The evolution of biogeochemical recycling by persistence-based selection

Richard A. Boyle¹✉ & Timothy M. Lenton¹ ¹

Darwinian evolution operates at more restricted scales than the feedback processes within the Earth system, precluding the development of any systematic relationship between the organism-level traits favored by natural selection and the impact of these traits upon Earth's long-term average habitability for life. "It's-the-song-not-the-singer" theory proposes an extended understanding of natural selection to encompass differential persistence of non-replicating entities, potentially allowing for a quasi-Darwinian understanding of biogeochemical cycles. Here we use a simple stochastic model to demonstrate how persistence selection of the form invoked by "It's-the-song-not-the-singer" can stabilize a generic nutrient recycling loop, despite its dependence upon genotypes with relatively low organism-level fitness. We present an evolutionary trajectory plausibly representative of aspects of Precambrian biogeochemical cycles, involving persistence-based selection for recycling via fluctuations in abiotic boundary conditions and strong genetic drift. We illustrate how self-perpetuating life-environment correlation patterns, as opposed to specific state-values, may help empirically distinguish "It's-the-song-not-the-singer" from conventional Earth-system feedbacks.

¹Global Systems Institute, College of Life and Environment, University of Exeter, Exeter, UK. ✉email: richboyleacademic@gmail.com

Lovelock's "Gaia" hypothesis suggested that a planetary-scale homeostatic system emerges from life's interaction with the abiotic environment, which tends to maintain Earth's average long-term habitability for life¹. This idea stimulated a long-running debate concerning whether there is any legitimate sense in which life keeps the Earth in a state that is good for life, and how the biota's global-scale impact relates to the local, future-insensitive scope of natural selection^{2–4}. The basic homeostatic properties that Gaia invokes (e.g. resistance to, and resilient recovery from, external perturbation) are uncontroversial features of the internal physiology and biochemistry of living organisms⁵. But such intra-organism features are adaptations derived from natural selection: the result of heritable variation causing the non-random differential survival and reproduction of a subset of a population of replicating entities, relative to the rest of that population⁶. Neither the biosphere as a whole, nor any planetary scale process or feedback, exhibits discrete replication analogous to that of organisms, or occurs in an interacting population of comparable entities. This has led many to the conclusion that any "Gaian" habitability-promoting influence cannot be the product of natural selection and is therefore either non-existent², or somehow life-specific without being particularly Darwinian⁷. The most generous Darwinism-focused interpretation of Gaia is along the lines that *selection* of some degree of climatic/geochemical homeostasis might emerge as a by-product^{3,8} of *selection for*⁹ conventional organism-level adaptations at a lower scale, in conjunction with physical constraints on adaptation¹⁰ and anthropic luck^{11,12}. The least generous such interpretation characterizes Gaia as a vague and insubstantial metaphorical appeal to nature's interconnectedness⁹.

However, the basic observations that stimulated Lovelock's hypothesis remain legitimate. Dynamically sustained atmospheric chemical disequilibrium is both a key contributor to long-term habitability and a fundamental distinguishing factor that separates Earth from comparable lifeless planets, such as Mars or Venus^{13,14}. Some statistical evidence is suggestive of homeostasis-like properties at the planetary-climatic and ecosystem scales^{15,16}. The Earth system is characterized by global-scale biologically driven recycling of essential elements, such that the flux into photo-synthesizers of several essential elements greatly exceeds their influx into the biosphere from abiotic sources¹⁷ (for carbon by a factor of ~200, for nitrogen ~500–1300, for phosphorus >1000¹⁸, and for sulfur ~10^{8,19}). Evidence of this sort begs the question of precisely how and why the by-products of life appear to be more conducive to the persistence of habitability than by-products of abiotic processes on lifeless planets. Despite years of polarized debate, no consensus has emerged as to what (if anything) life's putative habitability promoting influence is supposed to be, or how any such influence relates to natural selection²⁰.

Ford Doolittle's "It's-the-song-not-the-singer" (ITSNTS) theory²¹ aims to move forward from this unproductive confusion, proposing that a "Darwinized Gaia" is achievable through an emphasis on the differential persistence of non-replicating, non-interacting entities²², including the global-scale biogeochemical cycles and processes upon which Earth system science focuses^{23,24}. The idea of natural selection based on differential persistence alone builds on moves by philosophers of biology towards a more relaxed understanding of heredity that does not require conventional replication^{25,26} and is consequently applicable to (for instance) ecological functions^{27,28}. It is uncontroversial that a biological phenotype may impact its abiotic environment in a way that feeds back upon the fitness and relative frequency of the allele corresponding to that phenotype. Existing bodies of theory within the extended neo-Darwinian synthesis, such as niche construction²⁹, interpret such feedbacks as maximal phenotypic extensions of the relevant allele, which can be construed

as adaptations only insofar as they reliably map on to heritable biological variation. Geochemical perspectives tend to remain agnostic about the extent to which life's abiotic impact is legitimately interpretable as the result of Darwinian adaptation, instead focusing on the emergent environmental impacts of biological evolution's unselected by-products⁸.

Defining the problem. ITSNTS differs from the above views in suggesting that particular life-environment interaction states have something loosely comparable to a fitness of their own, derived from the potential for such states to persist and irreducible to either biotic or abiotic states in isolation²¹. A problem with invoking selection without replication is the lack of a continuous source of variation analogous to mutation³⁰. In the absence of such a source, persistence-selection is potentially restricted to a negative form of stability based sorting³¹, capable only of removing a subset of the variation initially present within the population of differentially persisting entities, rather than anything analogous to incremental evolution of complex adaptations. ITSNTS confronts this difficulty via "evolutionary recruitment" of genotypes by geochemical cycles, such that cycle-variants with a tendency to reinforce the growth/proliferation of the key taxa connected to them (by stabilizing the presence of the relevant niches²³) tend to out-persist comparable cycle-variants that do not^{23,24}. Persistence selection thus acts on variation derived from the biota but not reducible to organism-level phenotypes. The crux of ITSNTS is the proposition that novel biological variation arises within lineages that are present by virtue of their association with an environmental state, such that without the environmental interaction the novel variation would not have arisen, because the relevant lineage would be absent or fundamentally different.

We tentatively propose that the key to begin falsifying the subtle difference between ITSNTS and a more conventional view based on emergent by-products, feedbacks and/or niche construction, is to focus on the change over time in the correlation between biological traits and environmental states. We suggest that if a life-environment interaction state is a function of fixed time-invariant physical constraints on adaptation, any directional change in this interaction state will be restricted to that resulting from changing physical boundary conditions. By contrast, if such an interaction state is the result of some form of persistence-selection, it will (we argue) exhibit a directional change analogous to that observed in selected organism-level traits (see supplementary material, section S1). We aim to make these ideas concrete by formulating our arguments in terms of a simple model of the recycling of growth-limiting environmental substances. The basic structure of our model is already a feature of evolutionary-ecological thinking on syntrophy³², symbiosis³³ and multi-level selection^{34,35}—in the sense that there is an incompatibility between the individual-level dynamics and the population level features upon which the existence of these dynamics depend. Our aim here is to use this general model structure to conceptually relate ITSNTS to the issues connecting natural selection and the Gaia hypothesis.

Model structure. Key model processes are shown in Fig. 1, full model equations are given in the methods section, Tables 1 and 2 list model variables and parameters respectively. Two environmental substances R_1 and R_2 , analogous to nutrients or respiratory substrates, respectively limit the growth of two species S_1 and S_2 . Individuals assimilate growth-limiting substance at a genotype-specific rate to reproduce, which produces genotype-specific by-products. Within each species, the by-product of a "producer" allele is the environmental substance that limits the

growth of the other species, whereas a non-producer allele produces no metabolically accessible by-product. A phenotypically plastic allele exhibits an environmentally contingent switch in

phenotype, such that plastic individuals adopt the phenotype of the producer when the environmental level of the substance is above a critical (fixed) threshold and that of the non-producer below that threshold. Mutation causes the genotype of offspring individuals to differ from that of their parent at a fixed probability that is uniform across genotypes.

The potential for inconsistency between the equilibrium solutions of population-level natural selection and global geochemical dynamics is represented by a per-capita producer cost κ_{prod} , by which the producer's growth rate is negatively offset relative to that of the non-producer. (The plastic allele exhibits the growth rate of whichever phenotype it adopts, i.e. incurs the cost only when it adopts the producer phenotype). Simultaneous presence of both producer alleles gives rise to a recycling loop involving the sequential interconversion between R_1 and R_2 (Fig. 1). The strength of recycling of each environmental substance is measured by the cycling ratio, the ratio between a substance's flux through the biota and the entry of that substance into the biota from the outside³⁶ (methods, model description, Eq. (21)). Environmental substance is distributed evenly across all individuals without any spatial structure whatsoever, thus there is no possibility for producer individuals to disproportionately benefit from the recycling they support. Our results refer to separate implementations of single isolated populations, which for a given set of parameter choices differ between each other only via the model's stochastic functions.

Fig. 1 Feedback diagram of life environment interactions in our model system. Environmental substance pools (circles) have a positive influence (solid arrow-lines) on the growth and biomass of the biota (squares). Environmental substance levels are negatively influenced (dashed arrow-lines) by biotic assimilation and positively influenced by abiotic input fluxes (diamonds) as well as by-products of producer phenotypes, the latter creating the potential for net recycling.

Results

For a given environmental substance input, the population reaches a steady state at which reproductive growth rate balances death rate (methods, model description, Eq. (1)). Because all individuals have an equal death probability, but producers have a lower growth rate due to the per-capita cost, individual level selection favors non-producers in all circumstances. In a deterministic system or an idealized infinite population size, the only equilibrium solution, assuming a non-zero producer cost, would be competitive exclusion of producers by non-producers. In the finite populations used here, the population composition reaches a dynamic steady state at which the production of new producer genotypes by mutation balances their decline via individual level selection.

Table 1 Model variables.	
Model variable	Meaning and units
$N_{q,j}(t)$	Number of individuals of genotype j in species q at time t
$\rho_{q,j}(t)$	Number of pre-reproduction starvations within genotype j , species q at time t
$G_{q,j}(t)$	Growth rate (genotype j , species q at time t). Number of offspring individuals produced per parent individual.
$\Upsilon_{q,j}(t)$	Total mutations away from genotype j to other genotypes produced by genotype j , species q at time t . Number of offspring individuals.
$\Upsilon_{q,k\neq j}(t)$	Total mutations to genotype j from other genotypes, within species q at time t . Number of offspring individuals.
$\delta_{q,j}(t)$	Total deaths due to factors other than insufficient environmental substance genotype j , species q time t (i.e. “drift”). Number of individuals.
$S_q(t)$	Total number of individuals of species q (of any genotype).
$\varphi_{R_q}(t)$	Net abiotic influx of environmental substance R_q at time t . Units of environmental substance per generation.
$\frac{\partial t_{geo}}{\partial t_{bio}}(t)$	Model forcing mapping between abiotic/geological time scales and evolutionary/biological timescales. For simulations shown either $\frac{\partial t_{geo}}{\partial t_{bio}}(t) = 1$ or $\frac{\partial t_{geo}}{\partial t_{bio}}(t) = 0$, or discrete switching between the two. Dimensionless scaling factor.
$A_{R_q}(t)$	Total units of environmental substance R_q assimilated by the biota at time t .
$P_{q,prod}(t)$	Production rate of environmental substance R_q per producer-genotype individual within species q at time t .
$P_{R_q,NET}(t)$	Total units of environmental substance R_q produced by the biota at time t .
$T_{R_q}(t)$	Residence time of environmental substance R_q (average number of biological generations between entry to and exit from the pool of this substance for a typical unit of this substance).
$CR_{R_q}(t)$	Cycling ratio of environmental substance R_q (ratio between the flux of this substance through the biosphere and the entry of this substance into the biosphere from the outside)

Table 2 Model parameters.

Model parameter	Default value and units	Qualitative impact on results
K	Carrying capacity (maximum number of individuals of either species), $K = 10000$	Dictates maximum size of a population in the absence of limitation by the level of environmental substance
G_{qjPR}	Reproductive growth rate (genotype j , species q), per unit environmental substance R assimilated	Controls the time a population takes to reach equilibrium size and composition
f_{conv}	Conversion efficiency between assimilation of environmental substance R and production of offspring by parent individuals $f_{conv} = 0.5$ unless otherwise stated	Impacts the number of biological individuals a given environmental substance pool can support
G_0	Baseline growth rate (offspring individuals produced per parent individual, per time-step) $G_0 = 5$	Impacts rate at which population reaches equilibrium size and composition
$R_{assimMAX}$	Maximum quantity of environmental substance that can be assimilated by a single individual during a single time-step $R_{assimMAX} = 3$	Dictates rate of net population growth
κ_{prod}	Relative decrease in reproductive growth rate associated with the producer phenotype compared to the non-producer phenotype. Default value $\kappa_{prod} = 0.1$	Dictates range of boundary conditions over which the producer can coexist with the non-producer
$p_{k \rightarrow j}$	Probability of a mutant offspring from another genotype being the j th genotype, $p_{k \rightarrow j} = \frac{1}{J_{total}-1} = 0.5$.	Ensures unbiased mutation
μ_0	Baseline mutation probability per offspring produced. Default value $\mu_0 = 0.01$.	Sustains genotypic diversity delaying displacement of producers by non-producers for a given set of boundary conditions
δ_0	Death probability due to “random” factors unconnected with environmental substance levels, per individual per generation. Default value $\delta_0 = 0.1$.	Thins the population, ensures starvation is not the only cause of death
φ_{0R_q}	Baseline net abiotic influx of environmental substances R_1 and R_2 , units of substance per time step. Default value $\varphi_0 = 10$ for both substances.	Dictates net equilibrium population size, if this number is below carrying capacity
$R_{q,NET0}$	Scaling factor representing sensitivity of abiotic removal flux to existing size of environmental substance pool, corresponding to baseline abiotic steady state residence time $R_{q,NET0} = \frac{R_{q,NET(0)}}{\varphi_{0R_q}}$. Default value $R_{q,NET0} = 100$	Inversely related to the time an environmental substance pool will take to become exhausted in the absence of abiotic influx or biotic production
$f_{conv,PROD}$	Efficiency with which individuals with the producer phenotype convert the environmental substance that they assimilate to the by-product substance that they emit. Default assumption $f_{conv,PROD} = (1 - f_{conv}) \cdot (1 - f_{loss,prod})$	Dictates the net efficiency with which producers generate one environmental substance from the other, thus impacts net population size during shut-off intervals
$f_{loss,prod}$	Fraction of growth-by-product environmental substance lost during production flux $f_{loss,prod} = 0.4$	As $f_{conv,PROD}$
$f_{symmetry}$	Symmetry breaking parameter applicable in a subset of runs, causing a difference in physiological efficiency between the two distinct producer genotypes, implemented as: $f_{conv,PROD,S_1} = (1 - f_{conv}) \cdot (1 - f_{loss,prod})$ $f_{conv,PROD,S_2} = (1 - f_{conv}) \cdot (1 - f_{loss,prod}) \cdot f_{symmetry}$ Default value $f_{symmetry} = 1$	Deviation from 1 causes net production in one species to exceed that of the other, meaning, increasing the rate of incremental population decline when abiotic influx is switched off.
$R_{q,BIOAVAILABLE_{crit}}$	Bio-available environmental substance level at which the plastic genotype switches phenotype (producer above, non-producer below) $R_{X,BIOAVAILABLE_{crit}} = 1$	When non-zero ensures that plastic genotype can contribute to an existing recycling loop, but not sustain it in the absence of the producer genotype

Persistence/viability selection for recycling. Figures 2 and 3 show how a transient shut-off in the external/abiotic influx of growth-limiting environmental substances imposes persistence selection for the existence of the recycling loop. The first quarter of each simulation involves a constant positive influx, followed by a shut-off event, during which population survival depends upon any environmental substance generated as a by-product of the opposite species, thus on the occurrence of net recycling. The stability of recycling is a function of the pre-shut-off producer frequency, which in turn is a function of the cost. As the producer genotype in each species approaches zero frequency, shut-off of the abiotic environmental substance influx leads to death of the whole population, because assimilation of environmental substance exceeds the size of bio-available substance pools, resulting in their exhaustion. Reduction in producer cost, under otherwise identical boundary conditions, allows a greater number of producers to coexist with non-producers prior to the shut-off, potentially allowing sufficiently strong recycling for the population to survive a given shut-off interval. In other words, there

exists a straightforward “tragedy of the commons”³⁷ in that the non-producer genotypes favored by individual-level selection result in collective extinction by exhausting the pool of environmental substance.

Net efficiency of recycling is less than 100%, due to the inclusion of parameters qualitatively representing metabolic heat loss (Eqs. (3) and (18)), meaning that persistence selection cannot support a population indefinitely. Extension of the influx shut-off interval eventually leads to a population crash under all parameter choices, but the occurrence of recycling due to the presence of producers delays such a crash. The size of the region of parameter space within which sufficient recycling occurs to allow population survival is a decreasing function of both producer cost and shut-off interval length, as illustrated by the replicate-averaged sensitivity analysis in Fig. 4. Note that the region of cost-versus-shut-off-interval parameter space that allows recycling and population survival is identical with that permitting survival of the producer genotype (the implications of which are discussed further below). This recycling-permitting

Fig. 2 Two model simulations illustrating the general principle of persistence selection for a recycling loop. Each column shows the result of a single 2000-generation model simulation, 500 generations into which the abiotic influx of both environmental substances is set to zero for 150 generations. Top row shows the frequencies of producer (blue), non-producer (red) and plastic (green) genotypes within species S_1 (solid lines) and S_2 (dashed lines). Bottom row shows abiotic input (purple) biotic assimilation (brown) and production (light blue) fluxes impacting the level of environmental substances R_1 (solid lines) and R_2 (dashed lines). Left column (A, C) shows a simulation in which the producer cost is relatively low $\kappa_{prod} = 0.001$, right column (B, D) a simulation in which it is relatively increased $\kappa_{prod} = 0.01$.

Fig. 3 Impact of the length of time during which abiotic influx of environmental substances is set to zero. Left column shows the results of a single simulation within which this shut-off event is 100 generations, right column one within which it is 150 generations. Panels A–D are equivalent to Fig. 2, showing genotypic composition and fluxes. E, F show the cycling ratios of both environmental substances, G, H the net level of these substances in the system. Shut-off length is the only difference between the two simulations, producer cost is uniformly $\kappa_{prod} = 0.01$.

Fig. 4 Analysis of the sensitivity of recycling and system composition to changes in producer cost and shut-off interval length. Each datapoint shows a 100 replicate average value of the variable given by the colored legend, at the end of a model run involving a single shut-off perturbation event, analogous to Figs. 2 and 3. Vertical axis shows the length of the environmental substance influx shut-off interval, horizontal axis shows producer cost. Individual graphs show the final values of **A** producer relative frequency, **B** total population (both given by a two species average), **C** environmental substance cycling ratio and **D** total assimilation rate of environmental substance per generation (both two substance averages). See also supplementary figure 1 for an expansion of this analysis along the X-axis.

region is also a small proportion of the parameter space examined (see also supplementary Fig. 1).

The effectiveness of persistence selection for recycling is a function of the physiological reciprocity between the two species. In addition to the producer cost κ_{prod} , the effectiveness and stability of recycling depends upon various parameters representative of fundamental constraints on organism physiology and ecology, the different impacts of which are illustrated in Fig. 5. These parameters are: (i) The baseline mutation rate μ_0 , which dictates the proportion of offspring with a different genotype to their parent. In the absence of mutation, non-producer fixation occurs more rapidly, because non-producer individuals do not give rise to producer offspring.

(ii) The inter-species physiological symmetry $f_{symmetry}$, which measures the relative effectiveness of the producer phenotype between the two species (see Eq. (18) and (19) and Table 2). Any substantial deviation from $f_{symmetry} = 1$ corresponds to a situation in which recycling in one species is more efficient than that in the other, which makes the low-efficiency species increasingly parasitic, translating into a reduced maximum tolerable shut-off interval.

(iii) The abiotic environmental substance residence time $R_{q,NETO}$, which dictates the time an average unit of environmental substance spends in the system before removal via metabolic heat loss (Eqs. (15) and (16)). A shorter abiotic residence time translates to more rapid depletion of the substance pool, thus a shorter maximum tolerable influx shut-off interval.

(iv) The efficiency of the producer by-product reaction, which is measured by $f_{loss,prod}$, the proportion of environmental substance lost during the conversion between the two environmental

substances (Fig. 1, Table 2 and Eq. (17)). The greater the value of $f_{loss,prod}$, the more environmental substance is lost at each iteration of the recycling loop, thus the shorter the maximum tolerable shut-off interval.

The remainder of our results focus on what we argue to be a plausible trajectory for the emergence of a biogeochemical cycle via persistence selection. This trajectory encompasses the key additional ideas with which we suggest ITSNTS be supplemented: (i) fluctuating abiotic boundary conditions, (ii) relaxation of individual level selection by increased genetic drift, and (iii) net genetic assimilation and entrenchment of ancestrally plastic recycling traits.

Population bottlenecks, net genetic assimilation and entrenchment provide a plausible evolutionary trajectory for a generic recycling loop. Figure 6 compares two single simulations so as to illustrate this biogeochemical/evolutionary trajectory. The model was initialized solely with individuals of the plastic genotype, thus any individuals of one of the other two genotypes subsequently present are the product of individual-level mutation and selection. In contrast to Figs. 2 and 3, repeated on-off fluctuations in abiotic substance influx were imposed, occurring over the middle half of the model run. These fluctuations occurred in conjunction with population bottlenecks, during which all but a random founder population was culled and the next generation seeded with this bottleneck-derived sub-population. Finally, during the second half of the simulation, mutation was set to zero for all genotypes, representing entrenchment of recycling phenotypes. The on-off influx fluctuations result in persistence selection for recycling, whilst the genetic drift resulting from repeated population bottlenecks corresponds to a relaxation of individual-level selection (which acts

Fig. 5 Sensitivity analysis to evolutionary and “eco-physiological” parameters affecting the maximum-tolerable length of an environmental substance influx shut-off. Color axes show the maximum influx shut-off interval for which the 100 replicate-averaged population size of both species is non-zero, given the other parameter choices, in the context of a 2000 generation single-shut-off simulation of the form of Figs. 2 and 3. Horizontal axes show the value of the producer cost κ_{prod} , vertical axes show the other parameter being examined, which is **A** baseline mutation probability μ_0 , **B** between-species producer efficiency ratio $f_{symmetry}$, **C** environmental substance abiotic residence time $R_{q,NETO}$, and **D** fraction of environmental substance lost during production reaction $f_{loss,prod}$. Expanded range of cost parameter values is given in Supplementary Fig. 2.

against the producer as a function of the cost, as described above). If fixation of producers by drift occurs whilst mutation is still present, re-emergence of all genotypes leads to a return to the original basic dynamics (non-producer dominance). However, where producer fixation occurs after the onset of genetic entrenchment, the result is a stabilized self-perpetuating recycling loop of the form shown in Fig. 6. The latter trajectory entails a cycle “lock-in” event involving producer fixation and stabilized recycling.

At the level of multiple modal replicates, the boundary conditions shown in Fig. 6 give rise to a strongly bimodal results distribution, in which the system undergoes either a recycling loop lock-in or complete population crash. This replicate-averaged results set is depicted in Fig. 7. The key parameters affecting the balance between these two extremes are producer cost, bottleneck-population size, and “entrenchment fraction” (fraction of the model run after which mutation is switched off). The proportion of model runs giving rise to dynamics of the form of Fig. 6 is larger within regions of parameter space at which entrenchment occurs early during the fluctuation interval, provided entrenchment overlaps with the imposition of bottlenecks and the chance for producers to reach fixation by drift. Cycle “lock-in” requires that producers reach fixation by the bottleneck after entrenchment (i.e. cessation of mutation) has occurred, which occurs more frequently at smaller bottleneck sizes, because the non-producer is always the most abundant genotype before the bottleneck.

Regions of parameter space associated with persistence selection for recycling exhibit correlation between cycling ratios and producer frequencies. Figure 8 illustrates how increasingly strong persistence selection for recycling is identifiable by a top-down pattern irreducible to either genotypic or environmental

properties in isolation: an increased correlation between the species-averaged relative producer frequency and the environmental-substance averaged cycling ratio. In the absence of persistence selection for recycling, neither the correlation between the two cycling ratios of each separate environmental substance, nor the correlation between the species-averaged producer relative frequency and the environmental-substance-averaged cycling ratio, approach 1 in a stable or consistent manner. Fig 8 illustrates how the strength of these correlations track a progressive increase in the strength of persistence selection for recycling, in the sequence of single shut-off, fluctuations only, and fluctuations with bottlenecks and entrenchment (see also Supplementary Figs. 3 and 4).

Discussion

Our model is a highly simplified representation of (what we argue to be) the functionally essential features of biogeochemical cycles: growth-limitation by the ambient environmental level of essential elements/substances, cyclical inter-conversion between variants of these substances, and irreducibility of some properties of the cycle as a whole to its individual (biotic and abiotic) constituents. Recent investigations into Gaia have focused on the relative balance between chance and stabilizing mechanisms in the sustenance of habitability over the course of Earth’s history³⁸. We agree that an element of anthropic luck has probably been critical for the continuation of habitable conditions but argue that the central question raised by the Gaia hypothesis concerns the comparison between life and non-life, rather than that between chance and stabilizing feedbacks. ITSNTS confronts the puzzle posed by Gaia insofar as biogeochemical cycles that regenerate some essential element (e.g. N, P, etc.) inherently maintain a degree of homeostasis with respect to the level of that element at

Fig. 6 Comparison between two single simulations illustrating a biogeochemical-cycle “lock-in” trajectory. The two columns show two single model simulations under the same parameter choices, which differ from each other only in terms of the model’s stochastic processes. Both runs show a 2000 generation simulation, 600 generations into which sustained on-off fluctuations are induced in the influxes of both environmental substances. The start of each “off” interval involves an extreme population bottleneck during which only one randomly selected individual survives. Halfway through each simulation, mutation is switched off in all genotypes, mimicking genetic entrenchment. The left column shows a cycle “lock-in” event, the right a (more typical) population crash. Panels **A, B** show genotype frequencies, **C, D** fluxes, **E, F** cycling ratios, and **G, H** net quantity of environmental substances. Mutation rate $\mu_0 = 0.01$, producer cost $\kappa_{prod} = 0.01$, shut-off interval during fluctuations 100 generations.

Fig. 7 Factors affecting the likelihood of cycle “lock-in” during fluctuation runs. Color contour shows the % of runs (100 replicates per parameter combination), each involving influx fluctuations as in Fig. 6, which result in producer fixation and cycle “lock-in”. Horizontal axes show producer cost, vertical axes various parameters of relevance to this percentage: **A** entrenchment fraction (i.e. proportion of model run for which mutation is switched off), **B** length of shut-off interval, **C** size of (repeated) population bottlenecks, **D** mutation rate. (“Lock-in” event defined as non-zero population combined with producer fixation in both species).

Fig. 8 The impact of different contexts on the strength of persistence selection for recycling. Within each plot, color contour shows the value of the focal quantity (100 replicate average final value for each parameter choice), which is (top row **A–C**) the correlation between the cycling ratios of the two environmental substances, (middle row **D–F**) the correlation between the two-species averaged relative producer frequency and the two-substance averaged cycling ratio, and (bottom row **G–I**) the value of this two-species averaged relative producer frequency. Left column shows results from single-shut off simulations (as Figs. 2 and 3), middle column shows on-off influx fluctuations, right column shows fluctuations combined with population bottlenecks and genetic entrenchment (as Fig. 6).

the life-environment interface. ITSNTS argues that some cycle-variants out-persist others via an evolutionary synergy in which cycles and the niches of genotypes performing their constituent reactions mutually perpetuate each other. Our model thus represents simple limiting cases of this line of reasoning: a cycle-variant in which the two producer alleles interact can persist through an interval during which recycling is essential to survival, whereas a cycle-variant comprising the two non-producer alleles cannot. A non-zero cycling ratio cannot arise during a shut-off interval without the inter-species interaction leading to net regeneration of environmental substance, and the species cannot survive without this recycling: Because there are singers, there's a song, because there's a song, there are singers²³.

Relationship to multi-level selection. We have explicitly imposed a significant fitness penalty on the producer genotype, in order to represent the central neo-Darwinian concern with “Gaian” thinking: That the biological traits conducive to geochemical-climatic homeostasis and the persistence of habitability may not be favored by organism-level natural selection. In reality, several (but not all) of the links in closed recycling loops are energetically favorable¹⁴, therefore are not inherently associated with an individual-level growth penalty. In this sense, the cost we impose here represents a neo-Darwinian “worst-case-scenario” in that there is a significant inconsistency between the equilibrium solutions of evolutionary-population and planetary-climatic dynamics. Our core argument is that extreme environmental contexts (represented by the influx shut-off perturbations) create a survival-filter, such that only populations containing these biogeochemically relevant genotypes (represented by producers and the recycling they support) are capable of persistence. The existence of this filter compensates for the producer genotype's low relative fitness. We reiterate that

here we do not invoke any form of multi-level fecundity selection in favor of the producer allele: at no point do individual producers accrue any preferential benefit from the recycling that they support. In previous work³⁹ we have argued that nutrient recycling associated with an individual level fitness cost, may be stabilized by spatially structured “viscous” populations, which favor evolutionary “altruism” more broadly⁴⁰ but require restrictive assumptions about demographic structure. In contrast, the general mechanism suggested here could occur at any spatial scale.

Net genetic assimilation and entrenchment of biogeochemically important phenotypes. A key problem with ITSNTS is the lack of any analog to a genetic information pool at the cycle level, restricting the potential for cumulative evolutionary change due to the lack of sufficient variation for any such persistence-selection to act upon. To some extent this barrier can be surmounted via metabolic redundancy in terms of the number of species/genotypes capable of performing a given step in a cycle, combined with the idea that “everything is everywhere and the environment selects”⁴¹. More fundamentally however, the environmentally sensitive “plastic” genotype that we have sketched out here illustrates how the concept of genetic assimilation may, we argue, bridge this conceptual gap. The definition of genetic assimilation refers to the change between a situation in which a phenotype is initially produced only in response to an environmental stimulus, and a situation in which this phenotype is expressed constitutively^{42,43}. This is usually interpreted in terms of either epigenetics or (more relevantly here) natural selection for heritable variation in regulatory genes that modify the environmental sensitivity of the trait in question⁴⁴. Our general argument in this regard is that biologically driven closure (or not) of a given cycle begins as a

function of an environmentally sensitive trait (the plastic allele), but over evolutionary-geologic time mutates/co-evolves into a hard-wired phenotype expressed in all environmental contexts (the producer allele).

Because we pre-suppose, within our hypothesized evolutionary trajectory, that the ancestral difference between non-producer and producer phenotypes is an environmentally sensitive one, our model mutation rates should be equated to regulatory changes modifying the expression of pre-existing and coevolving phenotypes, rather than (much lower) microbial point mutation rates. The constraints imposed by our results upon the supply of variation by mutation rate are qualitative: During the nascent phase of the evolution of biogeochemical cycles, non-producers must give rise by mutation to sufficient producer offspring to allow at least some recycling to occur during an influx shut-off and prevent producers becoming so rare that the likelihood of their fixation by drift becomes vanishingly small. Subsequently, mutations in regulatory genes must cause the producer phenotype to evolve to become “hard-wired” rather than environmentally sensitive.

Distinction from the interpretation of habitability as an unselected by-product. The crux of the evidential distinction between the default feedbacks/by-products view and ITSNTS is, we argue, the existence of top-down correlations between biological and environmental patterns and their directional evolutionary change (species-averaged producer relative frequency and substance-averaged cycling ratio in this case). It is important to note that such correlations are not necessary consequences of external constraints on the model, but rather emerge most strongly under boundary conditions imposing persistence selection. Much the same can be said for the correlation between the individual cycling ratios of each environmental substance (Fig. 8). Thus, our model illustrates how natural selection by persistence alone, as invoked by ITSNTS is applicable to emergent environmental patterns irreducible to their biotic or abiotic constituents. In short, biogeochemical recycling does not “merely” emerge as a by-product of growth, it emerges, persists, and, due to this persistence, transitions to a “locked-in” state.

A key proviso is that our model and arguments rely on an inbuilt synergy between the physiology of two separate species (i.e. alleles in two species fortuitously assimilate the by-products of each other). Such reciprocity cannot result from persistence selection for recycling, because it is a pre-requisite for the operation of such selection. We nonetheless argue that such physiological complementarity is an obvious feature of real biogeochemical cycles at various scales, potentially down to generalized requirements for autocatalysis itself⁴⁵. Reciprocal self-stabilizing inter-species interactions also seems to readily emerge in other models^{46,47} and experimental systems^{48,49} of recycling, resource cross-feeding³⁵ and computational-genomic physiological trends³². Such reciprocity is reminiscent of the two-way integral rein control upon which concrete physiological applications of the principles of Daisyworld model have been based⁵⁰. It also converges with our own previous suggestions that such integral rein control is topologically similar to the physiological inter-dependence that characterizes symbiosis⁵¹. Thus, we argue that assuming that such reciprocal complementarity exists merely amounts to the extrapolation to the climatic-geochemical scale of what appears to be a generic ecological property. We hope that this work may motivate theoretical biologists to formulate the impact of such biogeochemical and ecological synergies in terms of the traditional population-genetic understanding of fitness.

Applicability to real biogeochemical history. It is uncontentious that biogeochemical cycles are both the product of a co-evolutionary history involving transition between distinct cycle-variants^{52,53} and exhibit emergent properties⁵⁴. Nevertheless, invoking persistence selection between cycle variants as a causal explanation for such properties is necessarily a more subtle and tentative move. Geochemical data from Earth’s early history is sparse, molecular phylogenetics is uncertain, and relaxed molecular clocks carry significant age uncertainties, making it challenging to make reliable inferences about the prevailing evolutionary ecology at the time at which real elemental cycles originated. Even defining such origin-events is non-trivial. Nevertheless, we argue that our model is representative of several features of the early Earth and the selection pressures that likely influenced life-environment coevolution within it.

Life is now estimated to have originated >3.9 Ga with crown clades of archaea and bacteria radiating in the Archean <3.4 Ga⁵⁵. Asteroid impacts were sufficient to boil the oceans and mix the crust up to ~4 Ga⁵⁶ and continued to be a sporadic, dangerous occurrence during the Archean⁵⁷. The time at which plate tectonics was established continues to be debated, but it was not until after ~3 Ga that familiar tectonics including a super-continent cycle arose. Beforehand, postulated ‘stagnant lid’ tectonics⁵⁸ would likely have lead to a more sporadic input of materials from the inner Earth to the surface⁵⁹. Furthermore, any early chemoautotrophic biosphere would have been limited by a free energy supply to be many orders of magnitude smaller than today⁶⁰. Such limitation would have persisted even after the origin of early forms of anoxygenic photosynthesis, especially in the absence of recycling of electron donors⁶¹. The conditions under which key anoxic biogeochemical cycles began to emerge were therefore plausibly subject to fluctuating boundary conditions and involved population sizes significantly smaller (thus more susceptible to genetic drift) than the modern biosphere.

The later biosphere became larger and more productive with the origin of oxygenic photosynthesis <3 Ga and the emergence of more elaborate anoxic-oxic biogeochemical cycles, but was still subject to extreme fluctuations, especially with the Great Oxidation ~2.4 Ga at least one ‘snowball Earth’ global glaciation⁶². Afterwards, some authors argue the Proterozoic biosphere remained significantly smaller than the modern one⁶³. Later still, fluctuating abiotic boundary conditions are suggested by global ocean redox-proxies to have been a consistent feature of the Neoproterozoic-Cambrian transition⁶⁴, along with an additional series of planetary-scale glaciations⁶⁵. Independent phylogenetic arguments from this time also support the general premise of elevated genetic drift^{38,66}.

These distinct lines of evidence converge to give a picture in which growth-permitting resources and nutrients were scarcely and erratically distributed in time and space. We argue that such extreme variability in environmental boundary conditions would have resulted in small populations subject to frequent and extreme bottlenecks. Thus, an evolutionary/ecological context in which producers reach fixation by drift is (we propose) one with qualitative relevance to early Earth history. We add that a lock-in event of the form of the one sketched out by this model need only occur once to persist robustly. Finally, it should be noted that if something along the lines of this model and its themes is relevant to the emergence of real biogeochemical cycles, a form of sequential selection⁶⁷ is required, involving multiple instances of “cycle-extinction” before robust recycling emerges, due to the unavoidably low probability of a cycle lock-in event. Although any such event need only occur once in order to persist, a hard limit upon the plausibility of any sequential selection process is that no “re-set” event is so severe as to permanently undermine subsequent habitability. The implication is that Earth’s biogeochemistry may be robust, but it is not un-breakable.

Methods

Model description. The model involves a discrete time, discrete valued stochastic Markov process. Model variables and parameters are given in Tables 1 and 2 respectively. Both time and the number of individuals of each type are constrained to be integer valued. Death and reproductive mutation are stochastic processes derived from sampling from binomial distributions given by the relevant probabilities. All ensemble results give the 100-replicate average for the parameter choices in question.

Growth of individuals from species S_1 and S_2 is proportional to the bio-available level of environmental substances R_1 and R_2 respectively. At time t (where time is in units of biological generations) the change in the number N_{qj} of individuals of genotype j (non-producer, producer, plastic) within species q (S_1 or S_2) can be written as a function of the state of the variables at the previous time-step:

$$N_{qj}(t+1) = \left((N_{qj}(t) - \rho_{qj}(t)) \cdot G_{qj}(t) - \Upsilon_{qj}(t) - \delta_{qj}(t) + \Upsilon_{q,x \neq j}(t) \right) \cdot \left(1 - \frac{S_q(t)}{K} \right) \quad (1)$$

The leftmost bracket on the right-hand side represents the number of individuals escaping starvation (death due to insufficient environmental substance) at the previous time-step and $G_{qj}(t)$ is the per capita reproductive growth rate. $\Upsilon_{qj}(t)$ gives the number of mutant offspring individuals produced during reproduction from parent individuals of genotype j . $\Upsilon_{q,x \neq j}(t)$ represents the number of j genotype individuals derived from mutation in parent individuals of other genotypes. $\delta_{qj}(t)$ is the number of individuals of genotype j lost to random death, and the rightmost bracket relates the total number $S_q(t)$ of individuals of species q to carrying capacity K , which represents limitation of growth by any factor other than the relevant environmental substance, e.g. space. (The steady state population size in all simulations shown is below K and limited by the environmental substance influx. The carrying capacity is included in the model for computational reasons and as a “crash preventer” but has no qualitative effect on the results).

The total number of individuals $S_q(t)$ in species q is the sum of the number of individuals of each genotype (producer, non-producer and plastic, as discussed in the main text):

$$S_q(t) = \sum_{j=1}^{j_{\text{total}}} N_{qj}(t) = N_{q,\text{prod}}(t) + N_{q,\text{non-prod}}(t) + N_{q,\text{plast}}(t) \quad (2)$$

The genotype-specific reproductive growth rate $G_{qj}(t)$ (again for genotype j within species q , time t), gives the number of offspring individuals produced per parent individual, per time-step. Growth rate is an increasing function of the bio-available level of environmental substance $R_{q,\text{BIOAVAILABLE}}$ (the subscript q being identical because species S_1 and S_2 assimilate substances R_1 and R_2 respectively). Growth rate also includes a substance-to-biomass conversion efficiency parameter f_{conv} and a genotype-specific per capita term $G_{qj,\text{PR}}$ (number of offspring per parent, per unit environmental substance assimilated, per unit time). In the absence of growth-limitation by environmental substance levels, growth rate is capped at a genotype-specific maximum $G_{qj,\text{MAX}}$:

$$G_{qj}(t) = \text{MIN}[G_{qj,\text{PR}} \cdot R_{q,\text{BIOAVAILABLE}}(t) \cdot f_{\text{conv}}, G_{qj,\text{MAX}}] \quad (3)$$

$$G_{q,\text{non-prod},\text{PR}} = G_0 \quad (4)$$

$$G_{q,\text{non-prod},\text{MAX}} = G_0 \cdot R_{\text{assimMAX}} \quad (5)$$

G_0 is the baseline number of offspring, per parent, per unit substance assimilated. R_{assimMAX} is a universal maximum potential number of units of environmental substance that can be assimilated by a single individual per time-step (i.e. representing basic physiological constraints on growth). The producer genotype incurs a per capita reproductive growth rate cost κ_{prod} relative to the non-producer:

$$G_{q,\text{prod},\text{PR}} = G_0 \cdot (1 - \kappa_{\text{prod}}) \quad (6)$$

$$G_{q,\text{prod},\text{MAX}} = G_0 \cdot (1 - \kappa_{\text{prod}}) \cdot R_{\text{assimMAX}} \quad (7)$$

This growth rate formulation is therefore a highly simplified linearization of the Michaelis-Menten kinetics normally used in models of resource and nutrient assimilation.

The plastic genotype switches phenotype depending upon the level of environmental substance relative to a fixed threshold $R_{q,\text{BIOAVAILABLE}_{\text{crit}}}$, in effect becoming a second non-producer genotype below this threshold and a second producer genotype above it:

$$\text{IF}[R_{q,\text{BIOAVAILABLE}}(t) \geq R_{q,\text{BIOAVAILABLE}_{\text{crit}}}], G_{q,\text{plast}}(t) = G_{q,\text{prod}}(t)$$

$$\text{ELSEIF}[R_{q,\text{BIOAVAILABLE}}(t) < R_{q,\text{BIOAVAILABLE}_{\text{crit}}}], G_{q,\text{plast}}(t) = G_{q,\text{non-prod}}(t) \quad (8)$$

There is no spatial structure whatsoever, thus access to environmental substance is uniform across individuals. The bioavailable quantity of each environmental substance is simply the total amount $R_{q,\text{NET}}(t)$ divided by the total number of

individuals assimilating it:

$$R_{q,\text{BIOAVAILABLE}}(t) = \frac{R_{q,\text{NET}}(t)}{S_q(t)} \quad (9)$$

We allow the per capita reproductive growth rate to fall below $G_{qj}(t) = 1$, which, if interpreted deterministically at the individual level would correspond to an individual failing to sustain its biomass to the next time-step and thus dying. However, a population-level average $G_{qj}(t) < 1$ is interpretable in terms of a thinning factor that maps between discretized individuals and continuously distributed environmental substance. Thus, a thinning factor of $(1 - G_{qj}(t))$ is used to calculate the total number of individuals dying of starvation ρ_{qj} (again genotype j , species q). This represents pre-reproduction deaths, corresponding to the difference between the actual population size and the population size that the environmental substance pool is capable of supporting. $\rho_{qj}(t)$ is constrained to be an integer and is zero for $G_{qj}(t) > 1$:

$$\rho_{qj}(t) = N_{qj}(t) \cdot \text{MAX}\left[0, \left(1 - G_{qj}(t)\right)\right] \quad (10)$$

A subset of offspring are a different genotype from their parent via mutation. For parent genotype j , the number $\Upsilon_{qj}(t)$ of mutant offspring with genotype $\neq j$ is calculated using baseline mutation probability per reproductive event μ_0 , with the total number of new individuals produced by the parent individuals surviving starvation $G_{qj}(t) \cdot (N_{qj}(t) - \rho_{qj}(t))$. The total number of mutant offspring $\Upsilon_{qj}(t)$ is thus a binomially distributed random variable with success probability μ_0 and number of trials $G_{qj}(t) \cdot (N_{qj}(t) - \rho_{qj}(t))$. The expected value $E[\Upsilon_{qj}(t)]$ is the product of these two numbers:

$$\Upsilon_{qj}(t) \sim B\left(G_{qj}(t) \cdot (N_{qj}(t) - \rho_{qj}(t)), \mu_0\right), \\ E[\Upsilon_{qj}(t)] = G_{qj}(t) \cdot (N_{qj}(t) - \rho_{qj}(t)) \cdot \mu_0 \quad (11)$$

Mutation to genotype j from the other genotypes is calculated in exactly the same way using the number and reproductive growth rates of the relevant (other) genotypes. Any particular mutant offspring is randomly allocated to one of the other genotypes with equal probability $p_{k \rightarrow j} = \frac{1}{j_{\text{total}} - 1} = 0.5$ (where $j_{\text{total}} = 3$ is the total number of genotypes per species). The expected number of offspring with genotype j produced by mutation within parent offspring of other genotypes $k \neq j$ is therefore:

$$E[\Upsilon_{q,x \neq j}(t)] = \left(\sum_{k \neq j}^{k_{\text{total}}} \Upsilon_{q,k}(t) \right) \cdot \frac{1}{j_{\text{total}} - 1} \quad (12)$$

Independently of reproduction and assimilation of environmental substance, any given individual has a probability δ_0 at each time point of death due to stochastic factors. The genotype/species specific number of such deaths is again a random sample from a binomial distribution, with success probability δ_0 :

$$\delta_{qj}(t) \sim B\left(N_{qj}(t), \delta_0\right), E[\delta_{qj}(t)] = N_{qj}(t) \cdot \delta_0 \quad (13)$$

The net quantity of growth-limiting environmental substance at each time-step is given by the difference between total biotic assimilation A_{R_q} and the production P_{R_q} and abiotic input φ_{R_q} fluxes:

$$R_{q,\text{NET}}(t+1) = \varphi_{R_q}(t) + P_{R_q}(t) - A_{R_q}(t) \quad (14)$$

The abiotic net influx is the sum of two fluxes. First, an input term that is the product of a baseline scaling factor φ_{0,R_q} and a model forcing $\frac{\partial t_{\text{geo}}}{\partial t_{\text{bio}}}$ representing the mapping between abiotic-geological and biotic-evolutionary timescales. In practice $\frac{\partial t_{\text{geo}}}{\partial t_{\text{bio}}}(t)$ was set to either 1 or 0 or (in fluctuation runs) a time-dependent switching between the two. (More sophisticated implementations of $\frac{\partial t_{\text{geo}}}{\partial t_{\text{bio}}}(t)$, e.g. sinusoidal oscillations and stochastic time dependence, were attempted but made little qualitative difference to the results). Second, an abiotic removal term that scales linearly with the quantity of environmental substance:

$$\varphi_{R_q}(t) = \varphi_{0,R_q} \cdot \frac{\partial t_{\text{geo}}}{\partial t_{\text{bio}}}(t) - \frac{R_{q,\text{NET}}(t)}{R_{q,\text{NET0}}} \quad (15)$$

where $R_{q,\text{NET0}}$ is a normalization factor representing the sensitivity of the abiotic efflux to the influx. In the absence of any biota and for $\frac{\partial t_{\text{geo}}}{\partial t_{\text{bio}}} = 1$ the steady state environmental substance level is immediately given by (15) as $R_{q,\text{NET}}(t) = R_{q,\text{NET0}} \cdot \varphi_{0,R_q}$, thus the numerical value of $R_{q,\text{NET0}}$ corresponds to the abiotic steady state residence time.

Total biotic assimilation A_R of each environmental substance is given by:

$$A_{R_q}(t) = \sum_{k=1}^{j_{\text{total}}} \frac{G_{q,k}(t) \cdot (N_{q,k}(t) - \rho_{q,k}(t))}{G_{q,k,\text{PR}}} \quad (16)$$

The numerator gives the total number of individuals produced as a result of biological assimilation of environmental substance R_q and the denominator is the

genotype specific number of individuals produced per unit substance assimilated, dividing through by which therefore converts to total units of substance assimilated by the population as a whole.

Net biotic $P_{q,NET}$ production of substance R_q by the producer genotype in the other species p is calculated equivalently, via the product of the per capita production rate $P_{q,prod}$ and the total number of reproducing individuals:

$$P_{q,prod}(t) = \frac{\min[G_{q,prod}(t) \cdot f_{conv,prod}, G_{q,prodMAX}]}{G_{p,PRODUCERPR}} \quad (17)$$

$$P_{q,NET}(t) = P_{q,prod}(t) \cdot (N_{p,PRODUCER}(t) - \rho_{p,PRODUCER}(t)) \quad (18)$$

Where $f_{conv,PROD}$ is the per capita efficiency by which producers convert the environmental substance that they assimilate into the by-product they produce (note that the equivalent conversion efficiency for assimilation f_{conv} already appears in the growth functions of each genotype, therefore does not appear in Eq. (17)).

The residence time T_{R_q} of each environmental substance is given by the net quantity of this substance divided by the influx, to give units of the average number of biological generations a unit of environmental substance spends in the relevant pool before being removed. In those simulations in which the abiotic influx $\varphi_{R_q}(t)$ was set to zero (i.e. during the shut-off intervals) production $P_{q,NET}(t)$ was used as an alternative denominator:

$$IF[\varphi_{R_q}(t) > 0, T_{R_q}(t) = \frac{R_{q,NET}(t)}{\varphi_{R_q}(t)}] IF\left[\left((\varphi_{R_q}(t) = 0) \& (P_{q,NET}(t) > 0)\right), T_{R_q}(t) = \frac{R_{q,NET}(t)}{P_{q,NET}(t)}\right] ELSE[T_{R_q}(t) = 0] \quad (19)$$

The cycling ratio CR_{R_q} of each substance is given by the ratio between net biotic assimilation of that substance $A_{R_q}(t)$ and the abiotic influx of that substance $\varphi_{R_q}(t)$. As with the residence time, when the abiotic influx was zero, the input from biological production was used as an alternative denominator:

$$IF[\varphi_{R_q}(t) > 0, CR_{R_q}(t) = \frac{A_{R_q}(t)}{\varphi_{R_q}(t)}] IF\left[\left((\varphi_{R_q}(t) = 0) \& (P_{q,NET}(t) > 0)\right), CR_{R_q}(t) = \frac{A_{R_q}(t)}{P_{q,NET}(t)}\right] ELSE[CR_{R_q}(t) = 0] \quad (20)$$

Deterministic approximation to steady state. Assume that at steady state substance assimilation will reach a maximal state such that the level of environmental substance is limiting to population size. Assume that such a state is below the level $R_{q,BIOAVAILABLE_crit}$ at which the plastic genotype effectively becomes a second non-producer genotype and can thus be subsumed into non-producer frequency, such that (2) becomes $S_q(t) = \sum_{j=1}^{j_{total}} N_{q,j}(t) = N_{q,prod}(t) + N_{q,non-prod}(t)$. Assume that there are non-zero starvations at each time-step for all genotypes, which implies growth rate $G_{q,j}(t) < 1, \ll G_{q,jMAX}, \forall j, q$, which gives by (3) $G_{q,j}(t) = G_{q,jPR} \cdot R_{q,BIOAVAILABLE}(t) \cdot f_{conv}$. Substituting this into (10), then the first bracketed term in (1), then labeling the post-starvation number of individuals as $(N_{q,j}(t))_{NET}$:

$$(N_{q,j}(t))_{NET} = (N_{q,j}(t) - \rho_{q,j}(t)) = N_{q,j}(t) \cdot \left(1 - \left(1 - G_{q,j}(t)\right)\right) = N_{q,j}(t) \cdot G_{q,j}(t) \quad (21)$$

Approximate (14) deterministically by a fixed fractional parameter corresponding to the baseline random death rate:

$$\delta_{q,j}(t) \approx N_{q,j}(t) \cdot \delta_0 \quad (22)$$

Doing the same for mutation:

$$\Upsilon_{q,j}(t) = G_{q,j}(t) \cdot (N_{q,j}(t))_{NET} \cdot \mu_0 = N_{q,j}(t) \cdot G_{q,j}(t)^2 \cdot \mu_0 \quad (23)$$

The term in (12) for mutation to j from other genotypes simplifies to

$$\Upsilon_{q,x \neq j}(t) = \sum_{k=1}^{j_{total}} \frac{G_{q,k}(t) \cdot (N_{q,k}(t))_{NET} \cdot \mu_0}{j_{total} - 1} = N_{q,k}(t) \cdot G_{q,k}(t)^2 \cdot \mu_0$$

Substituting Eqs. (21–23) into (1):

$$N_{q,j}(t) \cdot G_{q,j}(t)^2 \cdot (1 - \mu_0) - N_{q,j}(t) \cdot \delta_0 + N_{q,k}(t) \cdot G_{q,k}(t)^2 \cdot \mu_0 = 0 \quad (24)$$

Noting that by Eqs. (3)–(5) combined with the above assumptions, the growth rate of the producer can be written as:

$$G_{q,prod}(t) = G_{q,cheat}(t) \cdot (1 - \kappa_{prod,q}) \quad (25)$$

Writing (24) explicitly for each genotype:

$$N_{q,non-prod}(t) \cdot G_{q,non-prod}(t)^2 \cdot (1 - \mu_0) - N_{q,non-prod}(t) \cdot \delta_0 + N_{q,prod}(t) \cdot G_{q,prod}(t)^2 \cdot \mu_0 = 0$$

$$N_{q,prod}(t) \cdot G_{q,prod}(t)^2 \cdot (1 - \mu_0) - N_{q,prod}(t) \cdot \delta_0 + N_{q,non-prod}(t) \cdot G_{q,non-prod}(t)^2 \cdot \mu_0 = 0$$

Adding:

$$N_{q,non-prod}(t) \cdot G_{q,non-prod}(t)^2 \cdot (1 - \mu_0) - N_{q,non-prod}(t) \cdot \delta_0 + N_{q,prod}(t) \cdot (G_{q,non-prod}(t) \cdot (1 - \kappa_{prod,q}))^2 \cdot \mu_0 + N_{q,prod}(t) \cdot (G_{q,non-prod}(t) \cdot (1 - \kappa_{prod,q}))^2 \cdot (1 - \mu_0)$$

$$- N_{q,prod}(t) \cdot \delta_0 + N_{q,non-prod}(t) \cdot G_{q,non-prod}(t)^2 \cdot \mu_0 = 0$$

Because the mutation terms cancel:

$$(N_{q,non-prod}(t) + N_{q,prod}(t) \cdot (1 - \kappa_{prod,q})^2) \cdot (G_{q,non-prod}(t)^2 - \delta_0) = 0 \quad (26)$$

Substituting in for the growth rate terms (3–7) gives:

$$(N_{q,non-prod}(t) + N_{q,prod}(t) \cdot (1 - \kappa_{prod,q})^2) \cdot ((G_0 \cdot R_{q,BIOAVAILABLE}(t) \cdot f_{conv})^2 - \delta_0) = 0 \quad (27)$$

By (16), (4), (6) and the above, total steady state assimilation of the growth limiting environmental substance by species q is:

$$\begin{aligned} A_{R_q}(t) &= \sum_{k=1}^{j_{total}} \frac{G_{q,k}(t) \cdot (N_{q,k}(t) - \rho_{q,k}(t))}{G_{q,kPR}} \\ &= \frac{N_{q,non-prod}(t) \cdot (G_{q,non-prod}(t))^2}{G_0} + \frac{(N_{q,prod}(t) \cdot (1 - \kappa_{prod,q})^2) \cdot (G_{q,non-prod}(t))^2}{G_0 \cdot (1 - \kappa_{prod,q})} \\ &= (N_{q,non-prod}(t) + N_{q,prod}(t) \cdot (1 - \kappa_{prod,q})) \cdot G_0 \cdot (R_{q,BIOAVAILABLE}(t) \cdot f_{conv})^2 \end{aligned} \quad (28)$$

By (16–18), production of this substance by the producer allele in the other species $p \neq q$, assuming the various arguments above simultaneously apply to this species, is:

$$\begin{aligned} P_{R_q}(t) &= \frac{G_{p,prod}(t) \cdot (N_{p,prod}(t) - \rho_{p,PRODUCER}(t))}{G_{p,prodPR}} \cdot f_{conv,PROD} \\ &= N_{p,prod}(t) \cdot (1 - \kappa_{prod,p}) \cdot G_0 \cdot (R_{p,BIOAVAILABLE}(t) \cdot f_{conv})^2 \cdot f_{conv,PROD} \end{aligned} \quad (29)$$

Balance between input and output fluxes of each environmental substance requires $\varphi_{R_q}(t) + P_{R_q}(t) = A_{R_q}(t)$, meaning that by substituting in $A_{R_q}(t)$ from (28) it is possible to solve for bioavailable substance level, then substitute in the production flux of this substance derived from the producer allele in the other species $p \neq q$:

$$\begin{aligned} R_{q,BIOAVAILABLE}(t) &= \frac{1}{f_{conv}} \sqrt{\frac{\varphi_{R_q}(t) + P_{R_q}(t)}{(N_{q,non-prod}(t) + N_{q,prod}(t) \cdot (1 - \kappa_{prod,q})) \cdot G_0}} \\ &= \frac{1}{f_{conv}} \sqrt{\frac{\varphi_{R_q}(t) + N_{p,prod}(t) \cdot (1 - \kappa_{prod,p}) \cdot G_0 \cdot (R_{p,BIOAVAILABLE}(t) \cdot f_{conv})^2 \cdot f_{conv,PROD}}{(N_{q,non-prod}(t) + N_{q,prod}(t) \cdot (1 - \kappa_{prod,q})) \cdot G_0}} \end{aligned} \quad (30)$$

Substituting this into (27) gives a symmetrical condition for steady state genotype frequencies and substance levels across the system:

$$\begin{aligned} &(N_{q,non-prod}(t) + N_{q,prod}(t) \cdot (1 - \kappa_{prod,q})^2) \cdot \\ &\left(\frac{(\varphi_{R_q}(t) + N_{p,prod}(t) \cdot (1 - \kappa_{prod,p}) \cdot G_0 \cdot (R_{p,BIOAVAILABLE}(t) \cdot f_{conv})^2 \cdot f_{conv,PROD})}{(N_{q,non-prod}(t) + N_{q,prod}(t) \cdot (1 - \kappa_{prod,q})) \cdot G_0} - \delta_0 \right) = 0 \end{aligned} \quad (31)$$

This solution illustrates the intuitive ideas that growth and reproduction balance random death at steady state and that the associated producer frequency is lower than that of the non-producer by a factor of the cost. (This factor is of second order because the growth rate is used both directly and (by (10)) in the calculation of starvations). Because our model is a discrete stochastic process, (31) can be viewed as an approximation to a steady state condition, subject to the above assumptions combined with the continuous generation of producers by mutation at a sufficient rate to preclude their extinction. The key point is that over long timescales in the finite populations with which we deal, organism-level selection unavoidably favors the non-producer, with no possibility for multi-level fecundity selection. The producer's stable presence is thus attributable to the combination of mutation and cycle-level selection.

Data availability

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Code availability

Matlab code for the model is available at: <https://github.com/richboyle111/TTSNTS>.

Received: 11 June 2021; Accepted: 26 January 2022;

Published online: 04 March 2022

References

- Lovelock, J. E. M. Atmospheric homeostasis by and for the biosphere: the Gaia hypothesis. *Tellus* **26**, 2–9 (1974).
- Doolittle, W. F. *Is nature really motherly?* The coevolution quarterly, Spring 58–63. (1981).
- Tyrell, T. *On Gaia: A critical investigation of the relationship between life and Earth* (Princeton, 2013).
- Lenton, T. M. et al. Selection for Gaia across multiple scales. *Trends Ecol. Evol.* **23**, 633–645 (2018).
- Maynard-Smith, J. *The problems of biology* (Oxford University Press, 1986).
- Gildenhuys, P. *Natural Selection*. The Stanford Encyclopedia of Philosophy, Winter 2019 Edition. (2019).
- Saunders, P. T. Evolution without natural selection: further implications of the Daisyworld parable. *J. Theor. Biol.* **166**, 365–373 (1994).
- Volk, T. *Gaia's body: Towards a physiology of the Earth*. (MIT Press, 2003).
- Postgate, J. Gaia gets too big for her boots. *New Scientist* **118**, 60 (1988). **1607**.
- Lenton, T. M. & Lovelock, J. E. Daisyworld is Darwinian: constraints on adaptation are important for planetary self-regulation. *J. Theor. Biol.* **206**, 109–114 (2000).
- Watson, A. J., Coevolution of the Earth's environment and life: Goldilocks, Gaia and the anthropic principle. In: *James Hutton - Present and Future*. (ed. Craig, J. H.) (Geological Society, London, Special Publications, 1999).
- Watson, A. J. Implications of an anthropic model of evolution for emergence of complex life and intelligence. *Astrobiology* **8**, 175–185 (2008).
- Wayne, R. P. *Chemistry of Atmospheres*. 2nd ed. (Oxford University Press, 1991).
- Lenton, T. M. Gaia and natural selection. *Nature* **394**, 439–447 (1998).
- Leggett, L. M. W. & Ball, D. A. Observational evidence that a feedback control system with proportional-integral-derivative characteristics is operating on atmospheric surface temperature at global scale. *Tellus A Dyn. Meteorol. Oceanogr.* **72**, 1–14 (2020).
- Borrett, S. R., Carter, M. & D. E. Hines, D. E. Six general ecosystem properties are more intense in biogeochemical cycling networks than food webs. *J. Complex Netw.* **4**, 575–603 (2016).
- Volk, T. Gaia is life in a wasteworld of by-products. In *Scientists Debate Gaia: The Next Century*. 26–35. (MIT Press, 2004).
- Lenton, T. M., Dutreil, S. & Latour, B. Life on Earth is hard to spot. *Anthropocene Rev.* **7**, 248–272 (2020).
- Schlesinger, W. *Biogeochemistry: an analysis of global change*. (Boston Academic Press, 1997).
- Free, A. & Barton, N. H. Does evolution and ecology need the Gaia hypothesis? *Trends Ecol. Evol.* **22**, 611–619 (2007).
- Doolittle, W. F. & Inkpen, S. A. Processes and patterns of interaction as units of selection: an introduction to ITSNTS thinking. *Proc. Natl Acad Sci USA* **115**, 4006–4014 (2018).
- Doolittle, W. F. Making the most of clade selection. *Philos. Sci.* **84**, 275–295 (2016).
- Doolittle, W. F. Darwinizing Gaia. *J. Theor. Biol.* **434**, 11–19 (2017).
- Doolittle, W. F. Making evolutionary sense of Gaia. *Trends Ecol. Evol.* **24**, 889–894 (2019).
- Bouchard, F. Causal processes, fitness, and the differential persistence of lineages. *Philosoph. Sci.* **75**, 560–570 (2008).
- Bourrat, P. How to read heritability in the “recipe” approach to natural selection. *Brit. J. Phil. Sci.* **66**, 883–903 (2015).
- Dussault, A. C. & Bouchard, F. A persistence enhancing propensity account to explain ecosystem evolution. *Synthese* **194**, 1115–11145 (2017).
- Ibanez, S. The Evolution of ecosystem phenotypes. *Biol. Theory* **15**, 91–106 (2020).
- Laland, K., Matthews, B. & Feldman, M. W. An introduction to niche construction theory. *Evol. Ecol.* **30**, 191–202 (2016).
- Bouchard, F. Darwinism without populations: a more inclusive understanding of “Survival of the fittest”. *Studies in History and Philosophy of Biology and Biomedical Sciences* **42**, 106–114 (2010).
- Toman, J. & Flegr, J. Stability-based sorting: the forgotten process behind (not only) biological evolution. *J. Theor. Biol.* **435**, 29–41 (2017).
- Libby, E. et al. Syntrophy emerges spontaneously in complex metabolic systems. *PLoS Comput. Biol.* **15**, e1007169 (2019). <https://doi.org/10.1371/journal.pcbi.1007169>.
- Frank, S. A. The origin of synergistic symbiosis. *J. Theor. Biol.* **176**, 403–410 (1995).
- Okasha, S. *Evolution and the levels of selection*. (Oxford University Press, 2006).
- Butler, S. & James, P. O. 'Dwyer Co-operation and stability for complex systems in resource-limited environments. *Theor. Ecol.* **13**, 239–250 (2020).
- Volk, T. Gaia is life in a wasteworld of byproducts Scientists debate Gaia: the next century. 26–35. (MIT Press, 2004).
- Frank, S. A. *Foundations of social evolution* (Princeton University Press, 1998).
- Zhang, H. et al. Snowball Earths, population bottleneck and *Prochlorococcus* evolution. *Proc. Roy. Soc. B* **288**, 20211956 (2021).
- Boyle, R. A., Williams, H. T. P. & Lenton, T. M. Natural selection for costly nutrient recycling in simulated microbial meta-communities. *J. Theor. Biol.* **312**, 1–12 (2012).
- Schonmann, R. H., Vincente, R. & Caticha, N. Altruism can proliferate through population viscosity despite high random gene flow. *PLoS ONE* **8**, e72043 (2013).
- O'Malley, M. Everything is everywhere but the environment selects: ubiquitous distribution and environmental determinism in environmental biogeography. *Philos. Biol. Biomed. Sci.* **39**, 314–325 (2008).
- Pigliucci, M., Murren, C. J. & Schlichting, T. J. Phenotypic plasticity and evolution by genetic assimilation. *J. Exp. Biol.* **209**, 2362–2367 (2006).
- Ehrenreich, I. M. & Pfennig, D. W. Genetic assimilation: a review of its potential proximate causes and evolutionary consequences. *Ann. Bot.* **117**, 769–779 (2016).
- Nishikawa, K. & Kinjo, A. R. Mechanism of evolution by genetic assimilation: Equivalence and independence of genetic mutation and epigenetic modulation in phenotypic expression. *Biophys. Rev.* **10**, 667–676 (2018).
- Preiner, M. et al. Catalysis, autocatalysis and the origin of metabolism. *R. Soc. Interf. Focus* **9**, 20190072 (2019).
- Downing, K. & Zvirinsky, P. The simulated evolution of biochemical guilds: reconciling Gaia theory and natural selection. *Artif. Life* **5**, 291–398 (1999).
- Crombach, A. & Hogeweg, P. Evolution of resource recycling in ecosystems and individuals. *BMC Evol. Biol.* **9**, 122 (2009). <https://doi.org/10.1186/1471-2148-9-122>.
- Germerodt, S. et al. Pervasive selection for cooperative cross-feeding in bacterial communities. *PLoS Comput. Biol.* <https://doi.org/10.1371/journal.pcbi.1004986>. (2016).
- Preussger, D. et al. Reciprocal fitness feedbacks promote the evolution of mutualistic cooperation. *Curr. Biol.* **30**, 1–11 (2020).
- Saunders, P. T., Koeslag, J. H. & Wessels, J. A. Integral rein control in physiology. *J. Theor. Biol.* **194**, 163–173 (1998).
- Boyle, R. A., Lenton, T. M. & Watson, A. J. Symbiotic physiology promotes homeostasis in Daisyworld. *J. Theor. Biol.* **274**, 170–172 (2012).
- Reinhard, C. T. et al. Evolution of the global phosphorus cycle. *Nature* **541**, 386–389 (2016).
- Stueken, E. E. et al. The evolution of Earth's biogeochemical nitrogen cycle. *Earth Sci. Rev.* **160**, 220–239 (2016).
- Follows, M. J. et al. Emergent biogeography of microbial communities in a model. *Ocean. Sci.* **315**, 1843–1846 (2007).
- Betts, H. C. et al. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nat. Ecol. Evol.* **2**, 1556–1562 (2018).
- Marchi, S. et al. Widespread mixing and burial of Earth's Hadean crust by asteroid impacts. *Nature* **511**, 578–582 (2014).
- Davatzes, A., Goderis, S. and Simonson, B. M. Archean asteroid impacts on Earth: stratigraphic and isotopic age correlations and environmental consequences. In: *Earth's Oldest Rocks* (van Kranendonk, M. J., Bennett, V. C. and Hofman, J. E. eds.) 169–185. (Elsevier, 2019).
- Debaille, V. et al. Stagnant-lid tectonics in early Earth revealed by Nd variations in late Archean rocks. *Earth Planet. Sci. Lett.* **373**, 83–92 (2013).
- Lowell, R. P. & Keller, S. M. High-temperature seafloor hydrothermal circulation over geologic time and archean banded iron formations. *Geophys. Res. Lett.* **30**, 1391 (2003).
- Canfield, D. E., Rosing, M. T. & Bjerrum, C. Early anaerobic metabolisms. *Philos. Trans. R. Soc. B* **361**, 1819–1834 (2006).
- Kharecha, P., Kasting, J. & Siefert, J. A coupled atmosphere–ecosystem model of the early Archean Earth. *Geobiology* **3**, 53–76 (2005).
- Warke, M. R. et al. The great oxidation event preceded a Paleoproterozoic “snowball Earth”. *Proc. Natl Acad. Sci. USA* **117**, 13314–13320 (2020).
- Laakso, T. A. & Schrag, D. P. A small marine biosphere in the Proterozoic. *Geobiology* **17**, 161–171 (2018).
- Och, L. M. & Shields, G. A. The Neoproterozoic oxygenation event: environmental perturbations and biogeochemical cycling. *Earth Sci. Rev.* **110**, 26–57 (2012).
- Hoffman, P. F. et al. Snowball Earth climate dynamics and Cryogenian geology–geobiology. *Sci. Adv.* **3**, e1600983 (2017).
- del Cortona, A. et al. Neoproterozoic origin and multiple transitions to macroscopic growth in green seaweeds. *PNAS* **117**, 2551–2559 (2020).
- Lenton, T. M. & Wilkinson, D. M. Developing the Gaia theory: a response to the criticisms of Kirchner and Volk. *Climatic Change* **58**, 1–12 (2003).

Acknowledgements

This work was supported by the NERC-NSFC 'Biosphere Evolution, Transitions and Resilience' programme NERC (NE/P013651/1). We thank W. Ford Doolittle for stimulating our thinking, as well as all members of the Evolutionary Gaia Group <https://www.evolutionarygaia.org> for numerous useful discussions pertaining to this subject area.

Author contributions

R.B. designed and performed research, R.B. and T.L. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43247-022-00371-3>.

Correspondence and requests for materials should be addressed to Richard A. Boyle.

Peer review information *Communications Earth & Environment* thanks Frédéric Bouchard and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Joe Aslin, Heike Langenberg. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022