under a model and evolutionary time window.
- Perturbation experiments ask what happens when a sequence, transcript, or interaction is disrupted in a particular context.

The junk-DNA defense fails when it treats the limits of one measure as proof of global uselessness. Constraint is powerful but can miss recently acquired, lineage-specific, weakly selected, redundant, or structure-dependent roles. Knockouts are powerful but can miss conditional and distributed functions. Biochemical maps are sensitive but can include incidental activity. A mature science triangulates.

7. Functional Evidence in the Former Junk Drawer

7.1 Long noncoding RNAs: context-specific control

Long noncoding RNAs (lncRNAs) regulate transcription, chromatin state, nuclear organization, RNA processing, and development. Their functions are often cell-type-specific, which makes them easy to miss in bulk assays or under a single culture condition. In a CRISPR-interference screen of more than 16,000 lncRNA loci, Liu et al. (2017) identified 499 loci required for robust cellular growth; 89% of the functional hits acted in only one cell type. That pattern is precisely what a context dependent regulatory system predicts and precisely what a one condition dispensability test can overlook.

Not every annotated lncRNA is functional. The important result is that the category “noncoding transcript” cannot be equated with transcriptional noise. Mechanism and context must be tested.

7.2 Pseudogenes: the danger of naming before testing

Pseudogenes were defined as defective copies and were commonly treated as evolutionary relics. Some undoubtedly are damaged or dispensable. Others produce regulatory RNAs, small interfering RNAs, antisense transcripts, or competing endogenous RNAs. The PTENP1 pseudogene transcript can regulate the tumor suppressor PTEN through microRNA competition (Poliseno et al., 2010), while mouse pseudogene transcripts generate endogenous small interfering RNAs in oocytes (Tam et al., 2008).

The methodological consequence is documented. Cheetham, Faulkner, and Dinger (2020) noted that pseudogenes have typically been excluded from genome-wide functional screens and expression analyses and warned that pseudogene-annotated regions risk premature dismissal. Here the paradigm becomes experimentally self-sealing: label a sequence defective, exclude it from the screen, and later cite the absence of functional evidence as confirmation of the label.

7.3 Satellite and heterochromatic repeats: structure, segregation, and genome stability

Repetitive DNA at centromeres was once difficult to sequence and easy to classify as monotonous filler. Yet centromeric repeats participate in kinetochore assembly and chromosome segregation. In

Drosophila, depletion of repetitive centromeric satellite RNA disrupted kinetochore formation and cell division (Rošić et al., 2014). Such function may be collective: thousands of repeated units can create a chromosomal environment even when no single copy is individually indispensable. The telomere-to-telomere human genome assembly closed most of the remaining gaps in the reference sequence and added nearly 200 million base pairs, much of it repetitive (Nurk et al., 2022). The delay was heavily technological, short reads are intrinsically poor at resolving long repeats. It would be unfair to blame the junk concept alone. But the episode exposes the danger of conflating “hard to assemble” with “unimportant to understand.”

7.4 Repeats and transposable elements: distributed regulatory networks

Standard annotations often state that roughly half of the human genome is repeat-derived, while more sensitive estimates place the fraction at 66%–69% (de Koning et al., 2011). These sequences are not one uniform entity. They include SINEs, LINEs, DNA transposons, LTR elements, and ERV-derived fragments with diverse ages and states. Mechanistic studies show that transposable elements have rewired regulatory circuits. Kunarso et al. (2010) found that transposable elements contributed up to 25% of binding sites for key transcription factors in human and mouse embryonic stem cells and recruited new genes into the pluripotency network. Jacques, Jeyakani, and Bourque (2013) found that transposable elements account for nearly half of open chromatin regions and the majority of primate-specific regulatory elements in their analysis. These results do not make every transposon beneficial. They establish that the category cannot be dismissed as functionless merely because its members can copy, move, or cause disease.

7.5 Endogenous retroviruses: from “genomic fossils” to regulatory modules

ERV-derived sequence comprises roughly 8% of the human genome. Most loci are fragmented and incapable of producing an infectious virus, but infectious capacity is not the relevant test of host function. ERV-derived promoters, enhancers, proteins, boundary elements, and RNAs can serve the host at multiple biological levels.

Placental development

In 2000, the same year Walkup’s creationist review appeared, Mi et al. (2000) characterized syncytin, an ERV-envelope-derived protein expressed in human placenta that promotes trophoblast fusion. Subsequent work has established syncytins as central examples of ERV-derived proteins in mammalian placentation. Evolutionary accounts interpret these genes as independent viral captures and co-options; a design account interprets their highly coordinated deployment as evidence of intended biological function. Both accounts recognize the same decisive fact: sequence long treated as viral residue can be indispensable to reproduction.

Promoters and transcriptional initiation

Conley, Piriyapongsa, and Jordan (2008) identified 51,197 ERV-derived promoter sequences that initiate transcription in the human genome, including 1,743 located in proximal promoters or 5′ untranslated regions. Their paired-end transcript evidence supported 114 ERV-derived transcription start sites driving 97 human genes. This result is substantial, but it should not be inflated into the inaccurate claim that ERVs provide promoters for more than one-fifth of all human genes. Precision makes the argument stronger.

Innate immunity

Chuong, Elde, and Feschotte (2016) showed that MER41 elements supplied interferon-inducible regulatory sites to a network of human immune genes. Deleting a MER41 element near AIM2

impaired the gene's interferon response. This is causal evidence, not mere transcription. ERV-derived restriction factors also defend hosts directly. The mouse Fv1 protein, derived from retroviral gag sequence, blocks infection by susceptible retroviruses (Yap et al., 2014). A review of antiviral mechanisms concluded that ERVs, once disregarded as junk or dark matter, can regulate immune activation and directly target invading viral pathogens (Badarinarayan & Sauter, 2021).

Stem-cell identity and three-dimensional genome architecture

Kunarso et al. (2010) implicated repeat-derived binding sites in the core pluripotency network. Ito et al. (2017) systematically mapped hundreds of thousands of HERV-derived transcription-factor binding sites and identified thousands of shared regulatory elements associated with pluripotency, differentiation, CTCF occupancy, and immune regulation. Zhang et al. (2019) then demonstrated that transcriptionally active HERV-H elements demarcate topologically associating domains in human pluripotent stem cells; silencing or deleting selected HERV-H loci weakened or abolished nearby chromatin boundaries and changed gene expression. These findings justify the language now used even in conventional nomenclature reviews: ERVs are not merely junk (Gifford et al., 2018). They are heterogeneous. Some are damaged, some are harmful when misregulated, some may be neutral, and some are integrated components of development, immunity, and chromosome architecture. The scientifically responsible prior is therefore not *"viral-looking means useless."* It is "*test the locus, the cell state, and the mechanism."*

7.6 "Silent" substitutions: a parallel cautionary failure

Synonymous substitutions occur inside protein-coding DNA and are not junk DNA. They belong here as a cautionary parallel. Because they do not change the encoded amino acid, they were long called "*silent*." Yet synonymous changes can alter splicing, messenger-RNA stability and folding, translation rate, protein folding, and disease risk (Sauna & Kimchi-Sarfaty, 2011; Sharma et al., 2019). The lesson is methodological: a classification based on what a sequence does not obviously change can conceal what it actually controls.

Table 2. Direct evidence from categories once treated as junk or inert

Category	Demonstrated role	Evidence type	Representative source	Claim limit
lncRNA loci	Cell-type-specific growth and regulation	CRISPR interference	Liu et al. (2017)	Does not prove every lncRNA functional
Pseudogenes	Tumor-suppressor regulation; endogenous siRNA	Mechanism and perturbation	Poliseno et al. (2010); Tam et al. (2008)	Some pseudogenes remain damaged or dispensable
Satellite repeats	Kinetochore assembly and cell division	RNA depletion	Rošić et al. (2014)	Collective function need not make each repeat essential
Transposable elements	Pluripotency and species-specific regulatory wiring	Binding, enhancer assays, network analysis	Kunarso et al. (2010)	Many insertions can still be neutral or harmful
ERV-derived env	Placental cell fusion	Expression and fusion assays	Mi et al. (2000)	Origin interpretation remains contested
MER41 ERVs	Interferon-responsive immune regulation	Targeted deletion	Chuong et al. (2016)	Demonstrates selected loci, not all MER41 copies
HERV-H	Chromatin-domain boundaries in pluripotent cells	Silencing and deletion	Zhang et al. (2019)	Function is cell-state-specific
ERV-derived gag	Direct antiviral restriction	Infection experiments	Yap et al. (2014)	A specific exapted or designed factor, not a genome-wide count

8. Exactly How the Junk Paradigm Set Science Back

The charge that the junk paradigm hindered science can be made with evidence, not rhetoric. The cost can be stated in four documented mechanisms.

8.1 It normalized repeat masking

Repetitive sequence is technically difficult: short sequencing reads map ambiguously, and repeat-derived signals can confound gene-centered analyses. Masking repeats can therefore be a practical step. The problem is that a temporary computational convenience became a default biological filter. Slotkin (2018) documented that whole-genome studies commonly mask or discard repetitive regions, including transposable elements and ERVs, to focus on genes, and he urged researchers to stop narrowing their analyses to a minority of the genome. When the largest sequence classes are hidden before the analysis begins, their regulatory contributions cannot be discovered by that analysis.

8.2 It converted annotations into exclusions

“Pseudogene” often functioned as both a similarity annotation and a functional verdict. Cheetham et al. (2020) documented that pseudogenes were typically omitted from functional screens and expression analyses. This is a direct research cost: assays were designed in ways that systematically reduced the probability of finding the very functions under dispute.

8.3 It reinforced a protein-centric clinical pipeline

Clinical genetics historically concentrated on exons and canonical splice sites because their effects are easier to predict and test. Ellingford et al. (2022) reported that most clinical testing still focused almost exclusively on protein-coding regions even as penetrant disease mechanisms in noncoding regions accumulated. Whole-genome data often excluded noncoding variants or left them as variants of uncertain significance. The result is not an abstract philosophical loss. It is missed diagnoses, delayed mechanism discovery, and families left without answers.

8.4 It encouraged absence-of-knowledge reasoning

The most corrosive effect was epistemic. “No known function” was repeatedly shortened to “no function,” and the new label then reduced the incentive to investigate. This circularity is visible wherever an untested sequence is cited as evidence for its own uselessness. The problem was not that evolutionary biologists never studied noncoding DNA, many did, brilliantly. The problem was that an influential worldview made nonfunction the comfortable default and function the surprise.

The junk-DNA paradigm did not merely misname DNA; it misallocated scientific attention.

It told researchers which data to mask, which annotations to omit, which variants to deprioritize, and which discoveries to treat as exceptions. Technology, funding, and tractability also shaped those decisions, so the paradigm cannot be blamed for every delay. But the documented pattern is too specific to dismiss: repeats were filtered, pseudogenes were excluded, and noncoding clinical variants were sidelined. Science advanced fastest when investigators suspended the junk presumption and looked.

9. Evolutionary Rescue by Co-option: Valid Mechanism, Insufficient Defense

Evolutionary theory now explains many useful ERV and transposable-element functions through co-option or exaptation: an element enters or expands in a genome, and selection later preserves a

beneficial effect. These are empty phrases based on assertions. Individual cases are only ever supported by phylogenetic inference distribution and ancestral reconstruction. Like usual, after the fact they invoke different mechanisms to explain the very data that falsified them. On our end, the design model must explain both ordered use and real damage and we can do just that.

Co-option does not rescue the historical strong-junk DNA prediction for three reasons.

First, a mechanism proposed after functions are discovered does not change which framework forecasts the direction of discovery. Creationists Dembski, Walkup, and Wood urged researchers to expect function while selfish-DNA narratives still supplied a reason not to.

Second, “co-option” names a transition but does not explain its probability, timing, regulatory integration, or dependence on multiple cellular systems. An accurate or complete model must show how a sequence becomes expressed in the right cells, avoids destructive side effects, interacts with existing transcriptional and epigenetic machinery, and contributes to fitness. Just invoking co-option does none of these, it's just a filler word.

Third, recurrent functional recruitment across promoters, enhancers, boundary elements, immune circuits, placental proteins, and antiviral factors changes the prior probability assigned to untested elements. One is free to interpret these functions as evolved, designed, or some mixture of original purpose and subsequent degradation. But no responsible investigator should return to a default of uselessness.

The correct conclusion is therefore limited and forceful: modern evolutionary theory can accommodate functional noncoding DNA, but the historically influential junk expectation lost the prediction test. A theory’s flexibility after discovery must not be mistaken for prior success but reminded of its failure and asked; why can a theory get so many predictions wrong time and time again and never legitimately be falsified for good?

10. A Creation–Design Model for a Functional but Fallen Genome

A biblical creation framework expects the original genome to be purposeful, integrated, and capable of regulated variation. It also expects corruption, mutation, parasitism, and loss of function over time in a fallen world. The model therefore does not predict that every present-day nucleotide is pristine or indispensable. Walkup (2000) explicitly considered multiple possibilities: current designed function, designed elements subsequently damaged or silenced, and genuinely harmful sequence.

The design inference rests on system-level features: coordinated timing, modular reuse, information layering, robust control, and repeated incorporation of sequence classes into essential biological networks. These observations do not identify the Designer by molecular assay, and they do not by themselves establish a timescale. They are, however, consistent with a genome conceived as an engineered information system and were directionally anticipated by that framework.

To determine when this process occurred one needs to look no further than internal clocks of the mitochondria and Y chromosome. Pedigree substitution rate studies place a MRCA just thousands of years ago and the current Y chromosome studies that use high coverage scanning techniques do the same. To determine which Designer, one needs to look at the religious texts around the world and compare the claims with the science. The biblical model of ancestry stands alone and why it is Christians who are making the accurate testable predictions and not Hindus, Buddhists, or others.

11. Predictions for the Next Round of Testing

A productive creation–design model must risk even additional future falsifiable predictions. We believe a strong model should specify what investigators ought to find.

1. Conditional function will continue to emerge from untested noncoding loci. Effects should often appear only in particular cell types, developmental windows, stresses, infections, or environmental conditions.
2. Repetitive elements will prove to act in families and networks. Many functions will depend on copy number, spatial distribution, chromatin state, or collective transcription rather than on one indispensable repeat.
3. ERV-derived sequences will continue to supply promoters, enhancers, insulators, boundary elements, noncoding RNAs, and immune or developmental proteins. Locus-specific perturbation will be more informative than family-wide annotation.
4. Noncoding clinical analysis will raise diagnostic yield. Improved long-read sequencing, functional assays, and regulatory-variant interpretation should resolve cases missed by exon-centered pipelines.
5. Some alleged junk will prove genuinely damaged or harmful. A creation model predicts a mixed present genome and we discuss discriminating signatures of original function, subsequent corruption, and active genomic conflict. Some harmful ERVs we propose come from harmful foods in the past, that is why God gave dietary restrictions and laws to follow so we would be safe from such harm.
6. The Wood–Walkup AGE hypothesis should be tested at its own level of specificity. Within proposed created kinds, mobile-element distributions should correlate with trait variation in size, morphology, coloration, and environmental adaptation; insertion preferences and germline horizontal transfer should be demonstrable where the mechanism is invoked. Taxon-restricted elements alone are good but not enough, the predicted genotype-to-phenotype connection must be shown even further to validate their predictions even further.
7. Function will frequently be encoded above the single-base level. Spacing, composition, topology, redundancy, and three-dimensional organization should explain some sequences that appear weakly conserved or dispensable when tested in isolation.

These predictions make the current model falsifiable. If increasingly comprehensive, multi-condition, locus-specific perturbations repeatedly show that most noncoding sequence has no causal, structural, regulatory, or selected role; if repeat families remain biologically inert once technical limitations are removed; and if the AGE-specific trait associations fail, then the strong design expectation would be weakened.

Confidence invites testing, and we thrive on this and it is why our model is so strong and superior. For many more head to head predictions of creation vs evolution over the decades be sure to read our paper titled: Predictions on trial (Nailor, M., & Budinsky, D. 2026). The gold standard of science is accurate testable future predictions and the biblical creation model dominates this down the line going back decade after decade. The reason isn't because our model is better, it's because it is true.

13. Conclusion

The historical junk DNA thesis offered a clear picture: because genomes are products of undirected accumulation, large quantities of phenotypically inert, selfish, parasitic, or broken sequence should be expected. That picture shaped language, textbooks, computational pipelines, functional screens, and clinical priorities. It also proved to be a poor guide to discovery.

Modern genomics has not shown that every nucleotide is essential. It has shown something more consequential: noncoding and repetitive sequence cannot be dismissed by category. Pseudogenes regulate genes. Satellite transcripts help build kinetochores. Transposable elements wire stem-cell networks. ERVs provide placental proteins, immune enhancers, antiviral factors, promoters, and chromatin boundaries. Some noncoding variants cause disease, others transcription. Each finding pulls the term “junk” further out of the trash bin ignorance it was placed in by evolution theory.

Biblical creation and Intelligent-design advocates were the ones who were directionally right. Mims warned in 1994 that apparent filler could encode timing. He was rejected and ignored by peer review and ended up being correct. Dembski stated in 1998 that design predicts function and should motivate researchers to seek it. Wood and Walkup went further in 1999–2000, predicting regulated, trait-linked roles for mobile and repetitive elements and proposing concrete tests. They did not make every relevant discovery first, and not every detail of their models has been confirmed. But they recorded the better broad expectation before the genomic evidence matured.

The scientific setback regarding the genome as mostly junk can be seen in the way researchers approached the genome itself. Repetitive DNA was often filtered out, pseudogenes were frequently ignored, and medical genetics focused mainly on protein-coding regions. The idea of “junk DNA” made it easier to assume that much of the rest of the genome was unimportant and it was ignored for a long time. Progress accelerated when scientists finally found a function which they then began treating these neglected regions as possible sources of biological function rather than as useless genomic leftovers.

The DNA was not waiting to become functional. Science was waiting to stop calling the unknown useless based on evolutionary thinking.

For creationists, the lesson is not triumphalism to brag about but rather shows research courage and confidence and trying to make a difference. You see, we view disease as a consequence of a fallen world and something that will increase over time. This means we believe disease should be viewed and treated very differently than the evolutionary model does. We warn of a crumbling genome, a genome of genetic entropy where mutations build up over time that puts a shelf life on genomes. So while we believe and expect purpose without denying damage in junk DNA. We also believe we should test every claim at the locus, network, organism, and historical levels. State predictions before the results. For evolution, the junk-DNA paradigm failed its prediction test, while the design-guided expectation of extensive function was strikingly vindicated.

Now that we realize most of the genome is not junk at all, we can work on discovering its functions. But most importantly, what disease can be helped or cured from these different elements so we can utilize them in medicine and gene editing.

The future is not about creation vs evolution, it is about usefulness and accuracy. The biblical model of ancestry doesn't just answer the diversity of life better, it explains disease, purpose, function and diversity better all at the same time.

References

Badarinarayan, S. S., & Sauter, D. (2021). Switching sides: How endogenous retroviruses protect us from viral infections. Journal of Virology, 95(11), e02299-20. https://doi.org/10.1128/JVI.02299-20

Cheetham, S. W., Faulkner, G. J., & Dinger, M. E. (2020). Overcoming challenges and dogmas to understand the functions of pseudogenes. Nature Reviews Genetics, 21, 191–201. https://doi.org/10.1038/s41576-019-0196-1

Chuong, E. B., Elde, N. C., & Feschotte, C. (2016). Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science, 351(6277), 1083–1087. https://doi.org/10.1126/science.aad5497

Chuong, E. B., Elde, N. C., & Feschotte, C. (2017). Regulatory activities of transposable elements: From conflicts to benefits. Nature Reviews Genetics, 18, 71–86. https://doi.org/10.1038/nrg.2016.139

Conley, A. B., Piriyapongsa, J., & Jordan, I. K. (2008). Retroviral promoters in the human genome. Bioinformatics, 24(14), 1563–1567. https://doi.org/10.1093/bioinformatics/btn243

de Koning, A. P. J., Gu, W., Castoe, T. A., Batzer, M. A., & Pollock, D. D. (2011). Repetitive elements may comprise over two-thirds of the human genome. PLOS Genetics, 7(12), e1002384. https://doi.org/10.1371/journal.pgen.1002384

Dembski, W. A. (1998). Science and design. First Things, 86, 21–27. https://firstthings.com/science-and-design/

Dickel, D. E., Ypsilanti, A. R., Pla, R., et al. (2018). Ultraconserved enhancers are required for normal development. Cell, 172(3), 491–499.e15. https://doi.org/10.1016/j.cell.2017.12.017

Doolittle, W. F., & Sapienza, C. (1980). Selfish genes, the phenotype paradigm and genome evolution. Nature, 284, 601–603. https://doi.org/10.1038/284601a0

Ellingford, J. M., Ahn, J. W., Bagnall, R. D., et al. (2022). Recommendations for clinical interpretation of variants found in non-coding regions of the genome. Genome Medicine, 14, 73. https://doi.org/10.1186/s13073-022-01073-3

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

Gifford, R. J., Blomberg, J., Coffin, J. M., et al. (2018). Nomenclature for endogenous retrovirus (ERV) loci. Retrovirology, 15, 59. https://doi.org/10.1186/s12977-018-0442-1

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

Ito, J., Sugimoto, R., Nakaoka, H., et al. (2017). Systematic identification and characterization of regulatory elements derived from human endogenous retroviruses. PLOS Genetics, 13(7), e1006883. https://doi.org/10.1371/journal.pgen.1006883

Jacques, P.-É., Jeyakani, J., & Bourque, G. (2013). The majority of primate-specific regulatory sequences are derived from transposable elements. PLOS Genetics, 9(5), e1003504. https://doi.org/10.1371/journal.pgen.1003504

Kellis, M., Wold, B., Snyder, M. P., et al. (2014). Defining functional DNA elements in the human genome. Proceedings of the National Academy of Sciences, 111(17), 6131–6138. https://doi.org/10.1073/pnas.1318948111

King, J. L., & Jukes, T. H. (1969). Non-Darwinian evolution. Science, 164(3881), 788–798. https://doi.org/10.1126/science.164.3881.788

Kunarso, G., Chia, N.-Y., Jeyakani, J., et al. (2010). Transposable elements have rewired the core regulatory network of human embryonic stem cells. Nature Genetics, 42(7), 631–634. https://doi.org/10.1038/ng.600

Liu, S. J., Horlbeck, M. A., Cho, S. W., et al. (2017). CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. Science, 355(6320), eaah7111. https://doi.org/10.1126/science.aah7111

Mi, S., Lee, X., Li, X., et al. (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature, 403, 785–789. https://doi.org/10.1038/35001608

Mims, F. M., III. (1994). Hints of a language in junk DNA [Unpublished letter submitted to Science]. https://forrestmims.org/publications/

Nailor, M., & Budinsky, D. (2026). Predictions on trial: Comparing creation and evolutionary predictions across biology and development. Truth in Research, 1(1). https://doi.org/10.67690/p13n5z85

Nurk, S., Koren, S., Rhie, A., et al. (2022). The complete sequence of a human genome. Science, 376(6588), 44–53. https://doi.org/10.1126/science.abj6987

Ohno, S. (1972). So much “junk” DNA in our genome. Brookhaven Symposia in Biology, 23, 366–370.

Orgel, L. E., & Crick, F. H. C. (1980). Selfish DNA: The ultimate parasite. Nature, 284, 604–607. https://doi.org/10.1038/284604a0

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J., & Pandolfi, P. P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature, 465, 1033–1038. https://doi.org/10.1038/nature09144

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

Rebollo, R., Romanish, M. T., & Mager, D. L. (2012). Transposable elements: An abundant and natural source of regulatory sequences for host genes. Annual Review of Genetics, 46, 21–42. https://doi.org/10.1146/annurev-genet-110711-155621

Rošić, S., Köhler, F., & Erhardt, S. (2014). Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. Journal of Cell Biology, 207(3), 335–349. https://doi.org/10.1083/jcb.201404097

Sauna, Z. E., & Kimchi-Sarfaty, C. (2011). Understanding the contribution of synonymous mutations to human disease. Nature Reviews Genetics, 12(10), 683–691. https://doi.org/10.1038/nrg3051

Sharma, Y., Miladi, M., Dukare, S., et al. (2019). A pan-cancer analysis of synonymous mutations. Nature Communications, 10, 2569. https://doi.org/10.1038/s41467-019-10489-2

Slotkin, R. K. (2018). The case for not masking away repetitive DNA. Mobile DNA, 9, 15. https://doi.org/10.1186/s13100-018-0120-9

Sternberg, R. V. (2002). On the roles of repetitive DNA elements in the context of a unified genomic–epigenetic system. Annals of the New York Academy of Sciences, 981, 154–188. https://doi.org/10.1111/j.1749-6632.2002.tb04917.x

Tam, O. H., Aravin, A. A., Stein, P., et al. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature, 453, 534–538. https://doi.org/10.1038/nature06904

Walkup, L. K. (2000). “Junk” DNA: Evolutionary discards or God’s tools? Journal of Creation, 14(2), 18–30. https://creation.com/en/articles/junk-dna-evolutionary-discards-or-gods-tools

Wood, T. C. (1999, August 5–7). The AGEing process: Post-Flood intrabaraminic diversification caused by Altruistic Genetic Elements (AGEs) [Unpublished conference paper]. Baraminology ’99: Creation Biology for the 21st Century, Liberty University, Lynchburg, VA.

Wood, T. C. (2002). The AGEing process: Rapid post-Flood intrabaraminic diversification caused by Altruistic Genetic Elements (AGEs). Origins, 54, 5–34.

Yap, M. W., Colbeck, E., Ellis, S. A., & Stoye, J. P. (2014). Evolution of the retroviral restriction gene Fv1: Inhibition of non-MLV retroviruses. PLOS Pathogens, 10(3), e1003968. https://doi.org/10.1371/journal.ppat.1003968

Zhang, Y., Li, T., Preissl, S., et al. (2019). Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. Nature Genetics, 51(9), 1380–1388. https://doi.org/10.1038/s41588-019-0479-7