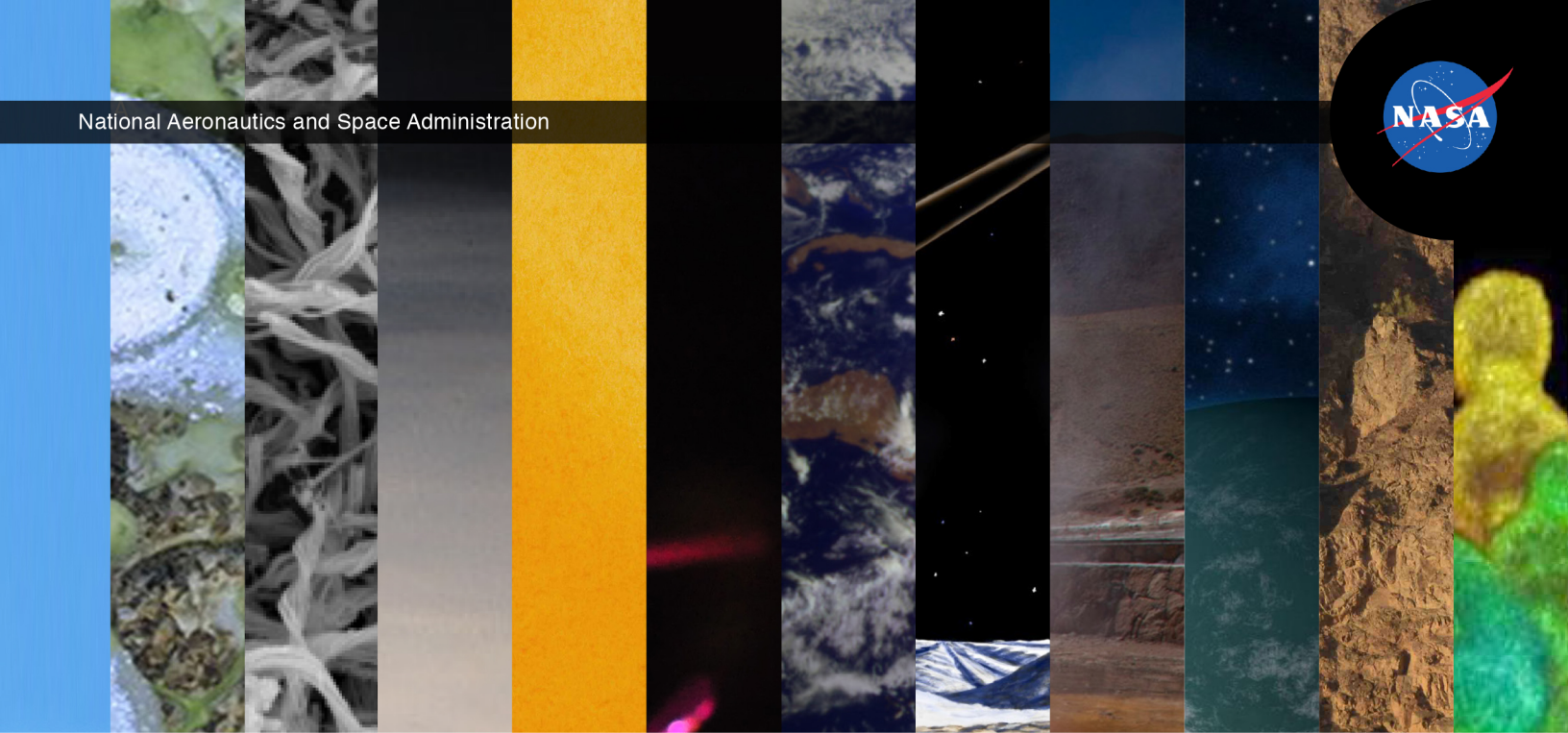
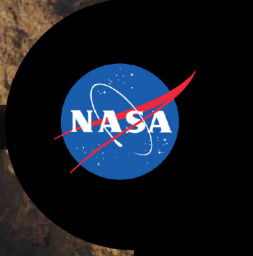


National Aeronautics and Space Administration

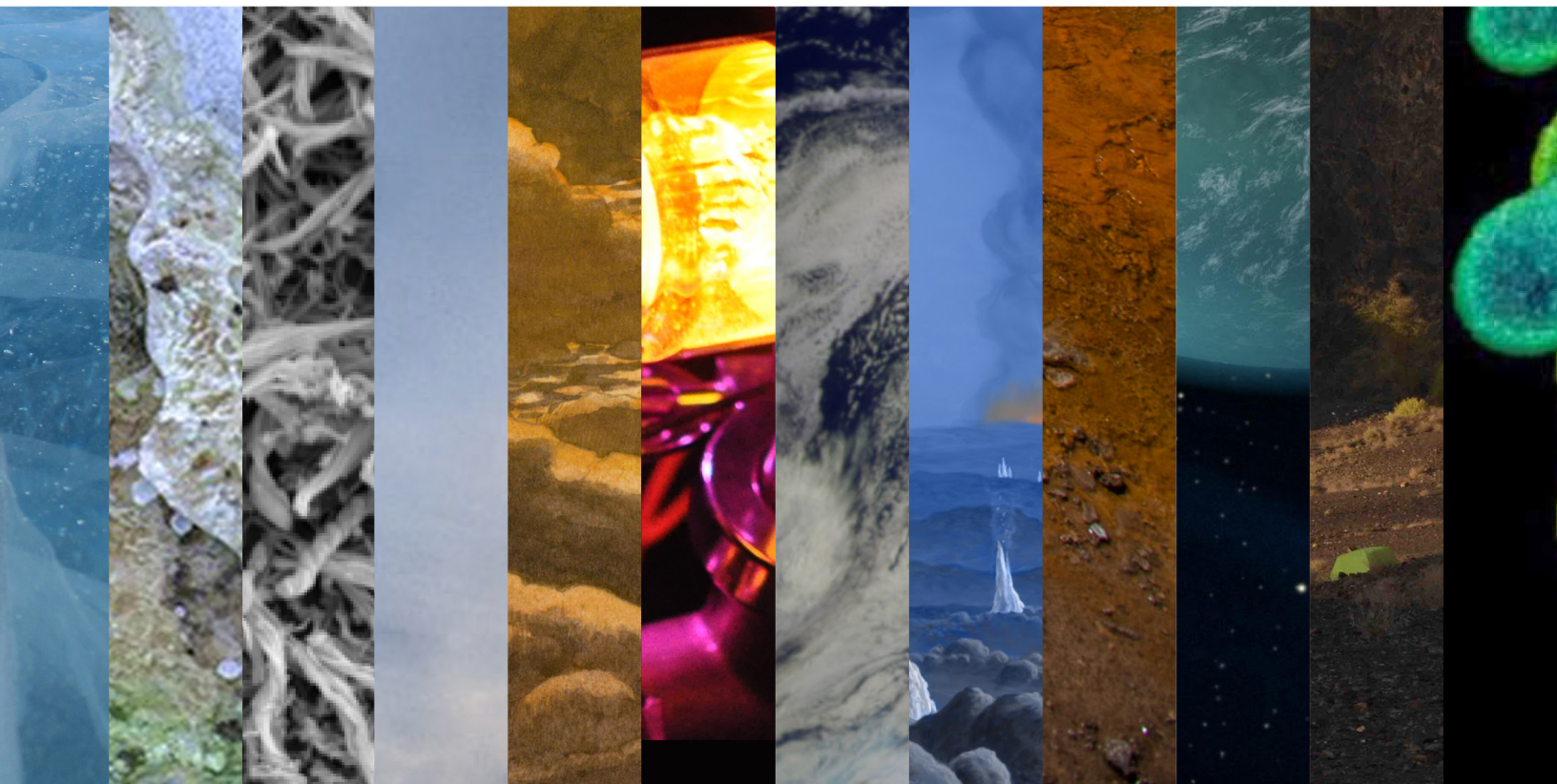


NASA ASTROBIOLOGY INSTITUTE ✨ ✨ ✨

2017 Annual Science Report

Institute for Universal Biology

University of Illinois at Urbana-Champaign





INSTITUTE FOR UNIVERSAL BIOLOGY

NASA ASTROBIOLOGY INSTITUTE

Lead Institution:

University of Illinois at Urbana-Champaign



Team Overview

"How does life begin and evolve?" This simple question, a fundamental topic within astrobiology, still today remains an enigma. The Towards Universal Biology Team is exploring how molecules come to life by investigating fundamental principles of biology from a multidisciplinary perspective – encompassing microbiology, geobiology, computational chemistry, genomics and even physics. A series of four innovative and cross-disciplinary research themes have been established by this Team in order to define their research program:

- Theoretical understanding of the universal features governing living systems, their operation, evolution and origin;
- Constraints on the nature of life before the Last Universal Common Ancestor (LUCA), in particular to identify new signatures of the collective state of life ("progenote") which enabled the evolution of the cell to occur so rapidly. In this early phase of life, believed to exist before about 3.8 billion years ago, there was neither a "family tree" of life nor species in the modern sense of the word.
- To explore the breakdown of the progenote and the transition to the current epoch of "vertical" evolution;
- Explore the interplay between biological and environmental determinants of the rate of evolution. How do environmental fluctuations influence the nature and speed of evolution?



Principal Investigator:
Nigel Goldenfeld

2017 Executive Summary

Key aspects of astrobiology that remain unresolved and poorly constrained are the location and manner of life's beginnings on Earth, its subsequent evolutionary transitions, and the most robust generic biosignatures of life.

The Illinois Team has strongly advocated the alkaline hydrothermal vent (AHV) theory for the origin of life, because this location is the only one known in which the puzzling direction of the proton motive force across extant cell walls can be naturally explained. The AHV theory is also of great current interest because of the strong circumstantial evidence from the Cassini space problem that the ice-lidded worlds Europa and Enceladus harbor the conditions that would support tectonics and vents. In the latter case, the chemical constituents of the fountain of Enceladus show strong evidence for hydrogen disequilibria, exactly as required for AHV life on Earth.

We have shown that the right way to interpret the energy flow powering cells is that they are processes mediated by macromolecular "mechanochemical" machines acting as dual-process 'stochastic escapement' engines – in which a successful downhill fluctuation in one reaction is made to be contingent on the quasi-coincident occurrence of an uphill fluctuation in another, where both up and down fluctuations are thermally induced. In this way, although no energy passes between the processes, if one is thermodynamically stronger than the other it then proceeds in the net downhill, and work is done via the other process being stochastically driven uphill by the relative irreversibility of the downhill process. The essential conceptual point is that what is taking place is the conversion of thermodynamic disequilibria, creating one by dissipating another. This mechanism has recently been understood to be instantiated in living cells through electron bifurcation, and our work provides a model framework in which to view the cell as a non-equilibrium thermodynamic engine.

Whatever the initial location and mechanism of biogenesis, it had to develop early on biological homochirality – the single-handedness of proteins and sugars – in order to build complex biochemical functional groups. Our stochastic mechanism for the non-equilibrium driving force behind homochirality requires the amplification of copy number fluctuations, and we recently showed how this arises through a novel pattern-formation mechanism that we were the first to identify. A strong implication of our work is that biological homochirality is an excellent biosignature, with no false positive, one especially useful for in situ measurements such as those to be performed on the Mars 2020 missions.



Fig. 1. Urbana High School students taking part in NASA Scholar-Athlete Astrobiology Workshop instructed by Dr. Bruce Fouke, University of Illinois Urbana-Champaign, Carl R. Woese Institute for Genomic Biology. Image Source: Dr. Bruce Fouke



Fig. 2. Filamentous mats of Sulfurihydrogenibium. Dr. Bruce Fouke and Dr. Charles Werth

In order for life to be detectable, it must evolve and develop planetary-wide ecologies. The speed of evolution and its dependence on environment has been tested in several laboratory settings, ranging from microfluidic devices for bacteria to live-cell imaging of transposon activity. We have studied theoretically the interplay between viruses and bacteria, highly plausible candidates for life on other worlds, and shown how co-evolutionary dynamics leads to both collective states that improve both bacteria and viruses, as well as leading to stabilized and highly communities with a well-characterized signature known as the “Kill the Winner” dynamics.

Outreach Nugget: Reinaldo Alcalde had the opportunity to teach a week-long workshop in Merida, MX in August 2017 through a program called Clubes de Ciencia. He designed a course entitled *Extremophiles: A Journey from Life Origins to Astrobiology*. The class consisted of 15 undergraduate students and two high school students.

Course Goal: The goal was to provide students the scientific posture of life’s origins by taking extremophiles as a model to understand the complexity and transcendence of this unique phenomenon. At the same time, we aim to motivate them to discover the connection between scientific knowledge and practical application.

Topics: 1) Origins of life and the phylogenetic tree of life, 2) nutrition, culture and growth, 3) bioenergetics, 4) bioremediation in toxic/extreme environments and 5) NASA astrobiology.



Fig. 3. The Emergence of Life display presented at the Field Museum in Chicago, Illinois, on May 18-21, 2017. University of Illinois Urbana-Champaign, Carl R. Woese Institute for Genomic Biology. Credit: Institute for Genomic Biology

Project Reports

The Evolution of Homochirality as a Universal Biosignature

The goal of this project is to understand one of the two most obvious universal signatures of life on Earth, the homochirality of amino acids and sugars. The only class of theories that has been proposed to explain the mystery of homochirality is spontaneous symmetry breaking as a result of autocatalysis. However, our analysis has shown that these theories in effect put in the answer by the assumptions they make about the other molecular processes accompanying autocatalysis. Faced with this fact, it remains to ask whether autocatalysis on its own is capable of explaining homochirality. Our breakthrough work, published in 2015 in *Physical Review Letters* (widely regarded as the premier broad impact journal in physics), answered this question in the affirmative. In that work, we showed how the property of life being far from equilibrium and the steady progression of the efficiency of autocatalysis inexorably leads to homochirality. The mechanism we discovered is intrinsically stochastic, and very different from the previous chiral symmetry breaking theories proposed previously. Moreover, it is an intrinsically non-equilibrium mechanism, and accordingly breaks detailed balance

and violates the fluctuation-dissipation theorem, as do all non-equilibrium steady states.

A remaining question that needs to be resolved is this: if a pattern is induced by noise, e.g. demographic stochasticity or copy number fluctuations in molecules, how can the pattern amplitude be larger than the noise, and thus grow to be macroscopically observable? We tackled this question by studying the linear stability of generic pattern-forming systems, and discovered a surprising mathematical property that seems to be generic: the eigenfunctions of the linear stability operator about the well-mixed phase exhibit “nonnormality”. This may sound very technical, and it is; but the consequence is that even when the well-mixed state seems to be formally linearly stable, long-lived, massive transients are almost certainly going to exist, and these will push the system out of the linear dynamics regime into a state where large pattern forming domains will form and grow.

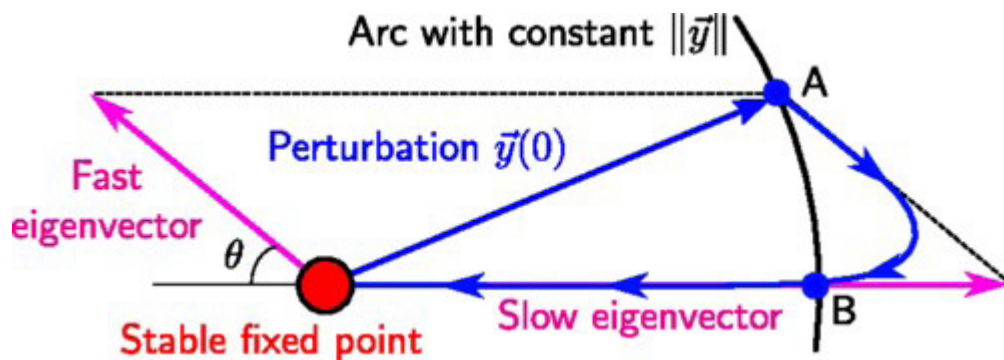


Fig. 4. Transient amplification is caused by nonorthogonal eigenvectors and a separation of time scales. The stable fixed point is subject to the perturbation $\rightarrow y(0)$. Because of the separation of time scales, the deterministic trajectory (blue arrowed line) is initially parallel to the fast eigenvector before relaxing to the slow manifold. From A to B, the trajectory has magnitude greater than $\|\rightarrow y(0)\|$. Credit: Dr. Nigel Goldenfeld

The Interplay Between Ecology and Evolution

Ecosystems are in general out of equilibrium. For example, the stability of highly diverse microbe-phage communities is difficult to explain by the simple Lotka-Volterra type predator-prey equations. In fact, stochastic complex interactions such as the coevolutionary arms race and horizontal gene transfer are the driving force for the evolution, dynamics and stability of bacteria-virus ecosystems. We have shown how stochasticity and co-evolution between bacteria can give rise to a characteristic pattern of boom-bust population cycles, poetically described as “Kill the Winner”. Another situation that is of great interest and importance to the Earth system is the tiny marine bacterium *Prochlorococcus* and its phages. The remarkable aspects of this system begin with the presence of photosystem II genes in the phage, and include niche stratification, a huge pan genome and the documented occurrence of multiple horizontal transfers of photosystem II genes back and forth between the bacteria and phage populations. Explaining these unusual features are the focus of our work, encompassing the stability and population dynamics of this system, and the role played by horizontal gene transfer. To this end, we developed a stochastic, spatially-extended individual-level model of this ecosystem, describing both the marine cyanobacteria and their predator

cyanophages that carry photosynthesis genes. The theory predicts how *Prochlorococcus* would diverge from its ancestor (*Synecococcus*, a low-light adapted bacterium) and eventually develops strains that adapt to higher light levels and different frequencies of light as the *Prochlorococcus* expands its range towards the surface. The way in which this microbial community develops is remarkable, because of the way in which evolution is accelerated due to the predation by viruses. This ecological example shows yet again how collective effects can be a major and poorly recognized driving force in the evolution of planetary-scale ecosystems.

These dynamics surely leave their trace on microbial genomes over evolutionary time. In a separate study, we traced the evolution of the core genome since the last universal common ancestor of cellular life (LUCA). Our phylogeny suggests that the evolution of the essential genome occurred through three major transitions encompassing a sequential evolution of LEACA (Last Eukaryote Archaeal Common Ancestor) from Bacteria followed by divergence of Archaea and finally the individuality of *Sulfolobus* species. Interestingly, patterns of gene orthology place the Asgard organisms clearly within the Archaeal phase according to the evolution of *Sulfolobus* essential gene components.

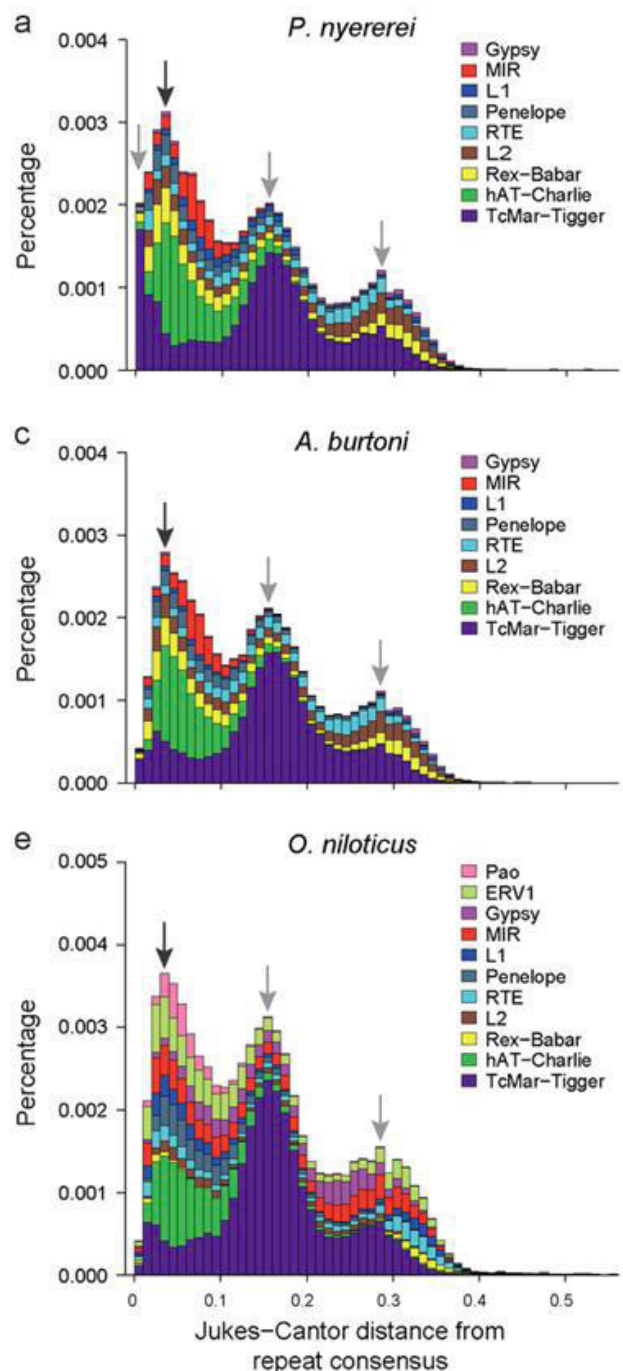
Background Image Credit: Theme Dr. Bruce Fouke and Dr. Charles Werth

The Rate of Evolution: Laboratory Experiments and Theory

We have been experimenting with bacteria in artificial microhabitats and stress gradients controlled by a microfluidic device known as the GeoBioCell. The implications of this study, in an astrobiological context, provide insight on how habitability can be used to assess microbial life in heterogeneous stress-inducing microenvironments. Beyond the underlying mechanism of evolution, group-level traits such as chemotactic motility may provide a competitive advantage that can allow us to predict microbial diversity, adaptation and evolution.

We have also been using tools from the physical sciences to quantify the evolution of populations, including a novel experiment that used live cell imaging techniques to observe the activity of transposable elements (TEs) in real time in living cells. TEs are sometimes called “jumping genes”. They are sequences that code for proteins that cut and paste or copy and paste the sequence into a different location of the genome, thus destabilizing the genome locally. As such, TEs are regarded as a major driving force for the evolution of genomes, in the same way that parasites and predators can drive the evolution of ecosystems.

To explore the idea that evolving genomes can usefully be considered as examples of evolving ecosystems, we studied a TE called an autonomous long interspersed nuclear element (LINE). These are the most abundant elements in the human genome for example. Another class is nonautonomous short interspersed nuclear elements (SINE). These TEs are a parasite of the LINE elements: they do not code for the enzyme which cuts them out of the genome, but instead recruit the proteins made by LINE during their activity, thus depriving the LINE elements of the ability to function. The interplay between these two types of retrotransposon turned out to be similar to a stochastic predator-prey population cycle, with a period of about 10 million years and thus potentially observable in genomes of ancient organisms. We have focused on the “fossil fish” *Coelecanth*, which has putatively lived in a stable marine environment for over 400 million years, attempting to find clear evidence that there are bursts of transposon activity that are driven by internal fluctuations and not by exogenous environmental driving forces. We certainly find



D. Brawand *et al.* *Nature* 513 375–381 (2014)

Fig. 5. Bursts of transposon activity as a function of evolutionary time (~50 million years) expressed in Jukes-Cantor distance units. The arrows indicate putative oscillatory bursts of transposon activity. Credit: Dr. Nigel Goldenfeld

evidence for bursts of activity, but we are currently engaged in trying to test whether or not we can conclude that these are the predicted “predator-prey” oscillations. Our work demonstrates the important role of the number fluctuations in the expression of mobile genetic elements, and shows explicitly how ecological concepts can be applied to the dynamics and fluctuations of living genomes.

Serpentinization and the Emergence of the Universal Engines of Terrestrial Life

The Illinois Team has strongly advocated the alkaline hydrothermal vent (AHV) theory for the origin of life, because this location is the only one known in which the puzzling direction of the proton motive force across extant cell walls can be naturally explained. The AHV theory is also of great current interest because of the strong circumstantial evidence from the Cassini space problem that the ice-lidded worlds Europa and Enceladus harbor the conditions that would support tectonics and vents. In the latter case, the chemical constituents of the fountain of Enceladus show strong evidence for hydrogen disequilibria, exactly as required for AHV life on Earth. A distinguishing property of the Alkaline Hydrothermal Vent (AHV) theory for biogenesis is that it posits, and largely explains the reasons for, a strict and detailed continuity between the specific thermodynamic mechanisms that power all extant life and those that powered its birth; indeed compelled its birth. Developing and explaining an understanding of how such processes could have been initially been carried out abiotically in order to drive the emergence of life, has been a central concern of our work thus far and of a few papers published before this year. But it is the exclusive focus of a major, review-length paper (60 pages) published in March of 2017 in *Physics Reports*.

We have shown that such processes can be mediated by macromolecular “mechanochemical” machines acting as dual-process ‘stochastic escapement’ engines – in which a successful downhill fluctuation in one reaction is made to be contingent on the quasi-coincident occurrence of an uphill fluctuation in another, where both up and down fluctuations are thermally induced. In this way, although no energy passes between the processes, if one is thermodynamically stronger than the other it then proceeds in the net downhill, and work is done via the other process being stochastically driven uphill by the relative irreversibility of the downhill process. The essential conceptual point is that what is taking place is the conversion of thermodynamic disequilibria, creating one by dissipating another. This mechanism has recently been understood to be instantiated in living cells through electron bifurcation, and our work provides a model framework in which to view the cell as a non-equilibrium thermodynamic engine.

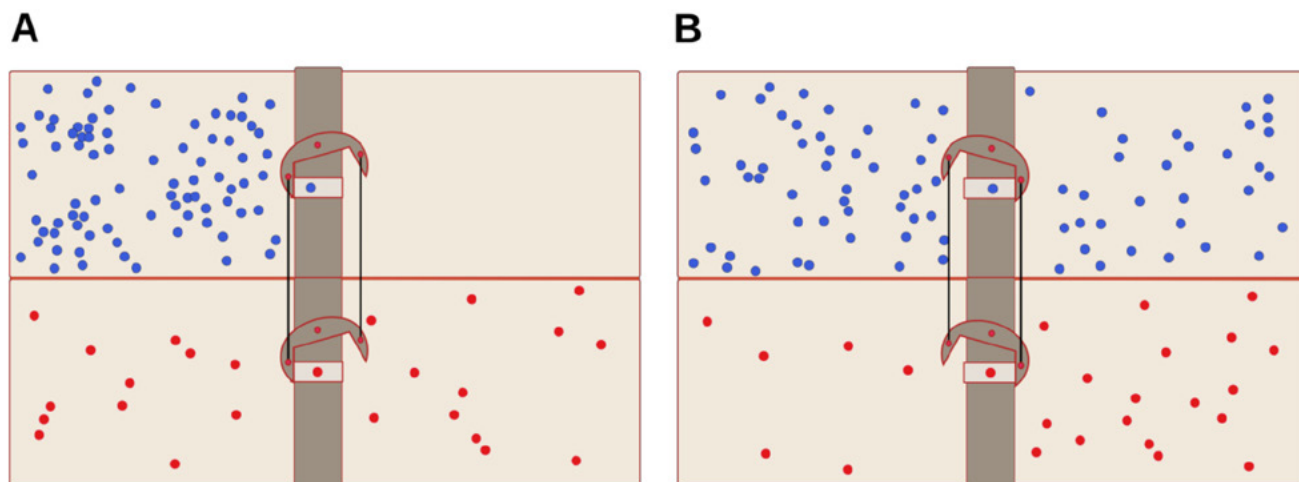


Fig. 6. Two diffusion chambers with linked ‘flip gates’. Gate flipping is stochastic, but blocked (in either direction) if only one port is loaded. A: Pictured state: blue chamber in a strong disequilibrium, red chamber in an approximate equilibrium. B: System in its end, ‘steady’ state; top chamber’s initial disequilibrium partly dissipated; bottom chamber’s initial equilibrium driven to a partial disequilibrium. The probability of both chambers being loaded from either side is now equal.

Field Work

How cells evolve in response to environmental gradients: Fieldwork at Mammoth Hot Springs, Yellowstone National Park

Modern geothermal systems represent ecological niches analogous to those colonized by ancestral life forms in early earth history. Their study is critical for reconstructing the possible functions of ancient microbial ecosystems preserved in the geological records. In the present study, a filamentous microbial mat inhabiting Mammoth Hot Spring in Yellowstone National Park was selected as the representative thermophilic community for a comprehensive study of ecophysiology and the interactive microbe-water-mineral relation at Mammoth. This geothermal niche exhibited turbulent water flow and dysoxic conditions (70°C, 6.2 pH, 0.3-10 cm/s flow rate) with abundant inorganic substrates (e.g., CO₂, sulphide and thiosulfate). CO₂ off gassing leads to rapid CaCO₃ (travertine) precipitation and deposition. Evidence from integrated meta-omics (metagenomics, metatranscriptomics and metaproteomics) of DNA, RNA, and proteins extracted from this community indicated that it contained more than 90% of single genus of *Sulfurihydrogenibium*. Cross-evaluation of the 1499 predicted non-rRNA open reading frames (ORFs), their transcripts and cognate proteins associated to the pan-genome of *Sulfurihydrogenibium* species suggested that they were finely tuned to the native environments. The *Sulfurihydrogenibium* spp. were chemoautotrophic and utilized CO₂ for carbon fixation. The chaperons may facilitate enzymatic stability and functionalities under elevated temperature. Meanwhile, high abundance of the transcripts and identified proteins for Type IV pili and exopolysaccharides (EPS) suggested the mechanisms for the *Sulfurihydrogenibium*



Fig. 7. Filamentous mats of *Sulfurihydrogenibium*.



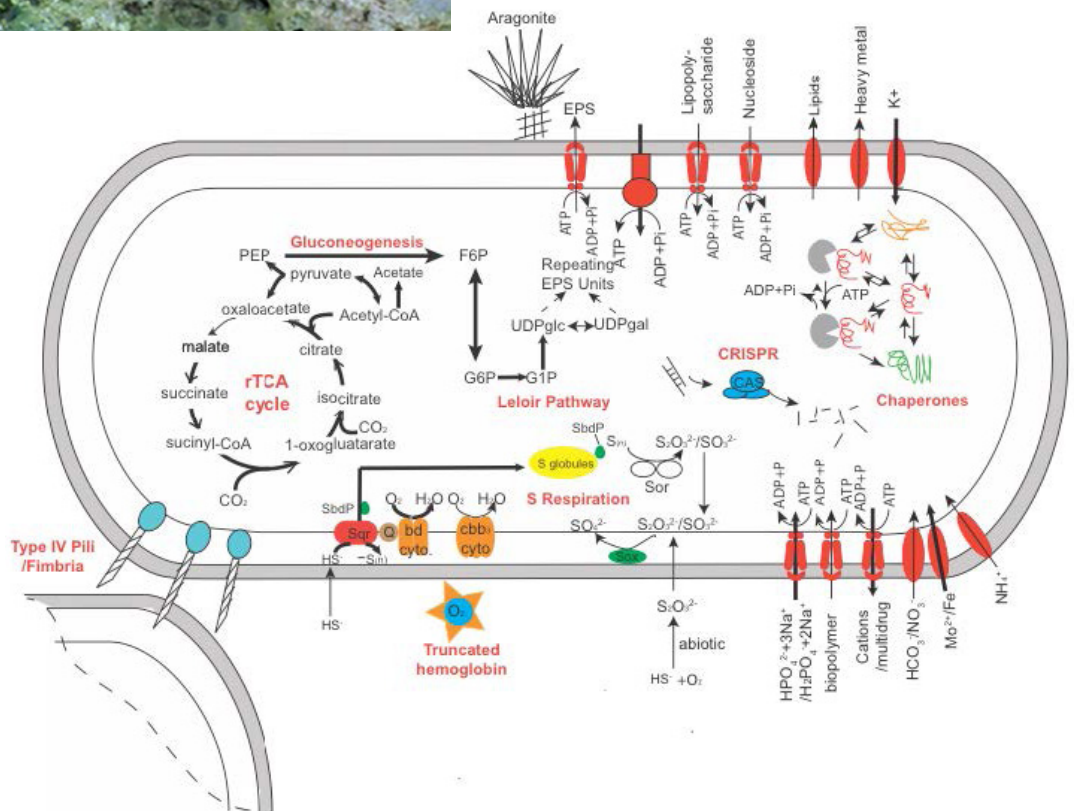
Fig. 8. Filamentous mats of *Sulfurihydrogenibium* entombing travertine terraces.



Fig. 9. (Above) Submerged cyanobacterial mats produce oxygen bubbles that buoyantly lift the microbial mats to the surface, where they become encrusted with disks of low-Mg calcite as the stretched mat is encrusted with aragonite. These biosignatures are instantly fossilized and are common in the geological record of hot-spring travertine.

drogenibium spp. to form strong intertwined microbial filaments that tolerated hydrodynamic shear force and accelerate rapid CaCO₃ travertine precipitation. Among the multiple previously proposed energy generation mechanisms, the major energy source for this group of organisms derived from oxidation of reduced sulfur (e.g., sulfide, sulfur or thiosulfate) and uptake of O₂ facilitated by bd-type cytochromes under the O₂-limiting environment. These environment-microbe mutual selective processes may explain dominance of this single genus of organisms in the ACF of the Mammoth Hot Springs and other geothermal environments with similar geochemical conditions around the world. This study can be applied as a template to understand environmental limits of life and microbial adaptive evolution on earth and in other planets.

Fig. 10. (Right) Reconstructed metabolic functions for the dominant *Sulfurihydrogenibium* spp. in the ACF Fettuuccini microbial community. The names of the metabolic pathways with significantly transcribed COGs and identified proteins were shown in red. The arrows connecting steps of the metabolic pathways were highlighted in thicker lines when both transcripts and proteins were identified.



Institute for Universal Biology: 2017 Publications

Anderson, Rika E., Kouris, Angela, Seward, Christopher H., Campbell, Kate M., and Whitaker, Rachel J. Structured Populations of *Sulfolobus Acidocaldarius* with Susceptibility to Mobile Genetic Elements. *Genome Biology and Evolution* 9, no. 6 (June 1, 2017): 1699–1710. DOI: 10.1093/gbe/evx104.

Biancalani, T., Jafarpour, F., Goldenfeld, N. (2017). Giant amplification of noise in fluctuation-induced pattern formation. *Phys. Rev. Lett.* 118, 018101 (5 pages). DOI: 10.1103/PhysRevLett.118.018101.

Branscomb, E., Biancalani, T., Goldenfeld, N., Russell, M. (2017). Escapement mechanisms and the conversion of disequilibria: the engines of creation. *Physics Reports* 677, 1-60. DOI: 10.1016/j.physrep.2017.02.001.

Campbell, Kate M., Kouris, Angela, England, Whitney, Anderson, Rika E., McCleskey, R. Blaine, Nordstrom, D. Kirk, and Whitaker, Rachel J. (2017). *Sulfolobus Islandicus* Meta-Populations in Yellowstone National Park Hot Springs. *Environmental Microbiology*. DOI: 10.1111/1462-2920.13728.

Fitzgerald, D., Hastings, P.J., Rosenberg, S.M. (2017). Stress-induced mutagenesis: implications in cancer and drug resistance. *Annual Reviews, Cancer Biology* 1, 6.1-6.22. DOI: 10.1146/annurev-cancerbio-050216-121919.

Moore, J. M., Correa, R., Rosenberg, S. M., Hastings, P. J. (2017). Persistent Damaged Bases in DNA allow Mutagenic Break Repair in *Escherichia coli*. *PLoS Genetics* 13:e1006733. DOI: 10.1371/journal.pgen.1006733.

Peterson, J. R., Cole, J. A., Luthey-Schulten, Z.. (2017). Parametric studies of metabolic cooperativity in *Escherichia coli* colonies: Strain and geometric confinement effects. *PLoS ONE*, 12(8): e0182570. DOI: 10.1371/journal.pone.0182570.

Thor, S., Peterson, J. R., Luthey-Schulten, Z. (2017). Genome-Scale Metabolic Modeling of Archaea Lends Insight into Diversity of Metabolic Function. *Archaea* vol. 2017, Article ID: 9763848, 18 pages. DOI: 10.1155/2017/9763848.

Xue, Chi and Goldenfeld, Nigel. (2017). Coevolution Maintains Diversity in the Stochastic “Kill the Winner” Model. *Phys. Rev. Lett.* 119, 268101. DOI: 10.1103/PhysRevLett.119.268101.

Team Members

Nigel Goldenfeld

Hongkai Bi
Elbert Branscomb
Isaac Cann
Krause Daniel
Scott Dawson
Lee DeVille
Bruce Fouke
Glen Fried
Philip Hastings
Yoshizumi Ishino
Neil Kim
Tom Kuhlman
Gloria Lee
Zaida Luthey-Schulten
Rod Mackie
Michael Martini, Jr.

Gary Olsen
Susan Rosenberg
Nick Sherer
Hong-Yan Shih
Rajveer Singh
Maria Spies
Sheng Shee Thor
Charles Werth
Rachel Whitaker
Chi Xue
Changyi Zhang
Lang Zhou