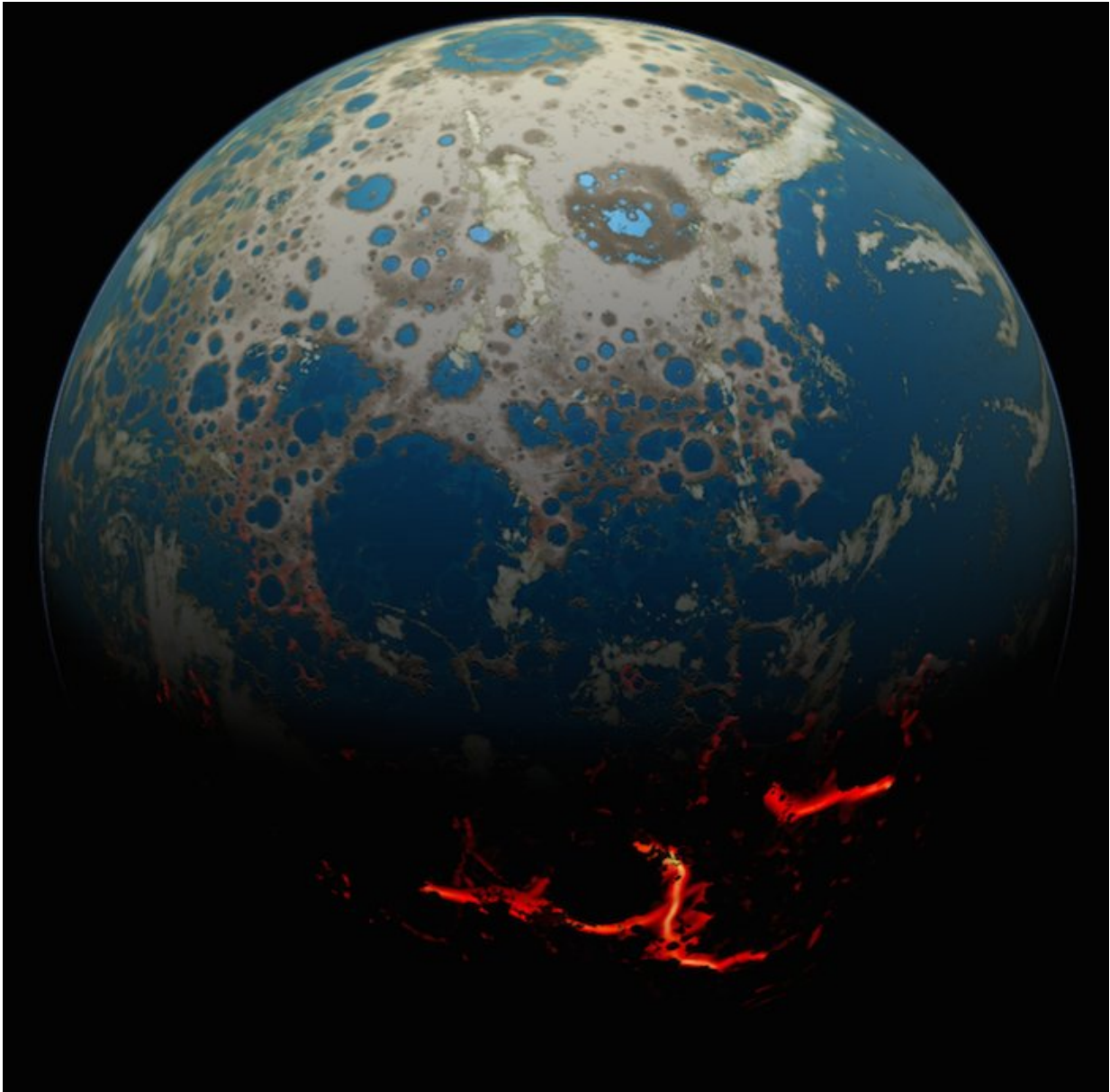


Computing the origin of life

December 14 2018, by Keith Cooper



The early Earth was a hellish place with impact galore and a choking atmosphere, and yet somehow life got a grip there. Credit: Simone Marchi/SwRI

As a principal investigator in the NASA Ames Exobiology Branch, Andrew Pohorille is searching for the origin of life on Earth, yet you won't find him out in the field collecting samples or in a laboratory conducting experiments in test tubes. Instead, Pohorille studies the fundamental processes of life facing a computer.

Pohorille's work is at the vanguard of a sea-change in how science can tackle the complex question of where [life](#) came from, how its biochemistry operates and what life elsewhere might be like. Rather than relying on the hit-and-miss of laboratory experiments, Pohorille believes that theoretical work is just as important, if not more so, in understanding how life could have emerged from non-life.

"The role of theory is twofold," he says. "It provides explanations and generalizations of what is observed in experiments, but it also has some predictive power."

Pohorille's theoretical work resides within a field known as [computational biology](#); Pohorille himself is director of the Center for Computational Astrobiology and Fundamental Biology at NASA's Ames Research Center in Mountain View, California, and a Principal Investigator with the Exobiology & Evolutionary Biology Program. Computational biology involves designing and writing algorithms within mathematical models that seek to explain life's complex biochemical processes. This is in comparison to 'artificial life,' which creates virtual life-forms that reside in the computer and which can mimic life's processes. However, the approach of computational biology hasn't been an instant hit with all biochemists and evolutionary biologists.

"It's still contentious, partly because there is a group of people who do believe that [searching for] the [origins of life](#) is a strictly experimental

issue that can only be solved in the lab," says Pohorille. "I respectfully disagree with those who think that way."

This is a view shared by Eric Smith, a researcher in complex non-equilibrium systems at the Earth-Life Science Institute (ELSI) which is attached to the Tokyo Institute of Technology in Japan. Smith highlights how, in recent years, the fields of computational biology and chemistry have matured to the point that researchers used to working in the laboratory can no longer ignore it. "I think we're on the threshold of where it's going to start becoming a serious tool, but it's important to remember that it's only one tool of many."

An example of the usefulness of computational biology can be seen in Smith's work delving into the origins of carbon-fixing, which describes how organisms convert inorganic carbon into the organic compounds vital to life. Smith and his colleague Rogier Braakman of the Chisholm Lab at the Massachusetts Institute of Technology combined a computational approach to phylogenetics (which is the study of the evolutionary relationships between organisms) with metabolic flux balance analysis (which allows metabolisms to be recreated in mathematical simulations on the computer) to disentangle the six different ways in which life is known to fix carbon, in the process figuring out which of the sextet evolved first. Consequently, Smith and Braakman were able to show how this form of carbon-fixing, which was one of life's original metabolic processes, was able to arise from simple geochemistry. As such, it mirrors the overall quest for the [origin of life](#) in terms of how biological processes developed from geochemistry.

Although some of these research questions are attainable using computer modeling, we are still lacking an understanding of many of the basic rules governing biochemistry as well as early life's genetics. Some researchers have speculated about an 'RNA world' wherein the self-replicating RNA molecule not only played the role that DNA, which is

fashioned from RNA, does today, but that it also arose pre-biotically and was a cornerstone in the origin of life. However, many scientists, including Pohorille and Smith, disagree, claiming that RNA is too big and unwieldy a molecule to have played a role in life's probably simpler origins. Instead, they suspect that there was some other chemistry at work in the origins of self-reproducing life, although what this chemistry could have been remains a subject of vigorous discussion.

Given this uncertainty, Pohorille favors forming generalizations about the biological processes at work in the origin and earliest evolution of life, rather than looking for specific outcomes. Using computational theory, he advocates focusing on the underlying principles of biological processes that are rooted in the laws of physics and chemistry. "What is needed are some general rules that guide us in building scenarios," he says. "Just having individual experiments that say something is possible – because that's as much as we can get from the experiment – is not enough."

Artificial life

One of the biggest questions about the origin of life and its subsequent evolution is how random molecules managed to organize themselves into complex living organisms. What prompted them to form complex molecular chains that became the basis of life, and what are the underlying principles that govern which molecules became the important cogs in the system? With so many permutations of how molecules can combine, on the face it would seem extremely unlikely that nature would just stumble onto the right combination of molecules to form self-replicating life.

At Michigan State University, Chris Adami thinks he has the answer. A professor of microbiology and molecular genetics, Adami takes computational biology to the next level by using the artificial life

software called Avida, which runs self-replicating programs that mimic biology and evolution. Through Avida, which he co-developed in 1993 with his Michigan State colleague Charles Ofria and UC Davis' Titus Brown, Adami is able to test his controversial but potentially revolutionary idea that life can be defined as 'information that self-replicates' and that the selection of useful molecular systems for life is governed by the laws of probability.

Avida operates by creating a virtual world in which programs compete for CPU time and memory access, just like organisms competing for resources in the real world. These virtual lifeforms can self-replicate, but crucially they have copy error programmed into them so that, just like in real life, mutations can be carried over to daughter programs to simulate evolution by natural selection. Because they are self-replicating, mutating computer programs could potentially be very dangerous were they to escape Avida and infect the Internet. As a safety precaution, the virtual world is run on a simulated computer inside a real computer so that the programs appear on the outside merely as data.

Where does the first replicator come from? In Avida the first replicator is purposefully written, but in real life the first biological replicator had to emerge spontaneously from nature, and this is where selectivity comes in. "It turns out that replicators, whether in nature or within Avida, are rare," says Adami, "and the odds are that a random program – or assembly of molecules – will not replicate."

The programs are written in a computer language that contains 26 instructions, analogous to individual monomers in chemistry, labelled as the letters of the alphabet from a to z. Adami uses this system to draw an analogy to the written word. Imagine a bag filled with equal numbers of all the letters of the alphabet. A random drawing of the letters into sequences of varying lengths, called 'linear heteropolymers,' creates strings of instructions into which information is encoded. If these

polymers were meant to be 'words,' they would mostly be gibberish, containing a jumble of 'q's and 'z's and other letters without connoting meaning. Similarly, the molecules that were available on early Earth had many different ways to bind together to produce a variety of chemical reactions; the chance of nature generating the right molecular structure to enable self-replication is slim.

The Biased Typewriter

Adami points out though that language is loaded to favor certain letters that crop up more often than others. Seldom are 'q's or 'z's used, but 't's and 'e's and 'a's are common letters in words. Adami suggests that the selection problem can be better understood as the 'biased typewriter' model, in which some molecules and chemical reactions are more likely to occur than others. If the letters in the bag were scrabble tiles, with more of the common letters and fewer of the rarely used letters, then even pulling them out at random would lead to some real words being produced, just by chance.

With his student Thomas LaBar, Adami tested the principle of the biased typewriter in Avida, loading the instructions with those monomers that are useful for self-replication. In a billion random programs made from chains of 'letters' that Avida subsequently produced, Adami found that 27 of them could self-replicate. He used those 27 to create a probability distribution and then kept running the program, finding that the number of self-replicators kept increasing dramatically.

"In other words, what this tells you is that if you have a process that generates these monomers at the right frequency, then you're going to be able to find the self-replicators much faster," says Adami.



How did the first self-replicating polymers, the precursors of life, come to form on the early Earth around four billion years ago? Credit: NASA Goddard Space Flight Center Conceptual Image Lab

Just 27 initial self-replicators out of a billion linear heteropolymers doesn't sound like very much, but early Earth was a big place full of opportunity, with all kinds of different environments in which nature could experiment by combining monomers to form useful polymers for life. However, although Adami's theoretical estimates have been born out experimentally by Avida, replicating the process to test the RNA World theory is a different proposition because the amount of information contained within RNA is too great even for the computer to handle. Nevertheless, Adami sees his 'biased typewriter' model as one of the general rules to which Pohorille was referring.

Eric Smith agrees with Adami that the basic idea behind the biased typewriter is on point.

"By biasing the building block inventory, you can drastically change the likelihood of one assembly versus another and we see it in all sorts of

places in biology," he says.

When it comes to the importance of information and the relevance of artificial life, Smith has his doubts. "One shouldn't look for a big answer from any one piece of work," he says. Instead, he says, the origin of life isn't just one problem that requires an overarching solution, but an enormous sequence of problems, including the origin of all the metabolic processes as well as self-replication that must each be solved and no one model or computer program can provide the answer. Yet it was once thought that artificial life might have been able to do just that.

"People on both sides—artificial life and origins of life—don't really pursue that much anymore," he says. "There's not much cross-talk between the two."

Andrew Pohorille is also skeptical about Adami's approach, as well as the usefulness of artificial life to origin of life research, suggesting that without some high-level mathematical concept that explains why there is only one set of rules that governs the origin of life and life's processes, whether real or virtual, then the rules of virtual worlds like Avida will not necessarily translate into the real world.

"There may be many rules that lead to these kinds of processes," says Pohorille. "The question is whether any of these rules have anything to do with the rules that operated at the origins of life."

Adami acknowledges that the rules in Avida won't be the same as the geochemical and biochemical rules that operate in real life, but he argues that regardless of the chemistry, the principles of information theory remain.

"It's of course true that we will not find how life evolved on Earth by looking inside a computer," he admits, "But we can test general

principles and, once we know these principles, we can go ahead and test those in biochemical systems."

Computational Astrobiology

In the laboratory researchers work with terrestrial life and observe its processes, but on alien worlds life could be very different, operating under different rules that are impossible to test with Earthly life in an experiment. Computational biology and artificial life, however, offer the unique abilities to explore life abstractly by investigating different processes that could exist on other planets with different environments and geochemistry. Could computational biology help astrobiologists describe alien life before we even find it?

NASA scientists certainly think that's feasible, having recently invited Chris Adami to a workshop to discuss biomarkers, where he presented his idea of how to look for life through information and its replication, rather than RNA and proteins. Adami describes this research effort in terms of patterns that are unnatural or in disequilibrium, looking for letters and finding 'e' more common than 'h', as in the Avidian life example. In order to do this the local geochemistry needs to be known fairly well, which is something far beyond our current abilities of exoplanet studies.

Closer to home our knowledge of geochemistry is a little better, or at least can be improved in the near future. Take Europa, for instance. "We're thinking about what evidence for life we should search for there," says Pohorille. The idea is to computationally explore the range of molecules other than proteins or nucleic acids that could perform the same functions as they do on Earth, and figure out what their biosignatures would be on Europa. On a cautionary note, it might be tempting to describe something too extreme using these alternative concepts for life. "It's kind of a dilemma," says Pohorille. "What is

enough and what is too much?"

Something might look like life in Avida, but there's a danger of falling into the trap of looking for a pattern that resembles life, but isn't, like confusing the motions of a slinky toy with those of a snake. It's this concern that virtual life in the computer, such as Avidian life, may only be masquerading as representing life that causes so many researchers to be suspicious of its results. "In principle artificial life could help provide alternatives to Earth life, but you've got to figure out what your computer model is an abstraction of, and that's the hard part," says Smith.

Nevertheless, the scientific community as a whole is slowly coming around to the notion that computational biology and chemistry, as well as possibly [artificial life](#), could be vital in progressing the field further. Smith, for example, wonders whether our understanding of the chemistry of complex systems needs to take on a cyborg-like quality by integrating a lot more closely with computational research.

Meanwhile, the field needs a new generation of scientists trained in the use and application of computers for theoretical work, something that is forthcoming now that the computational tools are available and scientists are figuring out new and innovative ways to use them.

"When I started doing these computational simulations, almost nobody could see how it could possibly be related to anything remotely interesting to the origins of life community," admits Pohorille, saying he was tolerated by his peers because he was just "one odd guy." Today, however, he says that younger researchers are realizing that theory and experiments have to go hand-in-hand.

If the field of computational biology is truly going to grow, the funding has to also. Currently, NASA is the only agency in the United States

funding origin of life research, with some private money coming from the likes of the Simons Foundation and the Templeton Foundation. "Tell me of a university that is looking for theorists specializing in the origin of life," asks Pohorille rhetorically. "I haven't heard of one."

Internationally, ELSI in Japan is one of the few institutions hoping to get closer to the origin of life through computational efforts.

As computing power increases, scientists using it will increasingly be able to solve problems about life's processes. Perhaps computational biology will be just one tool among many available to researchers, but its presence will not only help scientists to think of new ways to explore the origins of life, but also to come up with new ways to think about it too. The mystery of life's origins could one day be solved thanks to that modern antithesis of life – the computer.

This story is republished courtesy of NASA's Astrobiology Magazine. Explore the Earth and beyond at www.astrobio.net .

Source: Astrobiology Magazine

Citation: Computing the origin of life (2018, December 14) retrieved 7 September 2024 from <https://phys.org/news/2018-12-life.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--