

Open-Endedness for the Sake of Open-Endedness

Arend Hintze

Michigan State University
Department of Integrative Biology
Department of Computer Science
and Engineering
BEACON Center for the Study
of Evolution in Action
hintze@msu.edu

Abstract Natural evolution keeps inventing new complex and intricate forms and behaviors. Digital evolution and genetic algorithms fail to create the same kind of complexity, not just because we still lack the computational resources to rival nature, but because (it has been argued) we have not understood in principle how to create open-ended evolving systems. Much effort has been made to define such open-endedness so as to create forms of increasing complexity indefinitely. Here, however, a simple evolving computational system that satisfies all such requirements is presented. Doing so reveals a shortcoming in the definitions for open-ended evolution. The goal to create models that rival biological complexity remains. This work suggests that our current definitions allow for even simple models to pass as open-ended, and that our definitions of complexity and diversity are more important for the quest of open-ended evolution than the fact that something runs indefinitely.

Keywords

Open-ended evolution, computational model, complexity, diversity

I Introduction

Open-ended evolution has been identified as a key challenge in artificial life research [4]. Specifically, it has been acknowledged that there is a difference between an open-ended system and a system that just has a long run time. Large or slow systems will eventually converge on a solution, but open-ended systems will not and instead keep evolving. Interestingly, a system that oscillates or that is in a dynamic equilibrium [21] would continuously evolve, but does not count as an open-ended evolving system. While we seem to know well what an open-ended system is not, we struggle to find a positive definition of open-endedness [6, 7, 12, 18, 24, 26].¹ In addition, we keep finding evidence leading us to refine previous constraints or requirements [18, 24].

Across the literature, we find several conditions, or requirements, of an indefinitely² evolving system. Maley [18, pp. 3–8] defined a set of requirements for open-ended evolution:

1. An open-ended evolutionary system must demonstrate unbounded diversity during its growth phase.
2. An open-ended evolutionary system must embody selection.
3. An open-ended evolutionary system must exhibit continuing new adaptive activity.
4. An open-ended evolutionary system must have an endogenous implementation of niches.

¹ The term open-endedness will be used for the collection of properties that defines open-ended evolution.

² Indefinitely in this context does not mean infinitely long. Natural evolution, even though we think of it as open-ended, is still constrained by our physical universe, which will probably come to an end at some point. Similarly, we ignore the physical constraints of the computers we would run an open-ended evolving computer model on, and just ask if, in principle, it would be open-ended.

Soros and Stanley [24, pp. 2–3] proposed four necessary (but maybe not sufficient) conditions an open-ended system should satisfy:

5. A rule should be enforced that individuals must meet some minimal criterion (MC) before they can reproduce, and that criterion must be nontrivial.
6. The evolution of new individuals should create novel opportunities for satisfying the MC.
7. Decisions about how and where individuals interact with the world should be made by the individuals themselves.
8. The potential size and complexity of the individuals' phenotypes should be (in principle) unbounded.

Also, Taylor [26, p. 2] stated the requirements for an open-ended system as:

9. Robustly reproductive individuals.
10. A medium allowing the possible existence of a practically unlimited diversity of individuals and interactions, at various levels of complexity.
11. Individuals capable of producing more complex offspring.
12. An evolutionary search space that typically offers mutational pathways from one viable individual to other viable (and potentially fitter) individuals.
13. Drive for continued evolution.

A more general condition that should be sufficient to create an open-ended evolving system is “evolutionary dynamics in which new, surprising, and sometimes more complex organisms continue to appear” [26, p. 1; 27]. While some of the conditions above describe evolutionary dynamics themselves (2–5, 7, 9) and define what open means (8, 12, 13), other requirements pertain to *diversity* (1, 6, 10) and *complexity* (10, 11). Additionally, such an open-ended evolving system will create novelty by definition, since a repetition of prior inventions would not satisfy the requirements for open-endedness (8, 12, 13). Thus, we do not need to take more specific definitions of *novelty* into account, as they are an emergent property of open-ended evolution, even though novelty remains an important aspect in evolutionary computation [15].

Note that the requirements for the evolutionary dynamics and openness are very specific, whereas diversity and complexity themselves lack specificity. Let us elaborate on the example of biological diversity, which is greatly constrained by the genetic material itself: Only DNA or RNA (in RNA viruses) is used. Life mostly uses only 20 amino acids, even though synthetic biology has shown that life has many alternatives to those 20 [2]. Typically, metabolisms either bind carbon dioxide and produce carbon dioxide or oxidize carbohydrates using carbon dioxide (with exceptions, as for example *chemolithotrophs*, which oxidize non-organic compounds [19]). However, when considered in this way, diversity is minimal. On the other hand, in consideration of species diversity, we find millions of very different species, and their diversity massively changes over the course of our own evolutionary history [10]. This means that even in the context of our naturally evolving world, some aspects are diverse and some are not. Our requirement for open-endedness just asks for diversity of some sort and is not specific.

The same applies to complexity in a much more drastic fashion. As we know, science does not even have a single definition of complexity [11], and the requirements for open-endedness do not suggest a single *flavor* of complexity. Unfortunately, we cannot just use the “I know it when I see it” [14] approach to this matter, as it would leave each of us with our own subjective definitions, a fragmented bunch of mavericks, incapable of creating the objectivity science demands.

Since there is no specific definition or requirement within the context of open-endedness, we will instead use our own two specific metrics for diversity and complexity. As a proxy for diversity, we will define the difference between organisms by the Levenshtein distance [16] between their genetic sequences. This distance computes the number of sites that are different between two sequences of different length. Similarly, we will use Kolmogorov complexity [13] to assess the increase of complexity over evolutionary time. Kolmogorov complexity is defined as the length of the shortest description of a string. Here we will approximate Kolmogorov complexity by using compression [30] instead of the infinite search that would be required actually to prove that the shortest description

is used. Kolmogorov complexity and Levenshtein distance at least identify if the system itself is producing more complex and more diverse states than before.

The long-term behavior of a simple computational evolving system (as has been done before to challenge open-ended evolution specifically [5, 6, 12, 17, 24]) will be studied here. This system is used as an instance of a potential open-ended system, and serves as a basis for a thought experiment that we are going to undertake. The system presented here is an open-ended evolutionary system within our definitions of diversity and complexity and the requirements laid out before (1–13). I will then argue that the study of open-ended evolution, without a specific definition of complexity and diversity, becomes a study for its own sake. The exciting property of an evolving system is not its openness but instead the complexity of the actual evolved solutions. It is thus the definition of complexity and diversity that we should focus on.

2 The Ever-Evolving System

The system presented here consists of 100 organisms, each defined by a genome. The genome of each organism is a sequence made from three possible symbols: right, left, forward. Each sequence defines a path, or trajectory, through a 2D space. When given agency at the time of evaluation, organisms begin at the origin facing up (moving forward would now increment the location along the y -axis). The genetic sequence then defines a specific path. This path is evaluated to compute the fitness (the score that defines reproductive success). Along this path, each organism can collect points that determine its fitness. These points are higher for locations where fewer organisms have visited, and lower for highly visited locations. Specifically, all paths are generated at the beginning of a generation, and for each location, the number of organisms, V_i , that visited it is computed. The count for visitations accumulates over the course of the experiment. Each location i then rewards each visiting organism with $\frac{1}{V_i}$ points. The fitness (W) now becomes the sum of all visited locations:

$$W = \sum_i \frac{1}{V_i}. \quad (1)$$

Now that the score for each organism in the population has been computed, a Moran process is implemented by using roulette wheel selection [20]. This allows organisms with a higher score to transmit proportionally more offspring into the next generation. When the next generation is formed, each organism that propagates into the next generation has the chance to experience mutations. Each location in the sequence has a 1% chance to mutate by randomly drawing uniformly from the three symbols (allowing for neutral mutations), and each location also has a 1% chance to be deleted, and a 1% chance to experience an insertion of a random symbol. This scheme ensures that organism genomes can grow and evolve to become arbitrarily long sequences, which at least carries the opportunity to become arbitrarily complex. The next generation is formed entirely new from the previous one and has the same size. All organisms formed in the previous generation are discarded, and thus the population size is kept constant.

The experiment was run for 5 million generations in 100 replicate instances using different random number seeds. The starting sequences of the populations were all randomly generated with a length 10.

2.1 Zlib Compression

Kolmogorov complexity is the measure of a sequence by its shortest possible description length [13]. A convenient way to approximate the theoretical lower bound of Kolmogorov complexity is the compressibility of a sequence [30]. Here we use the zlib Python library [1, 8] to estimate the complexity of sequences by compression.

2.2 Diversity

The Levenshtein distance [16] is used to measure the diversity between the evolving populations. This measure aligns sequences of different length and counts the number of changes. However, this is an expensive computation. When measuring diversity, one random sequence from each experiment is chosen, and the mean Levenshtein distance between all pairwise comparisons across all 100 replicate experiments is computed.

3 Results

Let us first ask if our system satisfies all requirements and conditions for open-ended evolution listed above. The simpler requirements about the implementation of evolution (requirements 2, 5, 9) are satisfied, since the system has a criterion of replication (requirement 9) and selection (requirement 2, and maybe 5). Requirement 5 demands a minimal criterion for replication, which is implemented here by using a proportionally higher replication chance if sequences are longer and traverse a path that is less frequented by other organisms. This proportionality would become trivial and not limiting if all sequences had a length of 0. Then all sequences would replicate with the same likelihood. This is prevented by starting the experiment with sequences of length 10. The system is unbounded (requirement 8), since the 2D space is theoretically infinite and sequences are not limited. Organisms make individual decisions about their interactions, as determined by their genome (requirement 7). One might ask if organisms interact at different levels of complexity (requirement 10). I would argue that organisms in this system directly interact when their paths cross. These intersections can happen at any point on the paths (micro-scale interaction), but all earlier elements within each sequence contribute to where this interaction happens, leading the whole patterns to overlap or not (macro-scale interactions). This line of reasoning can be applied to ever-higher levels of organization within the structure of the path. One might also question the validity of this requirement itself. All interactions that organisms experience happen at the atomic level (between atoms); it is we humans that impose an order on top of that to describe macro phenomena more easily. A topic highly debated within this context [9, 22, 23] has even led to the notion that the term “hierarchical interactions” itself needs a better definition or to become an integral part of the definition of complexity itself [3]. Anyway, requirement 10 is satisfied in a general sense. The requirements for new niches (requirement 4) and novel opportunities for behavior (requirement 6) are both satisfied by the “fit when rare” rule, since new mutations are constantly required to prevent organisms from having overlapping paths. Secondly, new paths present new opportunities for other organisms to circumvent them, which in turn is an argument for requirements 3, 6, and 7 to be satisfied. In theory, mutational paths from every genotype to every other genotype exist, which satisfies requirement 12. What remains are the demands for increasing diversity (requirements 1, 11) or complexity, and that this increase should happen forever (requirement 13). As stated above, Kolmogorov complexity and Levenshtein diversity will be used to quantify complexity and diversity. However, how can we quantify whether something adapts indefinitely and whether complexity and diversity will increase or plateau at some point?

3.1 Is the System Open?

Applying the scientific method to this question demands of us to formulate a hypothesis, to test it vigorously with experiments, and potentially reject the hypothesis. Observe that we can only reject hypotheses. In this case, the hypothesis is that a system will run indefinitely, and we can only reject this hypothesis by finding a counterexample. Therefore, we cannot experimentally prove that a system will indefinitely keep evolving. Similarly, our system is a computational one, so maybe there is a mathematical or computational proof we could attempt to show indefinite evolution? Unfortunately, it has been shown that such an attempt is futile [28] and that there is no way to predict that sort of outcome without trying it out and actually running the system. Being unable to prove indefinite evolution, what remains then is the ability to make a prediction: We show that the system behaves in a certain way, and we infer that the behavior will not change. Fortunately, we can distinguish via

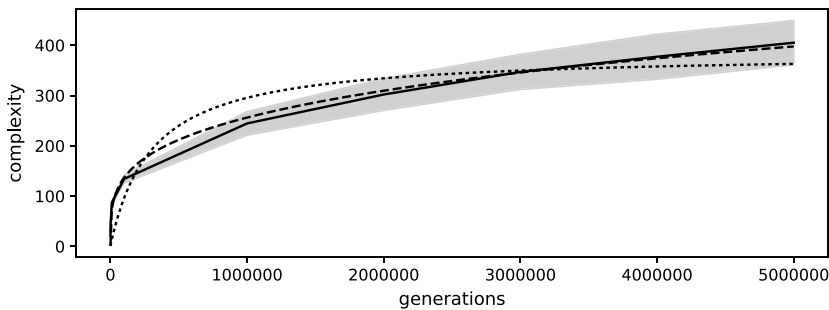


Figure 1. Kolmogorov complexity of genomes over time, averaged over 100 replicate experiments (solid black line); the gray shadow shows the standard deviation of the sample. The data was fitted against the hyperbolic (dotted line; see Equation 2) and the exponential (dashed line; see Equation 3) model. The standard deviations of the fit for the hyperbolic model for both parameters ($a3.5757 \text{ e} + 01$ and $b1.8312 \text{ e} + 05$) were greater than those for the exponential model ($a1.0149 \text{ e} - 02$ and $b5.2056 \text{ e} + 02$), indicating that the computational model is open-ended with respect to complexity.

evolutionary dynamics whether a system is going to plateau or potentially keep evolving forever [29]. This method has been proposed by Taylor et al. [2016; 27].

Simply put, if we collect data about the change in fitness of organisms over time, then we can fit either the hyperbolic model

$$\bar{W} = 1 + at/(t + b) \quad (2)$$

or an exponential model

$$\bar{W} = (bt + 1)^a. \quad (3)$$

The hyperbolic model describes systems that eventually plateau and thus are not open-ended, whereas the exponential model describes systems, like natural evolving systems, that are open-ended. Thus, the simple computational system used here must fit the exponential model and not the hyperbolic one in order for us to show that it is indeed open-ended. Of course, here we use complexity as a proxy for fitness, since fitness cannot be used within the context of open-ended evolution, as explained above.

As expected, the computational model fits the exponential model better than the hyperbolic model (see Figure 1). This suggests that complexity, or in this case, specifically, the Kolmogorov complexity approximated by the compressibility of the evolving genomes, will increase indefinitely. The limit to this is only set by the time or memory limitations of the computer running the computational model. This is a similar limit to the one that the finite universe imposes on natural evolution. This observation is also evidence that the sequences will increase in length (satisfying requirement 8, 11, and 13).

Similarly, when the measure of diversity over evolutionary time is fitted against the hyperbolic (Equation 2) or the exponential (Equation 3) model, the error is smaller with the exponential model. This supports the notion that diversity will also increase indefinitely and not plateau (see Figure 2).

4 Discussion

A simple computational model is capable of satisfying the criteria for open-ended evolution (requirements 1–13). It implements the necessary rules that specify how evolution should work (2–5, 7, 9). When running the model, diversity and complexity increase indefinitely, which satisfies the requirements for openness (8, 12, 13). What remains a crucial point of discussion is whether complexity (requirements 10, 11) and diversity (requirements 1, 6, 10) were measured properly. A more complex model might have satisfied some of the above points more obviously; however, this would have distracted from the scope of the model. The idea is to find a simple model that satisfies our criteria

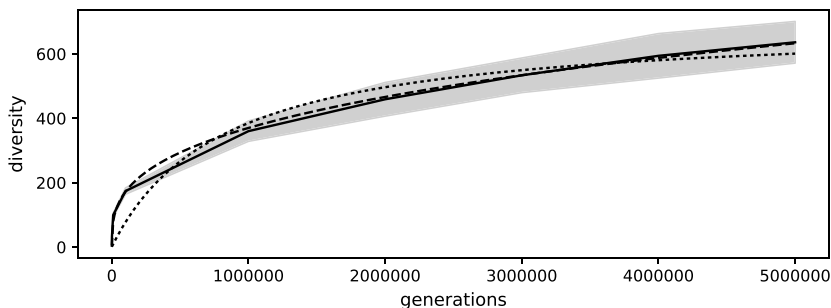


Figure 2. Levenshtein distance as a proxy for diversity of genomes over time, averaged over 100 replicate experiments (solid black line); the gray shadow shows the standard deviation of the sample. The data was fitted against the hyperbolic (dotted line; see Equation 2) and the exponential (dashed line; see Equation 3) model. The standard deviations of the fit for the hyperbolic model for both parameters ($a8.0069 \text{ e} + 01$ and $b3.7795 \text{ e} + 05$) were greater than those for the exponential model ($a9.8948 \text{ e} - 03$ and $b2.79423 \text{ e} + 01$), indicating that the computational model is open-ended with respect to diversity.

for open-endedness. We can now ask if the kind of complexity and diversity used here are meeting the expectations of open-ended evolution.

If the definitions used here (Kolmogorov complexity and Levenshtein distances) were satisfactory, then we could consider the quest for open-ended evolution as resolved. However, quite frankly, even the author (and hopefully you the reader) is disappointed by these results. The generated paths are interesting only to someone who finds meaning in seemingly random walks that seek to avoid each other (for examples of evolved paths see Figure 3). Instead of calling open-endedness a solved problem, we need to focus our attention on the definitions that allowed this simple model to be sufficient. The question now becomes, how can we define complexity and diversity in such a manner that one could have rejected this model with ease?

We are left with two options. Perhaps the current definitions of complexity and diversity that allowed for this simple model to be considered open-ended are insufficient. It would then not be a matter of creating more complex models, but of assessing the sufficiency of these models for open-endedness. We could then address this by augmenting our definitions. One obvious criticism of this model might be that internal rules for evolution or interaction do not emerge from within the system. Everything seems to be extrinsically defined, and thus evolution does not change intrinsically [25]. Including the idea that rules affecting the complexity of the system must themselves emerge from within the system might remedy this shortcoming.

Alternatively, we could agree that our assessment of whether or not a system is sufficiently open-ended will always be a consequence of the definitions of complexity and diversity. Then, open-endedness becomes secondary when it comes to assessing how meaningful a model actually is. The question would not be whether a system evolved open-endedly but whether the complexity and diversity are “interesting enough.” However, this might become highly subjective. I am interested in evolved artificial intelligence, and someone else might fancy evolved morphology. It is unscientific to reject proposed systems simply because they do not encourage the complexity and diversity in the context I care for. A more general criterion might be the scientific utility of the model. Does the model answer a particular scientific question? If so, it was sufficiently complex and it showed enough open-endedness. This approach would also avoid the following logical pitfall we encounter when demanding open-endedness:

- Biological complexity is observed in a seemingly open-ended system.
- Your computational model is not open-ended.
- You cannot make any inference about biological complexity, because your system is not open-ended.

This fallacy prevents us from focusing on the initial research question about the emergence of biological complexity, which could be answered without the necessity of open-endedness. Or ideally,

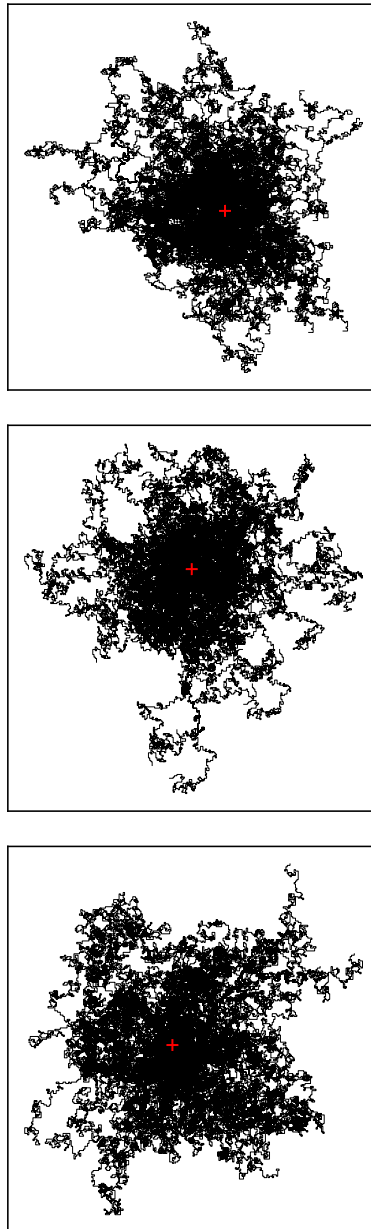


Figure 3. Three random examples of independently evolved populations at generation 1,000,000. The path of each individual in the population is superimposed in each panel. The red cross marks the origin.

with the proper definitions of complexity and diversity, open-endedness might be a natural consequence of these systems.

5 Conclusion

This model, even though seemingly open-ended, does not fulfill the quest for open-endedness. However, it suggests that we as a community fail to properly define what we want to study by using the term

open-ended evolution in lieu of other more specific and more productive terms. We want to create systems that are surprising, complex, and novel and that serve a scientific purpose. Our overly general definition of complexity and diversity within the requirements of open-ended evolution prevent us from achieving our goal. Including a testable definition of hierarchy and the nature of interactions might be another way to improve the requirements [3, 23] for open-ended evolution. This work aims to focus the attention of the field on what is crucial. A proper definition of what we mean by complexity and diversity is needed so that these definitions can become the basis of rigorous hypothesis testing. We should not engage in the quest for open-endedness for the sake of open-endedness.

Acknowledgments

I would like to thank Jory Schossau, Nitash C G, and Leonie Hintze for insightful comments, discussions, and help in editing this article. I would also like to thank the anonymous reviewers for their constructive criticism, which helped to improve the manuscript significantly. This work was in part sponsored by the BEACON Center for the Study of Evolution in Action under NSF Cooperative Agreement No. DBI-0939454.

References

- Adler, M. (2017). Zlib python compression library. <http://www.zlib.net/>. Accessed: 2018-09-22.
- Balaram, P. (1992). Non-standard amino acids in peptide design and protein engineering. *Current Opinion in Structural Biology*, 2(6), 845–851.
- Banzhaf, W., Baumgaertner, B., Beslon, G., Doursat, R., Foster, J. A., McMullin, B., De Melo, V. V., Miconi, T., Spector, L., Stepney, S., & White, R. (2016). Defining and simulating open-ended novelty: Requirements, guidelines, and challenges. *Theory in Biosciences*, 135(3), 131–161.
- Bedau, M. A., McCaskill, J. S., Packard, N. H., Rasmussen, S., Adami, C., Green, D. G., Ikegami, T., Kaneko, K., & Ray, T. S. (2000). Open problems in artificial life. *Artificial Life*, 6(4), 363–376.
- Channon, A. (2001). *Evolutionary emergence: The struggle for existence in artificial biota*. PhD thesis, Keele University.
- Channon, A. (2006). Unbounded evolutionary dynamics in a system of agents that actively process and transform their environment. *Genetic Programming and Evolvable Machines*, 7(3), 253–281.
- Channon, A., Maley, C., & Boudreau, E. (2000). Three evolvability requirements for open-ended evolution. In M. A. Bedau, J. S. McCaskill, & S. Rasmussen (Eds.), *Artificial Life VII Workshop proceedings* (pp. 39–40). Cambridge, MA: MIT Press.
- Deutsch, L. P. (1996). Gzip compressed data format specification. Available at <ftp://ftp.uu.net/pub/archiving/zip/doc>.
- Gross, D., & McMullin, B. (2001). Is it the right ansatz? *Artificial Life*, 7(4), 355–365.
- Huston, M. A. (1994). *Biological diversity: The coexistence of species*. Cambridge, UK: Cambridge University Press.
- Johnson, N. (2009). *Simply complexity: A clear guide to complexity theory*. London: Oneworld.
- Juric, M. (1994). An anti-adaptationist approach to genetic algorithms. In *IEEE World Congress on Evolutionary Computation, 1994, Proceedings of the First IEEE Conference on computational intelligence* (pp. 619–623). New York: IEEE Press.
- Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1(1), 1–7.
- Lattman, P. (2007). The origins of Justice Stewart’s “I know it when I see it.” Wall Street Journal Law Blog, 27 September.
- Lehman, J., & Stanley, K. O. (2011). Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary Computation*, 19(2), 189–223.
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10, 707–710.
- Maesani, A., Fernando, P. R., & Floreano, D. (2014). Artificial evolution by viability rather than competition. *PLoS ONE*, 9(1), e86831.

18. Maley, C. (1999). Four steps toward open-ended evolution. In W. Banzhaf, J. Daida, A. E. Eiben, M. H. Garzon, V. Honavar, M. Jakiela, & R. E. Smith (Eds.), *Proceedings of the Genetic and Evolutionary Computation Conference* (pp. 1336–1343). Orlando, FL: Morgan Kaufmann.
19. Matin, A. (1978). Organic nutrition of chemolithotrophic bacteria. *Annual Reviews in Microbiology*, 32(1), 433–468.
20. Moran, P. A. P. (1958). Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54(1), 60–71.
21. Olson, R. S., Mirmomeni, M., Brom, T., Bruger, E., Hintze, A., Knoester, D. B., & Adami, C. (2013). Evolved digital ecosystems: Dynamic steady state, not optimal fixed point. In P. Lio, O. Miglino, G. Nicosia, S. Nolfi, & M. Pavone (Eds.), *ECAL* (pp. 126–133). Cambridge, MA: MIT Press.
22. Rasmussen, S., Baas, N. A., Mayer, B., & Nilsson, M. (2001). Defense of the ansatz for dynamical hierarchies. *Artificial Life*, 7(4), 367–373.
23. Rasmussen, S., Baas, N. A., Mayer, B., Nilsson, M., & Olesen, M. W. (2001). Ansatz for dynamical hierarchies. *Artificial Life*, 7(4), 329–353.
24. Soros, L. B., & Stanley, K. O. (2014). Identifying necessary conditions for open-ended evolution through the artificial life world of chromaria. In H. Sayama, J. Rieffel, S. Risi, R. Doursat, & H. Lipson (Eds.), *Artificial life 14: Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems* (pp. 793–800). Citeseer.
25. Taylor, T. (2012). Exploring the concept of open-ended evolution. In C. Adami, D. M. Bryson, C. Ofria, & R. T. Pennock (Eds.), *Proceedings of the 13th International Conference on Artificial Life* (pp. 540–541). Cambridge, MA: MIT Press.
26. Taylor, T. (2015). Requirements for open-ended evolution in natural and artificial systems. arXiv preprint arXiv:1507.07403.
27. Taylor, T., Bedau, M., Channon, A., Ackley, D., Banzhaf, W., Beslon, G., Dolson, E., Froese, T., Hickenbotham, S., Ikegami, T., McMullin, B., Packard, N., Rasmussen, S., Virgo, N., Agmon, E., Clark, E., McGregor, S., Ofria, C., Ropella, G., Spector, L., Stanley, K. O., Stanton, A., Timperley, C., Vostinar, A., & Wisner, M. (2016). Open-ended evolution: Perspectives from the OEE workshop in York. *Artificial Life*, 22(3), 408–423.
28. Turing, A. M. (1937). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 2(1), 230–265.
29. Wisner, M. J., Ribbeck, N., & Lenski, R. E. (2013). Long-term dynamics of adaptation in asexual populations. *Science*, 1243357.
30. Ziv, J., & Lempel, A. (1978). Compression of individual sequences via variable-rate coding. *IEEE Transactions on Information Theory*, 24(5), 530–536.