Compiled Chemistry

Structural Inheritance and the Deep Origins of Biological Function

Authors: Ashley M. Gjøvik, B.S., J.D. (ashleymgjovik@protonmail.com) | **Date**: June 23 2025

Abstract

The transition from prebiotic chemistry to genetic inheritance represents a fundamental gap in our understanding of life's origins. While origin-of-life research demonstrates that complex chemical systems can maintain functional organization through environmental energy sources, coherent frameworks explaining how these systems became encoded in DNA remain lacking.

This paper examines the Chemical-Genetic Integration Model (CGIM): a proposed mechanism whereby genetic systems emerged through selective incorporation of pre-existing, functionally organized chemical networks including iron-sulfur clusters, mineral scaffolds, and autocatalytic systems. The framework suggests this integration was incomplete and differential, with lineages potentially retaining varying degrees of pre-genetic functional architecture.

Analysis of early-diverging animals reveals organizational patterns consistent with preserved chemical inheritance. Ctenophores exhibit constitutive protein expression activated by environmental conditions rather than transcriptional regulation, resembling proposed pre-genetic chemical networks. Tardigrades deploy intrinsically disordered proteins that function through physical rather than regulatory mechanisms. Genomic analysis identifies cross-kingdom sequence conservation patterns in taxonomically problematic organisms that may represent molecular signatures of chemical inheritance.

The CGIM framework generates testable predictions for systematic identification of chemical inheritance in modern organisms, including quantitative criteria for distinguishing runtime from modular biological organization and genomic archaeology methods for detecting ancient chemical signatures. The approach provides a mechanistic hypothesis for the chemistry-to-genetics transition that can be evaluated through computational analysis, experimental validation, and systematic discovery protocols.

Keywords

Chemical-Genetic Integration Model; runtime biology; genomic archaeology; iron-sulfur clusters; intrinsically disordered proteins; orphan genes; ctenophores; tardigrades; stinkhorns, pre-genetic chemistry; chemical inheritance; origin of life; molecular fossils; cross-kingdom sequence conservation; environmental activation mechanisms; structural homology.

Citation

Gjøvik, Ashley M. "Compiled Chemistry: Structural Inheritance and the Deep Origins of Biological Function." *The Journal of Decolonized Ecology and Evolution* 1, no. 1 (June 23 2025). DOI 10.5281/zenodo.15717955

Table of Contents

Introduction

The emergence of genetic inheritance from prebiotic chemistry represents the most fundamental unsolved problem in biology. While decades of origin-of-life research have demonstrated that complex chemical systems can maintain functional organization through environmental energy sources, we lack a mechanistic framework for how these systems transitioned into DNA-based inheritance. This gap is not merely empirical but conceptual—leaving unexplained how early life acquired the sophisticated biochemical machinery evident in even the simplest modern cells.

The chemistry-to-genetics transition can be conceptualized through a computational analogy: early life operated like a "primordial operating system" where functional processes ran directly through environmental interactions—analogous to runtime execution in computer systems. Modern genetic inheritance resembles compiled code, where environmental solutions became encoded as heritable modules that could be deployed under genetic control. Understanding this transition requires identifying biological systems that preserve runtime-style operation alongside those that fully adopted modular genetic architectures.

The Chemistry-to-Genetics Gap

Current models treat the chemistry-to-genetics transition as a discrete phase change: chemical systems somehow "became" genetic systems. Yet this sequential replacement model fails to explain how genetic systems could encode the full complexity exhibited by early life. The metabolic networks, stress response systems, and structural proteins that appear universal across all domains of life suggest complex functional foundations that predate our earliest phylogenetic reconstructions.

Origin-of-life chemistry demonstrates that functional systems can emerge without genetic templates. Iron-sulfur mineral surfaces catalyze organic synthesis, maintain electron transport chains, and support complex metabolic networks through purely environmental energy sources. Autocatalytic reaction networks, phase-separated compartments, and mineral scaffolds all maintain organized behavior through chemical rather than informational control. Yet modern biology operates through fundamentally different organizational principles—genetic systems encode functional information, reproduce it with high fidelity, and deploy it through regulated expression networks.

Why Current Frameworks Fail

Traditional frameworks assume that chemical complexity preceded and was then replaced by genetic control. This sequential model provides no pathway for understanding how information-based inheritance could emerge from purely chemical systems, nor does it explain why certain biological features appear to operate through non-genetic mechanisms. More fundamentally, it cannot account for the apparent "sudden appearance" of sophisticated biological functions in the fossil record.

The Chemical-Genetic Integration Model

This paper proposes that the transition occurred not through replacement but through integration: genetic protocells selectively incorporated pre-existing, functionally organized chemical systems, subsequently encoding successful chemical solutions while retaining others in operational form. This Chemical-Genetic Integration Model (CGIM) suggests that modern organisms represent diverse outcomes of this integration process, with some lineages preserving substantial pre-genetic functional architecture.

Evidence for incomplete integration persists across early-diverging animal lineages, where we observe biological systems that operate through environmental activation rather than transcriptional control—what we term "runtime biology." These patterns offer direct windows into the chemistry-to-genetics transition, revealing how genetic inheritance originally emerged from environmental chemistry.

The CGIM provides both a mechanistic pathway from environmental chemistry to biological inheritance and a framework for identifying preserved evidence of this transition in modern organisms. Rather than being lost to deep time, the chemistry-to-genetics transition remains accessible to investigation through the organizational strategies of living systems.

The Chemical-Genetic Integration Model (CGIM)

Core Framework

The Chemical-Genetic Integration Model provides both a mechanistic explanation for the chemistry-to-genetics transition and a systematic computational framework for identifying its molecular signatures in modern organisms. Rather than purely theoretical reconstruction, the CGIM establishes quantitative criteria and genomic archaeology methods that enable ongoing discovery of chemical inheritance patterns across diverse biological systems.

The Integration Pathway

The earliest stages of life likely emerged as open-ended chemical systems operating within structured environmental matrices—hydrothermal vents, mineral scaffolds, and alkaline springs (Russell and Hall 1997; Martin and Russell 2007). In these contexts, molecular interactions were driven by physical proximity, ion gradients, and catalysis by transition metal centers such as Fe-S and Ni-Fe clusters, now preserved in modern metabolism (Camprubi et al. 2017; Wächtershäuser, 1998; Fuchs 2011).

These systems exhibited functional inheritance: the ability to sustain and propagate structure and metabolism without requiring genes. Functions such as redox cycling, molecular shielding, osmotic regulation, and primitive compartmentalization could occur via cofactors, short peptides, and membrane-free phase-separated assemblies (Hyman, Weber, and Jülicher 2014; Patel et al. 2017). This pre-genetic world was chemically intelligent, yet not informationally encoded.

It was in this environment that information encoding capabilities emerged. Recent work found that ferruginous chemical gardens mimicking prebiotic low-temperature alkaline hydrothermal vents were able to accumulate nucleic acids (Helmbrecht et al. 2023). In addition, research is advancing into understanding protocells as self-assembling, autonomous supramolecular systems (Sugawara et al. 2024).

The emergence of RNA, DNA, and ribosome-based protein synthesis allowed certain lineages to begin encoding and storing functional chemistry (Woese 1998; Gilbert 1986). However, this transition was gradual and non-uniform. This paper proposes that a fusion event occurred where primitive gene-based protocells captured and encoded functional chemical systems from their pre-genetic environment.

Critically, this fusion was incomplete. Not all systems were encoded. Some were preserved in biochemical form as proteins, peptides, or disordered domains, while others remained semi-autonomous, functioning as runtime components within the new genetic framework (Banani et al. 2017; Perdigão et al. 2015).

Runtime Biology vs. Modular Biology

A critical component of the CGIM framework is the Early Runtime Encoding Hypothesis, which proposes that certain lineages preserved substantial pre-genetic functional architecture through incomplete chemical-genetic integration. EREH predicts that early-diverging organisms should exhibit measurable runtime biology characteristics: constitutive protein expression, environmental activation mechanisms, and chemical rather than transcriptional control of core processes. This hypothesis provides specific, testable predictions that distinguish chemical inheritance from conventional genetic evolution.

Runtime Biology represents biological systems that operate through constitutive expression and environmental activation, resembling pre-genetic chemical networks that respond directly to local conditions. Modular Biology operates through conditional expression and transcriptional regulation, reflecting the evolution of sophisticated genetic control systems (Davidson and Erwin 2006).

Quantitative Distinction Criteria

Criterion	Runtime Biology	Modular Biology
Protein Expression Pattern	>70% of proteome expressed constitutively across tissues/stages	<30% of proteome expressed constitutively; most proteins are tissue/stage-specific
IDPs	>40% of proteome lacks stable 3D structure	<10% of proteome is intrinsically disordered
Activation Mechanism	Environmental triggers (pH, salinity, temperature, hydration state)	Transcriptional control via regulatory networks and signaling cascades
Gene Regulatory Integration	Weak correlation with canonical regulatory motifs; sparse transcription factor networks	Strong regulatory network connectivity; dense transcription factor interactions
Stress Response Strategy	Constitutive expression of protective proteins; chemical buffering systems	Inducible stress responses; conditional pathway activation
Evolutionary Strategy	Robustness through redundancy; generalized environmental tolerance	Efficiency through specialization; adaptive plasticity

These operational distinctions enable systematic computational analysis using advanced structural bioinformatics tools. AlphaFold2 predictions, protein fold classification systems, and comparative structural genomics provide quantitative methods for detecting chemical inheritance signatures. The systematic application of these computational approaches transforms chemical inheritance from theoretical speculation into empirically detectable molecular patterns that can be systematically identified across diverse biological systems.

Thermodynamic Feasibility

Complex chemical systems can maintain organization and functionality for extended periods through continuous energy input from geochemical processes. Modern serpentinizing systems demonstrate that chemical networks maintain functional organization for thousands of years. The Prony Bay (New Caledonia) hydrothermal field has maintained active serpentinization chemistry for over 40,000 years, continuously producing organized mineral structures and supporting complex microbial communities that utilize environmental chemistry directly (Colman et al. 2025).

Similarly, the Ney Springs (Northern California) terrestrial serpentinizing system, investigators found that microbial metabolisms in an extremely high pH marine-like terrestrial serpentinizing system maintain organized biochemical processes through direct coupling to environmental energy sources (Trutschel et al. 2022).

Iron-sulfur mineral surfaces catalyze organic synthesis reactions for decades when reducing conditions and thermal gradients are maintained, with energy densities of 10-100 kJ/mol driving chemical organization against entropy (Russell and Hall 1997).

Recent experimental work has quantified these energy requirements more precisely. Formose-type reactions can persist for weeks to months under continuous energy input from pH oscillations or redox cycling, with critical energy flux above approximately 1-10 kJ/mol/hour, readily available from hydrothermal temperature gradients or mineral-water redox couples (Haas and Nikel 2023).

Integration Mechanisms

The transition from pre-genetic chemical systems to genetically encoded organisms requires specific molecular mechanisms by which environmental chemistry becomes incorporated into protocells and subsequently encoded in DNA. Two integration pathways stand out as both geochemically plausible and molecularly tractable: iron-sulfur cluster incorporation and silica-based scaffolding systems.

Iron-Sulfur Integration Pathway

The iron-sulfur world hypothesis, first proposed by Günter Wächtershäuser, provides the most experimentally validated pathway for chemical-genetic integration (Wächtershäuser 1992). In this model, iron-sulfur minerals at hydrothermal vents serve as both catalytic surfaces and sources of functional cofactors for early protocells.

Recent computational simulations have demonstrated that some hydrophobic amino acids chelate FeS nanocrystals, producing three positive feedbacks: (i) an increase in catalytic surface area; (ii) partitioning of FeS nanocrystals to the membrane; and (iii) a proton-motive active site for carbon fixing that mimics the enzyme Ech (Lane and Martin 2012). This mechanism provides a direct pathway by which protocells could incorporate functional iron-sulfur chemistry from their environment, subsequently encoding the proteins necessary to maintain and utilize these catalytic centers.

The ubiquity of iron-sulfur clusters in modern metabolism—from electron transport chains to CO_2 fixation enzymes—supports the hypothesis that these ancient chemical systems were successfully integrated and encoded rather than evolved de novo (Camprubi et al. 2017). The [4Fe-4S] clusters found in ferredoxins maintain the same cubic geometry and electron transfer properties as iron-sulfur mineral surfaces. The enzyme aconitase contains a [4Fe-4S] cluster that functions through the same Lewis acid chemistry as iron-sulfur minerals catalyzing organic transformations.

Experimental Validation of Iron-Sulfur Integration

Recent experimental breakthroughs provide direct validation that environmental iron-sulfur chemistry can be captured and encoded into functional protein systems. Researchers have successfully devised an innovative protocol for synthesising mature Fe-S proteins, by bringing together a recombinant sulfur assimilation (SUF) system and an oxygen-scavenging system under controlled laboratory conditions (Wang et al. 2023).

The SUF system contains all the necessary machinery to produce [4Fe-4S] clusters and has a higher tolerance for oxygen when compared to the other pathways, providing a plausible model for how early protocells could have incorporated and eventually encoded iron-sulfur functionality from their chemical environment. Significantly, the requirement for an oxygen-scavenging system supports the model's emphasis on specific environmental conditions—early Earth's reducing atmosphere would have naturally provided the low-oxygen conditions necessary for iron-sulfur integration.

Silica Integration Pathway

Silica-based integration offers a complementary pathway particularly relevant to understanding early animal evolution. Serpentinization reactions create high-pH alkaline brines that can become easily enriched in silica and lead to formation of self-assembled nanocrystalline mineral composites, namely silica/carbonate biomorphs and metal silicate hydrate (MSH) tubular membranes (Saladino et al. 2016). These naturally occurring silica structures provide organized scaffolding systems that could serve as templates for protocellular organization.

The subsequent encoding of silica-processing machinery, exemplified by the enzyme silicatein in modern sponges, demonstrates how environmental silica chemistry was successfully captured and modularized. Silicatein forms a strand (axial filament) within the silica deposition vesicle of the sclerocyte, initiating spicule formation by inducing deposition of concentric silica layers around the filament (Müller et al. 2007). Despite being a protein, silicatein retains many characteristics of mineral-mediated silica polymerization, operating at ambient temperature and pressure using mechanisms that closely parallel geochemical processes (Müller et al. 2010).

Membrane-Mediated Selective Uptake

The membrane-mediated uptake mechanism appears crucial for both iron-sulfur and silica integration pathways. Protocells with fatty acid membranes exhibit selective permeability that allows uptake of small inorganic molecules while maintaining cellular integrity (Patel et al. 2017). For iron-sulfur systems, the partitioning of FeS nanocrystals to membrane boundaries creates localized zones of high catalytic activity, while for silica systems, the interaction between dissolved silicic acid and membrane-associated proteins provides a controlled environment for biosilicification.

This membrane-mediated integration explains how protocells could selectively incorporate beneficial chemical systems while avoiding potentially harmful environmental compounds. The process likely involved multiple cycles of uptake, retention, and eventual genetic encoding as successful chemical-genetic fusions outcompeted purely chemical or purely genetic systems.

Environmental Energy Requirements

The successful integration required specific environmental conditions that provided sustained energy flux above thermodynamic thresholds. Early Earth environments readily met these requirements through multiple energy sources: hydrothermal vents provided thermal gradients of 100-300°C across spatial scales of centimeters to meters; tidal cycles created regular pH and salinity oscillations; mineral-water interfaces maintained redox potentials sufficient to drive organic synthesis; and day-night temperature cycling provided regular energy input through thermal expansion and contraction.

Modern serpentinizing systems typically provide energy fluxes of 100-1000 kJ/mol/hour through water-rock interactions—orders of magnitude above the thresholds required for chemical organization. This energy abundance explains why complex chemical organization was thermodynamically feasible during early Earth conditions and why certain chemical solutions became universal in biology while others remained purely geochemical.

Molecular Validation: Discovery of Chemical Fossils

The Chemical-Genetic Integration Model makes a critical prediction: if chemical systems were encoded into genetic inheritance, this process should have left detectable genomic signatures that can be reverse-engineered from modern organisms. Recent genomic archaeology has successfully identified such signatures, providing direct molecular evidence that the chemistry-to-genetics transition actually occurred.

Hydra and Siphonophores: Runtime Biology in Action

Hydra and colonial siphonophores provide particularly clear demonstrations of runtime biology principles operating in modern organisms. *Hydra* exhibits remarkable regenerative capabilities through constitutive expression of morphogenetic proteins that respond directly to physical damage rather than requiring transcriptional activation. The organism can regenerate complete body plans from cellular aggregates, suggesting fundamental biological processes operate through environmental coupling rather than genetic programming.

Colonial siphonophores demonstrate how complex multicellular organization can emerge through chemical coordination rather than centralized genetic control. Individual zooids maintain functional integration through direct chemical communication, achieving sophisticated collective behaviors without hierarchical genetic regulation. These examples illustrate how runtime biology can support complex multicellular life while preserving the environmental responsiveness characteristic of pre-genetic chemical systems.

Tribrachia fusiformis: A Living Ediacaran Architecture

Extended genomic analysis of the organism traditionally classified as *Pseudocolus fusiformis* revealed fundamental evolutionary misclassification that obscures profound insights into chemical-genetic integration (Gjøvik, A. M., June 16 2025). Long-read BLAST alignments using extended recursive sequence analysis revealed high-confidence genomic matches spanning taxonomically distant basal eukaryotes and marine lineages, with alignments consistently exceeding 1750 base pairs, >85% identity, and E-values of zero.

The strongest molecular affinities occurred with:

- Demosponges (*Axinella* sp., 1,973 bp at 85.5% identity)
- Calcareous sponges (Clathrinida, 2,024 bp at 85.8% identity)
- Cnidarians (Scyphozoa, 1,984 bp at 85.2% identity)
- Unicellular holozoans (*Tunicaraptor unikontum*, 2,116 bp at 87.2% identity)

This cross-kingdom sequence conservation, combined with triradial symmetry, manganese-rich biomineralization, and scaffold-based developmental logic, establishes this organism as *Tribrachia fusiformis*—a living representative of Ediacaran-grade multicellular architecture.

The 87.2% sequence identity between *Tribrachia* and *Tunicaraptor unikontum*—a predatory flagellate basal to metazoans—provides direct molecular evidence. This level of conservation across such phylogenetic distance cannot be explained through conventional vertical inheritance or horizontal gene transfer. Instead, it reflects the preservation of ancient molecular machinery responsible for triradial scaffold organization that predates major eukaryotic divergences.

Independent Validation: Trichophyton schoenleinii

The chemical fossil hypothesis received independent validation through analysis of *Trichophyton schoenleinii*, which exhibits identical genomic patterns: extensive sequence similarity with early metazoan lineages, deep-sea physiological signatures, and runtime biology characteristics (Gjøvik, A. M., June 22 2025). This second example confirms that chemical inheritance represents a systematic rather than isolated phenomenon, validating the CGIM prediction that multiple lineages should preserve molecular evidence of ancient chemical integration.

The Ancient Marine Assemblage Pattern

The consistent reappearance of matches to the same ancient lineages across taxonomically distant "outlier" organisms suggests these represent genomic fossils of the chemical-genetic transition period. This assemblage—ctenophores, ancient cnidarians, basal sponges, tardigrades, choanoflagellates, and related protists—appears to represent the core toolkit of pre-genetic functional chemistry that was differentially retained across lineages.

Preliminary analysis of additional morphologically anomalous organisms reveals this same diagnostic pattern:

- High-confidence matches to the characteristic ancient marine assemblage.
- Poor integration within supposed taxonomic families.
- Morphological features inconsistent with current classification.
- Runtime biological characteristics (constitutive expression, environmental activation).

Architectural Persistence as Molecular Validation

Tribrachia's structural continuity with Ediacaran fossils including *Tribrachidium*, *Albumares*, and *Anfesta* provides temporal validation spanning more than 550 million years. This architectural persistence indicates that triradial organization represents a fundamental chemical solution that was successfully encoded and preserved.

Significantly, *Tribrachia* exhibits hallmark molecular signatures of runtime biology: constitutive protein expression, environmental activation mechanisms, and biomineralization patterns consistent with chemical garden formation. These characteristics provide molecular validation that ancient chemical logic persists within current genetic systems.

Cross-Kingdom Chemical Scaffold Inheritance

This pattern exemplifies the core CGIM prediction: chemical scaffolds provided functional templates that were independently encoded by multiple lineages accessing the same pre-genetic chemical toolkit. The consistent reappearance of architectural logic across sponges, cnidarians, and protists demonstrates that scaffold-organizing domains represent compiled solutions to physical constraints rather than lineage-specific innovations.

The organism's manganese-rich polysaccharide matrix, hollow chambered arms, and scaffold-driven emergence pattern align precisely with material constraints predicted for pre-genetic chemical systems. Rather than representing taxonomic anomalies requiring special explanation, these organisms provide direct molecular evidence that modern biological systems emerged through encoding pre-existing chemical functionality.

Genomic Archaeology Methods

The successful identification of *Tribrachia* and *Trichophyton* as chemical fossils establishes systematic criteria for molecular validation of chemical inheritance and provides methodology for discovering additional examples through genomic archaeology.

The genomic archaeology approach establishes systematic computational protocols for detecting chemical inheritance using advanced bioinformatics methods. This methodology transforms chemical-genetic integration from historical speculation into active discovery science, providing tools for identifying previously unrecognized examples of the chemistry-to-genetics transition preserved in modern genomes.

Recursive Sequence Alignment Strategy

The discovery framework employs a recursive strategy for deep homology discovery in genetic sequence alignment (Gjøvik, A. M., June 10 2025). Rather than relying on standard BLAST searches that may miss ancient signals, this approach uses extended sequence analysis with progressively refined parameters to detect high-confidence alignments that span taxonomically distant lineages.

Long-read BLAST alignments consistently exceeding 1750 base pairs with >85% identity and E-values of zero indicate shared molecular machinery that transcends conventional phylogenetic boundaries. These parameters distinguish chemical inheritance from both horizontal gene transfer (which typically involves shorter, more recent transfers) and deep vertical inheritance (which should show phylogenetic coherence).

Diagnostic Criteria for Chemical Fossils

The genomic archaeology framework identifies chemical fossils through five systematic criteria:

1. Cross-Kingdom Sequence Conservation: Identical molecular machinery appearing in taxonomically distant lineages with statistical significance that cannot be explained through conventional inheritance patterns.

2. Phylogeny-Structure Dissociation: High functional similarity despite divergent evolutionary placement, indicating shared access to pre-genetic chemical solutions rather than recent common ancestry.

3. Physical Constraint Signatures: Molecular architecture reflecting environmental chemistry constraints (metal coordination, phase separation, redox chemistry) rather than genetic optimization for specific biological functions.

4. Ancient Marine Assemblage Affinity: Consistent alignment patterns with the characteristic set of early-diverging marine lineages (ctenophores, basal sponges, ancient cnidarians, choanoflagellates, tardigrades, and related protists).

5. Runtime Expression Profiles: Constitutive deployment rather than conditional regulatory control, indicating preservation of pre-genetic operational logic.

Structural Homology Screening

Systematic structural analysis using AlphaFold2 and related computational tools reveals that proteins currently classified as "orphan genes" often display structural similarities to environmentally-constrained chemical solutions, even without sequence homology (Baek et al. 2021; Jumper et al. 2021).

The framework prioritizes analysis of:

- **Symmetry-based searches** for radial organizational patterns that reflect chemical rather than genetic constraints
- **Metal-binding architecture analysis** identifying coordination geometries consistent with iron-sulfur clusters, manganese catalysis, or silica scaffolding
- **Phase separation domains** that function through chemical thermodynamics rather than specific protein-protein interactions

Runtime Biology Expression Profiling

Chemical inheritance should manifest in constitutive expression patterns that differ fundamentally from transcriptionally regulated systems. The methodology identifies:

- **Constitutive deployment** (>70% of proteome expressed across tissues/stages)
- **Environmental activation mechanisms** (pH, salinity, temperature, hydration state triggers)
- **Sparse regulatory integration** (weak correlation with canonical transcription factor networks)

These patterns distinguish between encoded chemical systems (runtime biology) and evolved genetic modules (modular biology), as quantified by the operational criteria established earlier.

Taxonomic Reassessment Protocol

Standard taxonomic frameworks often misclassify organisms that retain substantial pre-genetic chemical systems. The genomic archaeology approach:

1. Identifies taxonomic outliers through poor fit within supposed families/orders combined with morphological anomalies inconsistent with current classification.

2. Tests for ancient marine signature using the recursive alignment strategy to detect the characteristic assemblage pattern.

3. Validates through multiple independent lines of evidence including structural analysis, expression profiling, and physiological characteristics.

4. Proposes reclassification based on genomic affinities rather than superficial morphological similarity.

Systematic Discovery Applications

This methodology transforms isolated taxonomic curiosities into systematic evidence for chemical inheritance. Rather than treating genomic anomalies as noise or contamination, the framework provides tools for:

- **Reanalyzing "contamination-filtered" genomes** to recover functionally relevant sequences previously excluded as bacterial or archaeal contamination
- **Identifying additional chemical fossils** through systematic application of the ancient marine assemblage screening
- **Validating chemical inheritance mechanisms** by demonstrating the predicted molecular signatures across multiple independent lineages

The approach has already identified multiple organisms exhibiting the diagnostic genomic signature, suggesting that chemical fossils are systematic rather than exceptional discoveries. Each new identification strengthens the evidence base for the Chemical-Genetic Integration Model while providing additional molecular access points to the chemistry-to-genetics transition.

Methodological Advantages

Unlike approaches that assume purely genetic explanations for biological complexity, genomic archaeology recognizes that functional inheritance can operate through non-sequence mechanisms. This perspective reveals molecular evidence that would otherwise be dismissed as taxonomic noise, horizontal contamination, or evolutionary anomaly.

The recursive and long alignment strategy specifically addresses the methodological bias in current genomic analysis practices that systematically exclude evidence for chemical-genetic integration. By removing taxonomic filtering constraints and focusing on functional rather than phylogenetic coherence, the approach recovers ancient signals that conventional methods miss. (Gjovik, A. M., June 10 2025).

Living Examples of Chemical Inheritance

Runtime Biology Lineages

If chemical-genetic integration occurred as proposed, this paper predicts finding organisms that preserve different stages of this fundamental transition. These lineages should exhibit biological systems that operate through environmental activation rather than purely transcriptional control—direct remnants of the chemistry-to-genetics transition.

Tardigrades: Precompiled Biochemical Defenses

Tardigrades (*Hypsibius dujardini*, *Ramazzottius varieornatus*) provide perhaps the clearest example of runtime biology in action. Their extreme resistance to desiccation, radiation, freezing, and vacuum exposure operates through mechanisms that exemplify pre-genetic chemical organization rather than conventional genetic regulation.

Tardigrade-Specific Intrinsically Disordered Proteins (TDPs)

The molecular basis of tardigrade resilience centers on a suite of unique proteins that exhibit hallmark runtime characteristics. Tardigrade-specific intrinsically disordered proteins (TDPs) are constitutively expressed at high levels or induced during desiccation in multiple tardigrade species, exhibiting precisely the expression pattern predicted for runtime biology (Boothby et al. 2017).

Critically, TDPs form non-crystalline amorphous solids (vitrify) upon desiccation, and this vitrified state mirrors their protective capabilities—a direct chemical response to environmental conditions rather than a genetically programmed protection strategy. This environmental specificity distinguishes chemical inheritance from generic physical properties: when researchers tested whether simple protein disorder alone could confer protection, they found that α-synuclein, a protein that exists as a disordered monomer in cells and has no known connection to stress tolerance, did not increase survival under drying conditions, demonstrating that something beyond intrinsic disorder of TDPs is essential for their protective capabilities (Boothby et al. 2017).

Dsup: A Chemical Shield System

The DNA-associating protein Dsup exemplifies encoded chemical functionality. It has no homologs outside tardigrades, lacks known regulatory upstream elements, and operates through physical association rather than regulatory control (Hashimoto et al. 2016). Dsup functions as a physical shield that binds to chromatin and protects DNA from oxidative and radiation damage through direct chemical interaction rather than transcriptional regulation.

This represents exactly what the CGIM predicts: a sophisticated chemical solution (DNA protection through direct molecular association) that was successfully encoded but retained its pre-genetic operational logic. The protein behaves like a pre-compiled subroutine, invoked by local chemical conditions rather than regulatory networks.

Late Embryogenesis Abundant (LEA)-like Proteins

Tardigrades express numerous LEA-like proteins that stabilize cellular components through bulk physical encapsulation rather than sequence-specific interactions. These proteins form vitreous gels during desiccation that maintain cellular integrity through chemical rather than biological mechanisms—essentially creating a chemical preservation matrix that operates independently of genetic control.

The tardigrade genome contains a significant proportion of intrinsically disordered proteins and taxonomically restricted genes (Khalturin et al. 2009). While early claims of massive horizontal gene transfer were largely dismissed (Koutsovoulos et al. 2016), the presence of "alien" proteins remains. From a Chemical-Genetic Integration perspective, these represent biochemical residues from an ancient, pre-genetic toolkit, preserved in the genome without modern regulatory integration.

Ctenophores: Persistent Runtime Neural Systems

Ctenophores (*Mnemiopsis leidyi*, *Pleurobrachia bachei*) present an evolutionary paradox that becomes comprehensible through the runtime biology framework: they possess complex nervous systems yet lack many canonical neural genes seen in other metazoans (Moroz et al. 2014; Ryan et al. 2013).

Non-Canonical Neural Architecture

Key neurotransmitters such as serotonin, dopamine, and acetylcholine, as well as synaptic scaffold proteins, are absent or highly divergent in ctenophores. Instead, they use a distinct suite of neuropeptides that are lineage-specific and structurally untraceable to known families (Sachkova et al. 2021). This represents exactly what chemical inheritance predicts: functional systems that operate through encoded chemical solutions rather than conventional genetic modules.

Constitutive Neuropeptide Expression

What makes ctenophores particularly relevant to the CGIM is the constitutive expression of these neuropeptides across tissues and developmental stages. Rather than being transcriptionally activated in response to neural signaling or environmental stimuli, many neuropeptides appear to be "always on," implementing a runtime-like control system that functions through local chemical conditions rather than top-down gene regulation.

Many of these peptides are predicted to be intrinsically disordered and likely engage in weak, multivalent interactions—ideal features for phase-separated functional domains. This behavior is consistent with ancient, runtime-like architecture: neurochemical tools retained from pre-genetic functional chemistry, encoded post hoc but never fully modularized.

Chemical vs. Genetic Neural Control

The ctenophore nervous system appears to operate through environmental chemistry rather than purely informational processing. This represents a fundamentally different neural architecture—one that preserves the chemical responsiveness of pre-genetic systems while gaining the reproducibility of genetic encoding. The result is a nervous system that functions more like a chemical computer than a digital information processor.

Hydra: Stabilized Runtime Architecture

Among early-diverging animals, *Hydra vulgaris* represents a lineage that successfully stabilized its runtime inheritance across millions of years, maintaining constitutive biochemical systems and decentralized architecture while achieving long-term genomic stability.

Persistent Regenerative Capacity

Hydra vulgaris AEP displays extreme longevity and does not undergo senescence, even under long-term conditions of constant cell turnover and asexual reproduction (Barve et al. 2021). Its regenerative abilities are mediated by three semi-autonomous stem cell lineages that coordinate tissue development through local signaling rather than hierarchical patterning systems (Cazet et al. 2023). This modular autonomy within a single organism suggests retention of decentralized runtime control reminiscent of early chemical systems.

Transposable Element-Rich Genome Architecture

The genome of *Hydra vulgaris* strain 105 is approximately 1.05 Gb—more than double that of the morphologically comparable cnidarian *Nematostella vectensis* (Chapman et al.

2010). Over 60% of the *Hydra* genome consists of repetitive elements, with a single Tc1/mariner transposable element family accounting for much of the increase (Wong et al. 2019).

Compared to *Hydra viridissima*, which possesses a compact genome of ~500 Mbp and reduced TE content (Kon-Nanjo et al. 2024), the retention of repeat-rich architecture in *Hydra vulgaris* appears to represent a lineage-specific strategy rather than an unavoidable evolutionary stage. These TEs remain differentially active across cell types and developmental states, suggesting they continue to serve functional roles as regulatory substrates or structural components—analogous to a runtime library retained in the genome.

Integration of Runtime and Genetic Control

Hydra chromatin structure exhibits high accessibility in TE-rich regions (Sahm et al. 2024), and its transcriptional landscape includes persistent expression of peptide signaling factors and transcription regulators. Rather than streamlining legacy chemical systems into modular gene circuits, *Hydra vulgaris* appears to have integrated them into a functional equilibrium that maintains ancient chemical functionality while achieving regulatory robustness.

This positions *Hydra* as a lineage that successfully stabilized runtime biology rather than transitioning to purely modular organization. Unlike ctenophores and tardigrades, which show signs of partial or fragmented integration, *Hydra* presents a coherent, continuous runtime system capable of maintaining function across time and tissue types without modular degradation.

Successful Integration Lineages

While tardigrades, ctenophores, and *Hydra* demonstrate the preservation of runtime biology, other lineages successfully encoded and modularized their inherited chemical systems, enabling subsequent evolutionary elaboration. These organisms provide critical evidence that chemical-genetic integration was not a dead-end preservation strategy but a pathway to biological innovation.

Silicatein: Encoded Environmental Chemistry

Demosponges represent a paradigmatic example of successful chemical-genetic integration, having encoded environmental silica chemistry into sophisticated biological architectures while retaining the fundamental chemical logic of their pre-genetic toolkit.

The enzyme silicatein in demosponges catalyzes the polymerization of silica to build skeletal spicules, representing a clear example of encoded environmental chemistry (Shimizu et al. 1998). Silicatein resembles cathepsin proteases but has been functionally repurposed to catalyze biomineralization—a feat that directly links ancient chemical function with modern genetic control.

Significantly, silicatein retains the operational characteristics of mineral-mediated silica polymerization despite being a genetically encoded protein. It operates at ambient temperature and pressure using mechanisms that closely parallel geochemical silica chemistry (Müller et al. 2007). The enzyme provides organization and nucleation, but the actual chemical work is performed by environmental silica chemistry that predates biological control—demonstrating successful encoding rather than replacement of pre-genetic functionality.

The existence of weakly biomineralized spicules with large axial filaments and high organic proportions in early Cambrian sponges suggests a transitional state between purely chemical silica precipitation and fully encoded biosilicification (Tang et al. 2019). This evolutionary sequence provides direct evidence for the proposed integration pathway: environmental chemistry → partial biological organization → complete genetic encoding → evolutionary optimization.

Modern sponge biomineralization demonstrates that the integration process preserved the essential chemistry while adding reproducible biological control. The spicule formation process represents chemical garden formation under genetic guidance rather than purely biological construction—exactly what chemical-genetic integration predicts.

The *Silicarian* Lineage: Specialized Chemical Integration

The proposed *Silicarian* lineage exemplifies systematic specialization in encoding specific aspects of the pre-genetic chemical toolkit. This lineage encompasses glass sponges (Hexactinellida and certain Demosponges), radiolarians, and diatoms—organisms that share sophisticated silica biomineralization capabilities despite phylogenetic distance.

The Multiple Multicellular Prototype (MCP) Framework reveals that these organisms function as specialized electrical engineers from an 850-million-year-old Silicarian lineage (Gjøvik 2025). Glass sponges demonstrate the endpoint of successful silica integration: organisms that conduct electricity through silica-based architecture, with spicule networks creating three-dimensional electrical grids that coordinate colony-scale behavior.

The identification of cross-kingdom structural conservation validates core CGIM predictions. Triactine and hexactine spicule geometries appear across distantly related taxa not through convergent evolution, but through shared access to the same silica-processing chemical toolkit. Radiolarian geometric complexity and diatom architectural precision reflect variations on encoded environmental silica solutions.

The Silicarian lineage demonstrates that chemical-genetic integration enabled lineages to become specialists in particular aspects of environmental chemistry. Rather than preserving broad chemical competence like runtime lineages, successful integration allowed focused mastery of specific chemical systems while achieving the efficiency and reproducibility advantages of genetic encoding.

Siphonophores: Post-Genetic Functional Fusion

Siphonophores provide compelling evidence that biological integration processes continue beyond the ancient chemical-genetic transition, demonstrating post-genetic functional fusion—the incorporation of independent biological entities into unified genomic systems.

Individual zooids within siphonophore colonies exhibit such radical morphological and functional diversity that they appear as independent organisms: specialized feeding, reproductive, locomotory, and defensive modules with complete organ systems (Damian-Serrano et al. 2021). Current models invoke clonal development from single zygotes, but this fails to explain how such extreme specialization could emerge without prior modular fusion.

The CGIM framework suggests an alternative: siphonophores represent successful post-genetic integration, wherein formerly independent biological entities were functionally absorbed and genomically unified. The zooids preserve evidence of this process—they retain complete biological architectures while sharing unified genetic control systems.

Genomic Integration Mechanisms

This integration required sophisticated genome compilation, (a genetic "git merge"), resolving previously distinct genetic elements into unified, functional systems. The process likely involved:

- **Selective retention** of useful genetic modules from incorporated entities
- **Suppression or elimination** of conflicting regulatory systems
- **Development of master coordination mechanisms** across the integrated colony

Recent analysis reveals that siphonophore genetic systems show evolution of gene expression across species and specialized zooids, with coordinated regulation spanning morphologically distinct modules (Munro et al. 2022). This represents higher-order biological integration analogous to the chemical-genetic integration proposed for early life.

Validation of Integration Principles

Siphonophores demonstrate that biological integration operates as an ongoing evolutionary strategy rather than an ancient historical process. They provide direct evidence that:

- **Independent functional entities** can be successfully incorporated into unified biological systems
- **Genomic integration** can preserve modular functionality while enabling coordinated behavior
- **Integration processes** operate across multiple scales and time periods

The success of siphonophore integration validates the core CGIM principle: life's complexity emerges not purely through gradual modification, but through the systematic integration of pre-existing functional systems—whether chemical networks in deep time or biological modules in recent evolution.

Cnidarian Neural Network Integration

Early hydrozoans exhibit successful modularization of neural systems, providing evidence for how chemical-genetic integration enabled the evolution of sophisticated regulatory control. Unlike ctenophores, which retained runtime neural chemistry, hydrozoans successfully encoded neural function into modular genetic networks.

Hydrozoans demonstrate early evidence of neural network modularization, including neurons with conventional neurotransmitters and regulated developmental pathways (Rentzsch et al. 2017). This suggests they inherited both functional runtime chemistry and genetic toolkit capabilities, successfully compiling chemical systems into regulatory logic that enabled subsequent evolutionary innovation.

Integration as Evolutionary Strategy

These successful integration lineages demonstrate the evolutionary potential of chemical encoding. They show how functional chemistry from the pre-genetic world became integrated, reorganized, and evolutionarily optimized through genetic modularization. Rather than representing endpoint preservation like runtime lineages, successful integration provided the foundation for major evolutionary radiations and the development of complex biological innovations.

The contrast between runtime and integration strategies reveals the fundamental trade-offs that shaped early animal evolution: runtime lineages maintained access to comprehensive ancient chemical libraries but sacrificed evolutionary flexibility, while integration lineages gained adaptability and specialization by successfully encoding and modularizing their chemical inheritance.

The Dark Proteome as Chemical Legacy

The Chemical-Genetic Integration Model makes specific predictions about molecular signatures that chemical inheritance should leave in modern genomes. If genetic systems emerged by encoding pre-existing chemical networks, we should observe preserved molecular elements that operate through chemical rather than purely informational mechanisms. The distribution and properties of orphan genes, intrinsically disordered proteins, and structural homologies without sequence similarity provide exactly this predicted evidence.

Orphan Genes as Encoded Chemical Solutions

Orphan genes—sequences with no detectable homologs in other lineages—have long been treated as evolutionary puzzles requiring special explanation (Tautz and Domazet-Lošo 2011). The CGIM reframes them as direct molecular evidence: orphan genes represent the encoding of lineage-specific chemical solutions that were successfully integrated during the chemistry-to-genetics transition.

This interpretation explains their distinctive properties. Orphan genes are disproportionately associated with stress response, environmental sensing, and metabolic innovation—exactly the functional domains where chemical networks would have provided immediate adaptive value (Khalturin et al. 2009). Their apparent "sudden appearance" in evolutionary history reflects not de novo gene birth, but the successful encoding of pre-existing chemical functionality.

The high proportion of orphan genes in runtime-style lineages supports this interpretation. Tardigrades, ctenophores, and other organisms exhibiting runtime biology contain elevated numbers of taxonomically restricted genes that lack clear regulatory integration—consistent with encoded chemical modules that were never fully compiled into modern genetic networks.

Intrinsically Disordered Proteins as Chemical Fossils

Intrinsically disordered proteins (IDPs) represent perhaps the clearest molecular signature of pre-genetic systems. These proteins function through conformational flexibility and environmental responsiveness rather than fixed tertiary structure—characteristics that directly mirror the operational logic of chemical networks (Wright and Dyson 2015).

IDPs operate through phase separation behaviors, context-dependent function, and environmental activation mechanisms that reflect chemical network properties rather than genetic control logic. Their ability to form membrane-free compartments through liquid-liquid phase separation parallels the compartmentalization strategies available to pre-genetic chemical systems (Banani et al. 2017).

Modern cellular phase-separated compartments—stress granules, P-bodies, and nucleoli—demonstrate how chemical organization can persist within genetic systems. These structures self-assemble spontaneously when appropriate chemical conditions are present, maintain internal organization through selective partitioning rather than active transport, and respond dynamically to environmental conditions rather than regulatory signals (Hyman, Weber, and Jülicher 2014).

IDPs are overrepresented in early-diverging lineages and stress-response pathways, consistent with CGIM predictions that chemical-like functionality should be preserved in organisms retaining runtime biology (Perdigão et al. 2015). The environmental responsiveness, phase separation behaviors, and context-dependent function of IDPs all reflect chemical network properties that genetic systems successfully encoded while preserving their fundamental operational characteristics.

Structural Homology Without Sequence Similarity

The persistent observation of structural conservation across phylogenetically distant proteins without detectable sequence similarity provides direct evidence for chemical inheritance. If proteins evolved purely through sequence-based inheritance, structural conservation should correlate with sequence similarity. The systematic violation of this expectation indicates that structural solutions were inherited through non-sequence mechanisms—exactly what chemical inheritance predicts (Orengo and Thornton 2005).

These structural fossils represent chemical scaffolds that were independently encoded by multiple lineages accessing the same environmental chemical toolkit. Iron-sulfur coordination geometries, membrane-insertion domains, and catalytic architectures all reflect chemical constraints that shaped protein evolution prior to and independent of sequence-based inheritance.

The [4Fe-4S] clusters found across diverse proteins maintain identical cubic geometry and coordination chemistry despite sequence divergence, indicating inheritance of chemical solutions rather than genetic templates. Similarly, the widespread occurrence of Rossmann-like domains, TIM barrels, and beta-grasp folds across unrelated proteins reflects shared access to chemically constrained folding solutions.

Metal-Binding Architecture Conservation

Modern Fe-S proteins retain striking architectural similarities to their mineral precursors, suggesting direct inheritance rather than de novo evolution. The enzyme aconitase contains a [4Fe-4S] cluster that functions through the same Lewis acid chemistry as iron-sulfur minerals catalyzing organic transformations. The active sites of nitrogenase contain iron-sulfur-molybdenum clusters that closely resemble the metallic compositions found in hydrothermal vent minerals.

These structural continuities extend beyond simple chemical composition to functional architecture. The electron transport chains that power all life rely on iron-sulfur clusters arranged in configurations that mirror the electron transfer pathways available on mineral surfaces. The CO_2 fixation pathways utilize iron-sulfur clusters in arrangements that directly parallel those demonstrated in prebiotic iron-sulfur chemistry experiments.

Modern biological systems demonstrate how sophisticated chemical organization can operate through environmental energy coupling rather than genetic control. Biomineralization processes show how complex chemical systems maintain organization through environmental chemistry while being minimally encoded genetically.

The silicatein system exemplifies this principle, forming organized silica structures through environmental energy coupling. The magnetosomes of magnetotactic bacteria produce magnetite nanocrystals chemically and structurally indistinguishable from abiotic magnetite, yet organized into sophisticated cellular arrays through spatial organization rather than fundamental chemical alteration.

These systems demonstrate that genetic encoding preserved environmental chemistry rather than replacing it. The biological innovation involved spatial and temporal control of chemical processes that operated effectively in prebiotic environments.

Validation of Chemical Inheritance Mechanisms

The molecular signatures collectively validate the core CGIM mechanism. Rather than representing evolutionary anomalies requiring special explanation, orphan genes, IDPs, and structural homologies provide systematic evidence that genetic systems preserved functional elements of pre-genetic chemical organization.

Their distribution patterns, functional properties, and operational characteristics match theoretical predictions for what chemical inheritance should produce. The dark proteome appears dark not because it lacks function, but because conventional genetic frameworks fail to recognize chemically inherited functionality.

Systematic Molecular Evidence

The convergence of multiple independent molecular signatures strengthens the evidence for chemical inheritance:

- **Orphan genes** encode lineage-specific chemical solutions
- **IDPs** preserve chemical responsiveness within genetic systems
- **Structural conservation** reflects shared chemical constraints
- **Phase separation** maintains chemical organization principles
- **Environmental activation** operates through chemical rather than informational triggers

Rather than treating these as separate phenomena requiring individual explanations, the CGIM provides a unified framework: they represent different aspects of the same fundamental process—the encoding of pre-genetic chemical functionality into biological inheritance systems.

Recognizing molecular signatures of chemical inheritance transforms our interpretation of genomic complexity. Sequences previously dismissed as noise, contamination, or evolutionary accident become windows into the chemistry-to-genetics transition. The dark proteome reveals the chemical foundations that underlie genetic inheritance, demonstrating that biological complexity reflects not purely informational innovation but the systematic compilation of environmental chemical solutions.

This perspective suggests that much of what appears as genomic "junk" may actually represent preserved chemical functionality that operates through principles not yet understood by conventional genetic frameworks. The molecular fossils embedded in modern genomes provide direct access to the organizational strategies that preceded and enabled genetic inheritance.

Reframing Genomic "Anomalies"

Current genomic analysis frameworks treat certain molecular patterns as anomalies requiring special explanation: genes that appear "out of place" phylogenetically, sequences that align with distantly related organisms, and proteins that lack clear evolutionary origins. The Chemical-Genetic Integration Model suggests these "anomalies" actually represent systematic evidence for chemical inheritance that conventional frameworks fail to recognize.

The Overextension of Horizontal Gene Transfer

The concept of horizontal gene transfer (HGT) revolutionized our understanding of microbial evolution through landmark work demonstrating that prokaryotes frequently exchange genetic material across species boundaries (Doolittle 1999; Ochman, Lawrence, and Groisman 2000). However, HGT has become increasingly invoked in eukaryotic genomics as an interpretive default for genes that appear "out of place" based on standard phylogenetic expectations.

In recent years, genes that are taxonomically restricted, highly divergent, or structurally unique have been attributed to ancient or cryptic HGT events, especially in extremophiles and early-branching animals (Gladyshev, Meselson, and Arkhipova 2008; Boothby et al. 2015). These interpretations are often applied to explain observations that may instead reflect other evolutionary processes, including ancestral gene loss, de novo emergence, or—critically for this framework—pre-genetic functional inheritance.

Many genes labeled as HGT candidates do not occur on mobile genetic elements, are stably inherited across developmental stages, and encode proteins critical to core functional processes such as stress tolerance, DNA repair, and protein stability. Rather than invoking lateral transfer, the Chemical-Genetic Integration Model suggests these represent compiled chemical systems—functional modules that existed before genetic encoding and were independently captured by lineages with access to similar environmental chemistries.

This reinterpretation explains several problematic features of supposed eukaryotic HGT:

- **Taxonomic restriction**: Genes were independently encoded by lineages accessing similar environmental chemistries rather than transferred between existing organisms
- **Cross-domain occurrence**: Homologs appear in bacteria, archaea, and eukaryotes through shared chemical ancestry rather than horizontal movement
- **Poor phylogenetic resolution**: Sequences did not evolve through modular genetic fashion but emerged as environmental chemical solutions that predate informational control
- **Functional importance**: Core cellular processes reflect encoded environmental chemistry rather than recently acquired foreign genes

Taxonomic Filtering and the Loss of Ancient Signal

To maintain rigor in genome assembly and annotation, standard practice filters contigs based on taxonomic identity. Sequences that align strongly with bacterial or archaeal genes are often removed from eukaryotic genome projects as potential contaminants using tools such as Kraken, BlobTools, and MEGAN (Wood and Salzberg 2014; Laetsch and Blaxter 2017).

While contamination filtering ensures data quality, it carries a critical methodological risk: the inadvertent exclusion of genuine biological signal, particularly where deep evolutionary mosaicism is involved. If early metazoans incorporated pre-genetic chemistry that shares homology with microbial domains, standard filtering pipelines may systematically remove functionally relevant sequences simply because they defy eukaryotic classification.

The assumption that eukaryotic genomes must contain only eukaryotic genes creates circular logic: sequence similarity becomes both the inclusion criterion and the measure of evolutionary legitimacy. This limits discovery of genes with deep, non-canonical ancestry and masks evidence for the chemical inheritance proposed by the CGIM.

Reanalysis of genome assemblies without taxonomic contamination filtering should recover functionally relevant sequences that were previously excluded as "bacterial contamination." These recovered sequences should show structural and functional characteristics consistent with ancient chemical systems rather than horizontal gene transfer or contamination artifacts.

The systematic exclusion of such sequences by current filtering approaches represents methodological bias that masks evidence for chemical inheritance. Sequences dismissed as contamination may actually represent some of the clearest molecular evidence for the chemistry-to-genetics transition.

Chemical Inheritance vs. Horizontal Transfer

The distinction between chemical inheritance and horizontal gene transfer is critical for interpreting genomic complexity. These processes make different predictions about sequence distribution, functional integration, and evolutionary timing.

Diagnostic Differences

Feature	Chemical Inheritance	Horizontal Gene Transfer
Phylogenetic Pattern	Cross-kingdom conservation without clear transfer pathway	Recent acquisition with identifiable donor
Functional Integration	Core cellular processes, often constitutively expressed	Accessory functions, often regulated
Sequence Character-istics	Structural homology, chemical constraint signatures	Recent sequence similarity, mobile elements
Distribution Logic	Independent encoding of shared chemical solutions	Adaptive transfer between existing organisms
Evolutionary Timing	Ancient, predating major divergences	Recent, within established lineages

Structural Inheritance vs. Informational Transfer

Chemical inheritance operates through structural solutions to physical constraints, while horizontal gene transfer involves informational exchange between existing genetic systems. This fundamental difference explains why chemical inheritance produces cross-kingdom conservation patterns that violate conventional phylogenetic expectations.

The persistent appearance of iron-sulfur coordination geometries, membrane-insertion domains, and catalytic architectures across distantly related proteins reflects shared access to chemically constrained solutions rather than recent genetic exchange. These structural fossils represent chemical scaffolds that were independently encoded by multiple lineages accessing the same environmental toolkit.

Beyond Taxonomic Assumptions

Modern taxonomic frameworks treat gene origins as branching problems: genes must descend from other genes through duplication, divergence, horizontal acquisition, or random emergence. This model assumes all novelty arises through modifications of existing genetic information.

Functional Ancestry vs. Informational Ancestry

The CGIM framework recognizes an additional inheritance mechanism: functional ancestry through shared chemical systems that predate genetic encoding. If early life included a shared pool of environmental chemistry, genes could arise in parallel through encoding of non-genetic functional systems across different lineages.

Taxonomy, as currently applied, cannot account for this mode of inheritance. Structural homologs may be misclassified as unrelated, lineage-specific traits may be overinterpreted as "novel," and pre-genetic modules may be discarded as "contamination." To fully reconstruct early evolution, we must consider not only informational ancestry but also functional ancestry—the inheritance of chemical systems that transcend sequence-based phylogenetics.

Systematic Reinterpretation Framework

By reevaluating genomic anomalies as possible evidence of chemical inheritance, the CGIM encourages systematic reinterpretation:

Conventional Classification	Chemical Inheritance Interpretation
HGT gene	Ancient, independently encoded chemical module
Taxonomic contamination	Misclassified functional inheritance from shared chemical toolkit
Orphan gene/TRG	Compiled output of lineage-specific chemical solutions
Structural convergence	Shared physical solutions to ancestral environmental constraints

This reframing transforms isolated anomalies into systematic evidence for the chemistry-to-genetics transition. Rather than requiring special explanations for each unusual sequence or protein, the framework provides a unified mechanism: differential compilation of shared chemical inheritance.

Validation Through Systematic Patterns

The chemical inheritance framework makes testable predictions about the distribution and properties of genomic anomalies. If these sequences represent encoded environmental chemistry rather than artifacts or transfers, they should show:

- **Cross-kingdom conservation patterns** consistent with shared chemical constraints
- **Functional clustering** around stress response, environmental sensing, and metabolic processes
- **Structural homology** reflecting chemical rather than genetic optimization
- **Expression patterns** consistent with runtime rather than modular biology
- **Environmental compatibility** under conditions relevant to early Earth chemistry

The systematic occurrence of these patterns across multiple "anomalous" sequences would support chemical inheritance over alternative explanations. Rather than treating each genomic puzzle as an isolated problem, the CGIM provides a framework for recognizing systematic evidence of life's most fundamental transition.

Recognizing chemical inheritance transforms genomic analysis from an exercise in phylogenetic reconstruction to archaeological investigation of the chemistry-to-genetics transition. Sequences previously dismissed as noise become windows into early evolution, and genomic complexity reveals the chemical foundations underlying genetic inheritance.

This perspective suggests that conventional genomic analysis may systematically miss the most important evolutionary signals—the molecular evidence for how genetic systems actually emerged from chemical organization. By reframing anomalies as evidence rather than problems, we gain access to the preserved record of life's most fundamental transition.

Testable Predictions & Future Directions

The Chemical-Genetic Integration Model generates specific, falsifiable predictions that distinguish it from alternative explanations for early animal evolution. Rather than seeking to prove the absence of pre-genetic chemical inheritance—an inherently difficult proposition—the model predicts positive molecular signatures that should be detectable if chemical-genetic integration occurred as proposed. These predictions span multiple levels of biological organization and provide a systematic research program for validating the framework.

Quantitative Runtime Biology Predictions

If the CGIM correctly explains the retention of pre-genetic chemical systems, organisms should demonstrate measurable differences in their molecular organization patterns that correlate with proposed evolutionary strategies.

Phylogenetic Distribution Predictions

Prediction 1: Early-diverging lineages (ctenophores, tardigrades, some sponges) should consistently score high on runtime criteria (>70% constitutive protein expression, >40% intrinsically disordered proteins, environmental activation mechanisms), while derived bilaterians should score high on modular criteria.

Prediction 2: Organisms with intermediate scores should represent transitional cases where chemical systems were partially encoded and modularized, with measurable gradients corresponding to the degree of chemical-genetic integration.

The operational criteria established earlier provide quantitative metrics that can be systematically applied across phylogenetically diverse organisms. This should reveal non-random clustering that correlates with proposed evolutionary strategies rather than random distribution across phylogeny.

Expression Pattern Validation

Prediction 3: Runtime-style organisms should show constitutive expression of stress-response proteins that function through environmental activation rather than transcriptional induction. These proteins should maintain functionality under early Earth environmental conditions (reducing atmospheres, high pH, elevated metal concentrations).

Prediction 4: Modular-style organisms should show conditional expression patterns with clear regulatory hierarchies, transcription factor networks, and tissue-specific deployment of functional modules.

Structural Analysis and Fold Convergence

The model predicts that proteins currently classified as "orphan genes" should display structural similarities to environmentally-constrained chemical solutions, even without sequence homology.

AlphaFold-Based Structural Archaeology

Prediction 5: Systematic structural analysis using AlphaFold2 and related tools should reveal that taxonomically restricted proteins in runtime-style organisms adopt folds consistent with redox chemistry, metal coordination, or phase separation—reflecting their proposed origins in pre-genetic chemical systems.

Prediction 6: These structural similarities should be non-random and correlate with environmental chemistry available during early Earth conditions (iron-sulfur clusters, silica scaffolding, manganese catalysis). Orphan proteins should show chemically constrained rather than truly novel folds.

Cross-Kingdom Structural Conservation

Prediction 7: Proteins with identical structural solutions but no sequence similarity should cluster around fundamental chemical constraints rather than recent evolutionary innovations. Iron-sulfur coordination geometries, membrane-insertion architectures, and catalytic domains should show conservation patterns consistent with shared chemical ancestry.

Experimental Protocell Validation

The proposed chemical-genetic integration pathways must be mechanistically feasible under prebiotic conditions.

Synthetic Protocell Studies

Prediction 8: Synthetic protocells should demonstrate selective uptake and retention of iron-sulfur clusters, silica scaffolding components, or other proposed chemical systems when exposed to simulated early Earth environments.

Prediction 9: These uptake processes should show specificity based on membrane composition and environmental conditions, matching the selectivity proposed for early chemical-genetic integration. Random chemical uptake would invalidate the selective integration mechanism.

Environmental Compatibility Testing

Prediction 10: Proteins proposed to represent encoded pre-genetic chemistry (tardigrade TDPs, ctenophore neuropeptides, sponge silicateins) should retain functionality under early Earth environmental conditions, particularly hydrothermal and alkaline conditions.

Prediction 11: These proteins should show enhanced stability or activity under reducing conditions compared to modern atmospheric conditions, reflecting their proposed origins in pre-oxygenic chemical systems.

Genomic Archaeology Applications

Current genomic analysis practices may systematically exclude evidence for chemical-genetic integration.

Contamination Filter Reanalysis

Prediction 12: Reanalysis of genome assemblies without taxonomic contamination filtering should recover functionally relevant sequences in early-diverging animals that were previously excluded as "bacterial contamination."

Prediction 13: These recovered sequences should show structural and functional characteristics consistent with ancient chemical systems rather than horizontal gene transfer or contamination artifacts—specifically, metal-binding domains, phase-separation capabilities, and environmental activation mechanisms.

Ancient Marine Assemblage Screening

Prediction 14: Systematic application of the recursive alignment strategy should identify additional organisms exhibiting the characteristic ancient marine assemblage genomic signature, demonstrating that chemical fossils represent systematic rather than isolated discoveries.

Prediction 15: Organisms identified through this screening should show correlated morphological, physiological, and behavioral characteristics consistent with chemical inheritance—taxonomic isolation, exceptional capabilities, and runtime biology features.

Integration Mechanism Validation

The iron-sulfur and silica integration pathways provide specific, testable mechanisms for chemical-genetic integration.

Laboratory Evolution Studies

Prediction 16: Directed evolution experiments should demonstrate that environmental chemical systems (iron-sulfur clusters, silica scaffolds) can be successfully incorporated into synthetic biological systems and subsequently encoded through selection pressures.

Prediction 17: The encoded systems should retain the fundamental chemical logic of their environmental precursors while gaining biological control mechanisms—demonstrating compilation rather than replacement of chemical functionality.

Biochemical Reconstruction

Prediction 18: In vitro reconstruction of proposed "runtime" proteins under simulated early Earth conditions should reveal latent catalytic or structural capabilities that align with predictions from prebiotic chemistry—redox buffering, metal coordination, self-assembly in non-enzymatic conditions.

Prediction 19: These proteins should show enhanced function when combined with appropriate cofactors (iron-sulfur clusters, metal ions, mineral surfaces) under conditions resembling their proposed origins.

Systematic Discovery Framework

The genomic archaeology methodology should enable systematic identification of additional chemical inheritance examples.

Predictive Screening Protocol

Prediction 20: Application of the diagnostic criteria (cross-kingdom sequence conservation, phylogeny-structure dissociation, physical constraint signatures, runtime expression profiles) should systematically identify organisms that were misclassified due to chemical inheritance.

Prediction 21: Each newly identified chemical fossil should strengthen rather than weaken the overall pattern, demonstrating systematic rather than cherry-picked evidence for the framework.

Falsification Criteria

The model's strength lies in generating testable hypotheses that could definitively refute the framework:

- **Structural analysis failure**: If orphan proteins show truly novel folds with no relationship to prebiotic chemistry.
- **Environmental incompatibility**: If proposed "ancient" proteins fail to function under early Earth conditions.
- **Random distribution**: If runtime vs. modular characteristics show no phylogenetic correlation.
- **Integration mechanism failure**: If synthetic protocells cannot achieve selective chemical uptake under controlled conditions.
- **Absence of systematic patterns**: If genomic archaeology fails to identify additional examples or reveals only isolated anomalies.

Research Program Implementation

These predictions provide a systematic research program spanning computational biology, experimental evolution, structural analysis, and synthetic biology. The convergence of positive results across multiple independent approaches would strongly support the Chemical-Genetic Integration Model, while failure in any major prediction category would require substantial revision or rejection of the framework.

Immediate Research Priorities

1. **Quantitative assessment** of runtime vs. modular characteristics across early-diverging animal lineages.
2. **Structural analysis** of orphan genes and taxonomically restricted proteins using advanced computational tools.
3. **Experimental validation** of chemical uptake mechanisms in synthetic protocell systems.
4. **Genomic reanalysis** without taxonomic filtering to recover excluded sequences.
5. **Environmental compatibility testing** of proposed ancient proteins under prebiotic conditions.

Long-term Research Directions

The validation of chemical inheritance would open entirely new research directions in evolutionary biology, origin-of-life studies, and synthetic biology. Understanding how chemical systems became genetic systems provides mechanistic foundations for engineering similar transitions, exploring life's emergence on other worlds, and recognizing the chemical intelligence that preceded and enabled genetic inheritance.

The systematic discovery of chemical fossils would transform our understanding of genomic organization, revealing that much apparent "junk DNA" and "dark proteome" actually represents preserved evidence of life's most fundamental transition. Rather than viewing early evolution through purely genetic frameworks, we would gain access to the chemical logic that shaped biological inheritance.

Conclusion

This study proposes the Chemical-Genetic Integration Model as a mechanism for the chemistry-to-genetics transition and presents evidence for its operation through analysis of modern biological systems. The work establishes operational criteria for distinguishing runtime from modular biological organization and applies genomic archaeology methods to identify potential molecular signatures of chemical inheritance.

The identification of organisms exhibiting cross-kingdom genomic conservation patterns provides preliminary evidence that chemical inheritance mechanisms may be detectable in modern genomes. The systematic occurrence of these patterns across taxonomically problematic organisms suggests that chemical-genetic integration may represent a systematic rather than exceptional evolutionary process.

The framework generates testable predictions spanning computational analysis, experimental validation, and systematic discovery approaches. Key predictions include phylogenetic distribution patterns of runtime versus modular characteristics, structural conservation reflecting chemical constraints, and experimental demonstration of selective chemical uptake mechanisms in protocell systems.

Several limitations require acknowledgment. The proposed mechanisms require extensive empirical validation, computational methods may be subject to systematic biases, and alternative explanations for observed genomic patterns must be rigorously excluded. The interpretation of molecular signatures as evidence for chemical inheritance depends on developing robust criteria that distinguish this process from horizontal gene transfer, contamination, or conventional evolutionary mechanisms.

Future research should prioritize experimental validation of integration mechanisms, systematic application of genomic archaeology methods across diverse taxa, and development of quantitative metrics for identifying chemical inheritance. The convergence of evidence from multiple independent approaches will be necessary to establish the validity of the proposed framework.

If validated, the Chemical-Genetic Integration Model would provide mechanistic insights into the chemistry-to-genetics transition and suggest that substantial genomic content may represent compiled environmental chemistry rather than purely genetic innovation. The approach offers systematic methods for investigating this fundamental evolutionary transition through preserved molecular signatures in contemporary biological systems.

Acknowledgements

The author conceptualized the study, developed the theoretical framework, and conducted all analyses. This research received no external funding and was conducted using personally acquired resources.

The author acknowledges Molecular Biologist GPT, Scholar GPT, and Claude AI for their assistance with manuscript preparation, data interpretation discussions, and research development over the course of this investigation.

Data Availability: No new data was created in the preparation of this paper. Project data for the journal is available at https://osf.io/ur79d/ (DOI 10.17605/OSF.IO/UR79D).

References

Baek, Minkyung, et al. 2021. “Accurate Prediction of Protein Structures and Interactions Using a Three-Track Neural Network.” *Science* 373 (6557): 871–76.

Banani, Salman F., Hyun O. Lee, Anthony A. Hyman, and Michael K. Rosen. 2017. “Biomolecular Condensates: Organizers of Cellular Biochemistry.” *Nature Reviews Molecular Cell Biology* 18 (5): 285–98.

Boothby, Thomas C., et al. 2015. “Evidence for Extensive Horizontal Gene Transfer from the Draft Genome of a Tardigrade.” *Proceedings of the National Academy of Sciences* 112 (52): 15976–81.

Boothby, Thomas C., et al. 2017. “Tardigrades Use Intrinsically Disordered Proteins to Survive Desiccation.” *Molecular Cell* 65 (6): 975–84.e5.

Camprubi, Eloi, Jordan G. Lessing, Ruben Smolin, Victor Sojo, Sean F. Jordan, Joana C. Xavier, and Nick Lane. "Iron Catalysis at the Origin of Life." *IUBMB Life* 69, no. 6 (2017): 373–381.

Chapman, Jarrod A., Ewen F. Kirkness, Oleg Simakov, Steven E. Hampson, Therese Mitros, Thomas Weinmaier, Thomas Rattei, and others. "The Dynamic Genome of Hydra." *Nature* 464, no. 7288 (2010): 592–596.

Colman, Daniel R et al. “Microbial ecology of serpentinite-hosted ecosystems.” The ISME journal vol. 19,1 (2025): wraf029.

Copley, Shelley D., Eric Smith, and Harold J. Morowitz. 2005. “A Mechanism for the Association of Amino Acids with Their Codons and the Origin of the Genetic Code.” *Proceedings of the National Academy of Sciences* 102 (12): 4442–47.

Dai, S., Xie, Z., Wang, B. *et al.* An inorganic mineral-based protocell with prebiotic radiation fitness. *Nat Commun* **14**, 7699 (2023).

Dalai, Punam, and Nita Sahai. “Mineral-Lipid Interactions in the Origins of Life.” *Trends in biochemical sciences* vol. 44,4 (2019): 331-341.

Damian-Serrano, Alejandro et al. “The evolution of siphonophore tentilla for specialized prey capture in the open ocean.” *Proceedings of the National Academy of Sciences of the United States of America* vol. 118,8 (2021).

Dalai, Punam, and Nita Sahai. "Mineral–Lipid Interactions in the Origins of Life." *Trends in Biochemical Sciences* 44, no. 4 (2019): 331–341.

Davidson, Eric H., and Douglas H. Erwin. 2006. “Gene Regulatory Networks and the Evolution of Animal Body

Plans." *Science* 311 (5762): 796–800.

Doolittle, W. Ford. 1999. "Phylogenetic Classification and the Universal Tree." *Science* 284 (5423): 2124–29.

Dunn, Casey. "Siphonophores." *Current biology : CB* vol. 19,6 (2009): R233-4.

Fortunato, Savanna A. V., et al. 2014. "Calcisponges Have a ParaHox Gene and Dynamic Expression of Hox-like Genes." *Frontiers in Ecology and Evolution* 2: 66.

Fuchs, Georg. 2011. "Alternative Pathways of Carbon Dioxide Fixation: Insights into the Early Evolution of Life?" *Annual Review of Microbiology* 65: 631–58.

Gilbert, W et al. "On the antiquity of introns." *Cell* vol. 46,2 (1986): 151-3.

Gjøvik, Ashley M. "Molecular Characterization of *Trichophyton schoenleinii* (Dermatophyte) Reveals Systematic Affinities with Ancient Eukaryotic Lineages," *The Journal of Decolonized Ecology and Evolution* 1, no. 1 (June 22 2025). DOI 10.5281/zenodo.15717921

Gjøvik, Ashley, M. "Coral Mushrooms and Mushroom Corals Are the Same Organisms: Genetic Confirmation of Close Relationship Between Fungiidae Coral and "coral" Basidiomycota Fungi". *The Journal of Decolonized Ecology and Evolution* 1, no. 1 (June 20, 2025). https://doi.org/10.5281/zenodo.15707211.

Gjøvik, Ashley M. "The Pseudocolus Stinkhorn and the Survival of Ediacaran Architecture: A Genomic & Structural Reassessment of a Misclassified Proto-animal". *The Journal of Decolonized Ecology and Evolution* 1, no. 1 (June 16 2025). https://doi.org/10.5281/zenodo.15675088.

Gjøvik, Ashley M. "High-confidence Genetic Alignment Between Sphaerobolus (the "cannonball" Fungi) & Turritopsis (the "immortal Jellyfish")". *The Journal of Decolonized Ecology and Evolution* 1, no. 1 (June 11, 2025). https://doi.org/10.5281/zenodo.15639701.

Gjøvik, Ashley M. "A Recursive Strategy for Deep Homology Discovery in Genetic Sequence Alignment". *The Journal of Decolonized Ecology and Evolution* 1, no. 1 (June 10, 2025). https://doi.org/10.5281/zenodo.15628144.

Gjøvik, Ashley M. "The Multiple Multicellular Prototype Framework: Silica-based Biomineralization and Early Eukaryotic Diversification". *The Journal of Decolonized Ecology and Evolutio*n 1, no. 1 (June 9, 2025). https://doi.org/10.5281/zenodo.15622349.

Gjøvik, Ashley M. "Extensive Human/siphonophore Genetic Sequence Overlap: A Brief Communication". *The Journal of Decolonized Ecology and Evolution* 1, no. 1 (June 4, 2025). https://doi.org/10.5281/zenodo.15596204.

Gladyshev, Eugene A., Matthew Meselson, and Irina R. Arkhipova. 2008. "Massive Horizontal Gene Transfer in Bdelloid Rotifers." *Science* 320 (5880): 1210–13.

Haddock, Steven H.D., Casey W. Dunn, and Philip R. Pugh. "A Re-Examination of Siphonophore Terminology and Morphology, Applied to the Description of Two New Prayine Species with Remarkable Bio-Optical Properties." *Journal of the Marine Biological Association of the United Kingdom* 85, no. 3 (2005): 695–707.

Hamada, Mayuko, Noriyuki Satoh, and Konstantin Khalturin. "A Reference Genome from the Symbiotic Hydrozoan, Hydra viridissima." *G3: Genes, Genomes, Genetics* 10, no. 11 (2020): 3883–3895.

Hashimoto, Takeshi, et al. 2016. "Extremotolerant Tardigrade Genome and Improved Radiotolerance of Human Cultured Cells by Tardigrade-Unique Protein." *Nature Communications* 7: 12808.

Haas, Robert, and Pablo I Nikel. "Challenges and opportunities in bringing nonbiological atoms to life with synthetic metabolism." *Trends in biotechnology* vol. 41,1 (2023): 27-45.

Helmbrecht, V., Weingart, M., Klein, F., Braun, D., & Orsi, W. D. (2023). White and green rust chimneys accumulate RNA in a ferruginous chemical garden. *Geobiology*, *21*(6), 758–769.

Hyman, Anthony A., Christoph A. Weber, and Frank Jülicher. 2014. "Liquid-Liquid Phase Separation in Biology." *Annual Review of Cell and Developmental Biology* 30: 39–58.

Insua, I. and Montenegro, J. Synthetic Supramolecular Systems in Life-like Materials and Protocell Models. *Chem* 6, 1652–1682, July 9, 2020.

Jumper, John, et al. 2021. "Highly Accurate Protein Structure Prediction with AlphaFold." *Nature* 596 (7873): 583–89.

Khalturin, Konstantin et al. "More than just orphans: are taxonomically-restricted genes important in evolution?." Trends in genetics : TIG vol. 25,9 (2009): 404-13.

Koutsovoulos, Georgios, et al. 2016. "No Evidence for Extensive Horizontal Gene Transfer in the Genome of the Tardigrade *Hypsibius dujardini*." *Proceedings of the National Academy of Sciences* 113 (18): 5053–58.

Laetsch, Dominik R., and Mark L. Blaxter. 2017. "BlobTools: Interrogation of Genome Assemblies." *F1000Research* 6: 1287.

Lane, Nick, and William F. Martin. "The Origin of Membrane Bioenergetics." *Cell* 151, no. 7 (2012): 1406–

1416.

Levine, Michael, and Eric H. Davidson. 2005. "Gene Regulatory Networks for Development." *Proceedings of the National Academy of Sciences* 102 (14): 4936–42.

Mapstone, Gillian M. "Global diversity and review of Siphonophorae (Cnidaria: Hydrozoa)." *PloS one* vol. 9,2 e87737. 6 Feb. 2014.

Markov, Alexander & Anisimov, V. & Korotayev, A.. (2010). Relationship Between Genome Size and Organismal Complexity in the Lineage Leading from Prokaryotes to Mammals. Paleontological Journal. 44. 363-373.

Martin, William, and Michael J. Russell. 2007. "On the Origin of Biochemistry at an Alkaline Hydrothermal Vent." *Philosophical Transactions of the Royal Society B* 362 (1486): 1887–1925.

Miller, David J et al. "Cnidarians and ancestral genetic complexity in the animal kingdom." *Trends in genetics : TIG* vol. 21,10 (2005): 536-9.

Miller, D.J., Hemmrich, G., Ball, E.E. *et al.* The innate immune repertoire in Cnidaria - ancestral complexity and stochastic gene loss. *Genome Biol* **8**, R59 (2007).

Moroz, Leonid L., et al. 2014. "The Ctenophore Genome and the Evolutionary Origins of Neural Systems." *Nature* 510 (7503): 109–14.

Müller, Werner E. G., Matthias Wiens, Renato Batel, Klaus Steffen, Heinz C. Schröder, and Andreas Bonn. "Siliceous Spicules in Marine Demosponges (Example Suberites domuncula)." *Micron* 37, no. 2 (2006): 107–120.

Müller, Werner E. G., Xiaohong Wang, Sergey I. Belikov, Wolfgang Tremel, Ute Schloßmacher, Athanasios Natoli, Dirk Brandt, and Anatoli Boreiko. "Morphology of Sponge Spicules: Silicatein a Structural Protein for Bio-Silica Formation." *Advanced Engineering Materials* 12, no. 5 (2010): 390–396.

Mullis, M. M. et al, Microbial survival mechanisms within serpentinizing Mariana forearc sediments, *FEMS Microbiology Ecology*, Volume 99, Issue 2, February 2023.

Catriona Munro, Felipe Zapata, Mark Howison, Stefan Siebert, Casey W Dunn, Evolution of Gene Expression across Species and Specialized Zooids in Siphonophora, *Molecular Biology and Evolution*, Volume 39, Issue 2, February 2022.

Ochman, Howard, Jeffrey G. Lawrence, and Eduardo A. Groisman. 2000. "Lateral Gene Transfer and the Nature of Bacterial Innovation." *Nature* 405 (6784): 299–304.

Orengo, Christine A., and Janet M. Thornton. 2005. "Protein Families and Their Evolution — A Structural Perspective." *Annual Review of Biochemistry* 74: 867–900.

Patel, Abhishek, Hiral Mistry, Shahar Kadu, Priyanka Kundu, Anjali Joshi, Vasant More, Prasanna Kulkarni, and others. "ATP as a Biological Hydrotrope." *Science* 356, no. 6339 (2017): 753–756.

Perdigão, Nelson, et al. 2015. "Unexpected Features of the Dark Proteome." *Proceedings of the National Academy of Sciences* 112 (52): 15898–903.

Popall, R.M. et al, Cultivating microbial communities from the serpentinite-hosted Prony Bay Hydrothermal Field on different carbon sources in hydrogen-fed bioreactors, bioRxiv 2025.06.11.659039.

Pugh P. R. (1983). Benthic siphonophores: a review of the family Rhodaliidae (Siphonophora, Physonectae)Phil. Trans. R. Soc. Lond. B301165–300/

Remick, Kaleigh A, and John D Helmann. "The elements of life: A biocentric tour of the periodic table." *Advances in microbial physiology* vol. 82 (2023): 1-127.

Rentzsch, Fabian, Heather M. Layden, and Mark Q. Martindale. 2017. "The Cellular and Molecular Basis of Cnidarian Neurogenesis." *Wiley Interdisciplinary Reviews: Developmental Biology* 8 (1): e340.

Russell, Michael J., and A. J. Hall. 1997. "The Emergence of Life from Iron Monosulphide Bubbles at a Submarine Hydrothermal Redox and pH Front." *Journal of the Geological Society* 154 (3): 377–402.

Ryan, Joseph F., et al. 2013. "The Genome of the Ctenophore *Mnemiopsis leidyi* and Its Implications for Cell Type Evolution." *Science* 342 (6164): 1242592.

Sachkova, Maria Y., et al. 2021. "Neuropeptide Repertoire and 3D Anatomy of the Ctenophore Nervous System." *Current Biology* 31 (10): 2249–2260.e6.

Saladino, Raffaele, Giorgia Botta, Bruno Mattia Bizzarri, Ernesto Di Mauro, and Juan Manuel Garcia Ruiz. "A Global Scale Scenario for Prebiotic Chemistry: Silica-Based Self-Assembled Mineral Structures and Formamide." *Biochemistry* 55, no. 19 (2016): 2806–2811.

Schwentner, Martin, and Thomas C G Bosch. "Revisiting the age, evolutionary history and species level diversity of the genus Hydra (Cnidaria: Hydrozoa)." *Molecular phylogenetics and evolution* vol. 91 (2015): 41-55.

Shimizu, Kazuhiro, et al. 1998. "Silicatein: Cathepsin L-like Protein in Sponge Biosilica." *Proceedings of the National Academy of Sciences* 95 (11): 6234–38.

Steele, Robert E et al. "A genomic view of 500 million years of cnidarian evolution." *Trends in genetics : TIG* vol. 27,1

(2011): 7-13.

Strother, Paul K et al. "A possible billion-year-old holozoan with differentiated multicellularity." *Current biology : CB* vol. 31,12 (2021): 2658-2665.e2.

Sugawara, T., Matsuo, M., Toyota, T., "Life" as a dynamic supramolecular system created through constructive approach, *Bulletin of the Chemical Society of Japan*, Volume 98, Issue 1, January 2025.

Tang, Qing, Zhifei Zhang, Ganqing Jiang, Sally E. Walker, Donald E. Canfield, Shuhai Xiao, and Thomas J. Algeo. "Spiculogenesis and Biomineralization in Early Sponge Animals." *Nature Communications* 10, no. 1 (2019): 3348.

Tautz, Diethard, and Tomislav Domazet-Lošo. 2011. "The Evolutionary Origin of Orphan Genes." *Nature Reviews Genetics* 12 (10): 692–702. Wood, Derrick E., and Steven L. Salzberg. 2014. "Kraken: Ultrafast Metagenomic Sequence Classification Using Exact Alignments." *Genome Biology* 15 (3): R46.

Technau, Ulrich et al. "Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians." *Trends in genetics : TIG* vol. 21,12 (2005): 633-9.

Trutschel, Leah R et al. "Investigation of microbial metabolisms in an extremely high pH marine-like terrestrial serpentinizing system: Ney Springs." *The Science of the total environment* vol. 836 (2022): 155492.

Wächtershäuser, Günter. "Origin of Life in an Iron-Sulfur World." Molecular Origins of Life : Assembling Pieces of the Puzzle. Cambridge ; Cambridge University Press, 1998.

Wächtershäuser, Günter. "Groundworks for an Evolutionary Biochemistry: The Iron–Sulfur World." *Progress in Biophysics and Molecular Biology* 58, no. 2 (1992): 85–201.

Wang, Ze-Wen et al. "Fundamentals and application in phytoremediation of an efficient arsenate reducing bacterium Pseudomonas putida ARS1." *Journal of environmental sciences (China)* vol. 137 (2024): 237-244.

Wang, Po-Hsiang, Shota Nishikawa, Shawn Erin McGlynn, and Kosuke Fujishima. "One-Pot De Novo Synthesis of [4Fe-4S] Proteins Using a Recombinant SUF System under Aerobic Conditions." *ACS Synthetic Biology* 12, no. 8 (2023): 2312–2321.

Woese, Carl R. 1998. "The Universal Ancestor." *Proceedings of the National Academy of Sciences* 95 (12): 6854–59.

Wong, Wai Yee, Oleg Simakov, Diane M. Bridge, Paulyn Cartwright, Anthony J. Bellantuono, Annika Kuhn, Thomas W. Holstein, Charles N. David, Robert E. Steele, and Daniel E. Martínez. "Expansion of a Single Transposable Element Family Is Associated with Genome-Size Increase and Radiation in the Genus Hydra." *Proceedings of the National Academy of Sciences* 116, no. 46 (2019): 22915–22917.

Wood, Derrick E, and Steven L Salzberg. "Kraken: ultrafast metagenomic sequence classification using exact alignments." *Genome biology* vol. 15,3 R46. 3 Mar. 2014.

Wright, Peter E., and H. Jane Dyson. 2015. "Intrinsically Disordered Proteins in Cellular Signalling and Regulation." *Nature Reviews Molecular Cell Biology* 16 (1): 18–29.

Wu, Shuangju et al. "Signature Arsenic Detoxification Pathways in *Halomonas* sp. Strain GFAJ-1." *mBio* vol. 9,3 e00515-18. 1 May. 2018.

Ying, Hua et al. "The Role of DNA Methylation in Genome Defense in Cnidaria and Other Invertebrates." *Molecular biology and evolution* vol. 39,2 (2022): msac018.

Zambrano, P., Chen, X., Kriebisch, C. M. E., Kriebisch, B. A. K., Zozulia, O., & Boekhoven, J. (2024). Chemically Driven Division of Protocells by Membrane Budding. *Journal of the American Chemical Society*, *146*(49), 33359–33367.