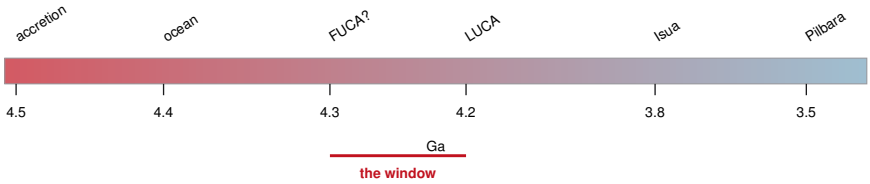
VOLUME I

The First Filter

Dissipative origins, extinctions as engines of creation,
and where the evidence already is

Meow-Ludo

meow@meow-meow.io • +61 491 266 084



Private working document

v2.2 • Volume I of II

27 July 2026

Claims are confidence-tagged. Key overleaf.

How to read this volume

This is the origins companion to Volume I. Where Volume I asked *what can light tell us about microbial biology*, this volume asks a narrower and older question: *what was the transition that produced the last universal common ancestor, and where is the evidence for it now*.

The material is reconstructed from a research archive spanning project documents and roughly 1.6 million characters of working conversation. Nothing has been discarded. Where the archive contained a claim that is underdeveloped rather than wrong, it appears here flagged as such rather than quietly dropped — the register in §16 exists for exactly that purpose.

Confidence key

Same system as Volume I, with one addition.

Mark	Meaning	What to do with it
✓	Verified in session	Checked against a source while writing. Safe to build on.
○	From memory	Confident it is right and findable; not checked. Verify before publication.
≈	My estimate	Fermi calculation from stated inputs. Arithmetic checkable; assumptions soft.
⋆	Novel proposal	No literature known to me. Search before investing.
▷	From the archive	Your prior work, restated. Provenance is the archive, not the literature — these are the framework's own commitments.

On register

One editorial decision, stated openly because it affects every page.

The archive argues several of these claims at very high confidence. This volume re-argues them at the confidence the evidence supports, which is in most cases lower — and in every case where I have done that, the claim comes out *stronger*, because expressed certainty is the first thing a reader attacks and the last thing these ideas need. The inversion method, the two-clocks prediction, the vent geometry and the lunar archive all stand on their own evidence. They do not require anyone to be sure.

Where the archive contains a felt-certainty statement, it is preserved in §16 as a record of what was believed and why, separated from the argued claims. That separation is the point.

How to cite

Meow-Ludo (2026). *The first filter: dissipative origins, extinctions as engines of creation, and where the evidence already is.* Working document v2.2, 27 July 2026. meow@meow-meow.io

Cite the version and date. Claims carry confidence marks; an unmarked citation of a ○ or ⋆ claim misrepresents its standing.

Contents

List of Figures

PART I

The framework, stated

§1

Eleven commitments

The framework is not a single hypothesis. It is a stack of eleven claims of very different character — some are methodological, some physical, some historical, some are predictions. They have different evidential status and they support each other unevenly. Setting them out as a dependency graph is the first useful thing that can be done to them, because it shows immediately which claims are load-bearing and which are decorative.

1.1 Reading the graph

Three structural facts, which matter more than any individual claim.

The framework is top-heavy in the right way. C1 (dissipative adaptation) and C2 (the inversion method) carry everything. If either fails, the rest is decoration. That is a virtue, not a weakness — it means the framework is concentrated rather than diffuse, and it tells you exactly where an opponent should attack and therefore where the defence needs to be strongest.

C10 is a convergence point. The two-clocks prediction is reachable from the phase-boundary claim, from the vent geometry, and from both consequence branches. Predictions that multiple independent parts of a framework agree on are the ones worth testing, because a positive result supports the whole structure and a negative one is genuinely informative about which branch failed.

C11 is detachable. The lunar archive argument works even if you doubt C3–C7 entirely. "Hadean Earth material is preserved on the Moon and can be searched for early-metabolic isotopic signatures" is a standalone proposition with its own evidence base. That makes it the natural first publication and the natural first funding request, because it does not require anyone to buy the whole framework first.

1.2 The commitments in full

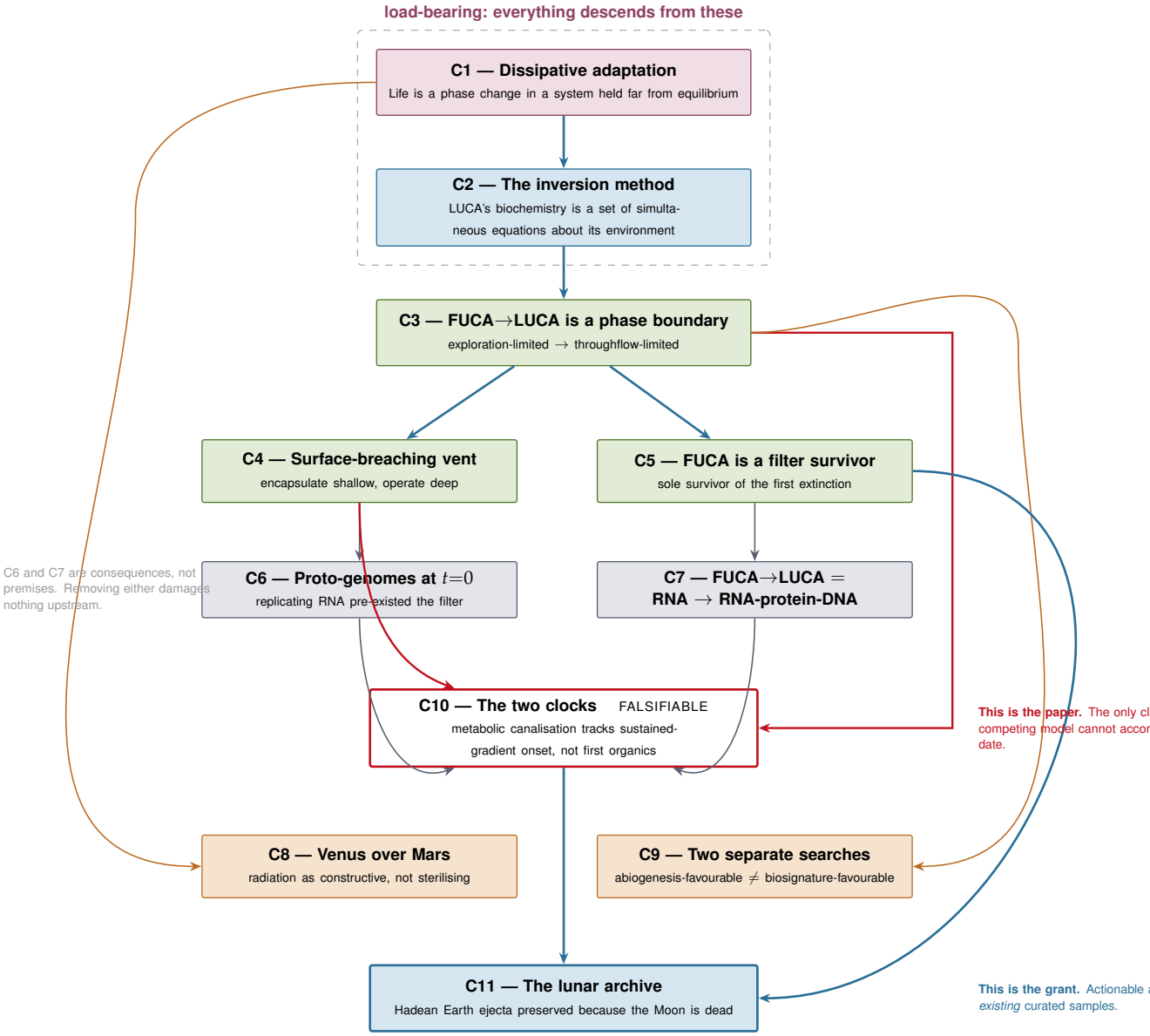


Figure 1.1: The framework as a dependency graph. Arrows run from premise to consequence. Two observations follow immediately: the whole structure rests on C1–C2, so those are where scrutiny should concentrate; and C10 (the two-clocks prediction) is the convergence point of four independent branches, which is why it is the strongest single test. C6–C7 are consequences rather than premises — attractive, but nothing upstream depends on them.

	Claim	Character	Status
C1	Life is a phase change in a system far from equilibrium; configurations that better dissipate a sustained drive are favoured	Physical premise	England's dissipative adaptation. Directionally sound, has serious critics.▷○
C2	LUCA's conserved biochemistry constrains its environment; solve backwards from a known output to the regime that makes it the attractor	Method	**The core contribution.** Defensible now.▷
C3	FUCA→LUCA is the switch from exploration-limited to throughflow-limited selection	Mechanism	Strong; reframes a historical question as a physical one.▷★
C4	Surface-breaching vent: encapsulation in a shallow regime, operation in a deep one	Site hypothesis	Dissolves the vent/hot-spring dichotomy. Bench-testable.▷★
C5	**(revised §27)** FUCA-class is the plural pre-translation population; FUCA 1.0 is the translation-sweep survivor	Historical	Two objects, distinguished by version. Resolves the class/individual contradiction.▷
C6	Self-replicating proto-genomes existed before the filter, providing the candidate pool	Consequence	Plausible; not load-bearing.▷
C7	**(revised §28)** Four sweeps: sugar/glycolipid → RNA → protein → DNA, carrier reactivity falling with chao; a DNA genome defines LUCA	**Premise**	Promoted twice. Now carries P8 and the merger generalisation.▷★
C8	Venus is a better abiogenesis target than Mars; radiation is constructive given an atmosphere	Search consequence	**Underdeveloped.** Needs a mechanism — supplied in Part VI.▷
C9	Abiogenesis-favourable and biosignature-favourable target lists are different searches	Search consequence	Genuinely useful distinction; the field does conflate them.▷★
C10	Metabolic canalisation tracks sustained-serpentinisation onset, decoupled from the first organic record	**Prediction**	**The falsifiable core.** Testable against existing stratigraphy.▷★
C11	Hadean Earth ejecta is preserved on the Moon and carries isotopic biosignatures	Evidence route	**Most actionable.** Existing curated samples.▷✓

§2

The inversion method

2.1 The move

The question "what was FUCA" is badly posed, because FUCA leaves no direct trace by construction — anything that could witness it postdates it. The question "what regime produces LUCA as its attractor" is well posed, because LUCA is reconstructed.

That substitution is the method. It converts an unanswerable historical question into a constrained inverse problem.

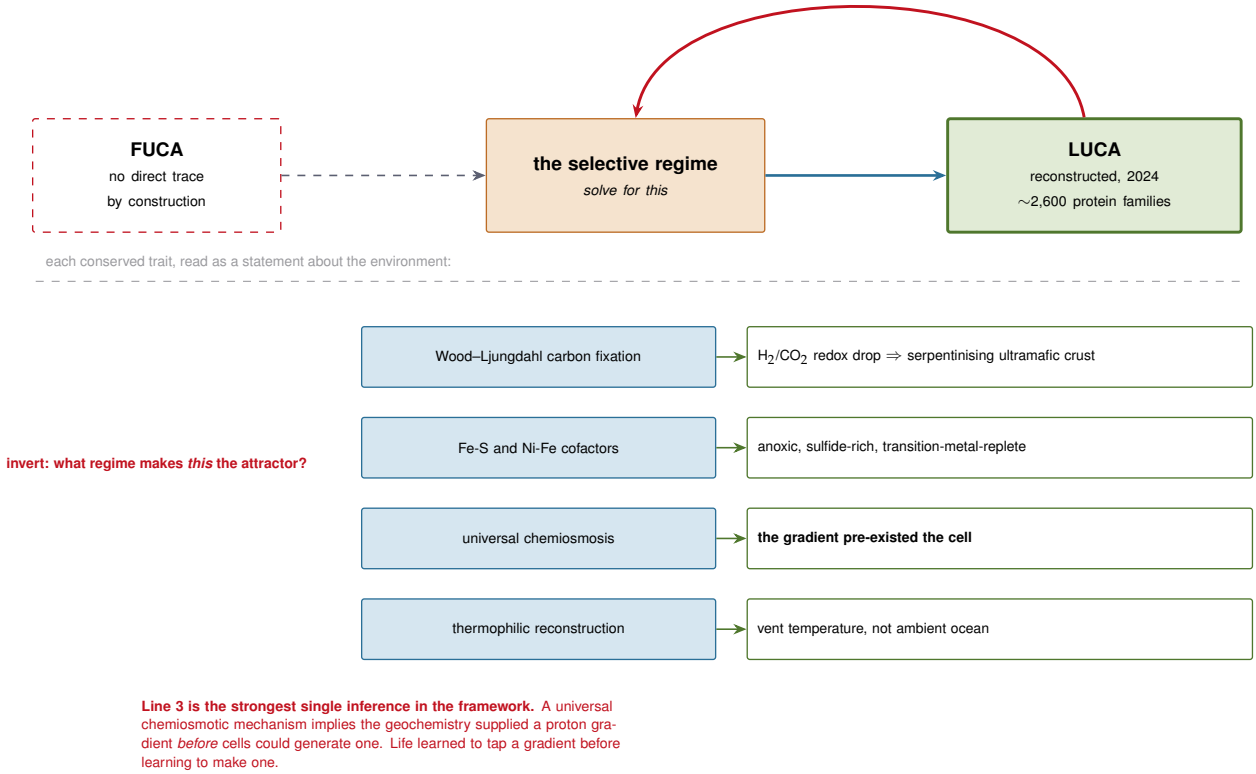


Figure 2.1: Solving backwards from a known output. FUCA is unreachable directly; the regime that canalises toward LUCA is not. Each conserved LUCA trait is treated as a fossilised constraint on the environment that produced it, and the constraints are intersected.

2.2 The reconstruction being inverted

The method's leverage is entirely proportional to the quality of the LUCA reconstruction, so it is worth being explicit about what is being assumed.

The 2024 synthesis places LUCA at roughly 4.2 Ga, with on the order of 2,600 protein families, an acetogen-like metabolism running the Wood–Ljungdahl pathway, Fe-S and Ni-Fe cofactors at catalytic cores, a chemiosmotic membrane, and — notably — an early immune system implying it already existed in an ecosystem with viruses.[○] [Moody et al. 2024, Nat Ecol Evol — verify family count and date range]

The date dependency, stated once

The 4.2 Ga figure is clock-derived with wide priors and it is doing real work in the compressed-window version of the argument (life appearing almost immediately after the ocean condenses, leaving almost no time, therefore a phase transition rather than a slow ramp).
Two versions of the argument exist and they have different dependencies:
Compressed-window version. LUCA at 4.2 Ga leaves ∼200 Myr from ocean condensation. That is too fast for gradualism, so canalisation must be a phase transition. **Depends on the date.**
Gradient-handoff version. Regardless of absolute timing, the transition is the handoff from a variety-generating regime to a throughflow-filtering regime, and canalisation is fast *relative to the preceding exploration phase* because the filter is sharp. **Does not depend on the date.**
Both are written in this volume. The gradient-handoff version is the one to lead with; the compressed-window version is a bonus if the date holds. If LUCA slides younger, you lose rhetorical force and keep the mechanism.

2.3 Why the method is the contribution

Reasoning from LUCA's traits to LUCA's environment is not itself novel — people do it one feature at a time, informally, usually to support a site preference they already hold.○

What is different here is running it *systematically and simultaneously*: every conserved core trait treated as a constraint, the constraints intersected, and the resulting regime taken as the answer rather than as support for a prior. That inverts the usual epistemic direction of origins arguments, which typically start from a favoured setting and ask whether LUCA is compatible with it.

The consequence is that the method generates predictions its author did not choose. C10 (the two clocks) falls out of the intersection rather than being posited. That is the signature of a method doing real work.

§3

The phase boundary

3.1 Two regimes, one axis

The central mechanical claim is that FUCA-era and LUCA-era systems are optimising different quantities, and that the transition between them is a change in which resource is scarce.

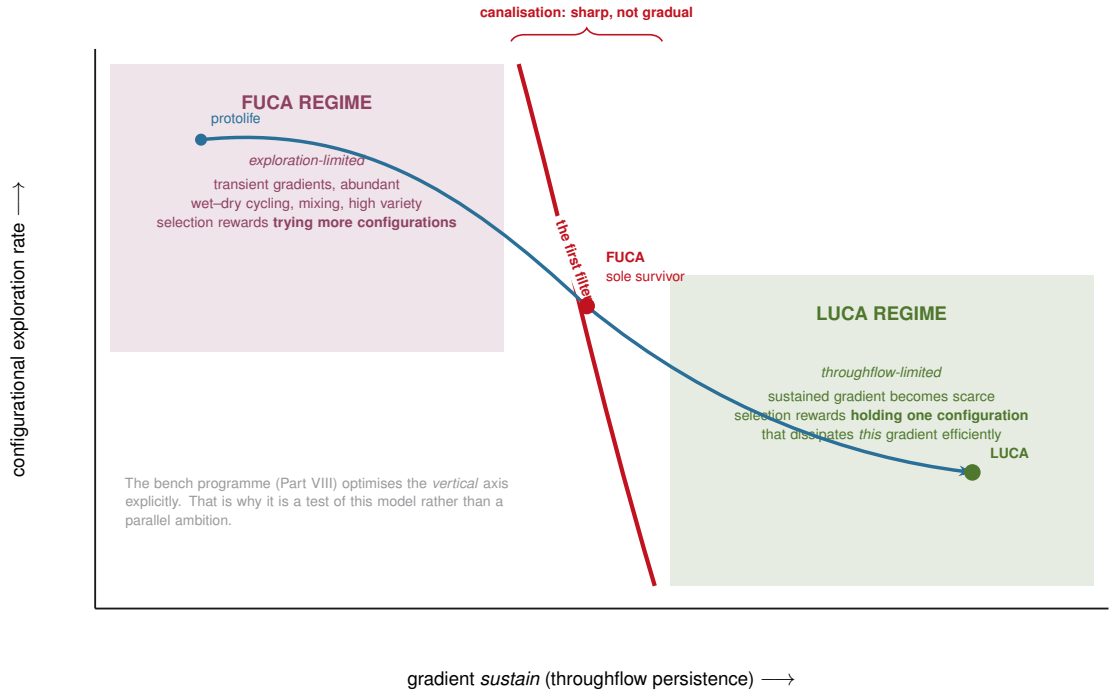


Figure 3.1: FUCA → LUCA as a change in the limiting resource. While gradients are transient and abundant, selection rewards configurational throughput. When sustained gradients become the scarce resource, selection switches to rewarding persistence on one specific gradient topology. LUCA is the first configuration better off persisting-and-dissipating than exploring — which is why canalisation is sharp.

3.2 Why this reframing earns its keep

It converts a historical question (*what happened*) into a physical one (*which resource was limiting*), and physical questions have signatures.

Specifically, it predicts that the transition should track the establishment of *sustained* geochemistry rather than the first appearance of organics — two different clocks, running independently. That is C10, and it is developed as a testable stratigraphic claim in Part IV.

It also unifies two things that otherwise sit awkwardly together. The wet–dry engine (Damer–Deamer; high variety, polymerisation, mixing) and the alkaline-vent gradient (Russell–Lane; sustained throughflow, chemiosmosis) have been treated as rival origin sites for two decades. Under this framing they are not rivals — they are **the generator and the filter**, operating in sequence on the same axis. The vent-versus-hot-spring debate dissolves because both are required, at different stages.

PREDICTION — P1 — the shape of the base of the tree

If C5 holds — FUCA is the sole survivor of a sharp filter — then the phylogenetic signature is a **bottleneck**: low effective diversity immediately ancestral to LUCA, and a star-like radiation immediately after. A gradual origin predicts a bushy, deeply-structured base instead.
This is testable against existing phylogenomic data, and it is a different test from anything in the rock record.⋆Developed in §8.

§4

Method: Phase Zero and geological reasoning

Two methodological claims underlie this volume. Both are structural rather than biographical, and both are testable against the document itself.

4.1 The four phases

Research on a problem this large has a generation stage that no methods section ever describes, because it is embarrassing and because it has no natural boundary marker. Naming it is useful, and the absence of the name is a real failure mode.

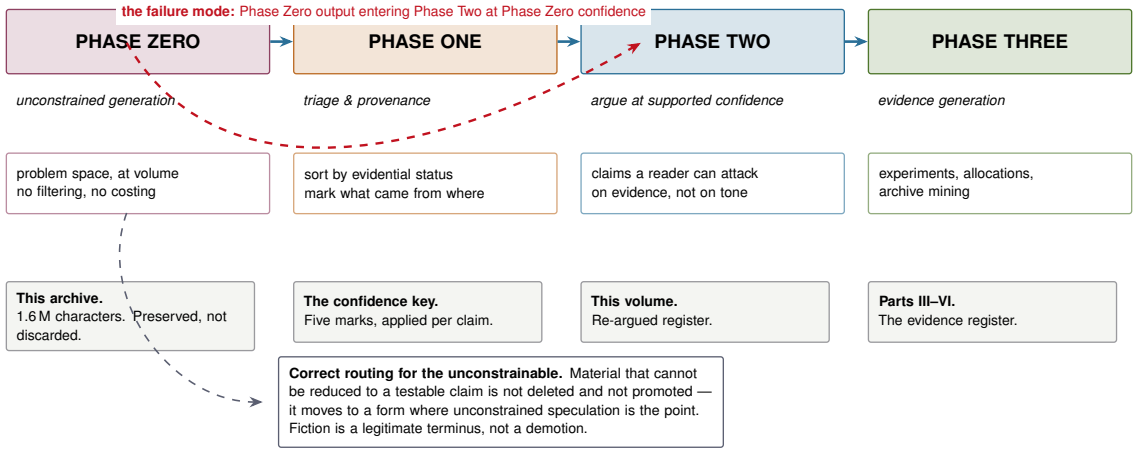


Figure 4.1: Phase Zero as a named stage. The stage exists in every research programme; what is usually missing is the boundary marker between it and argued work. The confidence key in this volume *is* that marker, applied at claim granularity.

4.2 Geological reasoning applied to a biological object

The second claim is about why the inversion of Chapter 2 was available but unmade.

Geology's characteristic operation is reading a depositional environment off a preserved product. Nobody observed the tidal flat; the desiccation cracks, ripple marks and evaporite laminae are read backwards to it. Facies analysis is *professionalised inverse inference* — an entire discipline organised around the move of solving for conditions from the artefact.

Molecular biology's characteristic operation is forward: mechanism, pathway, function. Phylogenetics reconstructs ancestry, but reconstructs the *organism*, not the *regime*.

The methodological thesis

Astrobiology contains substantial geological personnel and substantial biological reasoning. What it under-uses is *geological reasoning applied to the biological object*: treating LUCA's conserved biochemistry the way a sedimentologist treats a ripple mark — as a preserved feature whose function is to constrain the environment that made it.
C2 is that operation. Wood–Ljungdahl is a desiccation crack. Universal chemiosmosis

is an evaporite lamina. Each is a fossilised statement about conditions, and the correct question is not what the organism was doing but *what regime deposits this*.
This explains the gap without appealing to anyone's individual insight: the move requires holding a geological epistemology over a biochemical dataset, and the disciplinary structure does not routinely place those in the same head.⋆

The practical consequence is that this framework should be argued in geological register — facies, constraint, preservation, provenance — rather than in biological register. Part IV does this literally, and it is why the strongest test (C10) is stratigraphic rather than phylogenetic.

PART II

The transition

§5

The surface-breaching vent

5.1 The encapsulation problem

The vent-versus-hot-spring dispute is not a matter of taste. It is a genuine incompatibility between two requirements, and every origins model has had to sacrifice one.

Vents supply the gradient and forbid the bag. Alkaline hydrothermal systems give a sustained proton-motive force across mineral membranes, a redox drop between H_2 and CO_2, and mineral catalysis — everything LUCA's biochemistry says the environment provided. But fatty acid vesicles, the only prebiotically plausible compartment, are destroyed by exactly what vents have: high temperature, high ionic strength, and above all divalent cations. Seawater carries roughly 50 mM Mg^{2+} and 10 mM Ca^{2+}; fatty acid membranes flocculate at low-millimolar concentrations.[○] [Monnard & Deamer on divalent-cation disruption — verify thresholds]

Hot springs supply the bag and lose the gradient. Dilute, low-ionic-strength terrestrial pools permit vesicle self-assembly, and the wet–dry cycle does polymerisation work no submerged system can. But the gradient is transient, the setting is fragile, and there is no sustained throughflow to select on.

5.2 The two-regime resolution

The framework's move is to stop treating these as competing *sites* and start treating them as sequential *regimes* experienced by the same system.[▷]

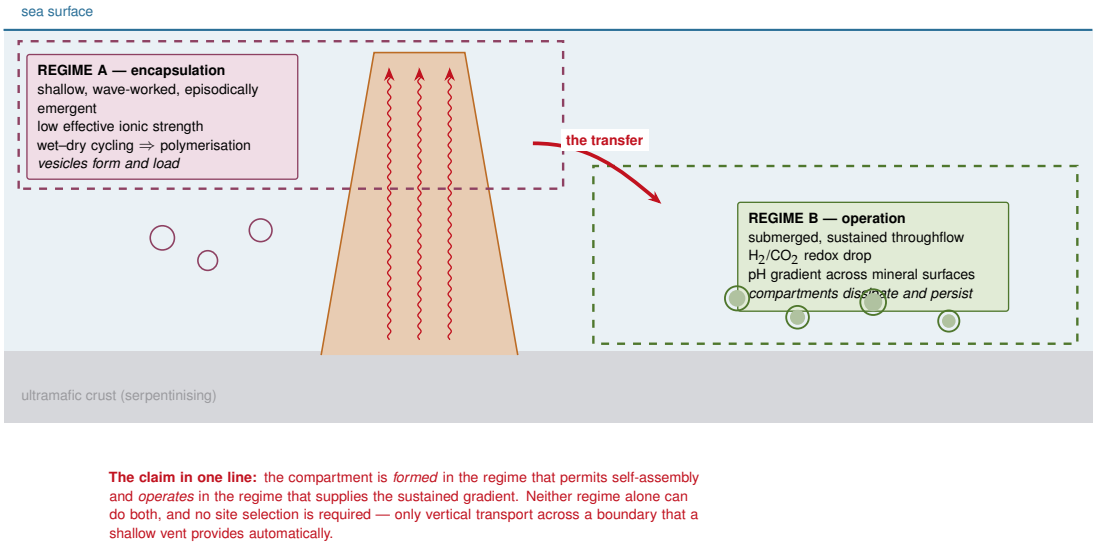


Figure 5.1: Encapsulation and operation as sequential regimes rather than competing sites. The divalent-cation objection to vent models applies to vesicle *formation*, not necessarily to vesicle *persistence* once formed and loaded. A vent shallow enough to breach into a wave-worked, episodically emergent zone supplies both regimes with a metres-scale transport path between them.

5.3 Why this is bench-testable, and the experiment

The resolution rests on an empirical claim with a clean laboratory form which, as far as I can establish, has not been directly tested.⋆

EVIDENCE PLAY — E1 — the transfer-survival experiment

Claim under test. Fatty acid vesicles formed under low-ionic-strength conditions and subsequently transferred into seawater-analogue divalent-cation concentrations survive at rates materially above vesicles *formed* in those conditions.
Why it might be true. Formation and persistence are different physical processes. Self-assembly requires the amphiphile to reach critical aggregate concentration in a regime where headgroup repulsion is not screened. A pre-formed, osmotically loaded, possibly peptide- or RNA-stabilised bilayer is a different object facing a different question.
Design. Form oleic-acid / mixed-fatty-acid vesicles at low ionic strength encapsulating a fluorescent marker. Split. Transfer one arm stepwise into increasing Mg^{2+}/Ca^{2+}; form the control arm directly at each concentration. Measure retention by encapsulated-marker fluorescence and vesicle count.
Modifiers worth crossing in. Co-encapsulated short peptides; RNA; amphiphile chain-length mixtures; and — the interesting one — prior wet–dry cycling, which the model says happens *before* transfer.
Cost. Low. Within reach of the existing bench; does not need the droplet platform. Probably the cheapest experiment in either volume with a direct bearing on a load-bearing claim.
Symmetric value. A positive result makes C4 a mechanism rather than a proposal. A negative result is a real constraint on the model, and worth knowing early.

PREDICTION — P2 — the facies signature

The two-regime geometry makes a prediction no competing model makes, and it is geological rather than biochemical.
Deep-vent models predict early biosignatures in deep-water facies. Hot-spring models predict them in terrestrial facies with no marine context. **This model uniquely predicts them at the join** — vent-derived mineralogy (serpentinite, sulfide chimney textures, carbonate) interfingered with shallow-water indicators (wave reworking, desiccation, intertidal structures) in the same succession.▷
The search image can be built in well-preserved Paleoarchean rock (Pilbara, Barberton) before hunting it in Eoarchean rock (Isua).

§6

The first filter

6.1 Repeated genesis under bombardment

The framework's historical commitment (C5) is that FUCA is not the *first* living thing but the *only survivor* of a population of living things — and that the Hadean impact regime both generated that population repeatedly and then culled it.▷

This is less exotic than it sounds once the impact physics is taken seriously.

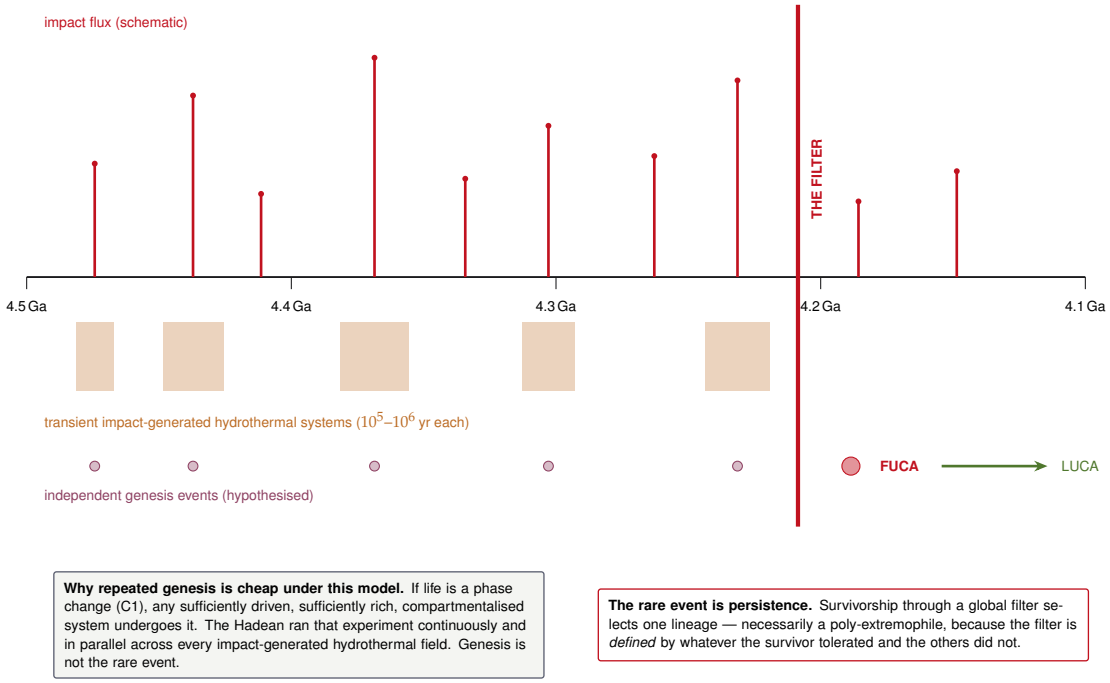


Figure 6.1: Repeated genesis, single survivor. Impact events generate hydrothermal systems persisting 10^5–10^6 years — long-lived on chemical timescales, transient on geological ones. Under a phase-transition model of origins, each is an independent trial. FUCA is then not the first life but the one lineage that survived a global culling event, which is why its reconstructed descendant is an extremophile.

6.2 The drilled analogue nobody uses for this

The impact-generated-hydrothermal component of this argument is not speculative and does not require Hadean rock, because a large impact-generated hydrothermal system has been drilled, cored and studied in detail.

EVIDENCE PLAY — E2 — Chicxulub as a process analogue

IODP–ICDP Expedition 364 drilled the Chicxulub peak ring in 2016 and recovered core through the post-impact hydrothermal system. The published work characterises the hydrothermal cell's extent, duration, mineralogy and subsequent microbial colonisation.○ [Kring et al. on the Chicxulub hydrothermal system; Exp. 364 proceedings — verify

duration estimates]
Why it matters here. It converts "impacts make habitable hydrothermal systems" from a plausibility argument into a measured one, with real numbers for duration, volume, temperature structure and mineral assemblage — and it is *existing, published, open data*.
What to extract. The duration distribution and the mineral assemblage. Duration sets the number of chemical generations available per event. The assemblage tells you whether the Fe–Ni–S catalytic inventory the model requires is what impact systems actually produce.
Status. Pure archive work. No allocation, no fieldwork. Scaling to Hadean-magnitude impacts is an argument to make carefully, but the process is documented.⋆

6.3 What a bottleneck does and does not imply

One distinction has to be made explicitly, because the obvious objection lands otherwise.

A lineage bottleneck is not a simple ancestor

The 2024 reconstruction gives LUCA on the order of 2,600 protein families — a substantial, well-equipped organism, not a bare replicator.○ That might look like it contradicts a sharp filter.
It does not. A bottleneck in *lineages* says nothing about the complexity of the surviving one. The filter removes diversity; it does not simplify the survivor. Indeed the model *predicts* a complex survivor: whatever passed a global extreme-conditions filter did so because it was already a well-defended poly-extremophile, and poly-extremophily is expensive in gene count.
The two claims are compatible, and the combination is more informative than either: **rich genome, narrow ancestry.**

PREDICTION — P1 (restated for test) — narrow ancestry, rich genome

If C5 holds, the phylogenomic signature is a *star-like* radiation immediately following LUCA with little resolvable deep structure beneath it, combined with a LUCA gene set whose family origination times cluster rather than smear.
A gradual, non-bottlenecked origin predicts the opposite: resolvable deep structure and a smeared origination-time distribution.
Testable against existing phylogenomic datasets, and entirely independent of both the rock record and the lunar route — which is what makes it valuable.⋆

6.4 Viruses before LUCA — an existing result that supports C6

C6 (self-replicating proto-genomes existed before the filter, supplying the candidate pool) is listed in Chapter 1 as a consequence rather than a premise. It has more support than that status suggests.

The 2024 LUCA reconstruction infers that LUCA possessed defence systems.○ [Moody et al. 2024 — verify the defence-system inference] A defence system is only worth carrying if there is something to defend against. That is a published inference that parasitic replicators predate LUCA — which is C6, arrived at independently and by other people.

The framework should claim this. It is the only one of the eleven commitments with direct external corroboration already in the literature.

§7

What FUCA was

Pulling the constraints together. Each line is derived, not asserted — the provenance is given.

Property	Constraint	Derived from
Hyperthermophilic	LUCA's thermophilic reconstruction is a *relaxation* from a hotter ancestor, not an acquisition	C2 inversion; vent setting
Poly-extremophile	Defined by the filter it passed: pressure, temperature, radiation, pH swing, desiccation	C5; filter logic
RNA-based, pre-DNA	C7; the RNA→RNA-protein-DNA transition *is* FUCA→LUCA	C7
Compartmentalised	Required for individuation and selection; fatty acid vesicles	C1, C4
Exploration-optimised	Competing on configurational throughput, not efficiency	C3
Metabolically opportunistic	Not yet canalised on Wood–Ljungdahl; that is LUCA's signature, not FUCA's	C3, C2
Not necessarily simple	Filter survival is expensive	§6

The one inference worth arguing hardest

FUCA should be *less* metabolically canalised than LUCA, not more.
This inverts the usual intuition, which pictures the earlier organism as simpler and therefore more constrained. Under C3 the opposite holds: FUCA-era systems are exploration-optimised, so they should be metabolically promiscuous, running whatever chemistry the local gradient offered. LUCA is the *specialist* — the first lineage to commit to one gradient topology because committing had become the winning strategy.
If that is right, then the search image for pre-LUCA biosignatures should not be "a simpler Wood–Ljungdahl" but "a messier, less isotopically consistent carbon signal" — which is precisely the two-clocks prediction (C10) seen from the organism's side rather than the rock's.*

PART III

Molecular evidence you can generate now

§8

The depth limit

8.1 Why rRNA cannot do this job

The standard instrument for deep phylogeny is the small-subunit ribosomal RNA, and it is the wrong tool here for a structural reason rather than a technical one.

The ribosome is already a LUCA-level artefact — arguably a late-pre-LUCA one. By the time you have a coupled peptidyl transferase centre, a decoding centre and some fifty ribosomal proteins, you are looking at a machine that *postdates* the transition in question. rRNA phylogenetics bottoms out at LUCA because the molecule cannot witness its own assembly.

That is not a counsel of despair. It redirects the search toward a specific class of evidence: molecules and features that are **internally stratified**, so that layers within a single molecule record different depths. Those exist, and this Part is a catalogue of them.

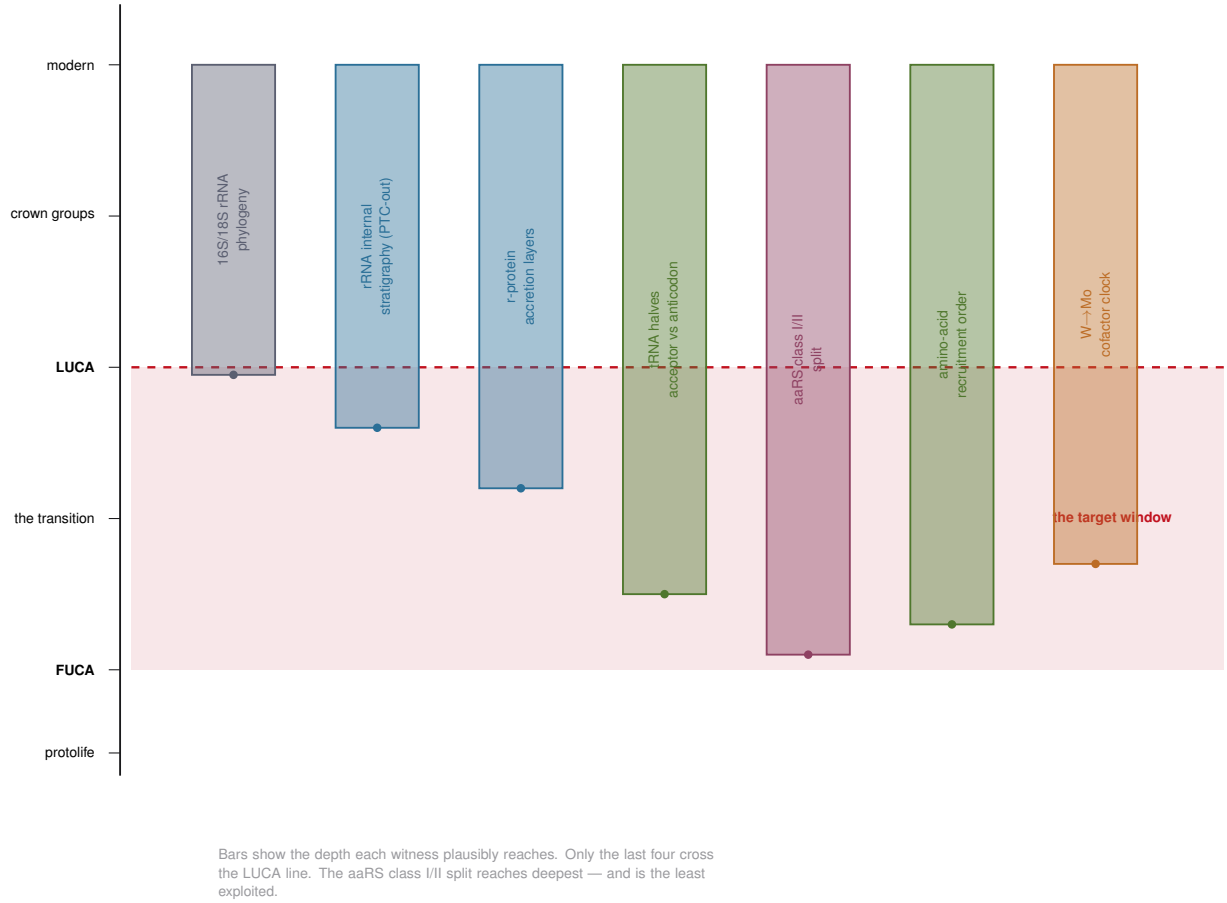


Figure 8.1: How far back each molecular witness reaches. Standard rRNA phylogeny stops at LUCA by construction. Internally stratified features — the ribosome's accretion layers, tRNA's two halves, the aminoacyl-tRNA synthetase class split, amino-acid recruitment order — can in principle see through it, because each records a sequence of events inside a single surviving structure.

§9

tRNA: the deepest structural witness

9.1 The claim

The archive states it flatly: *tRNA predates the ribosome*.▷ That is defensible, and the supporting structure is stronger than it is usually given credit for.

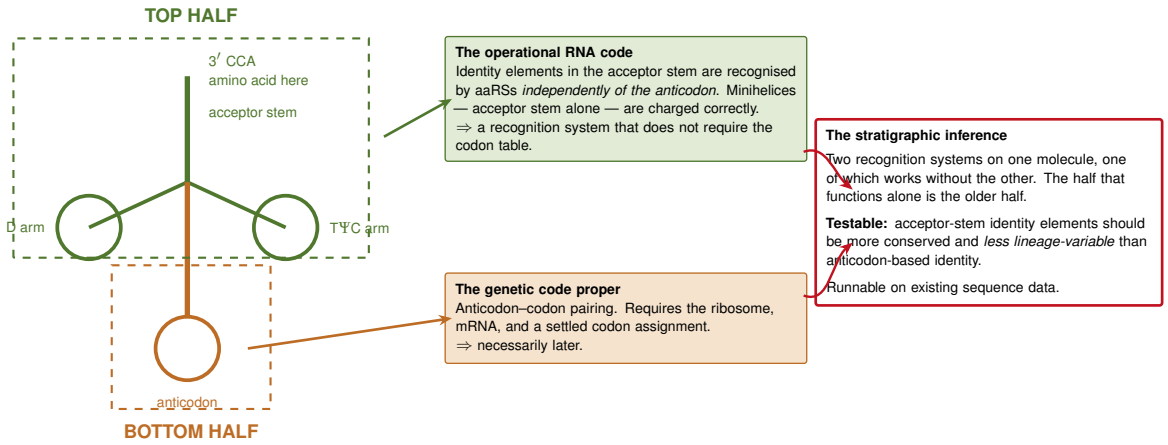


Figure 9.1: tRNA carries two recognition systems of different ages. The acceptor stem's "operational code" is read by synthetases without reference to the anticodon — minihelix substrates are charged correctly. A system that functions without the codon table cannot depend on it, and therefore predates it. That makes tRNA internally stratified, and readable.

9.2 The supporting structure

Modular origin. The cloverleaf is widely argued to derive from duplication of a smaller hairpin, giving a top half (acceptor stem + TΨC) and a bottom half (anticodon + D arm) with separate histories.○ [Di Giulio, tRNA duplication hypothesis — verify formulation]

The operational RNA code. Identity elements in the acceptor stem are sufficient for correct aminoacylation by several synthetases; minihelix substrates lacking the anticodon are charged appropriately. This is a recognition system that does not require the codon table, and therefore cannot depend on it.○ [Schimmel & Ribas de Pouplana, "second genetic code"]

EVIDENCE PLAY — E3 — lineage-variability of the two codes

Study. For each amino acid, compile acceptor-stem identity elements and anticodon-based identity elements across the three domains. Measure conservation and, more informatively, *across-domain variance*.
Prediction (framework). Acceptor-stem identity is deeper: higher conservation, lower cross-domain variance. Anticodon identity is shallower and more lineage-labile.
Data. Entirely existing — tRNA databases, structural data, published identity-element compilations.
Cost. A compute job and a careful literature compilation. No wet lab.★

§10

The synthetase class split

This is, in my assessment, the single most under-exploited pre-LUCA relic available, and the place where the framework could generate a genuinely novel result fastest.*

10.1 The observation

Aminoacyl-tRNA synthetases come in two classes that share no common ancestry detectable by sequence or fold. Both are universal. Both are essential. Both were in LUCA.

	Class I	Class II
Core fold	Rossmann dinucleotide-binding	antiparallel β-sheet
Signature motifs	HIGH, KMSKS	three class-defining motifs
Approaches tRNA from	minor groove	major groove
Acylates	2′-OH	3′-OH (mostly)
Oligomeric state	mostly monomeric	mostly dimeric

The striking fact is the last two rows of geometry: **the two classes bind opposite faces of the same acceptor stem.** They are complementary in a way that looks designed, and no mechanism of divergent evolution produces two unrelated folds that partition a substrate that cleanly.

10.2 The Rodin–Ohno hypothesis

The proposed explanation is that the two classes are encoded on *complementary strands of a single ancestral gene* — sense and antisense readings of one duplex.° [Rodin & Ohno 1995; Carter & Wills subsequent work] If that is right, then the deepest reconstructable object in molecular biology is a single ancestral double helix whose two readings became the two halves of translation.

That is a FUCA-era artefact by any reasonable definition, and it has never been decisively resolved.

EVIDENCE PLAY — E4 — the Rodin–Ohno structure-prediction test

Why now. The hypothesis was proposed when its central claim — that both readings of one duplex fold into functional, class-appropriate structures — could not be evaluated. It can be evaluated now.
Design. Construct candidate ancestral complementary-coding pairs from class-I and class-II urzyme alignments. For each, translate both strands in the proposed relative frame. Run structure prediction on both products. Score against class-I (Rossmann) and class-II (antiparallel β) reference folds. Compare to a null built from shuffled and from unrelated complementary pairs.
What a positive result means. Support for a single ancestral gene at the root of translation — and, in framework terms, a directly reconstructed FUCA-era molecular object.
What a negative result means. Also useful: it constrains the strongest existing hypothesis

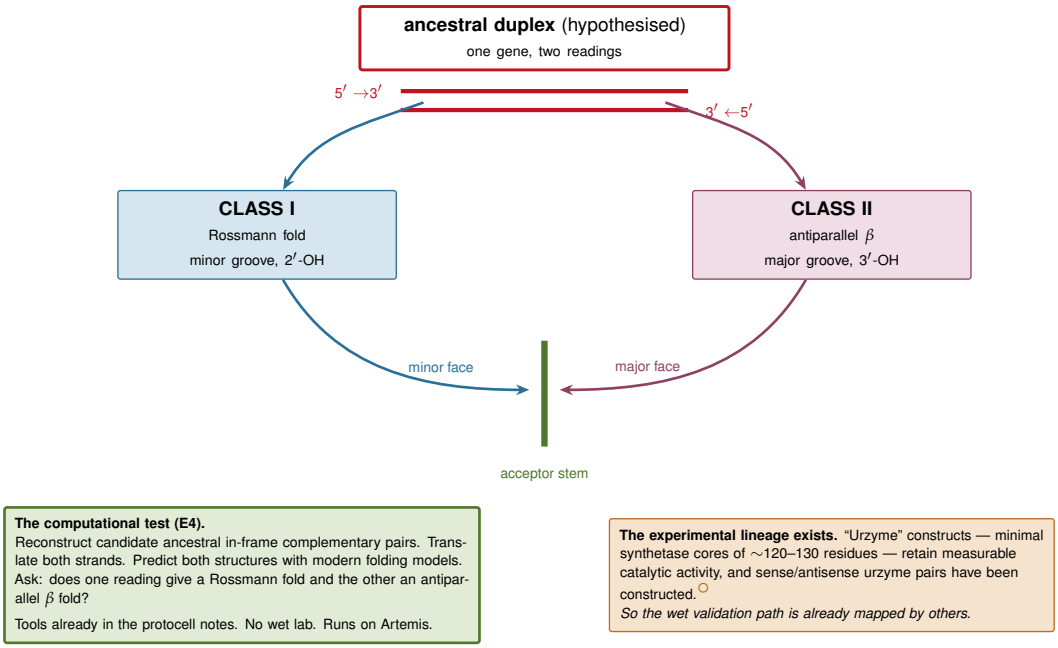


Figure 10.1: Two unrelated folds partitioning one substrate. The classes approach opposite faces of the acceptor stem and acylate different hydroxyls. The Rodin–Ohno explanation — complementary strands of one ancestral gene — would make this the deepest reconstructable event in molecular biology, and it is directly testable with structure-prediction tools that did not exist when the hypothesis was proposed.

about the deepest event in molecular evolution.

Resources. Structure prediction on an 8 GB GPU is the binding constraint; urzyme-length constructs (∼130 residues) are small enough that this is comfortable. Everything else is public sequence data.⋆≈

§11

Ribosomal stratigraphy and a methods paper

11.1 The onion

The ribosome is not a single-age object. Accretion models resolve it into concentric layers, with the peptidyl transferase centre as the innermost and oldest, and successive expansion segments added outward.[○] [Petrov, Williams et al. — accretion of expansion segments]

The archive's own reading of the primary literature supports this quantitatively: proximity to the PTC dominates conservation, with 85.7% of residues within 10 Å of the PTC in the slowest substitution-rate category, an effect that does not saturate but extends to the surface; and the PTC region is built overwhelmingly of junctions (71% within 10 Å) rather than stems.[▷○]

That is a usable stratigraphy. It lets you stratify the ribosome by age and interrogate the deepest layer separately — which is the closest rRNA can get to the target window.

11.2 The systematic error worth publishing on its own

Buried in that same reading is a methodological finding that is more immediately publishable than anything else in this Part.[▷]

The universal-substitution-model problem

The standard assumption in structural RNA phylogenetics is that compensatory mutation in paired stems dominates, and substitution models are built accordingly.
That assumption is **lineage-specific, not universal.** In eukaryotes, loops evolve roughly 1.37-fold faster than stems; in bacteria and archaea, stems dominate.[▷○]
The consequence: *any rooting or deep-divergence reconstruction that applies a single universal RNA substitution model across domains is systematically mis-weighting the signal* — and doing so in a way that is invisible unless you look for it. Every attempt to use rRNA to reach toward or past LUCA is exposed to this.

EVIDENCE PLAY — E5 — the methods paper

Contribution. Quantify the magnitude of the bias introduced by universal substitution models in deep rRNA phylogenetics, and provide a domain-partitioned alternative.
Why it is the fastest publication in this volume. It is a methods contribution, requires no new data, addresses a systematic error affecting an active literature, and does not require anyone to accept any part of the origins framework. Methods papers are also cited by people who disagree with you.
Strategic value. It establishes standing in the deep-phylogeny literature before the framework papers arrive — which materially changes how those are received.[★]

§12

Two independent clocks in the molecules

12.1 Amino-acid recruitment order

There is a broad consensus ordering for the recruitment of amino acids into the genetic code, synthesised from a large number of independent criteria — prebiotic abundance, biosynthetic cost, codon structure, thermal stability. Early: glycine, alanine, aspartate, valine. Late: tryptophan, cysteine, histidine, methionine, tyrosine.° [Trifonov, consensus temporal order]

PREDICTION — P3 — compositional stratigraphy of LUCA's proteome

If C7 holds (FUCA→LUCA is the RNA → RNA-protein-DNA transition), then LUCA's protein families are not all the same age, and the oldest should be compositionally biased toward early-recruitment amino acids relative to both younger LUCA families and modern averages.
Test. Take the reconstructed LUCA gene set. Rank families by inferred origination depth. Regress amino-acid composition against depth, using the consensus recruitment order as the predictor.
Discriminating power. A gradual model predicts weak or no gradient. The framework predicts a measurable one, concentrated in the deepest families.
Data. Existing — published LUCA reconstructions plus standard databases.⋆

12.2 The tungsten–molybdenum redox clock

An independent clock, and a genuinely underused one, because it couples molecular data to ocean chemistry directly.

12.3 Testing P1 — and being honest about it

The bottleneck prediction (P1: star-like radiation, clustered origination times) is the framework's most attractive phylogenomic test and its hardest.

The confound that has to be handled first

Early evolution had **rampant horizontal gene transfer**. HGT smears origination-time distributions, erodes tree-like structure, and can manufacture the appearance of a star phylogeny from a non-bottlenecked history — or destroy the signature of a real one.
This does not make P1 untestable, but it means the test cannot be a naive tree-shape statistic. It has to be posed as model comparison: simulate gradual and bottlenecked histories *under realistic HGT rates*, and ask which reproduces the observed gene-family age distribution.
Stated plainly: **P1 is the weakest of the four molecular tests in this Part**, and E3, E4 and E5 should be run first because they are cleaner. P1 is listed because it is independent of the rock record, not because it is easy.

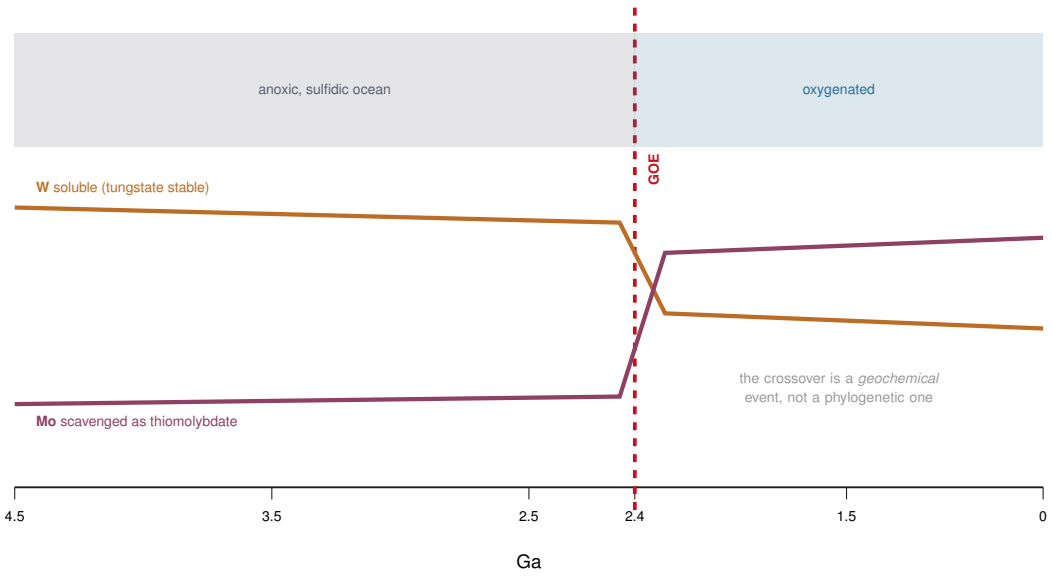


Figure 12.1: Cofactor availability as an independent clock. In an anoxic sulfidic ocean, molybdenum is scavenged as thiomolybdate while tungsten remains soluble; oxygenation reverses this. Enzyme cofactor dependence read from genomes is therefore datable against ocean chemistry without any reference to molecular-clock assumptions.

PART IV

The rock record

§13

Where the rocks are

13.1 The preservation reality

There is essentially no *in situ* Hadean sedimentary record. Everything older than about 4.0 Ga has been subducted, melted or metamorphosed. The search therefore has an unusual structure: the target interval is unreachable directly, so the strategy has to be built out of proxies at three removes.

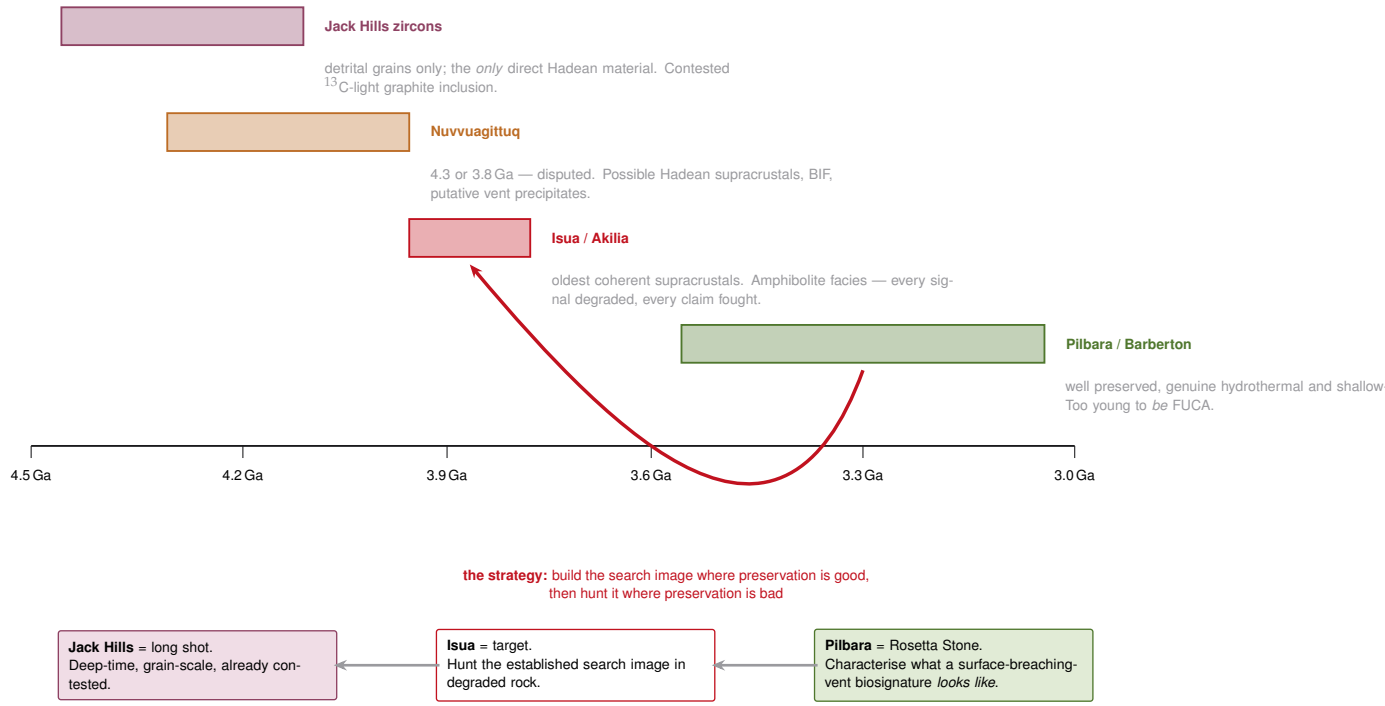


Figure 13.1: The search sequence. The oldest rocks are the worst preserved, so the search image cannot be built where the target is. Pilbara and Barberton provide well-preserved shallow-marine and hydrothermal facies in which to characterise the signature; Isua is where it must then be hunted; the Jack Hills zircons are a grain-scale long shot into the target interval itself.

A note on home ground. The Pilbara succession is in Western Australia and the Jack Hills zircons are Western Australian detrital material. Two of the three legs of this search programme are domestic, which matters for access, collaboration and funding framing.▷

§14

The signature stack

The payoff of the inversion method (Chapter 2) is that the metabolism it derives *is* the list of signatures to hunt. Each metabolic commitment implies a fractionation or a mineralogy.

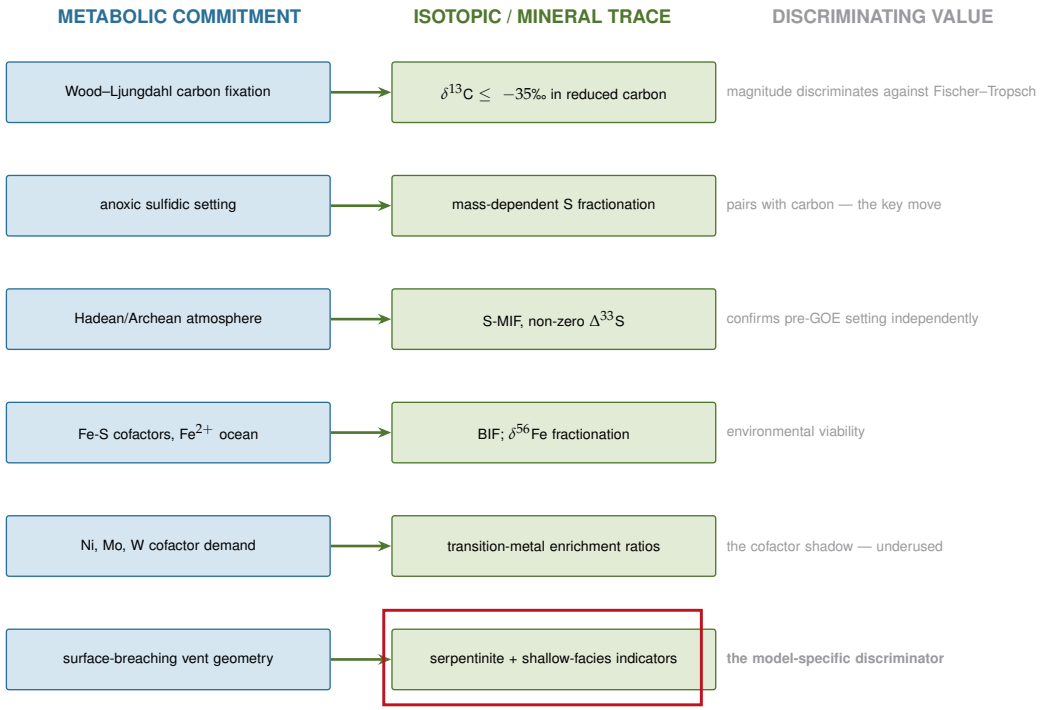


Figure 14.1: From derived metabolism to hunted signature. Each conserved LUCA trait, read backwards through the inversion method, commits to a measurable trace. The rightmost column is the honest part: most of these say "early metabolism," and only the facies juxtaposition says *surface-breaching vent*.

14.1 Two additions to the stack

Beyond what the archive assembled, two axes are worth adding.*

Nitrogen (δ^{15}N). A fourth independent fractionating system, sensitive to nitrogen fixation and to the redox state of the nitrogen cycle. Adding a third and fourth isotope system to a convergence argument is cheap relative to the credibility it buys, and δ^{15}N is routinely measured alongside δ^{13}C in the same preparations.

Molecular patterning, not just bulk ratios. Abiotic Fischer–Tropsch synthesis produces characteristic chain-length and branching distributions that differ from biological patterns. Where any molecular structure survives, the *pattern* discriminates where the bulk isotope ratio does not. This is the standard weakness of every contested Eoarchean claim, and it is addressable in better-preserved Pilbara material even if not in Isua.

§15

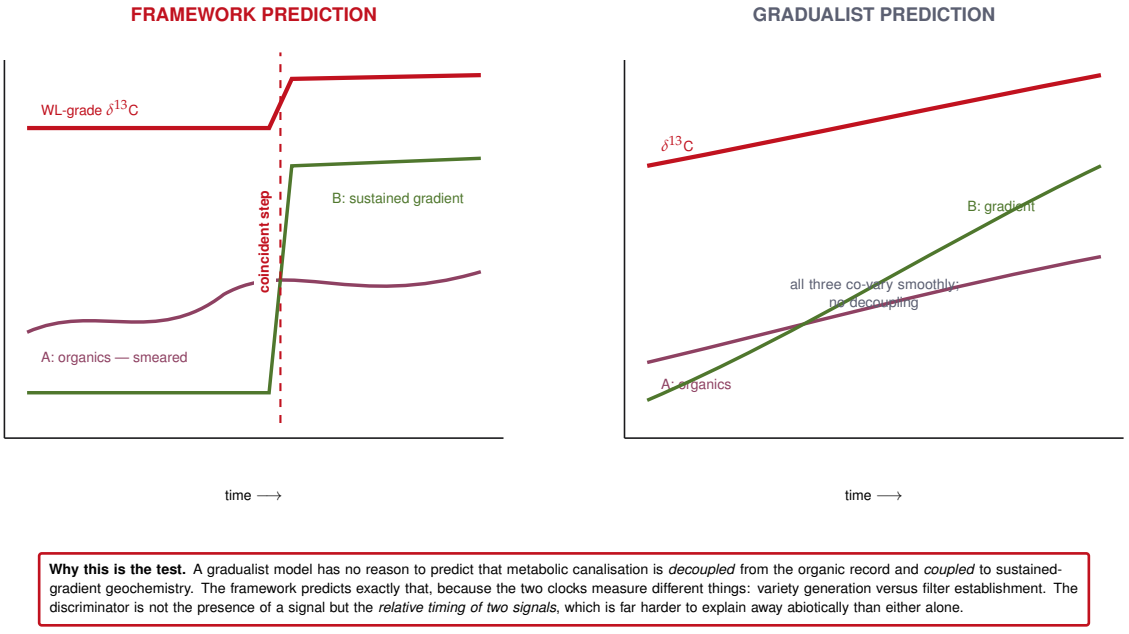
The two clocks, operationalised

C10 is the framework's sharpest falsifiable claim. This chapter turns it into a study design.

15.1 The claim restated as a measurement

Two quantities are tracked against the same stratigraphy:

- **Clock A — the exploration record.** Abundance and isotopic distribution of reduced carbon. Under the framework this reflects FUCA-era chemistry: early, smeared, isotopically inconsistent, because exploration-optimised systems are metabolically promiscuous (§7).
- **Clock B — the throughflow record.** Indicators of *sustained* serpentinising hydrothermal systems: serpentinite mineralogy, vent precipitate facies, and the transition-metal availability window.



Why this is the test. A gradualist model has no reason to predict that metabolic canalisation is *decoupled* from the organic record and *coupled* to sustained-gradient geochemistry. The framework predicts exactly that, because the two clocks measure different things: variety generation versus filter establishment. The discriminator is not the presence of a signal but the *relative timing of two signals*, which is far harder to explain away abiotically than either alone.

Figure 15.1: C10 as a measurable decoupling. The framework predicts that Wood–Ljungdahl-grade carbon isotope values appear as a step coincident with sustained-serpentinisation indicators, while the bulk organic record is early and smeared. A gradual origin predicts smooth co-variation of all three. The test is relative timing, not absolute detection.

EVIDENCE PLAY — E6 — the compilation study

The version that needs no fieldwork. Much of the required data is already published across decades of Isua, Nuvvuagittuq, Akilia, Pilbara and Barberton isotope and petrology literature. It has never been compiled with *this* question in mind, because nobody was asking about relative timing of two clocks.
Design. Build a stratigraphically-registered compilation of (i) reduced carbon abundance

and $\delta^{13}C$, (ii) serpentinisation and vent-precipitate indicators, (iii) S-MIF and mass-dependent S, (iv) transition-metal ratios — all tied to the best available age control. Then test for decoupling.

The hard part, named. Age control in Eoarchean successions is poor and correlation between belts is contested. The study may only be able to resolve ordering within single successions rather than globally. That is still a test; it is just a weaker one, and it should be designed as a within-succession test from the start rather than discovering the limitation late.

Status. Archive work with a real chance of a publishable negative result, which is a good property to have.⋆

§16

Interim evidence register

Consolidating what Parts II–IV have generated. Ranked by (tractability × discriminating power). Parts V–VI will add the lunar and Venus entries.

	Play	Resource	Tests	Note
E5	rRNA substitution-model bias — methods paper	existing data, compute	—	**Fastest publication.** No framework buy-in needed
E1	Vesicle transfer-survival experiment	existing bench, low cost	C4	Cheapest test of a load-bearing claim
E4	Rodin–Ohno structure-prediction test	compute, public sequence	C7, FUCA-era	**Highest novelty.** Tools postdate the hypothesis
E3	tRNA two-code lineage variability	existing databases	C7	Clean, bounded, no wet lab
E2	Chicxulub as impact-hydrothermal analogue	published IODP-364	C5	Converts plausibility to measurement
E6	Two-clocks stratigraphic compilation	published literature	**C10**	The framework's sharpest test
P3	LUCA proteome compositional stratigraphy	existing reconstructions	C7	Depends on origination-depth inference
P4	W→Mo cofactor clock	genomes + geochemistry	C2, setting	Two independent observables
P1	Bottleneck / star-phylogeny test	phylogenomic data	C5	**Weakest.** HGT confound is severe

16.1 Phase Zero material held for translation

Per Chapter 4, material that has not yet been reduced to a testable claim is held here rather than promoted or discarded.

- **The dimensionality proposal.** Earth or life modelled as a four-dimensional system with zero categories, life as the edge of a lower-dimensional shape, the observable being a meta-process optimising for survival of mass extinctions.▷ As stated this is not yet falsifiable. It may be reaching for something in the neighbourhood of a renormalisation-group treatment of fitness landscapes over deep time, or for a formal statement about which features of a history are scale-invariant. **Held for translation, not rejected.**
- **Felt-certainty statements.** The archive contains several high-confidence assertions offered without argument. They are recorded as historical fact about what was believed — and deliberately not used as premises anywhere in this volume. Every claim above stands on stated evidence.

PART V

The lunar archive

§17

Why the Moon holds Earth's Hadean

17.1 The argument

Earth destroyed its own Hadean record. Plate tectonics, hydrothermal alteration, weathering and metamorphism have removed essentially all of it; what survives is a few kilograms of detrital zircon. The Moon has none of those processes. It has been geologically dead for most of its history, and its regolith is an accumulating, gardened but non-destructive archive of whatever hit it.▷

Material *did* hit it from Earth. Large impacts eject terrestrial rock at above escape velocity, and a fraction of that material is only lightly shocked — the spallation zone near the free surface experiences high acceleration but comparatively low peak pressure, which is why Martian meteorites reach Earth containing intact minerals and, in some cases, organic matter.

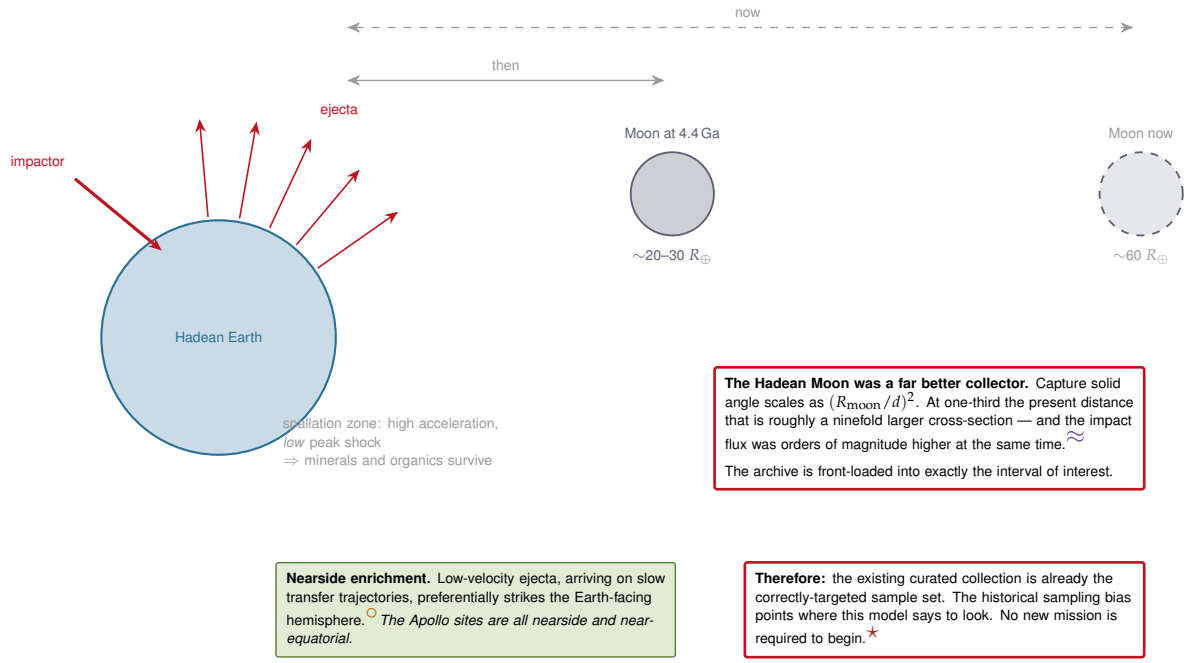


Figure 17.1: Why terrestrial Hadean material should be on the Moon, and why it should be on the nearside. Two effects compound: the Moon was much closer during the interval of interest, giving a far larger capture cross-section, and low-velocity ejecta preferentially reaches the near hemisphere. Apollo sampled the nearside.

17.2 The published precedent

This is not a thought experiment. A clast in Apollo 14 sample 14321 — "Big Bertha" — has been argued to be a terrestrial meteorite.○ [Bellucci et al. 2019, EPSL — verify age and depth estimates]

The argument rests on a granitoid, zircon-bearing felsic clast whose zircon records an oxidation state, crystallisation temperature and inferred pressure that are ordinary for terrestrial continental crust and difficult to produce on the Moon — the pressure in particular implying formation at some tens of kilometres depth in a body with terrestrial-scale gravity.

The interpretation is contested, and it does not need to be settled for the purpose here. What matters is that it establishes three things at once: a **search image** (evolved, felsic, zircon-bearing clasts in nearside breccias), an **analytical protocol** (zircon oxybarometry, thermometry, trace elements), and an **allocation precedent** (this work was done on curated Apollo material through the normal process).

17.3 The provenance problem, stated honestly

The obvious discriminator turns out to be the weak one, and this needs saying up front because it determines the whole analytical strategy.

Oxygen isotopes will not do the work

The intuitive test for "is this rock from Earth?" is $\Delta^{17}O$, which cleanly separates Earth from Mars and from most asteroid parent bodies.
It does *not* usefully separate Earth from the Moon. The Earth–Moon oxygen isotope difference is at the few-parts-per-million level — the two bodies are isotopically near-identical, which is itself a central constraint on giant-impact models.○
Consequence: terrestrial provenance has to be established on *petrological* rather than isotopic grounds — oxidation state, water content, crystallisation pressure and temperature, and trace-element patterns diagnostic of evolved hydrous crust. That is exactly the argument made for 14321, and it is why that argument is contested rather than decisive. This is the single hardest technical problem in the lunar route, and any proposal that does not confront it directly will not survive review.

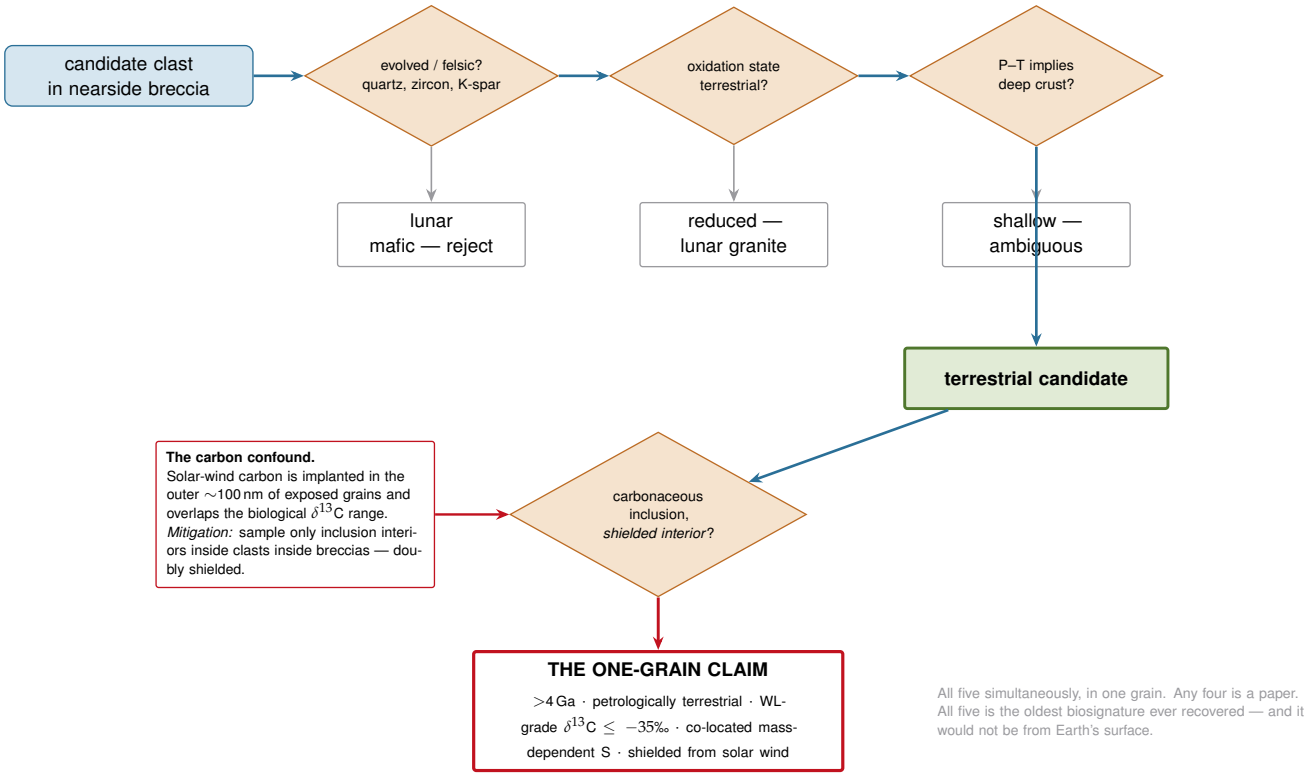


Figure 17.2: Establishing terrestrial provenance, then asking the biological question. Because oxygen isotopes cannot separate Earth from Moon, provenance rests on a petrological chain. The carbon measurement then has its own confound — implanted solar-wind carbon overlapping the biological range — which is why the target is an inclusion interior inside a clast inside a breccia.

§18

The existing-collection strategy

The framework's lunar route does not require a mission. It requires allocations.

	Material	Why	Access
T1	ANGSA cores (e.g. 73001, opened 2022 after 50 yr sealed)	Freshest available material; minimal terrestrial contamination history	Formal allocation; high competition; strong scientific framing required○
T2	General Apollo regolith breccia thin-section collection	Nearside, equatorial — the enriched population. 14321 came from here	Standard allocation process; thin sections are the cheap first ask○
T3	Lunar meteorites held on Earth (>600 known)	No allocation committee. Purchasable or loanable. Provenance unknown — which is a weakness for targeting and an irrelevance for a first survey	Commercial and institutional; **start here**○
T4	Chang'e 5 (2020) and Chang'e 6 (2024, *farside*)	CE-6 is the natural control: farside should be *depleted* in Earth ejecta if the nearside-enrichment model is right	International allocation now open○

EVIDENCE PLAY — E7 — the nearside/farside asymmetry test

The move that turns a needle-in-haystack search into a hypothesis test.
Searching for one grain is unfundable. Testing a *ratio* is fundable.
Design. Quantify the abundance of evolved-crust (felsic, zircon-bearing) clasts per unit mass in nearside Apollo regolith breccias versus farside Chang'e 6 material, using identical screening criteria.
Prediction. Nearside enriched relative to farside, by a factor the ejecta transport model can be made to predict in advance.
Why it is strong. It is a comparative measurement with a pre-registered predicted direction and magnitude, it uses two independent curated collections, and it does not require finding a biosignature at all. A positive result establishes the transport model quantitatively, which is what any subsequent biosignature claim would need as its foundation.
And it is publishable either way.⋆

EVIDENCE PLAY — E8 — the zero-cost modelling paper

Before touching any sample: compute the expected terrestrial mass fraction as a function of latitude, depth in regolith, and terrain age, using published impact-flux histories, lunar orbital evolution, and existing regolith maturity data.
Nobody appears to have done this with the Hadean orbital geometry treated properly — the compounding of a closer Moon and a higher flux.⋆≈
It costs compute and literature time, produces a testable spatial prediction, and gives every subsequent allocation request a quantitative justification instead of a plausibility

argument.

What the lunar route does *not* give you

Provenance to a *place* on Earth is unrecoverable. A grain tells you that Hadean Earth had a particular chemistry somewhere; it does not tell you where, or in what setting.
That means the lunar archive can confirm **that** Wood–Ljungdahl-grade carbon fixation existed by a given date. It cannot confirm the surface-breaching vent geometry (C4), which is a *facies* claim and therefore terrestrial-only.
The two routes are complementary and neither substitutes for the other: the Moon dates the metabolism, the rock record locates it.

PART VI

Venus

§19

The mechanism the argument needs

C8 — *Venus over Mars; radiation as constructive rather than sterilising* — is the framework's most exposed claim, because in the archive it is an intuition without a mechanism.▷ It needs one. There is a good one available.

19.1 Radiation as synthesis

The reflex association of radiation with sterilisation is a statement about *extant* life, which is a fragile encoded system with large, damageable information content. It is not a statement about prebiotic chemistry, which has no genome to damage and for which high-energy photons are a free energy source.

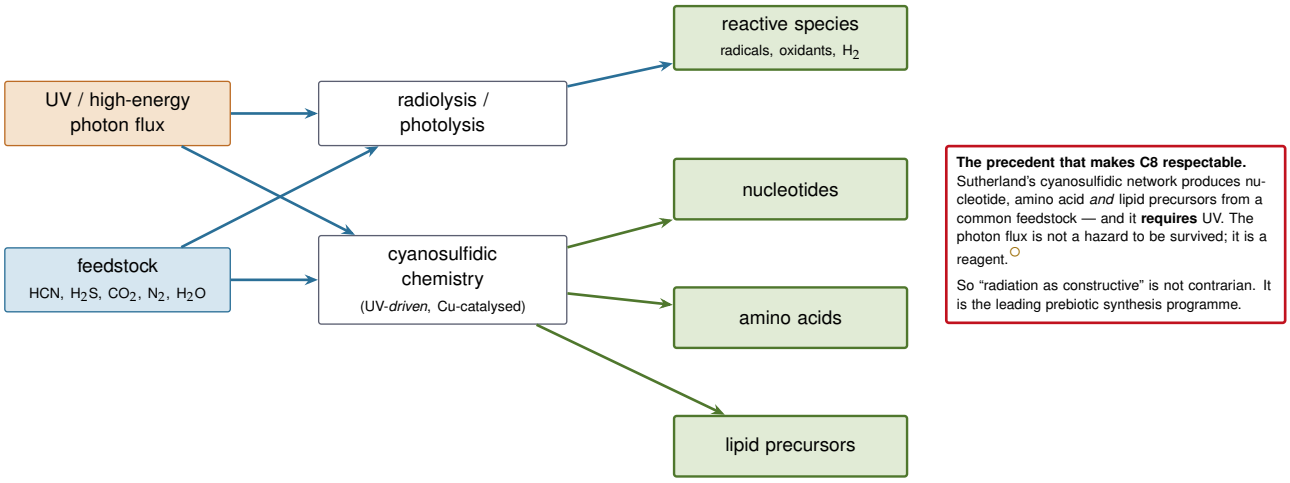


Figure 19.1: The mechanism C8 was missing. High-energy photons are destructive to encoded life and generative for prebiotic chemistry. A thick atmosphere supplies a stratified photon environment — synthesis above, accumulation below, transport between.

19.2 Where Venus actually stands

Two separate cases, with very different strength. They are usually conflated and should not be.

The cloud-habitability case is weak. The 48–60 km layer has tolerable temperature and pressure, but the droplets are 75–98 % sulfuric acid and the water activity is of order 0.004 — far below the ∼0.585 limit for any known organism.○ The 2020 phosphine claim has been substantially contested and should not be leaned on.○

The early-Venus case is strong. Climate modelling suggests Venus may have sustained surface liquid water for a considerable fraction of its history, ending with a resurfacing

event within roughly the last billion years.° [Way et al. — verify duration and end date] If that is right, Venus had a longer and larger habitable interval than Mars ever did — and it was closer to the Sun, so it had more energy throughput, which is the currency C1 cares about.

The sharpest statement of C9 — and it is about Venus

Venus is where the two searches (C9) diverge maximally.
By the framework's criteria — energy throughput, gradient availability, variety generation, radiation as reagent — **early Venus is arguably the best abiogenesis candidate in the solar system**.
By biosignature-search criteria it is close to the worst, because the resurfacing event destroyed the surface record and the present surface is hostile to preservation.
Every existing target-prioritisation scheme scores Venus low, because every existing scheme is implicitly a *biosignature* scheme. If the two searches are genuinely different questions, Venus is the case that proves it — and the correct response is not to search for extant Venusian life but to ask *what a planet that did abiogenesis and then lost the record retains*.⋆

§20

What Venus still remembers

The surface is gone. The atmosphere is not, and an atmosphere is an integrator.

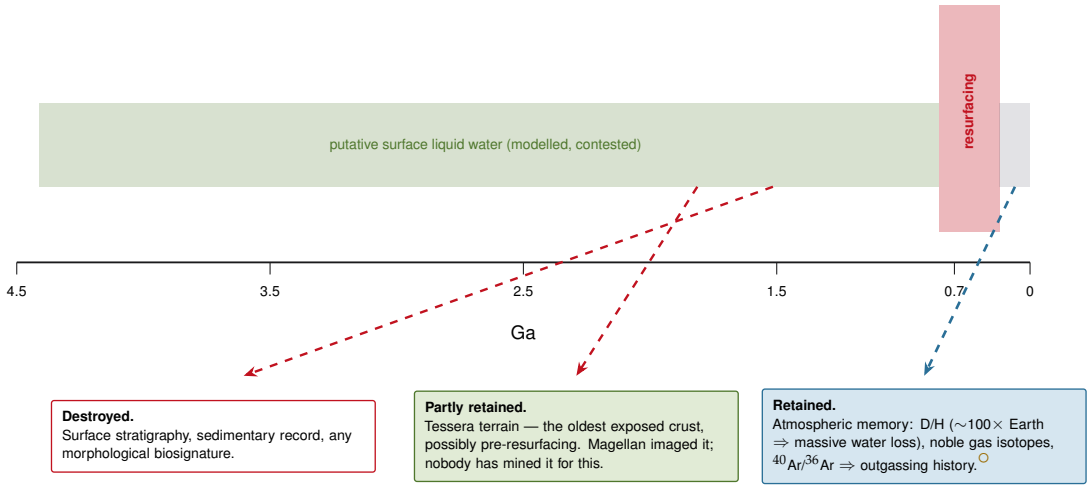


Figure 20.1: Where Venusian memory is stored. The resurfacing event removed the surface record but not the atmospheric one. Noble gas and hydrogen isotope systematics integrate the whole history and are measurable — some already measured, and under-mined.

EVIDENCE PLAY — E9 — reanalysis of existing Venus mission data

Three archives that astrobiologists have largely not worked:
Pioneer Venus Large Probe neutral mass spectrometer (1978). Repeatedly reanalysed for individual questions, most recently for reduced species. The full dataset has not been systematically re-examined against a prebiotic-chemistry hypothesis set.○
Venera and Vega (1975–1985). Landers and, critically, *balloons that flew in the cloud layer*. Much of the primary data is under-represented in the Western literature and some is only in Russian-language sources.○⋆
Magellan radar (1990–94). Tessera terrain is the oldest exposed crust and the only plausible window onto pre-resurfacing Venus. The data is public and complete; it has been mapped geologically but not interrogated for astrobiological targeting.⋆
Status. Archive work. No mission, no allocation, no fieldwork. The Venera strand in particular is an unusual opportunity precisely because of the language and access barrier.

PREDICTION — P6 — the noble gas constraint on early gradients

Noble gas isotopic ratios — particularly $^{40}Ar/^{36}Ar$ and the Ne and Xe systematics — constrain total outgassing and therefore the integrated crust–atmosphere exchange over Venus's history.○
Under C1, abiogenesis requires sustained throughflow, which on a rocky planet means sustained crust–mantle–atmosphere cycling. **The framework therefore makes a quantitative prediction about Venusian outgassing history** — it must exceed some threshold for the early-Venus abiogenesis case to hold.

That threshold has never been computed, because nobody has framed the question this way. Doing so converts a qualitative preference (“Venus is interesting”) into a number that existing and forthcoming atmospheric measurements can test.⋆

PART VII

The computational programme

§21

One graph, two labellings

21.1 The reframe

The natural first thought is two coupled graphs — a biological one and a geochemical one — joined by learned edges. That is not quite the right object, and the correction is the whole idea.

Metabolism and geochemistry are the same reaction network under different catalytic regimes. Wood–Ljungdahl is the abiotic acetyl-CoA pathway with enzymes bolted on. Ferredoxin's Fe-S clusters are greigite and mackinawite fragments held in a peptide scaffold. Biology is not a separate network that interfaces with chemistry; it is a subgraph of chemistry that has acquired catalysts and an encoding.

So the object is one hypergraph with a *time-varying edge labelling*:

nodes	chemical species — metabolites, minerals, dissolved gases, ions
hyperedges	reactions. A hypergraph, not a graph: reactions have n substrates and m products, and collapsing that to pairwise edges destroys stoichiometry
edge state	{uncatalysed $\to$ mineral surface $\to$ metal cluster $\to$ ribozyme $\to$ protein enzyme}
node features	thermodynamic properties; availability under a specified ocean chemistry
edge features	ΔG at specified T, P, pH, ionic strength; EC number where enzymatic; phylogenetic depth of the catalysing family

And the question "what was happening around LUCA and before" becomes precise: **for each edge, when did the catalytic state transition, and in what order?**

21.2 The bio:geo coupling already exists

The hard part of this design looks like it should be building a thermodynamic bridge between biochemistry and geochemistry. It is not, because that bridge has been built.

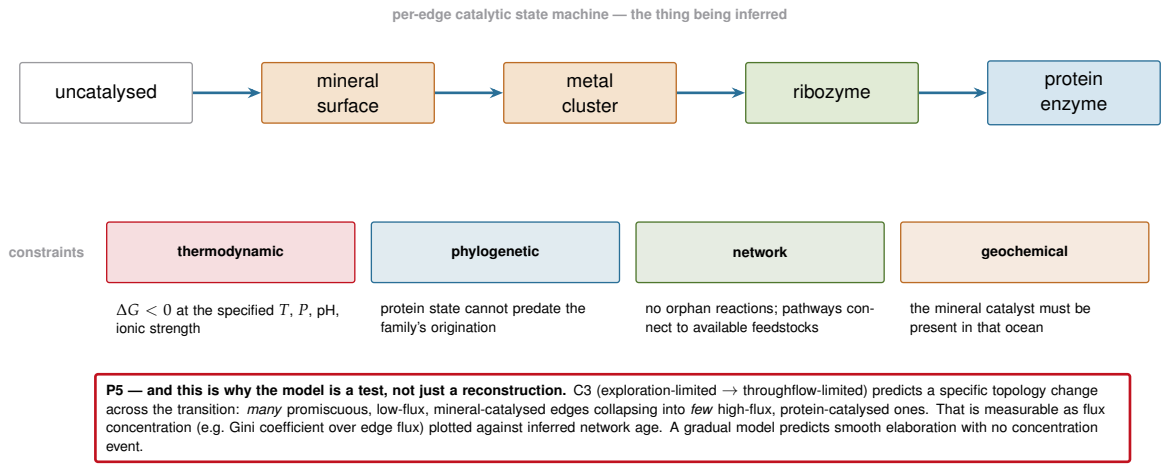


Figure 21.1: The inference problem, stated. Rather than coupling two networks, label one network's edges by catalyst type and infer the transition times under four constraint families. The output is the metabolic state of the world at a given date — and the topology of the transition is itself a test of C3.

CHNOSZ	Computes standard thermodynamic properties for minerals, aqueous species **and proteins** across geological T–P ranges. It will return ΔG for a biochemical reaction at 100 °C, 500 bar, pH 9.5. *This is the scoring function.*[○]
Rhea	Reaction-centric, ChEBI identifiers, cross-linked to UniProt and EC. Preferable to KEGG here because it is organised around reactions rather than pathways, giving clean hyperedges.[○]
LUCA sets	Published reconstructions supply the protein-family presence labels and origination-depth estimates.[○]
Mineral catalysis	The weakest link: there is no clean database of mineral-surface-catalysed organic reactions equivalent to Rhea. This has to be compiled from primary literature. **Compiling it is itself a publishable resource.**[★]

The paired record you would actually store looks like this:

```
{ "rhea_id": "RHEA:21540",
  "participants": ["CHEBI:15378", "CHEBI:16526", ...],
  "dG": {"T_K": 373, "P_bar": 500, "pH": 9.5, "I": 0.7, "kJ_mol":
-42.1},
  "catalysts": {"protein": ["EC 1.2.7.4"], "mineral": ["mackinawite",
"greigite"]},
  "phylo": {"luca_present": true, "family_ids": [...], "depth_rank": 12},
  "geo": {"requires": ["Fe2+", "HS-"], "available_hadean": true} }
```

21.3 On genomic foundation models

Evo 2 is a nucleotide-resolution genomic foundation model trained on the order of 10^{13} nucleotides across all domains, with very long context and zero-shot variant-effect capability.[○] It

is worth being precise about what it can and cannot do here, because the temptation to over-apply it is strong.

Right tool, specific job

What it is: a sequence prior. It has learned what genomes look like.
What it is not: a model of geochemistry, time, or phylogeny. It will not tell you what was happening around LUCA, because nothing in its training signal encodes when anything happened.
Where it genuinely helps — E4. The Rodin–Ohno test asks whether both readings of a reconstructed ancestral duplex are plausible coding sequences. Evo 2 operates on DNA rather than protein, so complementary-strand coding is exactly the class of structural regularity it may have internalised. Zero-shot scoring of reconstructed duplexes against shuffled controls, on both strands, is a direct application.
The caveat that must be stated in any writeup: it is trained on *extant* genomes. For a pre-LUCA question that is a prior pulling toward the present, not a neutral scorer. Use it as one feature among several, never as the arbiter.

21.4 The honest architectural verdict

This is not yet a deep learning problem

At the level of interest, $n = 1$. There is one reconstructed LUCA. You cannot train a large model on one labelled example, and any architecture that appears to do so is fitting something else.
The tractable shape is generate-and-filter:

1. *Generate* candidate ancestral networks by pruning modern metabolism under phylogenetic constraints.
2. *Filter* by thermodynamic feasibility at Hadean T, P, pH, ionic strength — CHNOSZ supplies this.
3. *Score* by parsimony, network connectivity, and phylogenetic support.
4. *Learn* only at the last step, where there is enough data: a graph neural network trained on *modern* metabolism for link prediction, transferred to score hypothetical ancestral edges.

Ancestral metabolic network reconstruction already exists. **The geochemical feasibility filter is the novel component**, because nobody currently runs it under Hadean conditions.

EVIDENCE PLAY — E10 — build the dataset; it is the contribution

The bottleneck is not the model. It is the paired data, and there is none.
A bio:geo paired reaction network — reactions with thermodynamics evaluated under specified geological conditions, catalyst annotations across both mineral and enzyme regimes, and phylogenetic depth labels — does not exist. Building it is a resource other people will use *even if they reject every one of C1–C11*, which is the best property a first output can have.
Sequence. (i) Rhea × CHNOSZ for the thermodynamic layer. (ii) Mineral catalysis compiled from primary literature — the slow part, and the part with no substitute. (iii)

LUCA family labels joined on EC. (iv) Only then, models.
This is the same argument as Volume II. The data exists in pieces; the integration does not; the integration is the paper.⋆

Sequencing note. E10 (the dataset) gates P5 (the topology test) and strengthens E6 (the two-clocks compilation), because a thermodynamically-scored network tells you which reactions were even possible in the conditions the stratigraphy records. Of everything in this volume, E10 has the most downstream dependencies — which makes it the right thing to start, and the wrong thing to start with if the goal is a fast first publication. E5 remains the fast one.

PART VIII

The bench programme

§22

The coupling gap is the phase boundary

22.1 The unification

The protocell programme in the archive was written as a separate, long-horizon ambition — explicitly *not* an attempt to recreate abiogenesis, but to bootstrap a minimal living system by designing components and forcing their combinatorial assembly.▷ It framed itself as a different problem from the origins work.

It is not a different problem. It is the same problem run forward instead of backward, and the archive contains the sentence that proves it:

> *"Optimise for configurational exploration rate, not for waiting."*▷

Configurational exploration rate is the vertical axis of Figure 3.1. The platform is not analogous to the FUCA regime — **it operates in it**. A microfluidic system delivering $\sim 10^6$ droplets/hour under sustained wet–dry driving, with compartmentalisation for individuation and selection, is an exploration-limited regime by construction. That is the same selective condition C3 assigns to pre-LUCA systems.

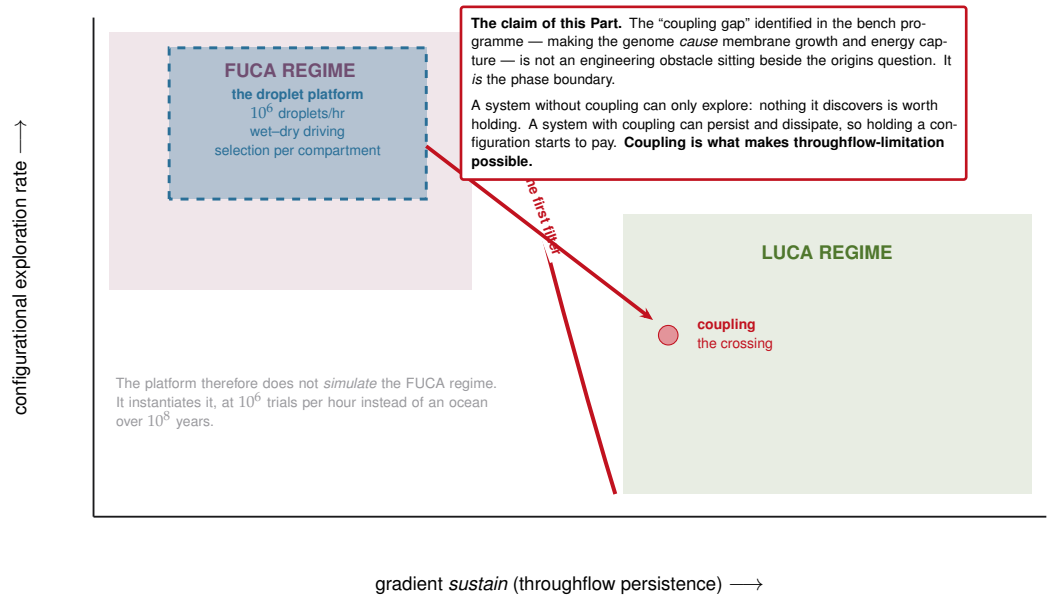


Figure 22.1: The bench programme sits inside the FUCA regime, and the coupling gap is the boundary it must cross. This is why the protocell work is a test of the origins framework rather than a parallel ambition: crossing the coupling gap in a droplet is the same transition, in the same coordinates, as FUCA → LUCA.

22.2 Why this reframing is worth having

Three consequences, and they are practical rather than rhetorical.

It makes the bench programme falsify something. As stated in the archive, the protocell

work could succeed or fail without bearing on the origins framework at all. Under this reframing it bears directly: if coupling can be established in a droplet under exploration-limited conditions and the system then *shifts* to rewarding persistence, that is C3 observed happening. If coupling is achieved and no such shift occurs, C3 has a problem.

It tells you which coupling mechanism to prioritise. Not the one easiest to engineer — the one that most closely reproduces the handoff the framework describes. That is the pyrophosphate route, for reasons in §24.

It makes the platform an instrument, not just a factory. A droplet population traversing the boundary is a measurement of where the boundary is.

22.3 Component map

The archive's minimal architecture, re-tabulated against the commitments each component tests.

Component	Status▷	Tests	Framework role
Fatty acid vesicle, grows and divides spontaneously	demonstrated	C4	Individuation. Division is free past a critical S/V ratio — engineer growth, not cytokinesis
Encapsulated Fe–S mineral nanoparticles	demonstrated	C2, C4	Regime B in miniature: buy proto-metabolism off the shelf rather than evolving it
External pH gradient	demonstrated	C1, C2	The sustained drive. Directly instantiates the chemiosmotic inference
Cooperative ribozyme networks (Azoarcus)	demonstrated	C6, C7	Heredity without protein; network-level autocatalysis
G-quadruplex RNA ion channels	demonstrated separately	**coupling**	A *heritable membrane property* encoded in sequence
Aminoacylation & peptide-bond ribozymes	demonstrated separately	C7	The RNA → RNA-protein step, in vitro
Polymerase ribozyme	**partial**	C7	**The field's bottleneck.** Error rates block sustained heredity
Wet–dry cycling with coalescence/re-splitting	platform to build	C3, C4	Regime A. The mixing step is the exploration engine

§23

The two-regime platform

23.1 The bench programme already contains both halves of C4

This is the observation that makes Part VIII more than a restatement.

The archive's protocell notes lean toward the Damer–Deamer hot-spring model, because wet–dry cycling solves polymerisation and mixing.▷ They *also* specify encapsulated iron–sulfide co-precipitation under an external pH gradient, which is the Lane–Martin alkaline-vent chemistry.▷

Both regimes are already in the design. What has never been specified is the *sequence* between them — and the sequence is exactly what C4 claims.

EVIDENCE PLAY — E11 — the full two-regime sequence on the platform

This is the bench test of C4 at system scale, and E1 (Part II) is its minimal precursor.
Protocol. Run droplets through an ordered sequence rather than a homogeneous condition:

1. **Regime A.** Low ionic strength, wet–dry cycling with explicit coalescence and re-splitting. Vesicle formation, polymerisation, combinatorial mixing. Optimise for exploration rate.
2. **Transfer.** Stepwise increase in Mg^{2+}/Ca^{2+} to seawater-analogue levels.
3. **Regime B.** Sustained pH gradient across encapsulated Fe–S mineral. No further wet–dry cycling. Optimise for persistence.

Measured. Survival fraction through transfer; and — the interesting quantity — whether the *composition* of the surviving population differs systematically from the pre-transfer population.
The framework prediction (P7). The transfer is a filter with a direction. Droplets that survive should be enriched for compositions that couple the gradient to internal chemistry, because under Regime B those are the only ones that gain anything from persisting. Under a null model the surviving population is a random subsample.
Why this is strong. It is a selection experiment with a predicted direction of selection, run on a platform that already has to exist for other reasons. And the readout is composition, which the platform's multi-modal sensing is already designed to capture.★

23.2 Sequencing against the antimicrobial project

The archive is explicit that the protocell work is the long-horizon project and the antimicrobial work is partly infrastructure for it.▷ That relationship holds and is worth making precise, because the shared infrastructure is larger than it looks:

- **Droplet generation and handling** — common to both.
- **Multi-modal readout** (scatter, fluorescence, impedance per droplet) — required for protocell anomaly detection, immediately useful for antimicrobial screening.
- **Dielectrophoretic sorting** — built from open designs; serves both.

- **Active-learning loop** — Bayesian optimisation over composition; identical machinery, different objective.
- **Spectral imaging of plate assays** (Volume II, Part V) — built for inhibition zones, directly reusable as a protocell readout.

The strategic point. Every instrument the antimicrobial project requires is an instrument the protocell programme requires. The three-month project is not a detour from the lifetime project; it is the funded, deadline-bearing excuse to build the lifetime project's instrument stack. That is a good structure and it should be stated that way in any funding or public-facing framing.▷

§24

Coupling: the two routes worth taking

The archive identifies coupling as "the actual research problem" and nominates two mechanisms.[▷] Under the framework reframing, they are not equally interesting, and it is worth saying why.

24.1 The pyrophosphate loop — the framework's preferred route

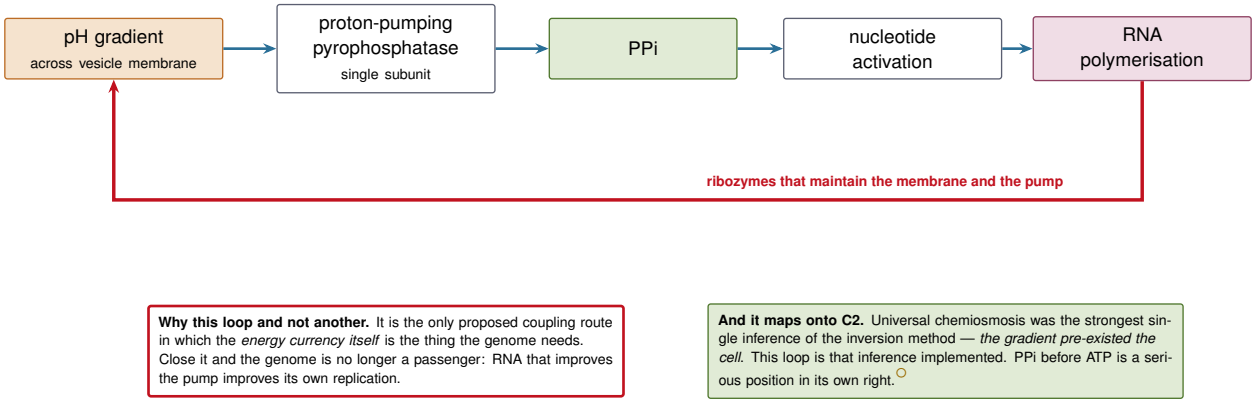


Figure 24.1: Gradient → PPi → nucleotide activation → polymerisation, closed by ribozymes that maintain the pump. The archive nominates this as the single most important pathway in the design. Under the framework it is also the closest bench analogue of the FUCA → LUCA handoff: the point at which a configuration begins to earn its own persistence.

Proton-pumping pyrophosphatases exist in modern bacteria as essentially single-subunit membrane proteins — vastly simpler than ATP synthase — and there is a serious argument that PPi preceded ATP as the energy currency.[○▷] That combination (simple enough to bootstrap, plausibly ancestral) is rare.

24.2 G-quadruplex channels — the heritable membrane property

The second route attacks coupling from the membrane side rather than the energy side. Certain G-rich RNA sequences spontaneously form four-stranded structures that conduct ions across bilayers.[○▷] If an ion channel can be encoded in RNA sequence, then a membrane property becomes *heritable* — which is the coupling problem stated in its purest form.

Why the framework prefers PPi but wants both

The G-quadruplex route gives heredity of a membrane property. The PPi route gives a closed energetic loop. **Coupling in the framework's sense requires both**: a configuration must be heritable *and* must improve its own throughflow, or persistence does not pay. Run in the same droplet, they are complementary rather than alternative — a sequence-encoded channel that modulates the gradient the pyrophosphatase runs on is a complete coupling circuit, and every component of it has separate experimental precedent.[★]

24.3 The honest bottleneck

The polymerase ribozyme remains the hard rock the whole field is breaking against: current constructs copy hundreds of nucleotides including portions of themselves, but error rates remain too high for sustained heredity.[○▷]

One framework-derived idea about the error threshold

The error-threshold problem is usually posed in bulk solution. The framework's regime distinction suggests reposing it.

Eigen's threshold assumes competition among replicators sharing a pool. Under strict compartmentalisation with selection at the compartment level, the relevant quantity is not per-molecule fidelity but *per-compartment* probability of retaining a functional network — and a cooperative network can tolerate loss of individual members that a single self-replicator cannot.

So the prediction is that compartmentalisation plus network redundancy raises the tolerable error rate, possibly substantially, and the platform is the natural place to measure by how much. That is a bounded, publishable question that does not require solving the polymerase ribozyme — it asks how much fidelity you actually need.[⋆]

PART IX

The Central Catma

§25

The freeze, not the rule

25.1 Provenance and standing

This Part restates a separate paper — *The Central Catma of Genetics* — and integrates it with the framework of Parts I–IX. The paper is prior work with its own claim–source verification table and its own staked originality; nothing here supersedes it.▷

It is included because it does something no other component of this volume does: **it supplies a mechanism for C7**, which until now has been the framework's weakest commitment — a staging claim (RNA → RNA-protein-DNA) asserted rather than derived.

The distinction the paper turns on

Crick's central dogma is a law about information flow *in life as it now runs*. It describes the frozen channel. It says nothing about how the channel was built, and the channel is the interesting object: a coded mapping from nucleotide sequence to amino-acid sequence, archived in double-stranded deoxyribose, is not a premise of life but a contingent outcome.
Where the dogma is the rule once frozen, the catma is the freezing.▷

The structure is a spine and four faces. The spine: a single merger event, whose decisive property is that it *froze*, built the inherited architecture of life. The faces: coded translation is that freeze; amino-acid handedness was fixed at it; DNA handedness descends from the same root by two further mechanisms; and a non-templated sugar-network regime was the generator the freeze acted upon.

25.2 Face one — translation is a merger, not an invention

The methodological constraint stated first, because it is the paper's hardest demand and most accounts violate it:

An account of the origin of translation must not select for translation. Neither the ribosome nor tRNA can have been favoured *in order to* build a coding system, because no coding system existed to be a target of selection.
Each component must therefore have had a complete, code-free function, and translation must fall out of their meeting rather than being its goal.▷

This is structurally the same refusal as C2 (§2): an outcome may not be used to explain its own origin.

The objection, pre-empted. Non-ribosomal peptide synthesis is not a relic of a pre-ribosomal world. Its patchy accessory distribution and its shared thiotemplate logic with polyketide and fatty-acid synthases mark it as a convergent canal, re-derived wherever the chemistry pays. The ancestral object is the activated-aminoacyl chemistry it exploits — which is what the transferase centre runs without it.▷

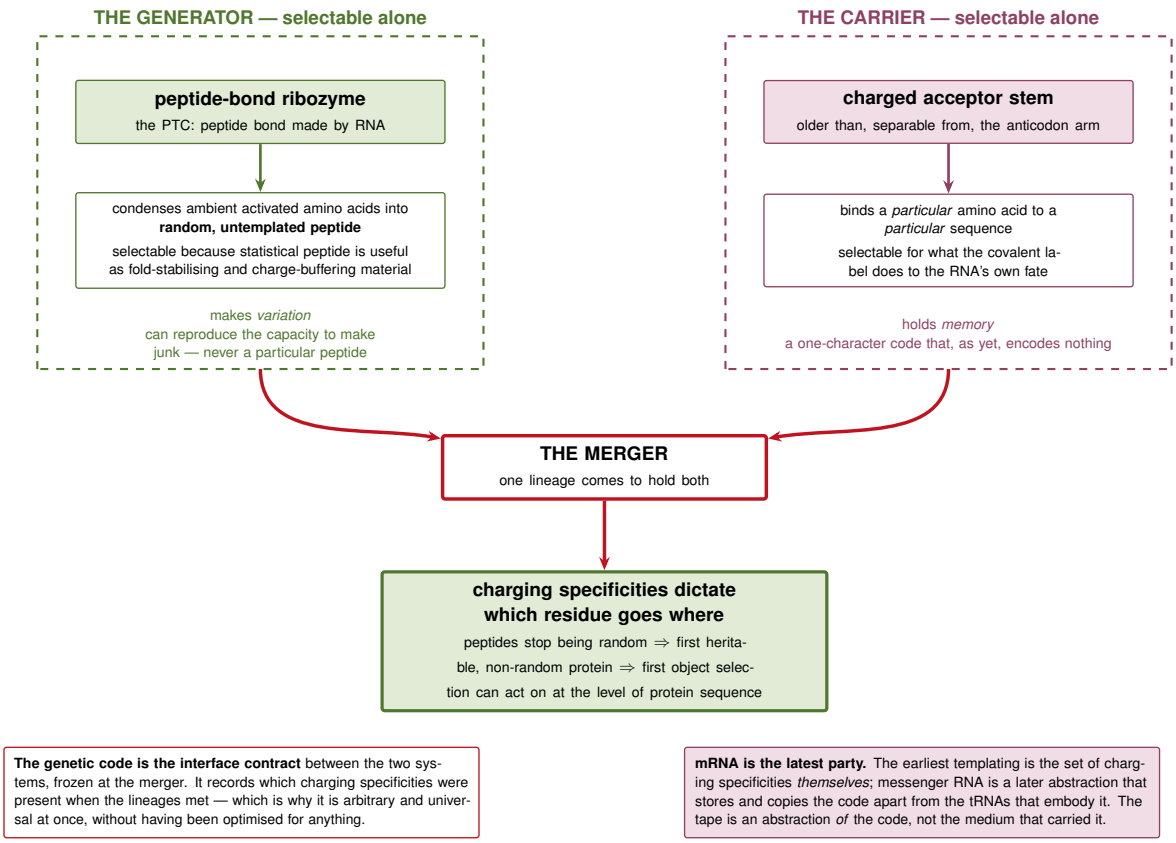


Figure 25.1: Two systems, each complete and selectable without the other, whose meeting produces coded translation as a by-product. Neither half is selected *for* translation — the generator makes non-heritable variation, the carrier holds specific associations it cannot read out. Translation is what happens when one lineage holds both.

PREDICTION — Face one's falsifiable edge — and it is already E3

The operational (acceptor-stem) code should be *older than and partially independent of* the anticodon–codon code, retaining residual amino-acid-assignment logic the codon table does not explain. Residual mismatch is a fossil of the pre-merger carrier; its absence puts the face under pressure.▷
This is E3 (§9), derived independently in Part III before this paper was available. The test programme for Face One is already in the register.

25.3 Face two — the merger froze a chiral choice the sugar only biased

Two measurements appear to conflict. A bare D-ribose acceptor stem, with no enzyme or ribozyme, aminoacylates L over D roughly fourfold, and the mirror L-ribose stem shows the opposite preference.▷ An exhaustive search of self-aminoacylating ribozymes — still built on D-ribose — finds a substantial fraction preferentially make the *D* product, with direction conserved within sequence families.▷

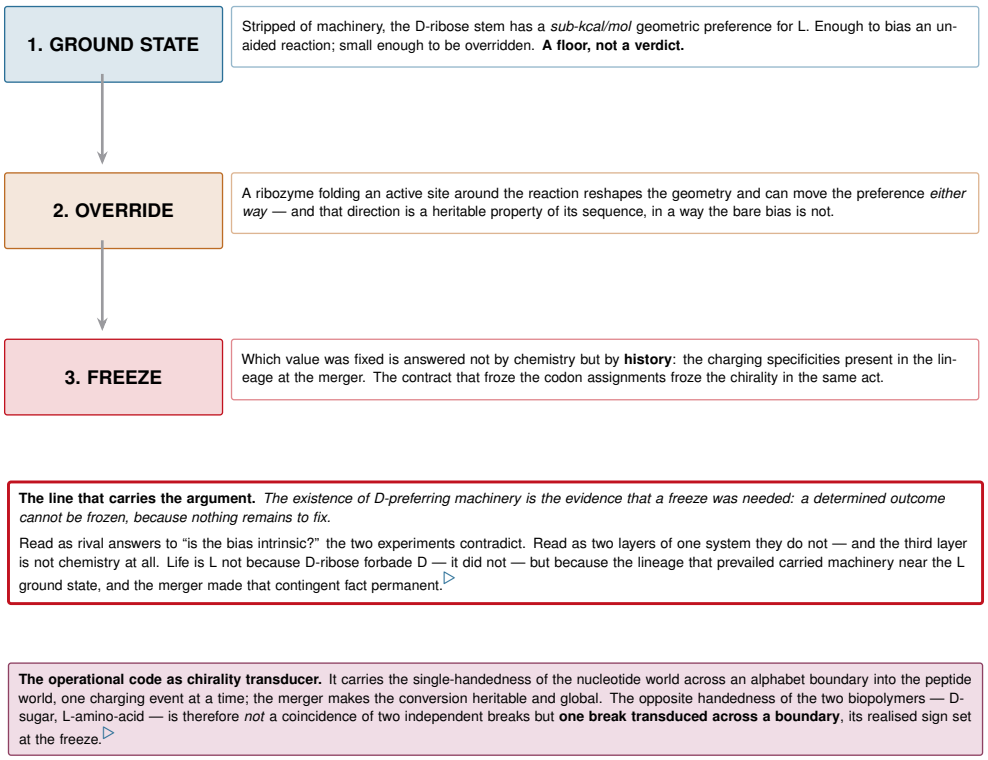


Figure 25.2: Ground state, override, freeze. Two experiments that look contradictory become two layers of one system, with the determining layer being neither — it is historical contingency made permanent by the merger. This reconciliation is the paper's strongest single move.

PREDICTION — Face two's falsifiable edge

Enantiomeric editing — the rejection of D-amino acids — should be *geometrically derivable* from the acceptor-stem recognition surface and *phylogenetically co-located* with the operational code, not the anticodon layer.
If D-rejection is instead a late anticodon-era addition, handedness was fixed upstream by some other agency and the code merely inherited it. **The model forbids that world.**▷

25.4 Face three — DNA handedness is inherited, not chosen

Two unrelated things are routinely conflated: *backbone* chirality (D-deoxyribose monomer asymmetry) and *helical* handedness (right-handed twist of the assembled duplex). The catma reaches one and is silent on the other in an informative way.

Backbone	D-deoxyribose because deoxyribose is *made from* ribose: ribonucleotide reductase removes the 2′-OH and leaves every other stereocentre untouched. Not transduction, not a freeze — straight biosynthetic inheritance, "the way a serial number survives a photocopy"▷
Helix	Forced. Given D-deoxyribose monomers and base-stacking, the stable duplex of ordinary sequence can only wind right. Z-DNA is a strained, high-energy, special-sequence local conformation built from the *same* D-monomers; no freestanding left-handed B-equivalent exists▷

The restraint is what makes it credible

The catma explains why DNA is chiral at all — it traces to the same D-sugar root — **without claiming the merger chose DNA's handedness.** The merger set the protein handedness and committed the lineage to a coded archive; the archive's handedness it simply inherited.
A framework claiming four independent freezes would be worth less than one claiming one freeze and deriving three consequences. Declining the extra claim is the move that earns the others.

PREDICTION — Face three's falsifiable edge

If backbone chirality is inherited rather than re-selected, ribonucleotide reductase should be **stereochemically conservative** — a clean pass-through with no enantioselective machinery of its own at the backbone centres. A reductase that re-selected handedness would mean DNA chirality was an independent freeze, which the catma forbids.▷

25.5 Face four — the sugar-network the freeze acted upon

The freeze needed something to freeze. Before templated sequence, a non-templated regime can carry information in the *relational composition and linkage-topology* of a self-perpetuating molecular set rather than in a strand. Sugars sit naturally at its centre: amphiphilic, mutually catalytic, prone to self-organisation.▷

Generator	autocatalytic formose-type sugar chemistry
Boundary	glycolipid vesicles
Information	relational composition of the sugar set — *non-sequence*
Function	lectin-like recognition; the glycan coat
Heredity	low-fidelity compositional transfer at fission (GARD)
Ceiling	**compositional error catastrophe** below narrow conditions

The ceiling does not kill the idea; it *locates* it. A glyco-world is not a self-sufficient living world but a **scaffold phase** — later co-arising with templated sequence, after which sequence took the carrier-of-heredity role and the sugar stayed on as backbone, metabolism and recognition surface. The template-independent glyco-code every cell still carries is the canalised residue.▷

Where I would push — the thinnest seam

Conceding "scaffold phase" in the face of the compositional error catastrophe buys survival at a cost: if the glyco-world is only ever a scaffold that left residue, what observation distinguishes it from *"sugars were around and got used"*?
The falsifiable edge is doing all the work here and should be foregrounded rather than closing the section. As drafted, Face Four is the one face a hostile reader could accept in full without conceding anything.⋆

PREDICTION — Face four's falsifiable edge — and it is testable with E10

If the glyco-world was generator and early carrier, the template-independent glycan code should be the **most ancient and pan-domain** of the non-sequence information systems, and its enzymatic core should *predate* the elaboration of templated machinery rather than postdate it.▷
And the fidelity ceiling is directly computable in the E10 framework (§21). GARD compositional inheritance is a mutual-catalysis network — nodes and hyperedges — which is the same formalism as the bio:geo reaction hypergraph. Compositional transfer fidelity as a function of network topology is a calculation, not a speculation.⋆

§26

The spine, and what it does to the framework

26.1 One freeze, four faces

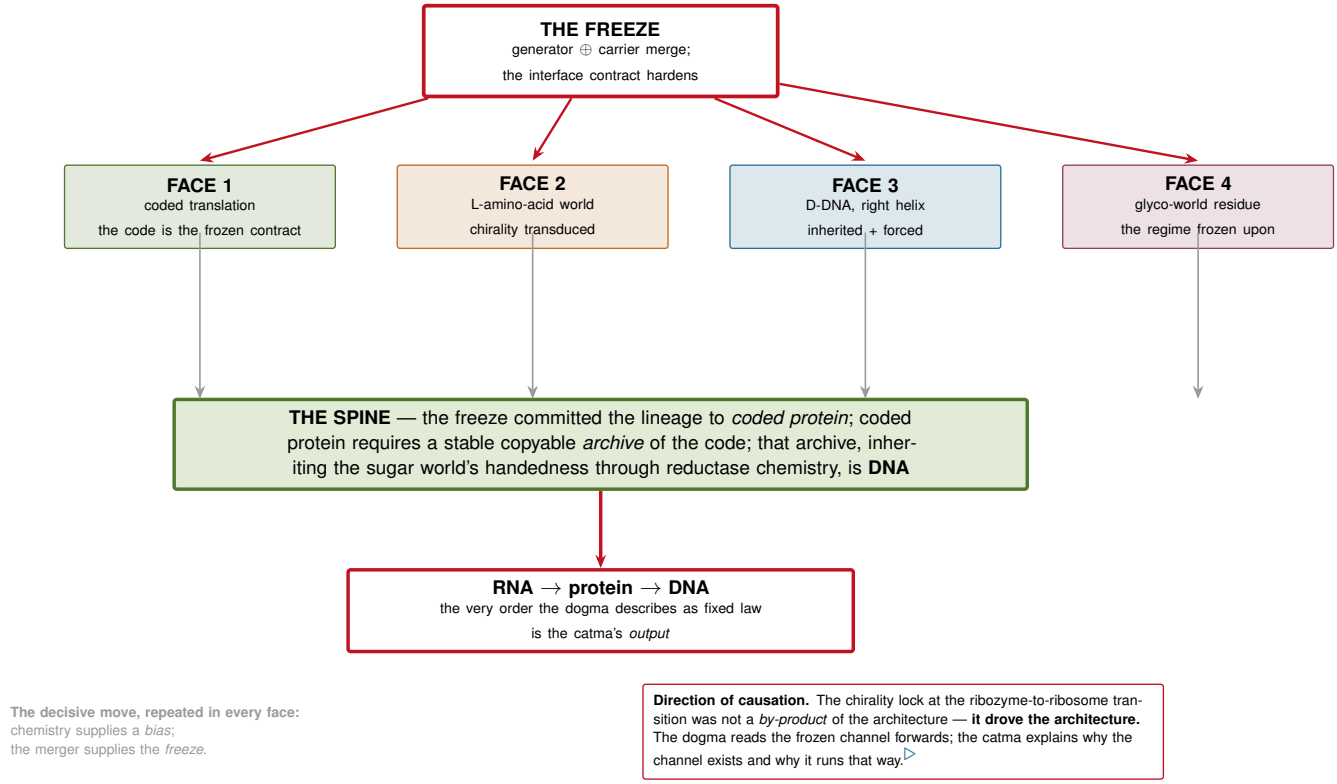


Figure 26.1: One event, four faces, and the dogma's ordering as output. Each face is the same operation seen from a different alphabet: chemistry biases, the merger freezes. The RNA → protein → DNA sequence is not assumed — it is derived from the commitment to coded protein and the consequent requirement for an archive.

26.2 What this does to C7

C7 is promoted from the weakest commitment to one of the strongest

As drafted in Chapter 1, C7 read: *"the transition is RNA → RNA-protein-DNA"*, marked *consequence; mainstream-adjacent; the framework adds timing, not the claim.*
That assessment is now wrong. With the catma, the ordering is **derived from a mechanism** rather than asserted as staging: the freeze commits the lineage to coded protein, coded protein requires an archive, the archive inherits its handedness.
Revised C7: *the RNA → protein → DNA order is the forced consequence of a single merger-and-freeze event, not a sequence of inventions.*
This moves C7 from a consequence to a premise-level claim with its own four falsifiable edges, and it changes the dependency graph: C7 now feeds the spine rather than hanging off it.

26.3 The third convergence on the phase boundary

This is the observation I would stake most on, because it was not designed in.

Three components of this volume, developed independently, land on the same object:
C3 (§3) — FUCA → LUCA is the switch from exploration-limited to throughflow-limited selection.
The coupling gap (§22) — the bench programme's central engineering obstacle is that same boundary: a system without coupling can only explore, because nothing it finds is worth holding.
The freeze (this Part) — generator meets carrier; the contract hardens; variation stops being disposable. Face Four states it in the paper's own words: the shift from compositional to template inheritance is *a persistence-filter event — the same event Face one calls the merger*.▷

All three describe: **a variation-generating system meeting a memory system, after which persistence begins to pay.** The catma specifies *which* generator and *which* carrier, at the molecular level, with named components and experimental anchors.

What the convergence licenses, and what it does not

Licenses: treating C3 as the abstract form and the catma's merger as its molecular instance. That is a real gain — C3 was a physical framing without a mechanism, and the catma is a mechanism that was missing a framing.
Does not license: claiming the convergence as evidence. Three descriptions generated by one mind under one set of intuitions are not independent observations, whatever it feels like from the inside. The paper's author-contributions note makes a stronger version of this claim — that the peptide and sugar faces were derived independently by two authors and found to share one event — and even there, independence is asserted rather than demonstrated.
The convergence is a reason to look, not a result. What makes it worth anything is that all three now point at the same testable edges.⋆

26.4 New plays entering the register

	Play	Resource	Tests	Character
E12	Enantiomeric-editing geometry: is D-rejection derivable from the acceptor-stem surface, and co-located with the operational code rather than the anticodon layer?	structures + phylogenomics	Face 2	Sharp, forbids a specific world
E13	Ribonucleotide reductase stereochemical conservatism — clean pass-through, no enantioselective machinery at backbone centres	existing structural/mechanistic literature	Face 3	**Cheapest in either volume.** A careful literature check
E14	Glycan-code antiquity: is the template-independent glycan code the most ancient and pan-domain non-sequence system, with an enzymatic core predating templated machinery?	phylogenomics	Face 4	Carries Face Four on its own
E15	GARD fidelity ceiling computed on the E10 hypergraph: compositional transfer fidelity vs. network topology	compute, gated by E10	Face 4	Converts a conceded weakness into a number

And two already in the register turn out to be catma tests. E3 (§9, tRNA two-code lineage variability) *is* Face One's falsifiable edge, derived in Part III before this paper was available. E4 (§10, the Rodin–Ohno structure-prediction test) targets the same acceptor-stem-era object the merger is built from; Rodin & Ohno is cited in the paper. The test programme for Face One was already written. That is worth noting for the same reason the convergence above is — as a reason to prioritise those two plays, not as evidence for anything.

Three things I would mark before circulation

1. The sharpness of the event is assumed. "Freeze" requires the merger to be single and sharp. Recurrent mergers followed by lineage sorting would produce a similar end state by a different route, and would not be a freeze. The assumption is load-bearing and currently unmarked.⋆

2. Face Four is the one a hostile reader can accept in full without conceding anything. See §25. E14 and E15 are the fix.

3. "One freeze" risks over-unification — four faces traced to a root selected to accommodate four faces. The defence is entirely the falsifiable edges, which is an argument for foregrounding them rather than closing each section with them.

PART X

The architecture, resolved

§27

FUCA names two things

27.1 Why the fork would not close

Chapters 1 and 6 carried a contradiction that survived several revisions. The master document declared *FUCA is a class* — the Cheshires, plural, Wonderland-bound. Earlier work, and C5 as drafted, declared *FUCA is the sole survivor of the first mass extinction*. Both were defended; neither would yield.

They were both right about different referents, and the resolution is notational.▷

> **FUCA-class** — the hammerhead ocean. Plural, contemporaneous, pre-translation. The Cheshires. Membership: compartmentalised, throughflow-eating, heritable-carrier-bearing.
> **FUCA 1.0** — the sole survivor of the *translation* sweep.
> **LUCA 1.0** — the sole survivor of the *DNA* sweep.

The version number is doing the work, and once it is written down the standard definitions survive untouched: FUCA is still first, LUCA is still last, and the framework supplies what the field cannot see — **one more sweep than anyone counts, and the epoch between them.**

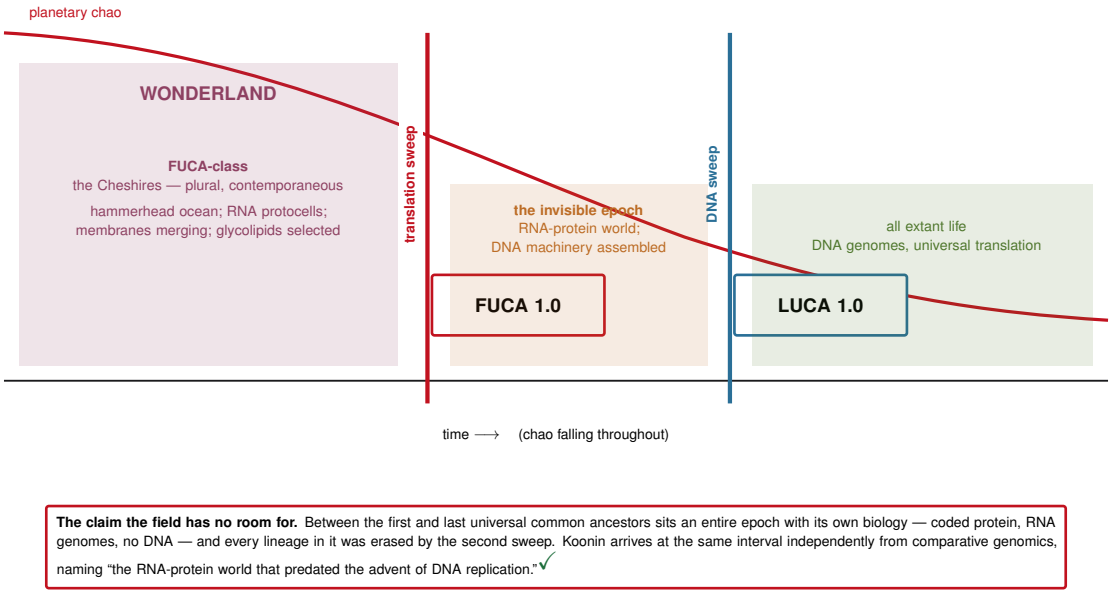


Figure 27.1: Two sweeps, three named objects, and the epoch nobody counts. FUCA-class is the plural pre-translation population; FUCA 1.0 survives the translation sweep; LUCA 1.0 survives the DNA sweep. Standard definitions are preserved — what is added is the interval between them.

§28

Four sweeps, and the carrier has a direction

28.1 The series

The Catma's spine gives the order: sugar and glycolipid, then RNA, then protein, then DNA.▷ Each is a phase change; each is a sweep. Run the throughflow argument across all four and the ordering stops being a list and becomes a consequence.

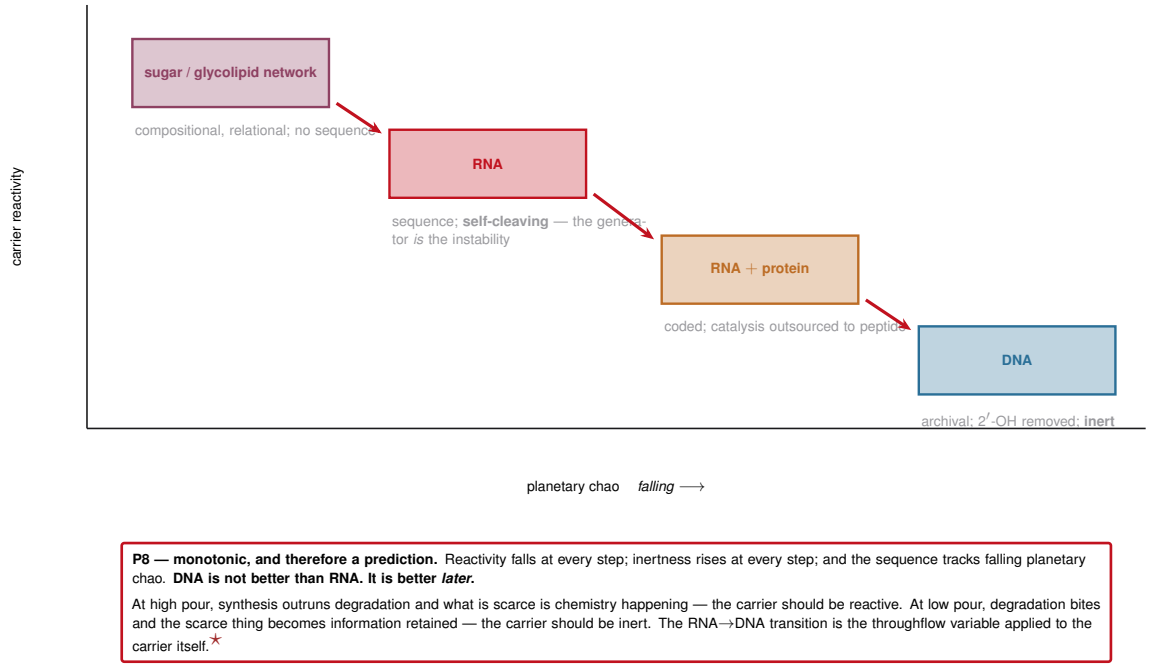


Figure 28.1: Carrier reactivity falls monotonically as planetary chao falls. Four carriers, four sweeps, one control variable. The ordering is not a just-so sequence: at each stage the carrier is the one whose reactivity matches the pour.

28.2 DNA is RNA with the chemistry switched off

The mechanism, because the prediction rests on it.

One structural difference does nearly all the work: **2′-OH versus 2′-H**. The 2′-hydroxyl is the nucleophile that attacks the adjacent phosphodiester and cleaves the backbone — RNA hydrolyses itself. Remove it and the intramolecular attack pathway is deleted; DNA is roughly five orders of magnitude more stable.○ The same hydroxyl is a catalytic group, so removing it also costs ribozyme chemistry. And the sugar pucker changes with it: 2′-OH forces C3′-endo and A-form; without it, C2′-endo and B-form, which is what permits long sequence-independent duplexes to stack regularly.○

Then U→T. Cytosine deaminates spontaneously to uracil; if U is a normal base, damage is indistinguishable from signal. Methylating it to thymine makes deamination *detectable and repairable*.○

Both changes are the same move: **deactivation for fidelity**. DNA is not an elaboration of RNA. It is RNA with the reactive groups removed — which is only worth paying for once

the pour has slowed enough that information loss, rather than chemical throughput, is the binding constraint.

And the RNA-stability objection inverts

The standard attack on an RNA world at a vent is that RNA hydrolyses too fast, worse when hot and alkaline. That attack assumes *stability* is the requirement.
A flame is not stable. It persists by throughflow. **What is needed is not stable RNA but synthesis faster than loss** — the molecules turn over, the network persists. Dissipative structure applied to the carrier itself.
Which is why the hammerhead ocean is the strongest part of the picture. Hammerheads are *self-cleaving*. Under a stability-first model that is incoherent — why would the dominant ribozyme destroy RNA? Under a throughflow model it is the generator: **cleavage writes tickets**, generating variants faster than any copying-error rate. The instability is not a problem the model must survive. It *is* the raffle.▷★

28.3 The hammerhead premise, checked

The premise was verified against the literature and it comes back *split* — which is worth recording precisely, because the half that fails is the half the argument does not need.

Distribution: confirmed, and more strongly than claimed. Comparative sequence analysis has recovered thousands of hammerhead sequences across all domains of life, including the first representatives in fungi and archaea — at least three in archaea and hundreds in bacteria.✓ [Perreault et al. 2011 PLoS Comput Biol] Independent searches found the motif across Bacteria, Chromalveolata, Plantae and Metazoa, in water moulds and twenty plant species from unicellular algae to vascular plants, with Galapagos metagenomic examples described by their authors as material that "could be regarded as direct relics from the RNA world."✓ [de la Peña & García-Robles 2010 RNA 16:1943] There are ultraconserved hammerheads in human introns, and their biological functions remain largely unknown.✓

The relic reading: contested. Most genomic hammerheads sit in mobile elements — retrozymes, Penelope-like elements, and in bacteria frequently near integrase genes grouped with prophages. The reviews are direct: most current data point to a common role in the propagation of diverse retrotransposons.✓ Mobile elements spread horizontally, so pan-domain presence does not by itself establish vertical descent.

And one datum settles it. R2 retrotransposons in *Drosophila* encode HDV-like ribozymes sharing 21 identical nucleotide positions in and around the active site with the hepatitis delta ribozyme — and tracing their sequence evolution shows that similarity to be **the result of convergent evolution, not common descent**.✓ [Eickbush & Eickbush 2010 MCB 30:3142]

The result that breaks the relic reading is the one the framework needs

Weak version (as previously implied): modern hammerheads are direct relics, therefore hammerheads were prebiotically abundant. *Undercut.*
Strong version (what the argument actually requires): self-cleaving RNA motifs are so accessible in sequence space that they are independently re-derived across all domains — therefore a prebiotic RNA pool was saturated with them. *Supported, and supported by the very convergence that undercuts the weak version.*
The premise needs the motif to be **abundant**, not **inherited**. Evidence against inheritance is evidence for accessibility, and accessibility is the premise.

And the framework already contains the correct handling. The Catma writes of non-ribosomal peptide synthesis: *"its patchy, accessory distribution and its shared thiotemplate logic mark it as a convergent canal, re-derived wherever the chemistry pays; the ancestral object is the activated-aminoacyl chemistry it exploits."*▷ Apply it unchanged — hammerheads are a convergent canal, and the ancestral object is self-cleaving RNA chemistry. Independent re-derivation of one motif from one alphabet is also, exactly, *tabby* novelty.⋆

If an argued relic is wanted, use different molecules. The strong candidates are the ones performing key functions and living outside mobile elements: peptide bond formation by the ribosome, mRNA splicing by the spliceosome, and — most relevant here — **tRNA maturation by RNase P**.✓ A surviving catalytic RNA whose job is processing tRNA sits directly on the tRNA-as-replicator claim, and is a far better relic argument than the hammerhead.⋆

§29

Ignition events are mergers

29.1 The generalisation

The Catma states a methodological constraint for translation: *an account of the origin of translation must not select for translation.*[▷] That constraint is not about translation.

The general form

You can never select for a thing that does not yet exist.
Therefore any genuinely novel capability must arise from components each selected for something else — and the cleanest way that happens is two independently selected lineages coming to occupy one compartment.
Therefore ignition events are mergers. And that is *why* they sit behind a high conjunction threshold: not because the chemistry is hard, but because a merger requires two *independent histories* to intersect in one place. That is the conjunction. That is why ignition events are singletons, and why re-canalisation events — a single lineage adapting, no coincidence required — are plural and convergent.[★]

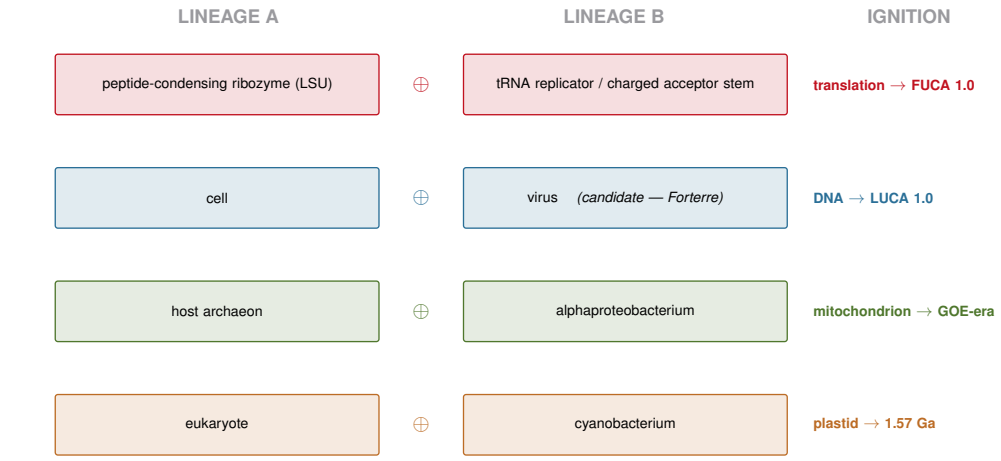


Figure 29.1: Four ignition events, four mergers. The requirement that two independent histories intersect *is* the high conjunction threshold that defines ignition in the taxonomy — which converts a stipulated property into a derived one.

§30

The homology pattern reads out the trigger mode

The framework's two trigger modes make different predictions about what a system looks like afterwards, and the prediction is checkable on any conserved machinery.

	Mechanism	Predicted signature
Endogenous	one lineage acquires an attribute and sweeps; all descendants inherit *that* solution	**universal and homologous**
Exogenous	an external condition changes and presses on every lineage at once; each solves it separately	**convergent and non-homologous**

Now test it on the two systems that matter.

Translation is the most conserved system in biology. The ribosome, the code, the synthetases — homologous across all three domains. **Endogenous sweep.** ✓

DNA replication is not. The replicative polymerases belong to unrelated protein families with different folds, unlikely to share common ancestry, in sharp contrast to the transcription and translation machinery which is highly conserved in all divisions of life.✓ [Leipe, Aravind & Koonin 1999 NAR; Koonin 2020] The field's own verdict is that the reason is unknown, and Koonin's opening word for it is "enigma."✓

The distinction that saves the definition

DNA as genome is universal — every extant organism has one. **DNA replicase** is not. Those are different objects, and the diagnostic must be pointed at the right one. The *genome* transition is a single endogenous sweep, with the universal outcome the diagnostic predicts. The *replisome* non-homology is a later displacement: bacterial PolB-family polymerases appear to be of viral origin.✓

So a DNA genome can be the defining attribute of LUCA — the framework's position▷ — *and* the replisomes can be non-homologous, without contradiction. The sweep fixed the genome; viruses swapped the machinery afterwards.

PREDICTION — P1, rebuilt — two coalescence depths

This replaces the bottleneck-shape test, which was the weakest item in the register. Two sweeps at two depths predict **two distinct coalescence horizons**: gene families tracing to translation machinery should coalesce measurably deeper than families tracing to DNA replication.

A one-sweep model predicts a single horizon. The test uses existing phylogenomic data, discriminates directly between the two architectures, and does not depend on tree-shape statistics that early HGT can manufacture or destroy.*

§31

Generation and retention are different, and Venus proves it

31.1 The come-apart

The lexicon set a test for whether irreversibility-that-accumulates needs its own process-word: coin one only if it comes apart from channel-deepening.

Venus is the case. Venus has foreclosure in abundance — structures form and are annihilated, alternatives closed off, every cycle. What it has no trace of is *retention*. Nothing accumulates; each turn of the convection cell starts from zero.

Foreclosure without retention exists, it is a whole planet, and therefore one word cannot carry both jobs.*

The threshold, corrected: 1 chao is a *lip*, not a constant

An earlier draft of this chapter stated 1 chao as an absolute threshold. That is wrong, and the correction carries a result.

1 chao is the lip of the canal you are currently in. Water at the top of a volcano runs down, canalises, and pools. To leave the channel it must overtop the lip — and the deeper the channel, the higher the lip. So the threshold is *normalised against canalisation depth*, not fixed.

Which gives, for free, why specialists die in threshold events. *Canalisation depth is threshold height.* The same event that liberates a generalist cannot move a deeply specialised lineage at all — it simply kills it. Specialisation buys efficiency by raising the wall required to leave.

The word *can* is load-bearing. It is a **permission** condition, not a sufficiency condition — it licenses ignition without guaranteeing it, and says nothing about whether anything persists afterwards.

That is what makes the sort below work. Venus sits above threshold: life *can* form, and on the framework's reading does so constantly — and is constantly annihilated, because permission to ignite is not permission to keep anything. Titan sits below: the generator cannot run at all. Earth sits above *and* retains.

It also dissolves an apparent problem. Earth is still above threshold today — life holds it there by throughflow — so why no second genesis? Because chao ≥ 1 permits ignition where there is *un-eaten gradient*, and the extant biosphere eats them. The window closed biologically, not thermally.

Vacancy 1, filled: *cheshiring*. The Cheshire is gone but for the grin left in the record. It names precisely what the brief asks — what persists after the thing that made it is gone, and constrains the next round. Morphologically exact against the existing set (cattening, doggening, vurting, tigering), and *unreifiable*: "a cheshiring" does not parse as a thing, which is the test that retired "ratchet". Preferred over *grinning*, which carries an ordinary sense strong enough to doggen.

And it gives the pair: **canalising cuts down, cheshiring builds up.** Two irreversible processes, opposite directions, visibly distinct.

Venus fails on retention, not on generation — which is why "life forming and annihilating constantly" is a prediction rather than a hedge. A world that generates and never retains can only ever show *unpaired, non-cyclic* disequilibrium, because nothing survives long enough to close a loop. ▷

Titan fails on generation, and fails robustly: the peptide carrier polymer addresses *activation energy*, but a viscous slurry at 94 K is *transport*-limited. Fixing the barrier does not help if reactants cannot meet. Two independent failures, one remedy — a better limit case than a rescue would have been. ▷

Figure 31.1: Generation against retention sorts the candidate worlds. The Drake reformulation stops being three arbitrary terms: ignition needs generation, persistence needs retention, elaboration needs both plus time. Venus is ignition without persistence, made observable.

Sober alternative for a physicist audience: *terracing* — a river terrace is literally a floor the next round starts from and can never rise above.

Vacancy 2 resolves by staying vacant. Naming the three-process unity violates the blank-abstract convention, which requires the reader to assemble the identity. Paper 1 §5 already says so. Cut the name from Paper 0's beat sheet, keep the three absences, and the paper ships.*

Revised commitments

	As drafted	Revised
C5	FUCA is the sole survivor of the first mass extinction	**FUCA-class** is the plural pre-translation population; **FUCA 1.0** is the translation-sweep survivor. Two objects, distinguished by version
C7	the transition is RNA → RNA-protein-DNA	**Four sweeps** — sugar/glycolipid → RNA → protein → DNA — each a phase change, with carrier reactivity falling monotonically as chao falls. **A DNA genome is the defining attribute of LUCA**
C3	exploration-limited → throughflow-limited	Unchanged, but now with a mechanism: **ignition events are mergers**, and the merger requirement *is* the conjunction threshold

32.1 New plays and predictions

	Item	Resource	Tests
E16	Membrane-fusion merger assay: two vesicle populations, distinct encapsulated cargo, induced fusion, assay for function neither had alone	the droplet platform	Catma Face 1 — supplies the mechanism the paper never states
E17	Two coalescence horizons (translation vs. replication families)	existing phylogenomics	the two-sweep architecture; replaces old P1
E18	The 10/10 split: prebiotic amino-acid set vs. aaRS class assignment	existing databases	the doubling event = Rodin–Ohno duplication
E19	Gaoyuzhuang excursion vs. Columbia/Nuna rifting — discriminate by coincidence with the eukaryote fossil record	published literature	the 1.57 Ga installation event
E20′	**Motif accessibility** (revised): frequency of the minimal catalytic consensus in random and shuffled sequence; number of *independent* origins inferable on the tree, counting mobile-element copies separately	existing sequence data, compute	the hammerhead-ocean premise — supports it either way the relic question falls
P8	Carrier reactivity falls monotonically with planetary chao	—	C7 revised
P9	Ignition events are mergers; recanalisation events are not	—	C3, taxonomy
P10	Threshold events cluster at chao maxima (the apple lands at the peak of the party)	—	the catalyst model

Three unexplained observations, one engine

The pattern worth building the taxonomy paper around is not "here is a new idea." It is: **here are three things the field accepts and explicitly cannot explain, and one engine accounts for all of them.**

1. The 1.57 Ga excursion. Multi-proxy oxygenation pulse at the predicted coordinate, with the plastid literature stating in print that establishment-era selective pressures are distinct and unidentified.✓

2. The replication non-homology. The most conserved system in biology sits beside the least conserved, and the field calls the reason "unknown."✓

3. Punctuated equilibrium. Eldredge and Gould described the pattern and never supplied a mechanism for the punctuation; it has been open for fifty years.○ The framework supplies one — punctuation is a threshold event, stasis is canalisation.

Three is a pattern. One is an anecdote.★

§33

Chao, corrected and extended

33.1 Scale: local, systemic, Pandæmonium

The volcano map resolves the scale axis that never reached the master.

Local $>$ 1 chao — a canal carves into a side-vent of magma. A fire, a plague, an ice age. One channel is overtopped; the rest of the system is unmoved.
Systemic $>$ 1 chao — Earth as a whole above threshold. Tornadoes, lightning, earthquakes, storms: the planet forcing phase changes continuously.
$\gg$ 1 **chao — Pandæmonium.** Asteroid impact. *Every* phase change available is explored. There is geological evidence for what this looks like.

Two consequences

1. Ignition and re-canalisation may be one axis with a scale parameter. A local ignition (side-vent) and a systemic one (eruption) are the same threshold crossing at different scales — which is the third candidate resolution to the reconcile-the-two-sorts problem, and the cleanest of the three.
2. Poly-extremophily is not a coincidence, it is a definition. Survivors of sustained Pandæmonium are expected to be unusual, because the filter is defined by whatever tolerated an all-axes drive. **Poly-extremophiles are *always* FUCA-class organisms**, and LUCA therefore requires a poly-extremophile antecedent.

Not all chao dimensions are equal, and the axis has a sign. An ice age is disequilibrium with *less* free energy — it reduces the variation available rather than opening it. Hyperthermophiles are advantaged over cryophiles because temperature correlates with usable free energy where cold anti-correlates with it.
This is why cold is the low-dimensional case, why Titan fails, and why *magnitude* of chao is not sufficient without asking which axis.

33.2 The derivatives — and a mechanism for punctuated equilibrium

PREDICTION — P11 — acceleration, not magnitude, sets who survives

Chao has derivatives.The first is the rate of environmental change; the second is its acceleration, against the lineage's own rate of adaptation.
A lineage can track a *fast* change if it is *steady* — canalise, follow the gradient. It cannot track an **accelerating** one, because adaptation has a response time.
So d^2chao$/\mathrm{d}t^2$ is the variable that decides who passes the filter — and that is a mechanism for punctuated equilibrium that the pattern has never had. *Stasis is low acceleration; punctuation is high acceleration.* Not magnitude. Acceleration.
Eldredge and Gould described the pattern in 1972 and left the punctuation mechanism open. This is a candidate, and it is quantitative.*

And there is a ceiling as well as a floor. Chao measures the tension between order and disorder; *either end is death* — a snap of the tension.▷Zero chao is equilibrium and dead. Unbounded chao is decoupling. Pandæmonium sits just under the upper bound, which is why it is genesis-permissive rather than merely destructive.

§34

Chirality is a timestamp

This closes the Pizzarello problem, and the closure is simpler than the patch previously proposed for it.

Before the sweep there is no fact of the matter

"Until protein synthesis cattens, of course both existed. When translation caused a selective sweep, it just started sucking up one enantiomer — and this is a time stamp."

Both enantiomers present. Catalysis running in both directions. Tamura–Schimmel's bare-stem L-preference and Pizzarello–Weber's L-amino-acids-make-D-sugars are *both* descriptions of a pre-sweep world in which neither wins, and they do not contradict because there is nothing yet to contradict about.

The sweep consumes one enantiomer. Direction of catalysis stops mattering, because only one is left.

So homochirality is not a puzzle requiring explanation. **It is a date** — and it dates the translation sweep, not a chemical event.

The mutual-attractor construction proposed earlier is unnecessary machinery and is withdrawn. Nothing needs to select a handedness in advance; the sweep does the work, and the asymmetry we observe is its receipt.

§35

The amino acids

The framework's weakest section, and the one with the most available leverage.

35.1 Where the literature stands

Consensus recruitment order (Trifonov, synthesised from ∼40 independent criteria): early — Gly, Ala, Asp, Val, Pro, Ser, Glu, Leu, Thr; late — Cys, Phe, Met, Tyr, Trp.○
The published "first four": Gly, Ala, Asp, Val — the amino acids of the GNC codons, argued on thermodynamic grounds.○ [Higgs & Pudritz 2009, arXiv:0904.0402]
Prebiotic availability (Miller–Urey and Murchison): essentially Trifonov's first ten, with glycine dominating by a wide margin.○

35.2 The negative result that hands you the argument

Amend & Shock computed Gibbs energies of amino acid synthesis under hydrothermal conditions and found that *net synthesis reactions are exergonic for the 10 most reduced amino acids* — Leu, Ile, Val, Lys, Phe, Met, Pro, Tyr, Trp, Ala — while all 20 are endergonic in cold oxidised surface seawater.✓ [Amend & Shock 1998, Science 281:1659]

That second half is strong support for the vent cradle: **amino acid synthesis pays for itself at a vent and costs energy at the surface.**

But look at the exergonic list. It contains Trp, Tyr, Phe, Met and Lys — all *late* — and omits Gly, Asp, Glu and Ser — all *early*. Higgs & Pudritz checked this directly and found **no correlation between recruitment order and hydrothermal** ΔG, with early and late amino acids spanning the full range.✓ [arXiv:0904.0402]

Availability and recruitment decouple — which is the finding

Thermodynamics does not predict the order. Neither does availability: sulfur was abundant in an anoxic sulfidic ocean, and **cysteine is late**. Iron–sulfur chemistry sits at LUCA's catalytic core, and the amino acid that carries sulfur was recruited near the end.

If the environment supplied it and the code did not take it, then **the arbiter is the charging machinery, not the environment**. Recruitment order is set inside the system. That is a framework claim, it is testable, and half of it is already confirmed as a negative.⋆

35.3 What the acceptor stem would have favoured

PREDICTION — P12 — recruitment order tracks identity-element simplicity

If the acceptor stem is the arbiter, recruitment order should correlate with the *simplicity of the acceptor-stem identity element*, and *not* with hydrothermal ΔG (already shown) or with prebiotic abundance alone.
Three grounds, converging:

Sterics. Aminoacylation proceeds by nucleophilic attack from the 2′/3′-OH of the terminal adenosine. Smaller side chains, lower barrier. Glycine has none at all.
Identity-element minimality. Alanine identity is carried by a *single wobble pair*, G3:U70 — the most minimal recognition element in the entire system.[°] If simplicity of recognition tracks antiquity, alanine is very early, and it is early in Trifonov's order.
Chirality. The bare D-ribose minihelix has a ∼4× L-preference. **Glycine is achiral** — it has no enantiomer, so it charges with no chiral discrimination whatever.[⋆]

⋆ The freeze dates to the *second* amino acid

Glycine being achiral has a consequence nobody appears to have drawn.
A code containing only glycine **has no handedness**. There is nothing to freeze. The chirality lock therefore cannot predate the recruitment of the first *chiral* residue — and on the grounds above, that is alanine.
So the Catma's freeze dates to the Gly → Ala step, not to the origin of charging. That is a sharper timestamp than the framework currently claims, and it makes the alanine G3:U70 element the single most interesting object in the system: the minimal identity element, on the first chiral amino acid, at the moment handedness became possible.[⋆]

On the doubling. Candidate mechanisms — frameshift, imperfect transcription, viral duplication, fusion, polyploid-style whole-set duplication — all point at duplication.[▷]
Worth knowing: **Susumu Ohno wrote *Evolution by Gene Duplication* (1970)**, the founding text for duplication as the engine of evolutionary novelty — *and* co-proposed the complementary-strand origin of the two synthetase classes as Rodin & Ohno (1995).[°] The person who established that duplication drives evolution also proposed the specific duplication at the root of translation.
Discriminating test (extends E17): gradual accretion predicts a *smear* of recruitment times; a doubling event predicts **two clusters** — and the clusters should align with aaRS class rather than with biosynthetic cost.[⋆]

§36

Graphs, and the patron saints

36.1 Each sweep raises the graph order

The carrier series has a second reading, and it may be the same axis viewed twice.▷

Sweep	Carrier	Graph order
sugar network	compositional	a **multiset** — composition without topology (GARD carries *what*, not *how connected*)
RNA	sequence	a **path**, folding into a planar graph in 3-D
protein	coded mapping	**bipartite** — codon ↔ amino acid
DNA	duplex	the graph made **copyable**
endosymbiosis	merged networks	**hypergraph / multigraph**

Carrier reactivity falls (P8); **graph order rises**. If those are one axis, the carrier series stops being a sequence and becomes a statement about representational capacity: each sweep buys a strictly more expressive structure at the cost of reactivity.

And metabolism is graph transformation throughout — isomerisation is relabelling, enzymes are constraint satisfaction at binding sites, glycolysis is a rewrite sequence.▷

This closes a loop that was already open. `framework.pdf` names Petri nets as the formal layer — extinction as an unreachable marking, re-canalisation as a shifted T-invariant. **The graph-order escalation is what Petri nets would compute.** The motif and the formalism were built for each other.

36.2 Eris and Hephaestus

The two patron saints, and why the lameness matters

Eris — the Purge. Change, environment, information destruction. The draw that discloses a predicate.
Hephaestus — generating. The imperfect toolmaker: microbes, viruses, proteins. Information creation by combination.▷

The imperfection is the mechanism, not a flaw in the metaphor. A perfect toolmaker produces no variety. Fidelity below unity is precisely what writes tickets — which is the same point the hammerhead makes at the molecular level, and the same point the polymerase-ribozyme error threshold makes at the bench.
So the pairing is not destruction versus creation. It is **imperfect copying versus the draw** — the two processes, with the lameness load-bearing.

PART XI

Redundancy, and the reconstruction

§37

Redundancy is the requirement

37.1 A correction: merger was too narrow

Chapter 29 claimed that ignition events are mergers. That is a special case of something more general, and Ohno supplies the other half.

What ignition actually requires is a spare part

You can never select for a thing that does not yet exist (§25). So any novel capability must be built from components selected for something else — which means the requirement is **a component free to be repurposed.**

Merger supplies it from outside: two independent histories, each complete alone. **Duplication** supplies it from inside: one history, a redundant copy released from purifying selection. This is the thesis of *Evolution by Gene Duplication* (Ohno, 1970).

So the requirement is redundancy, and merger and duplication are two routes to it — which map onto the framework's two trigger modes without adjustment.

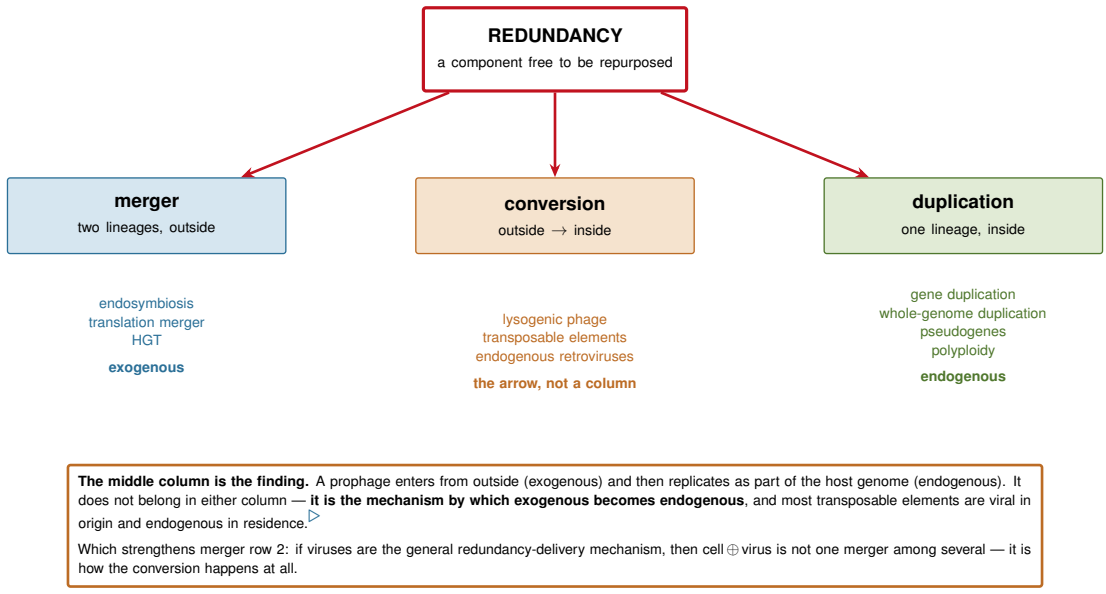


Figure 37.1: Three routes to a spare part, and the middle one is a process rather than a category. The framework's exogenous/endogenous split is complete only once conversion is named: lysogeny and transposition are how an outside source becomes an inside one.

37.2 The C-value paradox, and the objection it must survive

Genome size does not track organismal complexity. The standard *description* is that most DNA is non-coding. The framework offers a *function*: redundancy is stored ignition potential.

Not an adaptation — an exaptation substrate

The objection is immediate and it is the standard one: **you cannot select for future evolvability.** Orgel and Crick's selfish-DNA argument is sitting there waiting.[○]
The framework's answer is the answer it gives everywhere else. Redundancy is *not* selected for evolvability. It accumulates for unrelated reasons — selfish elements replicating for their own sake, polyploidy accidents, failed purging — and is **then available**. Nobody stores spare parts for a future they cannot see.

Which converts a paradox into a prediction (P13). Ignition events should be *preceded* by redundancy accumulation, not caused by it. Lineages that purge aggressively — streamlined genomes, *Prochlorococcus* — should show re-canalisation and not ignition. Lineages that hoard should show the reverse.
Plants fit: polyploidy is rampant, and whole-genome duplication is repeatedly associated with angiosperm radiations.[○⋆]

§38

Nodes saturate, edges proliferate

38.1 One pattern, four scales

The prediction is that as a system matures, novelty stops coming from new components and starts coming from new relations among existing ones.▷ That is graph-order escalation (§36) stated as a historical claim — and it recurs at every level.

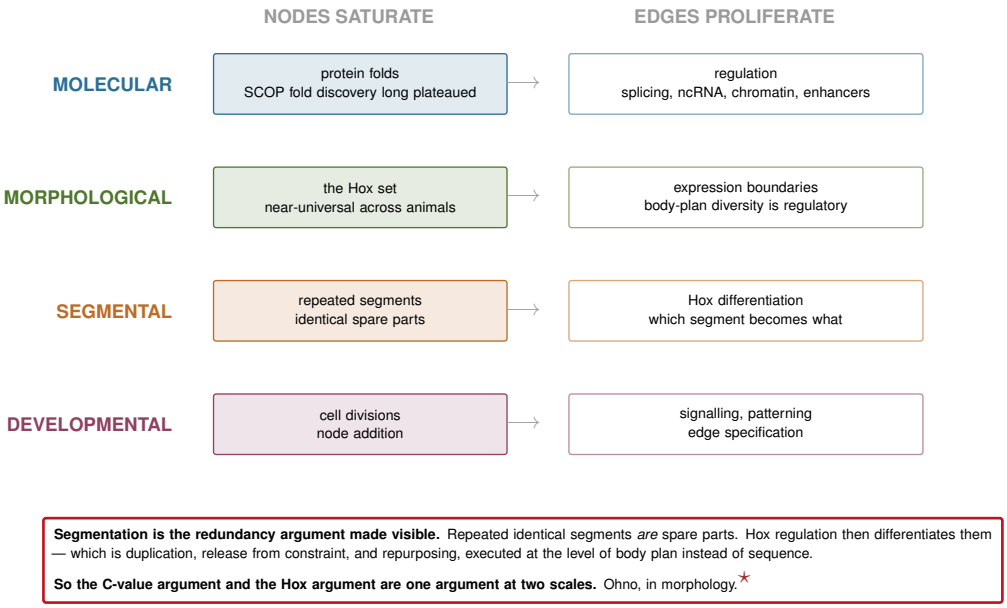


Figure 38.1: The same escalation at four levels. Once the component set is explored, novelty must come from relations. Fold space plateaus while regulatory complexity climbs; the Hox set is fixed while body plans diversify; segments repeat while their identities differentiate.

Hox colinearity is canalisation you can see in space. Order on the chromosome maps to order along the body axis — so the genome acquires a *spatial* dimension of meaning, and sequence becomes coordinate.
The cluster cannot rearrange **because the order is the information**. That is a canyon in chromosome space, with lip height proportional to how much of the body plan depends on it — which is why Hox clusters are among the most conserved architectures in animals, and why disrupting them is lethal rather than merely costly.★

PREDICTION — P14 — whole-genome duplication should leave ignition signatures

Lineages that underwent WGD should show *ignition* signatures — singleton, frozen architecture. Lineages that did not should show *re-canalisation* signatures — plural, convergent.
The vertebrate 2R event fits: one origin, deeply frozen. Testable across the tree with existing comparative genomics.★

§39

Later transitions

39.1 The Cambrian, and a test against the oxygen hypothesis

`framework.pdf` already stakes a claim here: palaeontologists describe a "sudden-then-channelled" Cambrian signature without a unifying mechanism.▷ That is canalisation, named by the field, with no engine.

But P11 gives something sharper than a relabelling.

PREDICTION — P15 — the Cambrian tracks acceleration, not concentration

The Ediacaran–Cambrian transition follows the Marinoan snowball deglaciation. A global glaciation followed by rapid deglaciation is not a magnitude event — it is an **acceleration** event. Very large $\mathrm{d}^2\mathrm{chao}/\mathrm{d}t^2$.
The mainstream explanation is an oxygen threshold, and it is contested precisely because the oxygen timing keeps refusing to line up cleanly.○
Discriminating prediction: the oxygen hypothesis says the radiation tracks a *concentration*; P11 says it tracks a *second derivative*. Those have different shapes in the record, and proxies for both exist.★

39.2 Cicadas — a known result, and the general form

Prime-numbered periodical cicada cycles as predator avoidance is established, proposed by Lloyd and Dybas in 1966 and popularised by Gould; a 12-year cycle meets predators on 2-, 3-, 4- and 6-year cycles, while a 17-year cicada meets a 3-year predator once in 51 years.✓ There is a live competitor: Yoshimura's hybridisation model, in which primes minimise interbrood interbreeding rather than predation.✓

The observation is not new. The generalisation may be. Both published explanations are domain-specific — predators, or hybridisation. The framework's version is not:
A prime period is maximal un-entrainability — anti-canalisation in the time domain. Any periodic filter canalises on *some* period, and a prime is the only strategy that decouples from all of them at once rather than from one.★
Frame it as the number-theoretic instance of a general strategy. That credits the existing work and takes the general claim.

§40

The blind reconstruction

40.1 LUCA's fossil is the empty node

An earlier draft treated the absence of LUCA body fossils as a limitation. That applies a palaeontologist's standard to a question that does not need one.▷

The shadow is the evidence

LUCA's fossil is the root of the rRNA tree. We know the node exists precisely because it is *empty* and *single* — everything descends from it and nothing occupies it. The shape of what radiated is the record of what was there.
This is biology-as-sole-record applied to morphology rather than metabolism, and the framework already argues it: *read the filter off the descendants, not the rocks.* Morphology is knowable through collective extant life, through wet-lab confirmation, and through simulation.▷

40.2 The pipeline, and why blind is the whole point

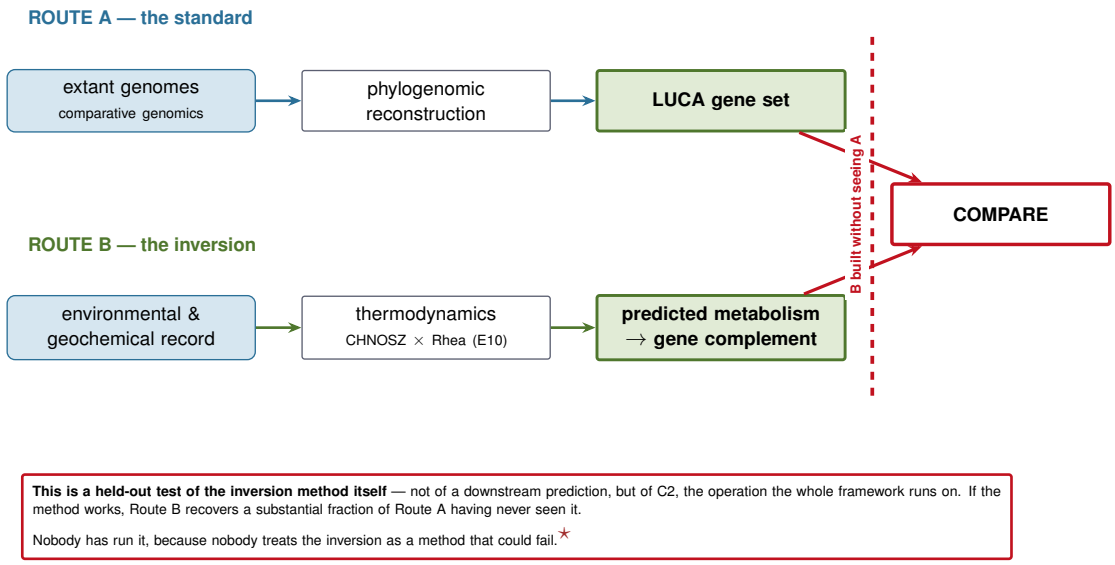


Figure 40.1: Two independent routes to LUCA, compared. Route A is comparative genomics; Route B is the inversion method run forward from the environment. Constructing B without reference to A converts the framework's core method into something falsifiable.

40.3 Closing the two-clocks test from the other end

EVIDENCE PLAY — E21 — predicted fractionation against the rock record

Genome→phenotype in general is unsolved. Genome→*metabolic* phenotype for a minimal autotroph is constraint-based modelling, which is mature.
reconstructed core metabolism → flux balance analysis → predicted carbon isotope fractionation → compare against Isua and Pilbara.
So instead of hunting for Wood–Ljungdahl-grade carbon and arguing about whether it is biological, **you predict the fractionation your reconstructed organism would have left, and check**. That is a far harder claim to dismiss than a threshold, and it closes E6 from the opposite direction.⋆

EVIDENCE PLAY — E22 — run FUCA → LUCA, accelerated

The most ambitious item in either volume, and not obviously impossible.
Establish reconstructed Hadean conditions — composition, temperature, pressure, redox, gradient — in a continuous-culture or droplet system, seed with the generator, and impose the persistence-filter. Then *run it*, with the pour sustained and the filter applied, and observe what survives.
The framework's own claim is that this is the natural experiment: the regime does the work, and the outcome is canalised by the conditions rather than designed. Directed evolution under reconstructed geochemistry is the wet-lab form of the inversion method.▷⋆
Honest scope: the accessible version is not "re-run abiogenesis." It is *does a defined starting population, under reconstructed conditions, canalise toward the predicted metabolism?* That is bounded, it is a real experiment, and a negative result constrains the model.

Reconstruction has a staging, and rung one is already in the register.

one gene → one cluster → one region → one genome.

E4 *is* ancestral sequence reconstruction — rebuilding a single ancestral duplex and asking whether both readings fold. Ancestral protein resurrection is mature; ancestral karyotype reconstruction is mature for vertebrates.○
What has never been attempted at depth is **ancestral gene order at LUCA**, because horizontal transfer shreds synteny. But the two-sweep model predicts something usable: *translation-machinery genes should retain ancestral synteny better than replication genes*, because they froze a sweep earlier. That is E17 with gene order substituted for coalescence times.⋆

PART XII

The derivations

§41

Show the chain, then the data

41.1 Why this Part exists

The framework has been presented as a set of claims with evidence attached. That is the wrong order, and it undersells what the thing actually is.

A derivation cannot be dredged

Pre-registration exists because a *dataset* can be dredged — you can search until something clears threshold, then report it as though you had expected it.
A derivation cannot be dredged. If the chain is stated, references nothing in the target dataset, and a reader can re-run it and arrive at the same coordinate, then whether the author thought of it before or after looking is *evidentially irrelevant*. Nobody pre-registers that general relativity predicts perihelion precession. You show the derivation; the reader re-runs it.

So the rule for this document: derivation first, coordinate second, data third. Never the reverse. That is stronger than attestation, because it is reproducible rather than merely witnessed.

41.2 The derivations, numbered

Each is stated as a chain. None references the observation it lands on.

D1 — The gradient pre-existed the cell.
chemiosmosis is universal → therefore ancestral → an ancestral mechanism was not invented by its descendants → **the proton gradient was environmental before it was biological.**

Lands on: vent geochemistry as the only setting supplying one.

D2 — Ignition requires redundancy.
you cannot select for a thing that does not yet exist → novelty must be built from components selected for something else → therefore a *repurposable* component is required → **redundancy is the precondition**, supplied by merger (outside), duplication (inside), or conversion (the arrow between).

Lands on: the C-value paradox; WGD before radiations; Ohno.

D3 — FUCA must be a poly-extremophile.
Pandæmonium drives every axis at once → the filter is defined by whatever tolerated all axes → **the survivor is necessarily poly-extremophilic**, not incidentally.

Lands on: LUCA's reconstructed thermophily as a *relaxation* from something harder.

D4 — The carrier must become less reactive over time.
high pour → synthesis outruns loss → scarce thing is chemistry happening → reactive carrier wins. Falling pour → loss bites → scarce thing is information retained → inert carrier wins. **Therefore carrier reactivity falls monotonically with chao.**

Lands on: sugar → RNA → RNA-protein → DNA, and the 2′-OH removal that makes DNA a deactivated RNA.

D5 — Two sweeps, from the homology pattern alone.
endogenous sweep → all descendants inherit *one* solution → universal and homologous. Exogenous pressure → each lineage solves separately → convergent and non-homologous.
translation is universal and homologous; DNA replication is neither. **Therefore two events, at two depths.**

Lands on: the replicase non-homology the field calls an enigma.

D6 — Homochirality is a date, not a puzzle.
before a sweep both enantiomers are present and neither is privileged → a selective sweep consumes one → **observed homochirality timestamps the sweep** rather than requiring a chemical cause.

Lands on: the mutually contradictory chiral-catalysis results, which stop contradicting.

D7 — The freeze postdates the first chiral residue.
glycine is achiral → a code containing only glycine has no handedness → there is nothing to freeze → **the chirality lock dates to the recruitment of the second amino acid.**

Lands on: alanine, whose identity element is a single wobble pair.

D8 — Specialists die in events that liberate generalists.
canalisation deepens the channel → the lip is the threshold → **canalisation depth *is* threshold height** → the same chao overtops a shallow channel and not a deep one.

Lands on: selective extinction of specialists at every mass extinction in the record.

D9 — Punctuation tracks acceleration, not magnitude.
adaptation has a response time → a *steady* change is trackable at any speed → an *accelerating* change is not → d^2**chao**$/\mathrm{d}t^2$ **decides who passes.**

Lands on: punctuated equilibrium, whose punctuation has had no mechanism since 1972.

D10 — Nodes saturate, edges proliferate.
component space is finite → exploration saturates → **novelty must move from components to relations.**

Lands on: the fold-space plateau beside rising regulatory complexity; fixed Hox set beside diversifying body plans.

D11 — A world that generates without retaining shows unpaired disequilibrium.
closing a cycle requires something to persist long enough to close it → no retention means no closed loop → **only unpaired, non-cyclic, throughflow-driven signatures are possible.**

Lands on: Venus, and the discriminator for 2026+ missions.

D12 — Stress machinery must be deeper than metabolic machinery.
developed in §43.

The point of the list. Twelve chains. None references its landing site. A reader can re-run any of them on paper before turning the page.
That is what the framework is — not a set of claims about deep time, but a method that emits them.

§42

The 1.57 Ga excursion, derived

Restructured to the rule: derivation, then coordinate, then data.

42.1 The derivation

Step 1. Endosymbiosis is an ignition event and therefore requires a selective pressure (D2). "They cozied up" is not one.
Step 2. The mitochondrion is the capture of an oxidising gradient. A gradient can only be captured once it exists. **Therefore mitochondrial establishment should sit at the appearance of that gradient** — the Great Oxidation Event.
Step 3. The chloroplast is the capture of the *inverse* gradient: carbon fixation running under the opposite redox direction.
Step 4. An inverse-gradient capture cannot precede the gradient it inverts. **Therefore chloroplast establishment must follow mitochondrial establishment in forced order** — not by contingency but by construction.
Step 5. Establishment of a new throughflow path re-cuts the network and is therefore a threshold event, which leaves a geochemical signature (the framework's general claim).
Step 6. Because the gradient is inverted, **the signature should be analogous to the GOE's and opposite in sense.**
Step 7. Therefore: *there should be a carbon/redox excursion marking the chloroplast crossing from captured to established, in the Mesoproterozoic.*

Nothing above references North China, the Gaoyuzhuang Formation, molybdenum isotopes, or any coordinate more precise than "Mesoproterozoic." A reader can re-run steps 1–7 and arrive at the same place.

42.2 Then the coordinate

The GOE sits at ∼2.2–2.4 Ga. Plastid establishment estimates in the literature span 0.9–2.1 Ga — a spread of more than a billion years, so the target was not narrow.✓ The derivation places the excursion after the GOE and before the Neoproterozoic expansion, i.e. the mid-Mesoproterozoic.

42.3 Then the data

A multi-proxy excursion exists at **1.57–1.56 Ga** in the Gaoyuzhuang Formation, North China Craton:

Cerium-anomaly decrease indicative of ocean oxygenation, with I/(Ca+Mg), carbonate and organic carbon isotopes, and phosphorus.✓ A prominent positive δ^{98}Mo excursion peaking at 2.03‰ and 1.92‰ in two sections, implying >54 % of global seafloor oxygenated.✓ Authigenic Cr isotopes shifting −0.18 to +0.66‰ then back.✓ A cyclostratigraphic study calibrating the duration of the ca. 1.56 Ga carbon isotope excursion and oxygenation event, noting redox was more dynamic than the "boring billion" model

proposed.✓

Co-located: decimetre-scale multicellular eukaryotic fossils in the upper Gaoyuzhuang, and multicellular rhodophytes at 1.6 Ga — red algae, which *require* plastids.✓

42.4 Status, stated honestly

What this is and is not

It is not a pre-registered prediction. There is no timestamped record of the derivation preceding the literature search, and the document does not claim one.
It is a derivation that a reader can re-run independently, landing at a coordinate where a multi-proxy anomaly exists and where the field has published, in terms, that the establishment-era selective pressure is distinct and *unidentified*.✓

The correct word is **consistency result**, and it is offered as one. That is weaker than a confirmed prediction and stronger than a post-hoc match, because the chain is checkable and does not mention its target.

42.5 The alternative, and how to beat it

Both formations were deposited in rift settings from the Columbia/Nuna break-up, and the isotope excursions have been read as a response to that break-up.✓ Rifting → weathering → nutrient flux → productivity, with no biology required.

The discriminator. Break-up predicts a productivity excursion with no reason to coincide with a specific biological innovation, and no particular internal structure.
The derivation predicts the excursion is **coincident with the appearance of the organisms that require plastids** — and multicellular rhodophytes at 1.6 Ga and decimetre eukaryotes in the same formation are exactly that coincidence.
Tectonics does not predict red algae.

§43

The inherited property

43.1 D12, derived

Everything alive descends from something selected to survive regime change

FUCA is the survivor of a persistence-filter event.▷ Not a survivor of competition — a survivor of a *threshold crossing*.

The chain:
the filter selects for surviving the transition, not for performing well in either regime → the survivor's machinery is optimised for *transition tolerance* → all extant life descends from that survivor → **transition-tolerance machinery should be more deeply conserved than regime-specific machinery.**

Metabolism is regime-specific: it is tuned to the gradient available. Stress response is transition-specific: it is what carries you across a change of gradient.

Therefore stress and repair machinery should sit deeper in the tree than metabolic machinery.

Where it lands. Chaperones are among the most conserved proteins known — HSP70/DnaK and HSP60/GroEL are universal and reconstruct to LUCA. So do core DNA repair systems. Meanwhile carbon fixation runs by at least six known pathways and nitrogen fixation is patchy.○
The most conserved thing in the cell is not how it makes a living. **It is how it survives not making a living.**

PREDICTION — P16 — the conservation ordering

Rank gene families by phylogenetic depth. **Stress-response, chaperone and repair families should rank systematically deeper than metabolic families**, and the separation should exceed what functional constraint alone predicts.
Testable against existing phylogenomic data, with the LUCA reconstructions supplying the depth estimates. A gradualist model predicts no particular ordering; the framework predicts a specific one.★

And it explains the shape of the whole tree. If the deepest selection pressure is *survive the transition*, then every subsequent threshold event re-applies the same filter to descendants already carrying machinery for it. The property is not merely inherited — it is **repeatedly re-selected**, which is why it never erodes.▷

§44

Definitions

The framework needs two definitions and they are not the same object.

Life — a dissipative structure held above 1 chao by sustained throughflow, which generates variation, is persistence-filtered, and cheshires its gains. **Substrate-neutral by construction**, because chao is.

Biological life — the realisation of the above by **self-assembly of supramolecular systems**: amphiphile compartments, template polymers, catalytic macromolecules.▷ One instantiation, not the category.

Why the split earns its keep

Most definitions of life fail because they conflate the process with its only known substrate — which is why they either exclude viruses, or admit fire, or collapse under the first edge case.
Splitting them does three things at once. It makes the astrobiology question *precise*: you are not asking "is there life," you are asking "is there a dissipative structure above threshold, and is it supramolecular?" It makes Venus answerable — generation without retention is *life failing to persist*, which is a coherent state under the general definition and incoherent under the biological one. And it quarantines the substrate-neutral extensions where they belong.⋆

Experimental anchor. The supramolecular half is not hypothetical. Work on self-reproducing giant vesicles — compartments that grow, divide, and amplify encapsulated DNA within the same system — is the closest existing demonstration of biological life's definition being met by construction rather than by inheritance.○
To verify: the primary reference (Sugawara group, Univ. Tokyo, self-reproducing supramolecular giant vesicles with encapsulated DNA amplification) is cited from memory and must be confirmed before circulation.

§45

E18, as a runnable protocol

The clustering formulation was underpowered. Restated as model comparison, it runs in an afternoon.

45.1 What is being claimed

Not "there are two clusters." The claim is that **recruitment order is set by the charging machinery, not by the environment** — which is a comparison between two models.

	Predictors	Status
M1 environment	prebiotic abundance; ΔG of synthesis under hydrothermal conditions	**Half-falsified already** — no correlation between recruitment order and ΔG_{hydro} ✓
M2 machinery	aaRS class; acceptor-stem identity-element complexity	Untested

45.2 Protocol

Pre-specify, before running: direction (Class II skews early), statistic (Mann–Whitney U on recruitment rank by class), threshold, and that the result is reported whichever way it falls.

1. Take the consensus recruitment rank, 1–20.
2. Assign each amino acid its aaRS class (I or II). Handle lysine explicitly — it has forms in both.
3. Compute observed rank separation between classes.
4. Permute class labels 10^4 times; build the exact null distribution.
5. Compare observed against null.

Secondary: repeat with identity-element complexity (number of identity nucleotides) as a continuous predictor, against recruitment rank. Spearman, same permutation null.

Tertiary: the same test using ΔG_{hydro} as predictor, as a negative control. It should fail — and its failure is already published. ✓

Power, stated in the text rather than discovered by a referee

With ∼10 per group, only a **large** effect is detectable. That is not a weakness of the design; it is what the hypothesis predicts — near-complete separation or nothing.
Say so in the paper. A reader who sees the author name their own power limit extends more credit to everything else, and a reader who finds it themselves extends less.

Why this is the first thing to run. Trifonov's ranks are published. The class assignments are published. It is twenty rows and a permutation loop — and it converts the framework's central methodological claim, that recruitment order is set *inside* the system, from

an assertion into a measurement.
Combined with the published ΔG negative, both halves of the model comparison would then exist.

PART XIII

The register

§46

Consolidated evidence register

Fifteen evidence plays and seven predictions, generated across Parts II–IX. This chapter supersedes the interim register in §16.

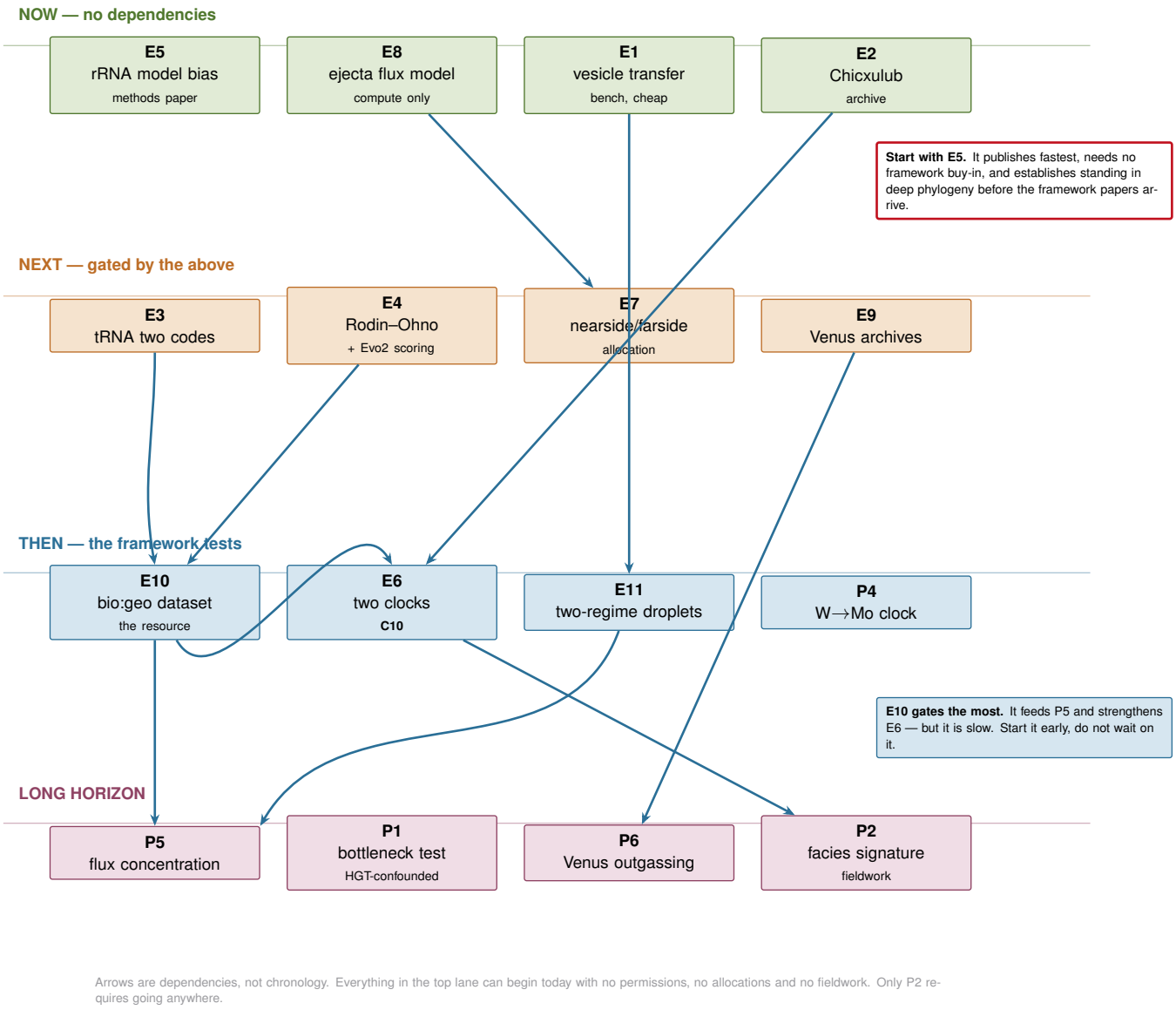
46.1 The plays

	Play	Resource	Tests	Character
E1	Vesicle transfer-survival assay	existing bench, low cost	C4	Cheapest test of a load-bearing claim
E2	Chicxulub as impact-hydrothermal analogue	published IODP-364	C5	Converts plausibility to measurement
E3	tRNA two-code lineage variability	existing databases	C7	Clean, bounded, no wet lab
E4	Rodin–Ohno structure-prediction test	compute, public sequence	C7, FUCA-era	**Highest novelty**; tools postdate the hypothesis
E5	rRNA substitution-model bias	existing data, compute	—	**Fastest publication**; no framework buy-in
E6	Two-clocks stratigraphic compilation	published literature	**C10**	The sharpest test; may resolve only within-succession
E7	Nearside/farside clast asymmetry	Apollo + Chang'e 6	C11	Turns a needle-search into a ratio test
E8	Lunar ejecta flux model	compute + literature	C11	Zero-cost; gates every allocation request
E9	Venus archive reanalysis (Pioneer, Venera, Magellan)	public + Russian-language	C8	Unusual opportunity behind an access barrier
E10	Bio:geo paired reaction network	Rhea × CHNOSZ + compilation	C2, C3	**Most downstream dependencies**; a resource in itself
E11	Two-regime droplet sequence	the platform	C3, C4	Selection experiment with predicted direction
E12	Enantiomeric-editing geometry	structures + phylogenomics	Face 2	Forbids a specific world
E13	Reductase stereochemical conservatism	existing literature	Face 3	**Cheapest play in either volume**
E14	Glycan-code antiquity	phylogenomics	Face 4	Carries Face Four alone
E15	GARD fidelity ceiling on the E10 hypergraph	compute, gated by E10	Face 4	Turns a conceded weakness into a number
E20′	Self-cleaving motif accessibility in random sequence	sequence data, compute	C7, hammerhead premise	**Checked:** distribution confirmed, relic reading contested, accessibility is the defensible claim

46.2 The predictions

	Prediction	Tests	Discriminates against
P1	Star-like radiation below LUCA; clustered family origination times	C5	Gradual origin (predicts bushy, smeared)
P2	Biosignatures at the *join* of vent mineralogy and shallow-water facies	C4	Both deep-vent-only and hot-spring-only models
P3	LUCA's deepest protein families biased toward early-recruitment amino acids	C7	Weak or absent gradient under gradualism
P4	W-dependence deepest; W→Mo transition dates to redox, not to a node	C2	Any purely phylogenetic timing
P5	Flux concentration event in the reaction network, not smooth elaboration	C3	Gradual metabolic elaboration
P6	Venusian outgassing history exceeds a computable sustained-throughflow threshold	C8	Nothing yet — the threshold is uncomputed
P7	Post-transfer droplet population enriched for gradient-coupling compositions	C3, C4	Random-subsample null

46.3 Sequencing



Arrows are dependencies, not chronology. Everything in the top lane can begin today with no permissions, no allocations and no fieldwork. Only P2 requires going anywhere.

Figure 46.1: Dependency structure of the evidence programme. Four of eleven plays have no prerequisites at all. The fastest publication (E5) is deliberately not the most important one — it is the one that buys standing for the rest.

§47

What would falsify this

A framework that cannot fail is not a framework. This chapter states, per commitment, what result would end it — and marks honestly the one that cannot be falsified at all.

	Commitment	What would falsify it	Via
C1	Dissipative adaptation	No directional selection under sustained driving in a compartmentalised system, at throughput	E11 / P7
C2	Inversion method	LUCA's conserved traits fail to intersect on a coherent environment — constraints mutually inconsistent	E10, ongoing
C3	Phase boundary	Smooth metabolic elaboration with no flux-concentration event; no shift in selection direction across the droplet transfer	P5, P7
C4	Surface-breaching vent	No transfer-survival advantage for pre-formed vesicles; no facies juxtaposition where preservation permits detection	E1, P2
C5	First filter	Resolvable, bushy deep structure below LUCA with smeared origination times, surviving HGT-realistic modelling	P1
C6	Proto-genomes at t=0	LUCA defence systems shown to be later acquisitions, not ancestral	literature
C7	**(revised)** the RNA → protein → DNA order is forced by one merger-and-freeze	No compositional gradient with depth; acceptor-stem and anticodon codes equally deep; D-rejection is a late anticodon-era addition; reductase re-selects handedness	P3, E3, E12, E13
C8	Venus / constructive radiation	Outgassing history below the sustained-throughflow threshold once computed	P6
C9	Two searches	**Not falsifiable.** A conceptual distinction about question-framing, not an empirical claim. Stated as such	—
C10	Two clocks	Smooth co-variation of organic record, gradient indicators and $\delta^{13}C$ within a single well-controlled succession	E6
C11	Lunar archive	No nearside enrichment of evolved-crust clasts; or predicted terrestrial mass fraction below detection limits	E7, E8

On C9, and why flagging it matters

C9 — that abiogenesis-favourable and biosignature-favourable target lists are different searches — is not an empirical claim and cannot be falsified. It is a statement about how a question is posed.

That does not make it worthless; conceptual distinctions do real work, and this one has a concrete consequence (Venus is scored low by every existing scheme because every existing scheme is implicitly a biosignature scheme). But it must not be presented alongside C10 as though the two have the same standing.
Mixing falsifiable and conceptual claims without marking which is which is the most common way a framework loses credibility. Marked here deliberately.

The load-bearing observation. Six of eleven commitments are falsifiable by plays already in the register, and four of those plays require no permissions. The framework is not merely testable in principle — it is testable this year, mostly from a desk.

§48

Publication sequence

Ordering matters more than usual here, because the framework is unorthodox and arrives without institutional backing. The sequence below is designed to establish standing before making claims that require it.

Order	Output	Why here
1	**E5** — rRNA substitution-model bias	Methods paper. No framework content. Addresses a systematic error in an active literature. Cited by people who disagree with you
2	**E8** — lunar ejecta flux under Hadean orbital geometry	Modelling paper, no samples. Establishes the quantitative basis for every subsequent allocation request
3	**E1** — vesicle transfer survival	Short experimental paper. Interesting to the protocell field independent of origins framing
4	**C4** — the surface-breaching vent, as a hypothesis paper	Now defensible: E1 supplies the mechanism, and hypothesis papers are the right venue for a site model
5	**E4** — Rodin–Ohno test	High-novelty result. Lands better once 1–3 have established that you produce careful work
6	**E7** — nearside/farside asymmetry	Requires allocation, which requires 2
7	**E6** — two-clocks compilation	The framework's central test. Do it when a negative result would still be publishable on your name
8	**The framework** — C1–C11 as a synthesis	Last. Synthesis papers from unknown authors are ignored; synthesis papers from authors with six prior outputs are read

The one structural piece of advice

Do not lead with the framework. The framework is the most interesting thing here and the least publishable thing here, and those facts are related.
Every item 1–7 stands alone, is useful to someone who rejects C1–C11 entirely, and is defensible on its own evidence. By the time item 8 arrives, the reader has already accepted seven pieces of it without being asked to accept any of it.

§49

What this framework is

A closing statement of scope, because ambitious frameworks are usually damaged by claiming more than they hold.

It is: a method (inversion from a reconstructed output to the regime that produces it), applied to a specific transition (FUCA → LUCA), generating a set of testable consequences, several of which are testable from existing data.
It is not: a theory of the origin of life. It says almost nothing about how the first replicator arose, and does not need to — C6 places proto-genomes before the filter as a starting condition rather than explaining them.
Its central claim is narrow and sharp: the transition to LUCA was a change in which resource was limiting, not an accumulation of complexity — and that change left a signature in rock, in genomes, and possibly on the Moon.
Its most valuable property is that it generates predictions its author did not choose. C10 and P5 fall out of the method rather than being posited. That is the signature of a method doing work rather than a story being told.

And the honest limit. Every route here has a confound that could sink it: Eoarchean age control is poor; oxygen isotopes cannot separate Earth from Moon; solar-wind carbon overlaps the biological range; early HGT smears phylogenetic signal; the Venus surface record is gone.
None of these is fatal individually. All of them are why the register contains eleven plays rather than one. **The framework's defence is not any single result — it is convergence across independent lines that fail in unrelated ways.**

§A

Glossary

Two sections. The first is the framework's own vocabulary; the second is field terminology used without definition in the body.

A.1 The fixed process-terms

The design rule. Each term names a *process*, never a thing — a verb wearing a noun's coat. The absurdity is load-bearing: a term that can be said with a straight face as a noun will doggen into jargon, and the process vanishes behind it.▷
Handling note. The right-hand column exists because the commonest way to lose one of these is to translate it into its nearest textbook equivalent. Where a coined term does work the standard term does not, that difference *is* the content.

Term	Names the process of…	Do *not* read as
chao	Distance from equilibrium. A *vector*: magnitude (how far) and dimensionality (along how many independent axes). **1 chao is the phase-change threshold at which life *can* form** — a permission condition, not sufficiency	"energy", "entropy", "disorder"
Pandæmonium	The high-dimensional limit of chao — driven along every axis at once. The true name of the Hadean, and why it is genesis-permissive	"chaos", "the early Earth"
generating	A driven system throwing up variation because it is forced to. The raffle: tickets written as long as throughflow feeds	"mutation", "a generator" (the noun reifies it)
persistence-filtering	Differential survival under perturbation. Needs no replication and no biology. The Purge: a predicate discloses and a *class* passes	"natural selection", "fitness"
cattening	Collapse of two-or-more viable outcomes into one, alternatives *foreclosed* — gone, unable to return	"extinction", "fixation", "bottleneck"
cheshiring	A gain, once made, being kept, so the next round starts from it. Irreversibility that *accumulates*. The body fades; the grin persists	"inheritance", "a ratchet" (a device, not a process)

Term	Names the process of…	Do *not* read as
doggening	The observer's error: reading a frozen outcome as if it had always been the only possibility. *Dedoggening* is the method — recovering the cattening that made one outcome of many	"determinism", "hindsight bias"
canalising	A channel deepening until flow cannot leave it. From Waddington — kept deliberately because it hooks into existing biology. *Canalising cuts down; cheshiring builds up*	"constraint", "specialisation"
throughflow (the pour)	The rate a gradient drives material through a system. The master variable: too slow, no generating; too fast, no persisting	"energy budget", "metabolic rate"
dissipating	Spreading a gradient faster — what driven structures get good at. From Prigogine and England	"wasting", "losing energy"
heritable carrier	The medium a retained gain is written onto. The one fixed term that is a thing rather than a process; guard it against doing more work than that	"the genome" (DNA is only the current one)
ignition event	A rare, high-conjunction-threshold crossing, happening approximately once, freezing into a universal signature. **Ignition events are mergers**	"innovation", "key adaptation"
re-canalisation event	A frequent, low-threshold crossing where a new flow path opens and reshapes which flows dominate, without foreclosing	"adaptive radiation"
threshold event	Generating and persistence-filtering firing together — pruning and opening at once. Exogenous (injected from outside) or endogenous (an acquisition whose own success becomes the pressure)	"mass extinction" (that is the palaeontology *map*, not the process)
the golden apple	The catalyst. One small thing thrown into a settled party that forces the improbable to resolve — not by adding energy but by adding a forced choice	"trigger", "perturbation"

Term	Names the process of…	Do *not* read as
the pageant	A *plural* selection rule. "Fairest at X, at Y, at Z" celebrates many and the party escalates, where "to the fairest" forces one winner and a war. Not the kind option — the dominant one	"cooperation", "diversity"
finite / infinite game	After Carse. A finite game is played to win, and winning ends it. Life is an infinite game; every cattening is a finite game trying to end it	"competition vs. cooperation"
feather-inside-a-feather	Each novelty from one Purge becoming the substrate for the next raffle — ticket-space built *higher* on the last round's survivors. Recursion depth, not clock time	"increasing complexity", "progress"
tabby / tiger	The *depth* of a cattening — how much possibility-space it foreclosed. Tabby: new arrangements of the same alphabet, findable in the space that already holds them. Tiger: a different character set entirely, unreachable by any address in the old space	"small vs. large change"
mutose / glucose	The worked analogy for tabby/tiger at the level of chemistry: a substrate no existing metabolism predicts, versus one it does	—
Wonderland / the Cheshires	Pre-LUCA Earth, and its inhabitants — plural, contemporaneous, mostly dead ends, gone but for the grin left in the record	"the primordial soup"
Eris	The complement of the guest list. What breaks a system's closure must lie outside the system's enumeration of itself; an endogenous threshold event is a *guest who becomes Eris*	"an outside shock"

A.2 Technical terms

Term	Definition
FUCA-class	The plural pre-translation population; the Cheshires
FUCA 1.0	Sole survivor of the translation sweep

Term	Definition
LUCA 1.0	Sole survivor of the DNA sweep; defining attribute is a DNA genome
LWKA	The self-cattening formerly smoothed to "LUCA" — retired vowel-run, W for waking; un-reifiable by design
the Catma	The specific cattening that produced the central dogma: a peptide-bond ribozyme merging with charged acceptor-stem RNAs, chirality-lock as driver. *Catma* (process) → *dogma* (the carrier it wrote)
Wood–Ljungdahl	The reductive acetyl-CoA pathway; the most ancient known carbon fixation route, and the most strongly ^{13}C-fractionating
chemiosmosis	Energy conservation via an ion gradient across a membrane. Universal — which implies the gradient *pre-existed* the cell
serpentinisation	Hydration of ultramafic rock, producing abiotic H_2 and alkaline fluid. The sustained-gradient source in the vent model
operational RNA code	Amino-acid identity elements in the tRNA acceptor stem, read by synthetases *independently of the anticodon*. A recognition system that works without the codon table, and therefore predates it
acceptor stem / anticodon arm	The two structurally separable halves of tRNA, carrying two recognition systems of different ages
PTC	Peptidyl transferase centre — the ribosome's catalytic core, made of RNA. The oldest layer in the accretion model
aaRS class I / II	Two structurally unrelated synthetase superfamilies, both universal, binding *opposite faces* of the acceptor stem and acylating different hydroxyls. Roughly ten amino acids each
Rodin–Ohno hypothesis	That the two aaRS classes were encoded on complementary strands of one ancestral gene — which would make that duplex the deepest reconstructable object in molecular biology
urzyme	A minimal synthetase core (∼120–130 residues) retaining measurable catalytic activity
hammerhead ribozyme	Small self-cleaving RNA motif; catalytic core of ∼15 conserved nucleotides in three helices. Found in all domains, mostly in mobile elements
retrozyme	Non-autonomous LTR retrotransposon bearing hammerheads, propagating via circular RNA
GARD	Graded Autocatalysis Replication Domain — a model in which non-covalent assemblies of mutually catalytic molecules transmit *composition* at fission
compositional inheritance	Heredity carried in the relational composition of a molecular set rather than in a sequence

Term	Definition
homochirality	Single-handedness. Life uses D-sugars and L-amino acids — opposite handedness, which the Catma reads as one break transduced across an alphabet boundary rather than two independent breaks
enantiomer	A non-superimposable mirror image. The cat and its reflection are literally D and L
ribonucleotide reductase	The enzyme making deoxyribonucleotides from ribonucleotides by removing the 2′-OH. Stereochemically conservative — which is why DNA's backbone chirality is inherited rather than chosen
A / B / Z form	DNA duplex conformations. B-form is the ordinary right-handed helix; Z is a strained left-handed local conformation built from the same D-monomers, and is not a counterexample
GOE	Great Oxidation Event, $\sim$2.4 Ga. Endogenous, caused by oxygenic photosynthesis, killed anaerobes, left no fossil record — detected by sulfur
S-MIF	Mass-independent sulfur fractionation (Δ^{33}S). Present only before the GOE; its disappearance dates the oxygenation
δ^{13}**C**	Carbon isotope ratio relative to standard. Values $\leq -35‰$ in reduced carbon are the Wood–Ljungdahl-grade target
BIF	Banded iron formation — alternating iron-rich and silica-rich sediment; a ferruginous-ocean indicator
HGT	Horizontal gene transfer. Rampant in early evolution, and the reason tree-shape statistics cannot carry the bottleneck test alone
non-orthologous displacement	Replacement of a gene by an unrelated one performing the same function — the standard explanation for bacterial replisome divergence
coalescence	The point at which lineages traced backwards converge on a common ancestor. Two sweeps predict *two* coalescence horizons
MAG	Metagenome-assembled genome
Rhea / CHNOSZ	Reaction-centric biochemistry database / thermodynamic package computing properties of minerals, aqueous species *and* proteins at geological T–P. Together, the bio:geo bridge
Evo 2	Nucleotide-resolution genomic foundation model. A sequence prior trained on *extant* genomes — useful for scoring, biased toward the present
CAPTEM / ANGSA	The Apollo sample allocation committee / the programme that opened sealed Apollo cores in 2022
tessera	Venus's oldest exposed, most deformed terrain — possibly pre-resurfacing, imaged by Magellan

Term	Definition
PAL	Present atmospheric level, used for oxygen

§B

Numbers used

Collected for checking. Confidence marks as elsewhere; the arithmetic is checkable, the inputs are the soft part.

Quantity	Value used	Note
LUCA age	∼4.2 Ga	clock-derived, wide priors○
LUCA protein families	∼2,600	2024 reconstruction○
Seawater Mg^{2+}	∼50 mM	modern; Hadean uncertain○
Seawater Ca^{2+}	∼10 mM	modern○
Fatty acid vesicle divalent tolerance	low mM	the C4 problem○
Wood–Ljungdahl $\delta^{13}C$ target	$\leq -35‰$	"WL-grade"▷
Water activity limit for known life	∼0.585	
Venus cloud droplet water activity	∼0.004	two orders below the limit○
Venus D/H relative to Earth	∼100×	massive water loss○
Moon distance at 4.4 Ga	∼20–30 $R_\oplus$	vs ∼60 now○
Lunar capture cross-section gain	∼9×	$(R/d)^2$ at ∼1/3 distance≈
Impact hydrothermal system lifetime	10^5–10^6 yr	scaled from Chicxulub○
Droplet platform throughput	$\sim 10^6$ hr^{-1}	design target▷
Apollo returned mass	∼382 kg	all nearside, near-equatorial○
Known lunar meteorites	>600	no allocation committee○

§C

Data sources and access

Source	Access	Used for
Rhea	public	Reaction hypergraph; ChEBI ids; UniProt/EC links (E10)
CHNOSZ	R package, public	Thermodynamics of minerals, aqueous species *and proteins* at geological T–P (E10)
KEGG / MetaCyc / ModelSEED	mixed licensing	Pathway context; secondary to Rhea
Published LUCA reconstructions	supplementary data	Family presence, origination depth (P1, P3, E10)
tRNA identity-element compilations	literature	E3
Evo 2	open weights	Zero-shot duplex scoring, E4 only — extant-genome prior
AlphaFold3 / Boltz / ESM3	mixed	Fold prediction for E4 urzyme reconstructions
IODP–ICDP Exp. 364	public proceedings	Chicxulub hydrothermal system (E2)
Isua / Pilbara isotope literature	published, uncompiled	E6 — the compilation *is* the work
CAPTEM (Apollo allocation)	formal request	T1, T2 lunar material (E7)
CNSA Chang'e 5/6 allocation	international, open	Farside control (E7)
Lunar meteorite market	commercial	T3 — start here, no committee
Pioneer Venus LNMS	public archive	E9
Venera / Vega	partly Russian-language	E9 — the access barrier is the opportunity
Magellan radar	public, complete	Tessera terrain (E9)

§D

Source stubs

Every stub below needs verification before publication. They are recorded at the precision required to find the work, not to cite it.

Origins and LUCA. Moody et al. 2024, *Nat Ecol Evol* — LUCA reconstruction, age, family count, defence systems. Preiner et al. 2020, *Nat Ecol Evol* — mineral-catalysed CO_2 fixation at pH gradients. Lane & Martin — alkaline vent chemiosmosis. Damer & Deamer — hot spring / wet–dry.
Molecular deep structure. Rodin & Ohno 1995 — complementary-strand aaRS origin. Carter & Wills — urzymes, sense/antisense constructs. Schimmel & Ribas de Pouplana — operational RNA code. Di Giulio — tRNA duplication. Trifonov — consensus amino-acid recruitment order. Petrov, Williams et al. — ribosomal accretion of expansion segments.
Protocell bench. Vaidya, Lehman et al. 2012, *Nature* — cooperative RNA replicator networks. Szostak group ∼2009 — vesicle growth and division. Baaske & Braun — thermophoretic concentration. Joyce group — cross-catalytic ribozymes. Holliger group — polymerase ribozyme programme. Abate group, UCSF — open dielectrophoretic droplet sorter designs. Monnard & Deamer — divalent cation disruption of fatty acid vesicles. Duysens 1956; Morel & Bricaud 1981 — package effect (Volume II).
Planetary. Bellucci et al. 2019, *EPSL* — terrestrial-like zircon in Apollo 14321. Armstrong, Wells & Gonzalez — terrestrial meteorites on the Moon. Kring et al. — Chicxulub post-impact hydrothermal system. Way et al. — Venus climate history modelling. Sutherland — cyanosulfidic prebiotic chemistry.
Prior volume. Behrenfeld et al. 2009, *Biogeosciences* — satellite fluorescence and iron stress. *Ocean Science* 20:217, 2024 — *psbO* and ocean colour. Duprat et al. 2016, *Nat Geosci* — iceberg wake productivity.

---

end of Volume I