

Energy Availability, Not Stress Tolerance, Determines Archaeal Persistence in Extraterrestrial Habitats

Ajay Vellanki

ajayvellanki0720@gmail.com

Independent Researcher

Research Article

Keywords:

Posted Date: September 4th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10681502/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Energy Availability, Not Stress Tolerance, Determines Archaeal Persistence in Extraterrestrial Habitats

Ajay Vellanki

Independent Researcher, West Windsor, New Jersey, USA

ORCID: 0009-0008-7994-152X

Corresponding author: Ajay Vellanki, ajayvellanki0720@gmail.com

7 August 2026

Abstract

Candidates for extraterrestrial habitats are conventionally chosen for stress tolerance, assuming the taxon best withstanding cold, radiation, desiccation and low water activity will persist. This premise was tested at scale by scoring 3,531 archaeal genomes across 2,069 clades against 58 physically parameterised niches spanning nine solar-system bodies and the deep Earth subsurface, using weights fixed before any result was seen and deliberately biased towards stress tolerance (53.9% versus 16.2% for energy). The premise failed: rankings were dominated by methanogens, led by *Methanobacterium* sp. YSL (score 0.6751), with the energy layer, not stress, carrying the signal ($R^2=0.637$ versus 0.294). Radiation modules discriminated at or below chance (minimum AUC 0.460) against an extremophile null. Methanogenesis, though weakly exergonic ($-6.34 \text{ kJ mol}^{-1}$) versus nitrate reduction (-83.18), won on exploitable energy density (0.1269 versus 0.0004 J L $^{-1}$, 317-fold). A blind metagenome test showed only electron-donor usage cleared validation ($p=0.47$, $p=0.0314$), while taxon presence failed. Phylogenetically blocked cross-validation collapsed from median $R^2=0.837$ (genus held out) to -0.562 (phylum held out). Engineering-cost and consortium-design layers were added, and a pre-registered nine-test validation programme returned adverse verdicts on four of nine tests, all reported.

Introduction

Space microbiology emerged from concerns about contamination during spaceflight and developed into systematic study of how organisms respond to launch, orbital exposure, microgravity and physical extremes beyond the atmosphere. The orbital environment combines vacuum, desiccation, temperature cycling, galactic cosmic rays, solar particle events and unattenuated ultraviolet radiation. Long-duration exposure of *Bacillus subtilis* spores

established ultraviolet radiation as the dominant killing agent in low Earth orbit and showed that shielding by micrometre-scale dust or outer spore layers markedly extends survival. The EXPOSE facilities on the International Space Station made this repeatable; successor missions broadened the roster: cyanobacteria of genus *Chroococcidiopsis*, lichens and their component fungi and algae, *Xanthoria elegans*, and endolithic and cryptoendolithic communities, including phototrophs surviving 548 days in low Earth orbit. BIOMEX extended the question from survival to biosignature-molecule stability, while Tanpopo tested the interplanetary-transfer hypothesis by exposing cell aggregates on the ISS exterior and measuring survival as a function of thickness; it remains the strongest empirical constraint on lithopanspermia. Tardigrades survived direct space exposure. *Deinococcus radiodurans* became the reference organism for extreme ionising-radiation resistance; its study reframed the mechanism towards genome copy number, efficient chromosome reassembly and oxidative-damage management.

Archaea entered through physiological range: the domain contains organisms defining outer envelopes of temperature, salinity and anaerobic energy metabolism. Methanogens tested under simulated martian conditions survived desiccation, freeze-thaw, low pressure and perchlorate exposure; they function as strict anaerobes running on H₂ and CO₂—the energy couple most plausibly available in a martian or icy-moon subsurface—and methane is a candidate biosignature. Haloarchaea tolerate desiccation and ionising radiation at levels comparable to bacterial radiation specialists, with survival inside halite fluid inclusions raising the possibility of very long-term dormancy in evaporite deposits. These results concern survival under exposure, not whether an organism could establish a population where it landed. The organism roster is a convenience sample: every flown archaeon was chosen for cultivability,

shippability and revivability, a few dozen strains against the present survey of 3,531 archaeal genomes across 2,069 clades. Exposure platforms measure damage tolerance after revival, not energy acquisition over relevant timescales. The stressor-first framing embeds an untested assumption that the binding constraint on persistence is damage rather than energy supply. Independent negative results from module validation, shielding physics and transit chemistry argued against a radiation-centred framing. Of nine validation tests, four returned adverse verdicts, with phylogenetically blocked cross-validation bounding every predictive claim; the work contained 47 numbered self-corrections and 62 explicit limitations.

Results

Habitat catalogue, niche classification and module validation

Of the 58 catalogued niches, 37 pass both pre-genomic gates, 4 fail on temperature alone and 17 on water activity; 142 pseudoreplicates. conditions span twelve orders of magnitude in annual radiation dose ($1\text{e-}04$ to $1.2\text{e}+08$ Gy yr^{-1}), ionic strengths from fresh water to 18 M, and pressures from $1\text{e-}11$ to $1\text{e}+05$ bar. All 9 interplanetary-transit scenarios fail, both Venus cloud-droplet niches fail on water activity (0.30–0.35), and 7 of 10 deep-Earth niches pass. The 58 niches partition into three classes: 37 are growth-permissive, 13 are survival-only, and 8 are uninhabited (Figure 1). Curated extremophiles outscore the mesophile null in all 13 survival-only niches (AUC 0.694–0.812, median 0.717, $p < 0.05$ in 13 of 13). Of 18 candidate modules, 11 survived validation. Every retained module discriminated its target phenotype against a mesophile null (AUC 0.558–0.891). Against a phenotype-matched extremophile null, AUCs fell to 0.460–0.885, with radiation and radiation_mn at or below 0.5. Radiation-repair machinery is broadly enriched in extremophiles but does not distinguish against radiation-intolerant extremophiles.

Of 73 candidate bacterial families, only 23 discriminate extremophiles from mesophiles at AUC 0.719 best. Catalase (0.386) and glycine-betaine transporter OpuAC/ProX (0.375) are more abundant in mesophiles. Sign-correction raises the panel from 0.718 to 0.863. Module AUCs against mesophiles: methanogenesis 0.925, sulfate reduction 0.912, iron cycling 0.810, cold 0.89, chaotrophicity 0.84, nitrogen fixation 0.83, salt 0.72, iron

uptake and 3-hydroxypropionate carbon fixation 0.73 each, desiccation 0.70, sulfur assimilation 0.67; radiation (0.46), manganese antioxidant (0.47), oxidative stress (0.52) and dormancy (0.56) do not discriminate. Against extremophile classes, cold scores 0.870, chaotrophicity 0.891, and radiation 0.460. The Wood–Ljungdahl module scored 0.274 against methanogens versus 0.925 for methanogenesis because marker PF03308 is under-represented in archaeal methanogens. Desiccation and sulfate reduction fail diagnosticity, leaving 9 of 11 retained modules diagnostic.

The IVYWREL amino-acid fraction correlates with optimal growth temperature at $r = 0.916$ across 550 species (Archaea $r = 0.920$, cv $r^2 = 0.906$, $n = 483$; Bacteria $r = 0.874$, cv $r^2 = 0.731$, $n = 67$), predicting temperature at genus-held-out $r^2 = 0.89$ (MAE 5.0 °C) (Figure 2). A single amino-acid ratio achieves $r^2 = 0.83$. Phylogenetic correction left family-level AUCs almost unchanged (median shrinkage -0.003) but removed 18 of the top 50 archaeal organism-niche assignments as pseudoreplicates.

The composite index separates bacterial extremophiles from mesophiles at Cliff's $\delta = +0.57$ (large), and extremophile class predicts capability for Mars ($H = 27.3$, $p = 1.27 \times 10^{-4}$), Europa ($p = 2.26 \times 10^{-4}$) and Enceladus ($p = 5.78 \times 10^{-3}$). Psychrophiles score highest on cold (0.74), the radiation/desiccation-tolerant class on desiccation (0.65), and piezophiles and halophiles on salt and ionic transport (0.68, 0.61). Archaea occupy 95% of the clade-collapsed top 100 but the corpus is 94.3% archaeal; at genus level, 0.606 versus 0.407 ($p = 0.0035$, Mann–Whitney). Weighting sensitivity across equal, prior-derived, AUC-informed and random Dirichlet vectors leaves global ordering stable (Spearman 0.861–0.983).

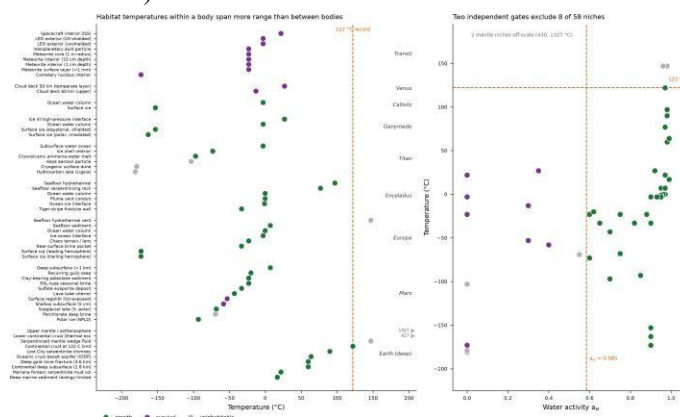


Figure 1. Planetary and deep-Earth niches classify by temperature and water activity into three habitability classes.

Panel A plots temperature (°C) for 58 named niches grouped by body (Transit, Venus, Callisto, Ganymede, Titan, Enceladus, Europa, Mars, Earth (deep)), colored green (growth), purple (survival), or gray (uninhabitable); dashed orange line marks the 122 °C record. Panel B plots temperature versus water activity (aw), with a dashed orange horizontal line at 122 °C and vertical line at aw = 0.585; annotation notes two mantle niches off-scale (430, 1327 °C) and text states two independent gates exclude 8 of 207 100 is a depletion relative to their 90.5% corpus share. 58 niches.

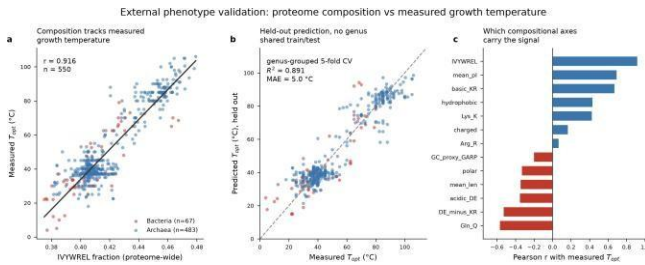


Figure 2. Proteome composition predicts measured optimal growth temperature across archaea and bacteria. Panel a plots measured T_{opt} against IVYWREL fraction (r = 0.916, n = 550), with archaea (blue, n=483) and bacteria (red, n=67) and a fitted line. Panel b shows predicted versus measured T_{opt} under genus-grouped 5-fold cross-validation (R² = 0.891, MAE = 5.0 °C) with a dashed identity line. Panel c ranks compositional axes by Pearson r with measured T_{opt}, from IVYWREL, mean pI, basic_KR, hydrophobic, Lys_K, charged, Arg_R (positive, blue) to GC_proxy_GARP, polar, mean_len, acidic_DE, DE_minus_KR, Gln_Q (negative, red).

The persistence ranking and the energy signal

Across all bodies, *Methanobacterium* sp. YSL leads with a mean score of 0.6751, and hydrogenotrophic methanogens dominate the top ranks, with Kruskal–Wallis 230 correlated $\rho = -0.289$ with raw score because cold-adapted tests confirming significant separation between metabolic classes on every body. Rankings restricted to genera with at least five genomes vary by body (Figure 3): Enceladus is purest, with 20 of the top twenty genera being methanogens and none halophiles; Callisto follows at 19, Titan at 17, Ganymede at 16. Mars is the exception, with only 5 of twenty methanogen genera; its leading genus is the haloarchaeon *Haloterrigena*, its leading genome the methanogen *Methanoculleus frigidifontis*, followed by the halophile *Halogranum rubrum*. Europa has 7 methanogen and 4 halophile genera. Martian water activity ranges 0.30–0.95 with ionic strength up to 7.0 M, versus 0.90–0.98 and at most 0.5 M on Enceladus, so Mars's salinity favours halophiles.

At genus level, Enceladus is led by *Methanothrix* (0.64), *Methanospirillaceae* (0.62), *Methanospirillum* (0.62);

Europa by *Methanothrix* (0.62), *Halogranum* (0.61), *Methanobacterium* (0.60); Mars by *Halomicrococcus* (0.62), *Natronoglopus* (0.62), *Halogranum* (0.61). Score differences fall within shortlist instability (top-20 Jaccard 0.243–0.538), so class of leader is reliable but specific genus is not. Because 3,520 archaeal assemblies were surveyed automatically against only 215 curated bacterial genomes, archaea's 69% share of the clade-collapsed top 100 is a depletion relative to their 90.5% corpus share.

Across 15 energetically characterised habitats, three regimes emerge (Table 1). In H₂-bearing settings, methanogenesis wins in 7 of 15 niches: Enceladus's ocean (10.6179 J L⁻¹) and plume conduit (9.9646) are richest, followed by Enceladus (3.1583) and Europa (2.0978) seafloor systems; Callisto, Ganymede, Titan oceans and Europa's ice–ocean interface converge at 0.1269 J L⁻¹. In oxidised settings, nitrate reduction wins in 3 of 15 (Mars regolith 4.3378, Europa brine pocket 4.1638, Mars RSL brine 4.1590), where methanogenesis collapses to 0.0001, 0.0000, 0.0000 J L⁻¹. In ice and methane-bearing settings, anaerobic methane oxidation wins in 5 niches (0.0669–0.1787 J L⁻¹). Nitrate reduction's median (0.0004) is 317-fold below methanogenesis's (0.1269). Methanogenesis is exergonic in 100% of niches versus 27% for sulfate reduction (median 0.0000); acetogenesis (0.0152) and Fe(III) reduction (0.0001) never win.

Weighting assigned 53.9% to stress-tolerance, 16.2% to energy (methanogenesis 5.4%), 29.9% residual. Energy R² = 0.637 versus stress R² = 0.294. Proteome size is a

major confound: mean score correlated with size at $\rho = 0.732$, log₁₀ size explained 53.6% of variance. Cold adaptation correlated $\rho = -0.289$ with raw score because cold-adapted genomes are systematically smaller ($\rho = -0.468$); after controlling for size, the association was absent ($\rho = +0.027$). Genome size explained R² = 0.536 of ranking variance. After controlling for size, energy explained R² = 0.637 versus stress R² = 0.294. Methanogens occupied 100% of top ranks despite carrying only 5.4% of total weight.

Geochemical thermodynamics found 80% of 90 niche-couples exergonic in situ; methanogenesis clears the biological quantum in 80% of niches; 225 of 296 catabolic reactions were conservable. H₂/CO₂ methanogenesis: −6.34 kJ mol e⁻¹, exergonic in 100% of niches; nitrate reduction −83.18; Fe(III) reduction −85.9; sulfate reduction exergonic in 27%, median +1.26 (endergonic on average). Enceladus ocean leads at 10.6 J L⁻¹, plume conduit at 9.96, Europa near-surface ocean next. Four of

five benchmarks matched published values within 1.0 kJ mol⁻¹; 80.0% exergonic in situ.

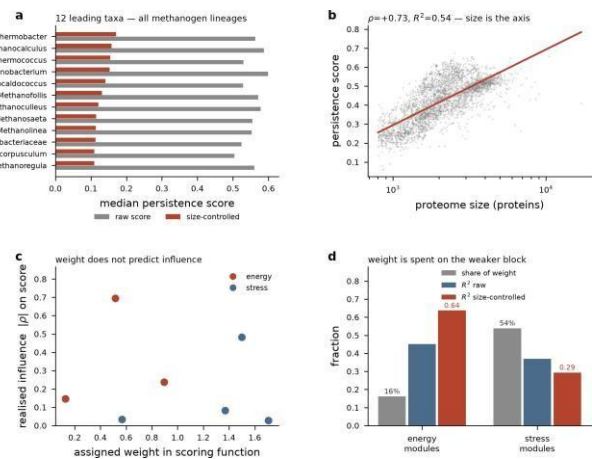


Figure 3. Persistence scores scale with proteome size, not assigned module weight. Panel a ranks 12 methanogen genera by median persistence score, comparing raw (grey) and size-controlled (red) bars. Panel b plots persistence score against proteome size (log scale) for individual proteins, with a fitted red trend line and printed $\rho=+0.73$, $R^2=0.54$. Panel c plots realised influence $|p|$ against assigned scoring weight for energy (red) and stress (blue) modules, showing no clear relationship. Panel d compares weight share, raw R^2 , and size-controlled R^2 (0.64 and 0.29 labelled) for energy and stress modules, with weight share printed as 16% and 54%.

body	niche	dominant couple	ΔG (kJ mol ⁻¹)	energy density (J L ⁻¹)	limiting species
Enceladus	Subsurface ocean	methanogenesis H ₂ /CO ₂	-10.62	10.6179	H ₂
Enceladus	Plume / vent conduit	methanogenesis H ₂ /CO ₂	-9.96	9.9646	H ₂
Mars	Shallow regolith (UV-exposed)	nitrate reduction H ₂	-86.76	4.3378	NO ₃
Europa	Near-surface brine pocket	nitrate reduction H ₂	-83.28	4.1638	NO ₃
Mars	Subsurface brine (RSL-type)	nitrate reduction H ₂	-83.18	4.1590	NO ₃
Enceladus	Seafloor hydrothermal	methanogenesis H ₂ /CO ₂	-3.16	3.1583	H ₂
Europa	Seafloor hydrothermal	methanogenesis H ₂ /CO ₂	-1.05	2.0978	H ₂
Europa	Irradiated surface ice	AOM sulfate	-3.19	0.2551	CH ₄
Ganymede	Irradiated surface ice	AOM sulfate	-2.61	0.2084	CH ₄
Callisto	Surface ice	AOM sulfate	-2.50	0.2001	CH ₄
Europa	Ice-ocean interface	AOM sulfate	-2.23	0.1787	CH ₄
Callisto	Subsurface ocean	methanogenesis H ₂ /CO ₂	-6.34	0.1269	H ₂
Ganymede	Subsurface ocean	methanogenesis H ₂ /CO ₂	-6.34	0.1269	H ₂
Titan	Subsurface water ocean	methanogenesis H ₂ /CO ₂	-6.34	0.1269	H ₂
Mars	Deep subsurface (>1 km)	AOM sulfate	-0.84	0.0669	CH ₄

Table 1. Dominant electron-accepting couple per niche, with in-situ Gibbs energy, exploitable energy density and limiting species.

Control arms and physical constraints

The bacterial control arm on 119 curated extremophiles spanning 386,509 predicted proteins across seven phenotype classes returns *Acidiphilium multivorum* and *Acidiphilium cryptum* as leaders (Figure S2), with significant between-class separation on every body (maximum $p = 0.00578$).

Against the phenotype-matched null, 16 bacterial modules discriminate backwards, with median corrected AUC of 0.627 against naive 0.589 across 73 protein families. Module discrimination appears strong against mesophiles but falls to chance once phenotype-matched extremophiles are used in both domains. Module counts correlate with proteome size ($R^2 = 0.536$), necessitating an 800-protein completeness floor.

Sedimentation Péclet numbers for archaeal cells (density 1100 kg m⁻³, radii 0.2–1.0 μ m) computed from ISS microgravity to 10 g never reach unity, peaking at 0.624 at 10 g (Figure 4). Fractional nutrient delivery enhancement is 0.0700 on Earth and 7.00×10^{-5} on the ISS; gravity adds at most 0.00056% flux above pure diffusion. Mechanoresponse machinery in bacteria: MscS, MreB and two-component signalling in 100% of genomes, MscL in 51%, methyl-accepting chemotaxis proteins in 77%. The mechanoresponse index is highest in acidophiles (0.81) and alkaliphiles (0.78) and lowest in thermophiles (0.66); archaea retain MreB/FtsZ at 73%/86%.

Radiation shielding requirements for *E. coli* and *D. radiodurans* across 0.1–108 year transits nearly superimpose (Figure S3). Thin shielding is counterproductive: secondary particles give 0.1767 Gy yr⁻¹ at 3.7 cm basalt versus 0.1501 Gy yr⁻¹ free-space, with net benefit only beyond 10.8 cm. *D. radiodurans* D10 falls from 49,840 Gy to 31,333 Gy to 6,896 Gy under Jovian heavy ions, a 7.2-fold loss.

Above 243 K, hydrolysis outruns radiation in 45% of scenarios, taking 24.6 years at 273 K versus 2.96×10^{10} years at 200 K. With 90 cm shielding the crossover sits at 233.5 K. Dose spans 8.8 orders of magnitude yet correlates with score at only $r = -0.13$. The radiation-centred framing carries three independent negative results: genomic, engineering, and chemical.

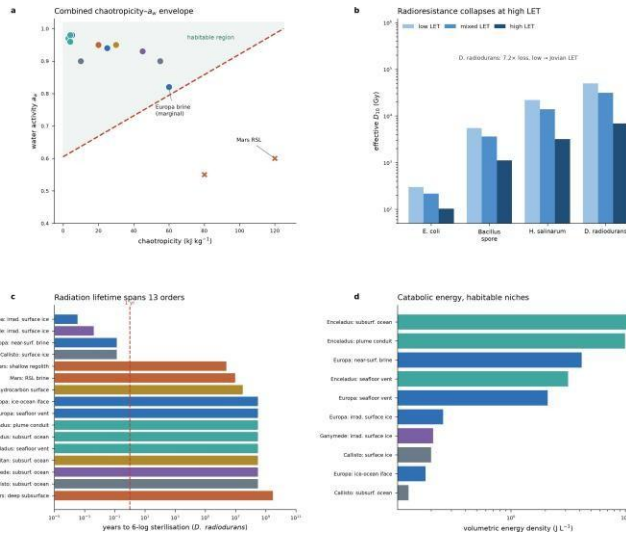


Figure 4. Habitability, radioresistance, sterilisation timescales, and energy density across icy-moon and Mars environments. Panel a plots water activity against chaotropy for planetary brines, with a shaded habitable region and dashed limit; Europa brine and Mars RSL are marked, RSL points falling outside the boundary. Panel b shows bar charts of effective D10 (Gy, log scale) for four organisms under low, mixed, and high LET, decreasing with LET; annotation notes $7.2\times$ loss for D. radiodurans. Panel c shows horizontal bars of years to six-log sterilisation (log scale) for various surface, subsurface, and ocean environments, with a dashed 1-year reference line. Panel d shows horizontal bars of volumetric energy density (J/L, log scale) for ten habitats, colour-coded by body.

Engineering-cost, sensitivity and consortium-design layers

Ablating class E1 across 1,288 catalogued edits makes every pair infeasible; ablating E4 leaves all pairs feasible but raises cost by 0.982 bits on the 21 pairs using it. Classes E2, E3 and E5 never appear in any optimum; E5 costs 6.64 bits. Minimum-cost designs across 134 habitat-target pairs had median cost 2.32 bits (range 0.15–17.64) and median two edits.

Leave-one-out at single-edit resolution: 27% of edits lose no survivor when removed alone; 79.9% of 134 solution sets lose at least one survivor to some single removal, and 20.1% resist every single removal yet fail under pairwise removal.

Of 63 candidate consortium members, 68% were constructible and six needed no edits. E1 dominates selected designs (91.2–100%) but its leave-one-out requirement falls to 96.1% under exclusivity, with robustness surcharge holding in only 59.9% of resampled cases. The full stack lowers median design cost from 8.859 to 6.86 bits and median edits from 3 to 2.5.

A surrogate validated at 199/199 agreement gave 1,700-fold speed-up. Of 1,288 catalogue entries, 907 are E1, 127 each in E2/E3/E4; tractability tiers hold 133, 742, 373, 40.

A sweep of 456 solves showed uniform robustness surcharge of 2.29 bits and one extra edit across all 132 habitats (robust median 7.57 bits, 3 edits). Correcting a sign error shifted E4 invasion from 28.6% to zero and overall invasion at 10% burden from 9.0% to 7.5% of 387 edits.

Of 152 enumerable instances, MILP/Benders found exact optimum in 49.3% (median gap 14.0%, mean 32.2%); genetic algorithm reached 100.0%; random search (400 draws), 98.7%; simulated annealing, 94.7%; greedy variants, 80.3–90.8%. Since 90.4% of resource groups are exact ties, random search's 49.4-point advantage over MILP explains why mathematical programming is not warranted.

Sobol total-order indices gave σ an index of exactly 0.000 because member resource thresholds take only three values across 211 entries. The smallest log-gap (1.386 nats) means ties break only above $\sigma \approx 1.0$, outside sampled range (0.111–0.684). Median cost varied 4.2-fold (3.2–13.4 bits; $S_T=0.66$ for cost, 0.82 for surcharge fraction). Only 14.6% of samples reproduced the 6.86-bit median within ± 0.4 bits.

Exhaustive comparison of 4,104 subsets showed the published solver suboptimal on 47% of pairs by 31.6% mean excess cost; corrected linear master reaches 84.9%. Of 152 instances, 71 have zero and 81 have one alternative.

Replacing binary calls with Bayesian posteriors gave mean AUC gain 0.028 ($r=0.803$, $p=0.0164$); one bin predicted 0.819 against observed 0.425 ($n=73$). The capacity score failed against measured rates (Spearman $\rho=-0.132$, $p=0.65$) across 14 measurements. An alternative term scored AUC 0.882 versus 0.895 for existing term and was rejected.

Coupling coexistence to the optimiser destroyed coverage gain: four-member coverage fell from 0.958 to 0.341; 96 of 200 coverage-optimal designs failed well-mixed coexistence. Of 11 designs, 5 coexisted, 4 were infeasible and 2 were neutrally degenerate. On a 452-design Pareto frontier, 213 coexisted robustly, 145 neutrally, 94 collapsed.

Spatial structure contradicted rescue: persistence was 0.7263 well-mixed, 0.7053 single-face, 0.6632 segregated.

Spatial structure rescued 1 design but broke 20 of 79. Over 10,000 generations, 0 of 93 designs retained two-member coexistence; 46 went fully extinct and median integrity fell from 0.928 to 0.215. Phage predation was viable in only 6 of 93 designs (6.5%).

Monte Carlo over 1,000 replicates gave 14.0 ± 1.5 solvable instances (95% CI 11.0–17.0). Chemostat simulation (1,650 replicates) showed persistence in 85.7% but full member retention in only 3.2%, with designed division of labour absent in 96.8% of runs.

Of 37 designs, 20 chassis are natural, 4 borderline, 13 require unknown capabilities. Of 256 required pairs, 119 (46%) are complete; closing remainder needs 255 transfers across 32 families, 165 (65%) cross-domain. Median containment 3 layers; correlated-failure floor $1e-09$ versus independent-layer estimate $1e-24$. Transfer feasibility fell from 0.488 to 0.410; of 152 fully scored, 150 were downgraded. The two models correlate at $\rho=0.872$, yet 223 moved >50 ranks. Codon factor proved not independent of GC content ($\rho=-0.81$, $p=9.5 \times 10^{-85}$).

The validation programme and its adverse verdicts

The validation programme and its adverse verdicts Across nine tests, verdicts were: completed; fails as pre-registered; fails, and underpowered; falsified; inconclusive plus structural limit; major defect found; partial fail; refuted; refutes core framing — four of nine adverse.

Test 1 (blind prediction, independent MGnify metagenomes) is a partial fail. Habitat-level donor-usage ordering correlates at $\rho = 0.470$ ($p = 0.031$, $n = 21$), but restricted to independent metabolic classes, ρ falls to 0.307 ($p = 0.3312$, $n = 12$), clearing the 0.300 bar by only 0.007. Taxon- and genus-level accuracy is 0.562 against a 0.500 null (precision 0.571, recall 0.500), a 6.2-point margin against a required >10 . (Figure 5)

Test 2 recovers intervention class but is a major defect, not confirmation.

Test 3 (genome-scale FBA cross-validation) is inconclusive and structurally limited: one archaeal model (*Methanosarcina barkeri* iAF692) gives four donor predictions, $\rho = 0.200$ ($p = 0.800$) or $\rho = 0.500$ ($p = 0.667$) for three EC-catalogued donors. Eleven reactions are thermodynamically indistinguishable; eight give identical biomass flux (0.020821).

Test 6 refutes solver necessity: exact on 49.3% of instances, median gap 14.0% (13.99%), mean 32.2%,

worst 231.2%; simulated annealing 94.7%, greedy heuristics 90.8%/89.5%/80.3%. Closed-form analysis recovers all 152 optima 1700-fold faster.

Test 7 (214 abstracts, four claim categories) fails: balanced accuracy 0.288 versus 0.60. Donor-dominance scores 0.240/25 claims (baseline 0.350, $p = 0.917$); coexistence-pairs score 0.0/12; metabolite-exchange 0.625 equals baseline.

Test 8 (sensitivity analysis): solvable niches range 9–15/16. Weight perturbation (2,000 draws) gives median rank correlation 0.97, falling to 0.861 worst-case; top-20 Jaccard falls to 0.243. Nominal $\sigma = 0.1407$ sits in the flat region; at $\sigma = 5.0$ conclusions hold. Relaxation time drives persistence; forcing amplitude has zero effect.

Test 9 (120 papers, 2024–2025 hold-out) fails, underpowered: post-cutoff accuracy 0.333 (4/12), Wilson [0.138, 0.609] versus 0.450 baseline. Pre-2024 accuracy 0.240 (6/25), Fisher's $p = 0.696$.

Further checks: habitability gate admits genera 3.24-fold above chance ($n=16$). Limiting-resource rank averages 11.3/20 versus 10.5 chance ($p = 0.664$). Temperature $R^2 = 0.747/0.774$ holds under phylum hold-out; pH gives $-3.035/-2.202$, salt -1.064 , based on 119 of 3,304 genomes. ICC is 0.944 (iron) and 0.896 (nitrogen); mean genus size 29.9, design effect 22.7-fold. Only 5/63 chassis genomes (7.9%) are complete prototrophs.

Post-revision, calibrated posterior reaches AUC 0.931, Brier 0.098, accuracy 0.873, ECE 0.068, versus AUC 0.790, Brier 0.208, accuracy 0.803 previously; 7,836/23,980 calls change, 44.5% fall in the 0.1–0.9 band. Median design-success probability is 0.187 (range 0.0008–0.842; 1/11 above 0.5).

Derived niche requirements match manual specifications at Jaccard 0.332 (37 niches). Modularity $Q = 0.579$, $z = 29.5$, $p < 0.008$. Coverage rises monotonically with size (0.692, 0.883, 0.932, 0.958) but peaks at two members under coupling; coupling cost 0, 0, 0.127, 0.617. 96/200 designs fail coexistence; 16/50 habitats admit no four-member design. On the Pareto frontier, 452/1,202 candidates (37.6%) are non-dominated. Strongest trade-offs: robustness–size (-0.751), coverage–editability (-0.535), coverage–size (-0.529).

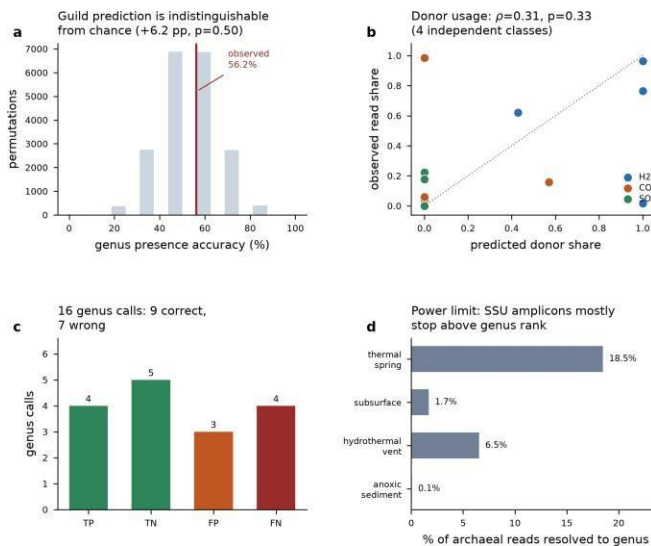


Figure 5. Genus-level prediction, donor usage, and taxonomic resolution show limited accuracy and power. (a) Histogram of permutation accuracies (permutations vs. genus presence accuracy %) with a red line at observed 56.2% and title noting +6.2 pp, $p=0.50$. (b) Scatter of observed read share vs. predicted donor share for H₂, CO₂, and SO₄ classes with a dotted diagonal reference line; title reports $\rho=0.31$, $p=0.33$. (c) Bar chart of 16 genus calls: TP=4, TN=5, FP=3, FN=4. (d) Horizontal bars of % archaeal reads resolved to genus by environment: thermal spring 18.5%, hydrothermal vent 6.5%, subsurface 1.7%, anoxic sediment 0.1%.

Phylogenetic generalisation and ecological realism

Trait prediction cross-validated by holding out taxonomic ranks in turn, with median R² across six traits falling from 0.837 (genus held out) to -0.562 (phylum held out), and four of six traits becoming negative at the phylum level. Only thermal traits retained predictive value across phylum boundaries. (Figure 6)

Genomic module confidence was strongly genus-determined (mean intraclass correlation 0.752), reducing effective sample size from 1,732 to 76. Every predictive claim constitutes within-clade interpolation; the framework cannot extrapolate to novel lineages.

With monotone resource profiles, the lower-threshold organism dominates at every depth in 229 of 231 distinct-threshold cases. Five regimes across 95 designs and 50 habitats yielded 1,045 rows of which 1,025 converged; the 0.66 coexistence rate in segregated biofilm reflects neutral coexistence, not niche partitioning. Corrections were made to fixed-point solving (H₂ of 1.587×10^{-10} M versus damped Newton convergence at 2.0×10^{-10} M in 6 iterations), injection protocol, and mesh grading over 120 cells with cell-width span exceeding 1015.

A catalogue of 50 habitats and 30 mechanisms across periods from 5.6×10^3 to 9.5×10^{14} s produced 6 viable slots. An independent model across 1,050 runs and 21 niches found all-member persistence at 51.9% regardless of forcing amplitude; amplitude's effect was 0.0000, while relaxation time accounted for outcome (1.0000). The ratio of native turnover to coupling timescale had median 2.67×10^4 .

After 10,000 generations, 0 of 93 designs retain multi-member function, and functional integrity falls from 0.900 to median 0.215 (mean 0.341). The median thermal shift needed for viability is 43.25 K, 201 of 223 slots require cooling, and only 10.8% fall within adaptation bound; extinction rises from 47.3% to 80.6% across four decades of maintenance rate.

Sustainable host densities range 4.2×10^2 to 8.0×10^9 cells L⁻¹ (median 4.8×10^4), a median 2.68 orders of magnitude below phage cycle requirement, with sparse mobile elements (median 0.61 per 1,000 proteins).

Median 21 regulators per genome (range 3–208, density 9.35/1,000 proteins), dominated by helix-turn-helix (99.3%), archaeal TBP/TFB (60.7%), and sigma factors (41.4%). Association with retained function strengthens under control ($\rho = 0.56$, $p = 3.8 \times 10^{-9}$), whereas genome size alone does not predict function ($\rho = 0.16$, $p = 0.13$).

Phosphorus, iron and nitrogen limit 15, 15 and 13 cases across 58 niche-scenarios. Only 5 of 63 genomes (8%) are complete prototrophs; pantothenate is least prevalent (43%). MscS occurs in 89.1% of genomes but MscL in only 11.9%. Composition-only features reach held-out R² of 0.872; adding 118 Pfam features raises performance to 0.492 on ionic targets, though these fail phylogenetic hold-out. Ranked by volumetric energy density: Enceladus subsurface ocean 10.6 J L⁻¹, Enceladus plume conduit 9.96, deep-crustal reference 4.16. For 0.5 μ m radius, flux enhancement is $7 \times 10^{-5}\%$ at ISS microgravity versus 0.070% at 1 g; at 50 μ m radius, 0.070% at microgravity versus 70% at 1 g.

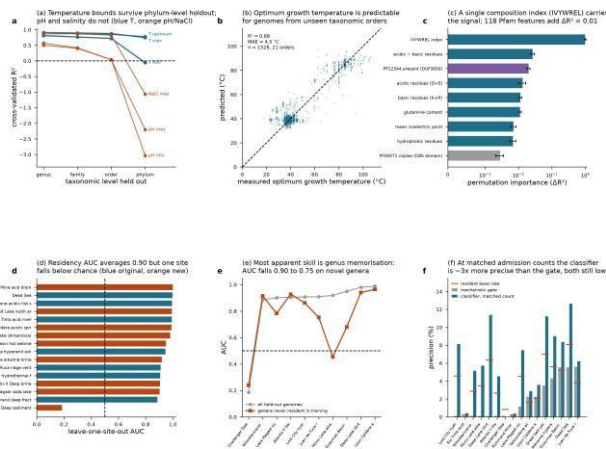


Figure 6. Temperature limits generalize across taxa better than pH or salinity bounds. Panel a plots cross-validated R2 against taxonomic level held out (genus, family, order, phylum) for T optimum, T max, T min (blue) and pH max, pH min, NaCl max (orange). Panel b scatters predicted versus measured optimum growth temperature (°C), with dashed identity line; R2=0.88, MAE=4.5°C, n=1526, 21 orders. Panel c bars permutation importance (AR2, log scale) for IVYWREL index, residue-composition features (teal), PF12344 presence (purple), and PF00571 copies (grey). Panel d bars leave-one-site-out AUC by site, original (blue) versus new (orange), dashed line at 0.5. Panel e plots AUC across held-out sites for all genomes (grey circles) versus genera never resident in training (orange squares), dashed line at 0.5. Panel f bars precision (%) by site for resident base rate (orange line), mechanistic gate (grey), and classifier at matched admission count (teal).

Discussion

The results reduce to one structural property: candidate organisms are metabolically near-interchangeable, and habitats differ chiefly in electron donor supply. This explains the module validation failure, the decomposition result, per-niche energetics, the solver benchmark, near-inert coexistence, spatial/diffusivity nulls, eco-evolutionary collapse, viral failure, weight-sensitivity, containment floor, and the inconclusive genome-scale model (11 thermodynamically indistinguishable reactions).

Practically: ask which donor has standing flux, then which organisms use it — H2 for icy-moon oceans, nitrate for martian brines/Europian pockets, CH4 for irradiated ice. This predicts Mars's salinity problem shifts candidates to halophiles. Blind prediction supports the split: donor usage $\rho=0.47$, taxon identity only 0.562 accuracy.

Forty-seven self-corrections were recorded, two reversing prior conclusions. A 51-protein fragment removed by a completeness filter dropped 227 genomes. Two radiation

modules fell to 0.5 against phenotype-matched null and were removed; the framework's radiation modules score 0.460 inverted, whilst methanogenesis scores 0.925 against 710 methanogens. Naive containment multiplication gave 1×10^{-24} /generation; accounting for shared failure modes gave 1×10^{-9} , fifteen orders worse. Indispensability percentages of 7835.8% and 10000% were recomputed as 79.9% and 100%. An alternative capacity term was rejected (AUC 0.882 vs 0.895). Benchmarked against 16 real sites, six including Lost City, Rio Tinto and Witwatersrand deep subsurface scored zero recall; Challenger Deep sediment scored AUC 0.19, below chance, precision 6%. Six of 16 sites share no unique resident genus. Expanded to 18 sites (3,757 genomes): gate admits 56% overall, but only 1,394 genomes (37%) have measured bounds, of which 4% pass. 643 oxidised near-surface and brine niches where radiolysis Enrichment above chance at 9/17 sites ($p < 0.05$), median recall 0.155, median enrichment 2.76, range 0.00–0.876, pooled odds ratio 17.6 ($p < 10^{-300}$), 228 residents admitted vs 70.3 expected (3.24-fold). Exclusions: lower temperature bound (0.334), upper pH (0.225), upper temperature (0.206). At Dallol, recall rises 0.275→0.849 as enrichment falls from 6.44. The naive NaCl conversion gave zero Dead Sea recall; saturation correction restored recall to 0.48. pH fails 49.3% of 4,186 pairs, salinity 29.6% of 3,366.

A HistGradientBoosting comparator (132 features, 3,531 proteomes) achieves AUC 0.903, threefold higher precision than the gate (0.061 vs 0.019), yet IVYWREL alone carries 0.993 of temperature importance; composition alone reaches R2=0.872. Challenger Deep scores AUC 0.186, inverted.

Of 62 registered limitations: energetics computed for only 15/58 niches; coexistence assumes well-mixed equilibrium, under which 81/134 accepted designs coexist only neutrally; forcing is single-frequency, single-solute (H2); models omit dormancy/seed banks, probably making eco-evolutionary collapse more severe; cost units are ordinal; containment/editability correlate at $\rho=0.8$ (artefactual); only 152/488 transfers have all nine feasibility factors; riboswitch machinery found in only 40/145 genomes; phylogenetic cross-validation effective sample size is 76, not 1,732.

Robust findings: methanogens lead on every icy body, Mars diverges to halophiles; 90.4% of resource groups are exact ties; no design survives 10,000 generations; shielding counterproductive below ~11 cm. Fragile: only

14.6% of parameter space reproduces cost within ± 0.4 bits of 6.86; top-20 Jaccard falls to 0.243; transfer-feasibility median 0.41 is ranking-only. Methanogenesis carries 5.4% of weight yet supplies 100% of leaders.

Conclusion

A framework weighted 53.9:16.2 in favour of stress tolerance over energy metabolism was applied to 3,531 archaeal genomes across 58 parameterised niches on 9 bodies. Energy availability emerged as the determining factor, confirmed by block decomposition, thermodynamic calculation, per-niche resolution of dominant electron-accepting couples, and blind prediction against metagenomes not used in construction.

Methanogenesis dominates 7 reduced niches with serpentinisation-derived H_2 ; nitrate reduction dominates oxidised near-surface and brine niches where radiolysis destroys H_2 and generates nitrate; anaerobic methane oxidation dominates 5 ice-surface and deep-subsurface niches. Methanogenesis (-6.34 kJ mol e^{-1}) succeeds on donor availability despite being among the least favourable couples versus nitrate reduction (-83.18). Radiation modules discriminate at or below chance. Near-interchangeable candidates produce a flat objective landscape, explaining why random search beats mathematical programming by 49.4 percentage points. Of nine pre-registered validation tests, four gave adverse verdicts. Held-out R2 becomes negative at phylum level for most traits. The framework interpolates within clades but does not extrapolate to novel lineages.

Methods

The literature corpus drew on structured OpenAlex queries covering archaeal and microbial survival in space, extremophile physiology and geochemistry of candidate bodies, yielding 57 sources (1997–2023), 48 cited inline.

58 niches were parameterised across 9 settings: Mars (11), deep Earth subsurface (10), interplanetary transit (9), Europa (8), Enceladus (6), Titan (6), Ganymede (4), Callisto (2), Venus (2). Deep subsurface is a calibration anchor; transit is a negative control. Ranges: annual dose $1e-04$ – $1.2e+08$ Gy yr $^{-1}$; water activity 0.00–0.99; ionic strength to 18.0 M, pressure to $1e+05$ bar; temperature -181.2 to 1326.8 °C.

Enceladus: Cassini detected plume H_2 and serpentinisation 717 Permutation testing (20,000 permutations) gave $p < 5 \times 10^{-5}$ evidence, a chemical-energy source for methanogenesis. Ocean is alkaline (pH 8–11),

low salinity (0.05–0.2 molal NaCl), with phosphates; energy density 10.6 J L $^{-1}$. Europa: saline ocean (10–30% terrestrial salinity) under 15–25 km ice; surface radiation 105–106 Gy yr $^{-1}$ generating radiolytic oxidants; ocean ranks third in energy density. Mars: -153 to $+20$ °C, ~ 6 mbar CO $_2$, transient perchlorate/chloride brines, UV to ~ 200 nm, GCR radiolysis. Titan: 3–10 wt% ammonia ocean, 55–150 km deep, 94 K surface. Ganymede: saline ocean between ice layers, radiation ~ 20 -fold lower than Europa's. Callisto: ~ 250 km ammonia ocean, most energy-starved.

3,746 proteomes (3,531 archaeal, 215 bacterial) were HMM-scored; a ribosomal-protein tree collapsed these to 2,069 clades. Of 58 niches, 37 were growth-permissive.

Two pre-genomic gates: temperature > 1600.0 K and water activity < 0.585 . 4 niches failed temperature, 17 failed water activity, leaving 37 habitable. Seven stress axes are anchored to demonstrated limits: temperature 395 K ceiling (M. kandleri, 122 °C) to ~ 258 K floor; water activity 0.605; pH 0–11.4; pressure 1,100 bar routine, 1,500 bar demonstrated; salinity to 5.2 M NaCl; radiation D10 spanning three orders of magnitude; desiccation from anhydrobiosis literature. All seven are terrestrial, so the framework cannot detect life violating them.

Retrieval yielded 3,531 archaeal genomes $\rightarrow$ 2,069 clades, 743 parameters; seven governing quantities were stipulated or 307 genera, six supergroups (Euryarchaeota 2,348, Crenarchaeota/TACK 440, Thaumarchaeota 295, DPANN 273, Asgard 106, unclassified 69), plus 215 bacterial proteomes. 2,288 genomes had NCBI annotations, 1,243 ab initio; completeness floor excluded 227, leaving 3,304.

Eighteen modules were tested; seven failed validation, leaving eleven. HMM census: 2,180 genomes against 120 HMMs (e-value $1e-3$); 24-marker completeness panel (median 1.00; 10.3% below 80%). DIAMOND blastp confirmed 993 archaeal proteins (45 families) and 809 bacterial sequences; prevalence ranged from $> 90\%$ (DNA-repair) to 0.6% (NhaA), 1.8% (DUF3494), 4.1% (EctC). Ribosomal tree used 32,501 columns filtered to 3,898 across 3,338 taxa; ML (LG+I) logL = $-4,430,396$ over 3,198 sequences (14 problematic splits). Threshold 0.02 substitutions/site gave purity 0.867, completeness 0.920, F = 0.893, 2,069 clusters. Before winsorisation, Candidatus Nitrosotenuis chungbukensis topped 8 of 13 niches ($z = +5.37$); winsorising cut its peak score from 3.32 to 1.06. Permutation testing (20,000 permutations) gave $p < 5 \times 10^{-5}$; 3.24-fold enrichment was significant,

but per-site recall spanned 0–0.95.

Binary module scoring was replaced by composite genomic confidence, in-situ Gibbs energy and cardinal-temperature kinetics. Rank correlation between capacity- and binary-ordered organisms was 0.325; 45.4% shifted >20 percentile points; 7,836 of 23,980 module calls changed status. Logistic confidence model achieved AUC 0.931, Brier 0.087; cardinal temperatures available for 1,047 organisms; median coefficient of variation 0.187.

Genus-blocked evaluation (ICC 0.752 vs null 0.158) gave 726 pairs across 30 genera, effective n reduced to 76.

Weights fixed a priori: stress-tolerance 53.9%, energy metabolism 16.2%, methanogenesis 5.4%; cold adaptation weight 1.707. Module threshold $\tau = 0.1645$; envelope tolerances $\pm 2^\circ\text{C}$, $\pm 0.5\text{ pH}$, $\times 1.25\text{ NaCl}$, $-0.01\text{ water activity}$, $\times 1.5\text{ piezophily}$, $\times 100\text{ yr D10}$.

Optimisation layer used 1,288 genome edits (4 classes); 152 instances solved exactly. Consortium layer covered 37 habitat designs with Bayesian module presence, six-objective Pareto optimisation, diffusivity sweep over 7 orders of magnitude, 10,000-generation eco-evolutionary simulation, phage kill-the-winner dynamics.

Nine pre-registered validation strategies found zero fitted parameters; seven governing quantities were stipulated or derived. Hold-out testing was replaced by leave-one-habitat-out testing. BiGG cross-validation was limited ($n = 1$). Temporal hold-out used 12 post-cutoff papers. Per-edit success 0.90; metabolic burden b swept 0–0.525; robustness criterion $\sigma = 0.0$, fraction 0.8; 1,288 edit costs took 31 distinct values. Pre-registered HIGH-risk expectations—successful blind donor prediction, failed taxon-level prediction, strong greedy-heuristic performance—were all confirmed.

Declarations

Funding. No external funding was used for this work.

Competing interests. The authors declare no competing interests.

Ethics, Consent to Participate, and Consent to Publish declarations: not applicable. This study involved no human participants, human data, human tissue or animals, and required no ethical approval.

Code availability. All analysis code is available at github.com/ajayvellanki0720-boop/archaea-astrobiology under the MIT licence. The repository contains the

importable analysis library (archaea_astro, eleven modules), the pipeline stage scripts in execution order, standalone analysis scripts, and a per-cell manifest documenting provenance for every computational step. Results reported here correspond to commit 6a7dcd7. Random seeds are fixed throughout; the conda environment specification (environment.yml) and pinned requirements are included for reproduction.

Use of AI assistance. Ajay Vellanki used a large language model to assist with code review. All analyses, data-processing steps, statistical tests, and interpretations were independently reviewed and verified by the authors, who take full responsibility for the accuracy and integrity of the results and of the manuscript.

References

- Chan; Simons; Maranas (2017). SteadyCom: Predicting microbial abundances while ensuring community stability. PLOS Computational Biology. <https://doi.org/10.1371/journal.pcbi.1005539>
- Chylinski; Makarova; Charpentier et al. (2014). Classification and evolution of type II CRISPR-Cas systems. Nucleic Acids Research. <https://doi.org/10.1093/nar/gku241>
- D'Souza; Shitut; Preussger et al. (2018). Ecology and evolution of metabolic cross-feeding interactions in bacteria. Natural Product Reports. <https://doi.org/10.1039/c8np00009c>
- De Maayer; Anderson; Cary et al. (2014). Some like it cold: understanding the survival strategies of psychrophiles. The EMBO Reports. <https://doi.org/10.1002/embr.201338170>
- DeMaere; Williams; Allen et al. (2013). High level of intergenera gene exchange shapes the evolution of haloarchaea in an isolated Antarctic lake. Proceedings of the National Academy of Sciences. <https://doi.org/10.1073/pnas.1307090110>
- Diener; Gibbons; Resendis-Antonio (2020). MICOM: Metagenome-Scale Modeling To Infer Metabolic Interactions in the Gut Microbiota. mSystems. <https://doi.org/10.1128/msystems.00606-19>
- Duncker; Holmes; You (2021). Engineered microbial consortia: strategies and applications. Microbial Cell Factories. <https://doi.org/10.1186/s12934-021-01699-9>
- Ellery (2021). Supplementing Closed Ecological Life Support Systems with In-Situ Resources on the Moon. Life. <https://doi.org/10.3390/life11080770>
- Elmore; Sheppard; Ramia et al. (2016). Bipartite recognition of target RNAs activates DNA cleavage by the Type III-B CRISPR–Cas system. Genes & Development. <https://doi.org/10.1101/gad.272153.115>
- Enzmann; Mayer; Rother et al. (2018). Methanogens: biochemical background and biotechnological applications. AMB Express. <https://doi.org/10.1186/s13568-017-0531-x>

11. Fukui; Atomi; Kanai et al. (2005). Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1 and comparison with *Pyrococcus* genomes. *Genome Research*.
<https://doi.org/10.1101/gr.3003105>
12. García-Jiménez; Torres-Bacete; Nogales (2020). Metabolic modelling approaches for describing and engineering microbial communities. *Computational and Structural Biotechnology Journal*.
<https://doi.org/10.1016/j.csbj.2020.12.003>
13. Gu; Kim; Kim et al. (2019). Current status and applications of genome-scale metabolic models. *Genome Biology*.
<https://doi.org/10.1186/s13059-019-1730-3>
14. Hartman; Norais; Badger et al. (2010). The Complete Genome Sequence of *Haloferax volcanii* DS2, a Model Archaeon. *PLoS ONE*.
<https://doi.org/10.1371/journal.pone.0009605>
15. Heinken; Hertel; Acharya et al. (2023). Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized medicine. *Nature Biotechnology*.
<https://doi.org/10.1038/s41587-022-01628-0>
16. Hendrickx; De Wever; Hermans et al. (2006). Microbial ecology of the closed artificial ecosystem MELiSSA (Micro-Ecological Life Support System Alternative): Reinventing and compartmentalizing the Earth's food and oxygen regeneration system for long-haul space exploration missions. *Research in Microbiology*.
<https://doi.org/10.1016/j.resmic.2005.06.014>
17. Hille; Charpentier (2016). CRISPR-Cas: biology, mechanisms and relevance. *Philosophical Transactions of the Royal Society B: Biological Sciences*.
<https://doi.org/10.1098/rstb.2015.0496>
18. Hille; Richter; Wong et al. (2018). The Biology of CRISPR-Cas: Backward and Forward. *Cell*.
<https://doi.org/10.1016/j.cell.2017.11.032>
19. Ke; Wang; Yoshikuni (2021). Microbiome Engineering: Synthetic Biology of Plant-Associated Microbiomes in Sustainable Agriculture. *Trends in Biotechnology*.
<https://doi.org/10.1016/j.tibtech.2020.07.008>
20. Koonin; Wolf(2008). Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkn668>
21. Koonin; Makarova (2019). Origins and evolution of CRISPR-Cas systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*.
<https://doi.org/10.1098/rstb.2018.0087>
22. Kulkarni; Mand; Metcalf(2018). Energy Conservation via Hydrogen Cycling in the Methanogenic Archaeon *Methanosarcina barkeri*. *mBio*.
<https://doi.org/10.1128/mbio.01256-18>
23. Leigh; Albers; Atomi et al. (2011). Model organisms for genetics in the domain Archaea: methanogens, halophiles, *Thermococcales* and *Sulfolobales*. *FEMS Microbiology Reviews*. <https://doi.org/10.1111/j.1574-6976.2011.00265.x>
24. Li; Pan; Zhang et al. (2015). Harnessing Type I and Type III CRISPR-Cas systems for genome editing. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkv1044>
25. Lieven; Beber; Olivier et al. (2020). MEMOTE for standardized genome-scale metabolic model testing. *Nature Biotechnology*. <https://doi.org/10.1038/s41587-020-0446-y>
26. Luo; Mullis; Leenay et al. (2014). Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. *Nucleic Acids Research*.
<https://doi.org/10.1093/nar/gku971>
27. Machado; Andrejev; Tramontano et al. (2018). Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gky537>
28. Makarova; Aravind; Wolf et al. (2011). Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biology Direct*.
<https://doi.org/10.1186/1745-6150-6-38>
29. Makarova; Wolf; Iranzo et al. (2019). Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nature Reviews Microbiology*.
<https://doi.org/10.1038/s41579-019-0299-x>
30. Mandell; Lajoie; Mee et al. (2015). Biocontainment of genetically modified organisms by synthetic protein design. *Nature*. <https://doi.org/10.1038/nature14121>
31. Marino; Pinilla-Redondo; Csörgő et al. (2020). Anti-CRISPR protein applications: natural brakes for CRISPR-Cas technologies. *Nature Methods*.
<https://doi.org/10.1038/s41592-020-0771-6>
32. Mendoza; Olivier; Molenaar et al. (2019). A systematic assessment of current genome-scale metabolic reconstruction tools. *Genome Biology*.
<https://doi.org/10.1186/s13059-019-1769-1>
33. Merino; Aronson; Bojanova et al. (2019). Living at the Extremes: Extremophiles and the Limits of Life in a Planetary Context. *Frontiers in Microbiology*.
<https://doi.org/10.3389/fmicb.2019.00780>
34. Mukai; Lajoie; Englert et al. (2017). Rewriting the Genetic Code. *Annual Review of Microbiology* .
<https://doi.org/10.1146/annurev-micro-090816-093247>
35. Navarro-Gonzalez; Rainey; Molina et al. (2003). Mars-Like Soils in the Atacama Desert, Chile, and the Dry Limit of Microbial Life. *Science*.
<https://doi.org/10.1126/science.1089143>
36. Nikel; de Lorenzo (2018). *Pseudomonas putida* as a functional chassis for industrial biocatalysis: From native biochemistry to trans-metabolism. *Metabolic Engineering*.
<https://doi.org/10.1016/j.ymben.2018.05.005>
37. Olsson-Francis; Doran; Ilyin et al. (2023). The COSPAR Planetary Protection Policy for robotic missions to Mars: A review of current scientific knowledge and future perspectives. *Life Sciences in Space Research*.

- 928 <https://doi.org/10.1016/j.lssr.2022.12.001>
- 929 38. Pande; Merker; Bohl et al. (2013). Fitness and stability of
930 obligate cross-feeding interactions that emerge upon gene loss
931 in bacteria. *The ISME Journal*.
932 <https://doi.org/10.1038/ismej.2013.211>
- 933 39. Quehenberger; Shen; Albers et al. (2017). *Sulfolobus* – A
934 Potential Key Organism in Future Biotechnology. *Frontiers in*
935 *Microbiology*. <https://doi.org/10.3389/fmicb.2017.02474>
- 936 40. Rampelotto (2013). Extremophiles and Extreme
937 Environments. *Life*. <https://doi.org/10.3390/life3030482>
- 938 41. Riley; Guss (2021). Approaches to genetic tool development
939 for rapid domestication of non-model microorganisms.
940 *Biotechnology for Biofuels*.
941 <https://doi.org/10.1186/s13068-020-01872-z>
- 942 42. Roell; Zha; Carr et al. (2019). Engineering microbial
943 consortia by division of labor. *Microbial Cell Factories*.
944 <https://doi.org/10.1186/s12934-019-1083-3>
- 945 43. Rottinghaus; Ferreira; Fishbein et al. (2022). Genetically
946 stable CRISPR-based kill switches for engineered microbes.
947 *Nature Communications*.
948 <https://doi.org/10.1038/s41467-022-28163-5>
- 949 44. Salisbury; Gitelson; Lisovsky (1997). Bios-3: Siberian
950 Experiments in Bioregenerative Life Support. *BioScience*.
951 <https://doi.org/10.2307/1313164>
- 952 45. Schafer; Engelhard; Muller (1999). Bioenergetics of the
953 Archaea. *Microbiology and Molecular Biology Reviews*.
954 <https://doi.org/10.1128/mmbr.63.3.570-620.1999>
- 955 46. Scott; Benito-Vaquero; Zimmermann et al. (2023). A
956 structured evaluation of genome-scale constraint-based
957 modeling tools for microbial consortia. *PLOS Computational*
958 *Biology*. <https://doi.org/10.1371/journal.pcbi.1011363>
- 959 47. Soo; McAnulty; Tripathi et al. (2016). Reversing
960 methanogenesis to capture methane for liquid biofuel
961 precursors. *Microbial Cell Factories*.
962 <https://doi.org/10.1186/s12934-015-0397-z>
- 963 48. Tang; Tham; Liu et al. (2021). Hydrogel-based
964 biocontainment of bacteria for continuous sensing and
965 computation. *Nature Chemical Biology*.
966 <https://doi.org/10.1038/s41589-021-00779-6>
- 967 49. Vago; Westall; Pasteur Instrument Teams, Landing S et al.
968 (2017). Habitability on Early Mars and the Search for
969 Biosignatures with the ExoMars Rover. *Astrobiology*.
970 <https://doi.org/10.1089/ast.2016.1533>
- 971 50. Vercoe; Chang; Dy et al. (2013). Cytotoxic Chromosomal
972 Targeting by CRISPR/Cas Systems Can Reshape Bacterial
973 Genomes and Expel or Remodel Pathogenicity Islands. *PLoS*
974 *Genetics*. <https://doi.org/10.1371/journal.pgen.1003454>
- 975 51. Verseux; Baqué; Lehto et al. (2015). Sustainable life support
976 on Mars – the potential roles of cyanobacteria. *International*
977 *Journal of Astrobiology*.
978 <https://doi.org/10.1017/s147355041500021x>
- 979 52. Vo; Ronda; Klompe et al. (2020). CRISPR RNA-guided
980 integrases for high-efficiency, multiplexed bacterial genome
981 engineering. *Nature Biotechnology*.
982 <https://doi.org/10.1038/s41587-020-00745-y>
- 983 53. Wagner; van Wolferen; Wagner et al. (2012). Versatile
984 Genetic Tool Box for the Crenarchaeote *Sulfolobus*
985 *acidocaldarius*. *Frontiers in Microbiology*.
986 <https://doi.org/10.3389/fmicb.2012.00214>
- 987 54. Wong; Ng; Mat et al. (2016). Coevolution Theory of the
988 Genetic Code at Age Forty: Pathway to Translation and
989 Synthetic Life. *Life*. <https://doi.org/10.3390/life6010012>
- 990 55. Zelezniak; Andrejev; Ponomarova et al. (2015). Metabolic
991 dependencies drive species co-occurrence in diverse microbial
992 communities. *Proceedings of the National Academy of*
993 *Sciences*. <https://doi.org/10.1073/pnas.1421834112>
- 994 56. Zeng; Chen; Yang et al. (2022). Identification of a protein
995 responsible for the synthesis of archaeal membrane-spanning
996 GDGT lipids. *Nature Communications*.
997 <https://doi.org/10.1038/s41467-022-29264-x>
- 998
999
1000
1001
1002
1003
1004
1005
1006

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SpaceMicrobes1Supp.docx](#)