Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 29, Issue Number 4, February 2012

Comment

Merits and Caveats of Using A Vocabulary Approach to Define Life

http://www.jbsdonline.com

Radu Popa

Portland State University,
P.O. Box 751, Portland,
OR, 97207, USA

The field of "defining life" is rich in artful wording yet lacking in cohesiveness. Opinions about the necessity and possibility to produce a definition of life range widely. At the extremes (and without necessarily being wrong) some authors may claim that a technically accurate definition of life is not needed or impossible. Assuming that a definition for life is needed and possible, it has to obey two basic requirements: to be coherent relative to what we already know about life, and the philosophy used to produce the definition has to be clear of systemic errors. The method proposed by Edward Trifonov to define life (1) is a minimalist vocabulary-based screening, combined with a personal interpretation of the meaning of the findings. Is this method consistent with the requirements listed above, and what novelties it introduces in our understanding of life?

The method proposed is seemingly simple. Take a large collection of definitions of life and calculate the frequency of different words (Table I in the original article). Identify the most common words and the words with similar meaning (Table II in the original article) and combine them in a first-hand definition. Then, shorten by eliminating terms and concepts that imply each other, in a way that allows essence and causes rather than trivia and consequences to be retained in a final definition.

Following this logic, the manuscript (1) should have ended with the analogy with Principal Component Analysis (p. 262). The author however finds more thrust to continue the manuscript past this point, by discussing issues such as thermo-stability, GC rich sequences, information complementarity and RNA-related "almost precise" replication. If the aim of this study (1) was to summarize why RNA-related molecules are important for biological life on Earth, then I find it insightful. If one however expected for this study to be a non-earth-centric attempt to define life, the focusing of the closing arguments on physical-chemical and informational properties of a particular class of molecules is distracting and unfulfilling.

It is also necessary to analyze the validity of the key premise of this paper. It is obvious that the scientific community cannot bring itself together to produce and support a singular definition for life. The motives and the diversity of various opinions are not discussed here, only their consequences on a vocabulary-based strategy to define life. The essence of a vocabulary method is that words and ideas that are the most common must also be the most important. This is true to a point. It does apply very well to fields where basic research has more or less ended, yet it makes it difficult for pioneers and novel theories to gain recognition, irrespective of how right they are. For example, if we promote a scientific model based

Corresponding author:
Radu Popa
E-mail: rpopa@pdx.edu

on popularity, then Alfred Wegener was correctly ignored in 1912, when he promoted the concept of continental drift. One example from the field of the origin of life is Kunin's contribution, which was used to explain the initiation of a living system as a symbiosis between two molecular networks (*e.g*. a protein-made RNA polymerase cooperating with an RNA-made protein polymerase) (2). This great contribution to understanding the essence of life would be considered "unworthy of mentioning" by a vocabulary-based method, simply because it is seldom cited in origin of life models.

If we leave it to a computer-selected one-word-at-a-time vocabulary to define the essence of life then some seldom, yet worthy, characteristics of life would be dropped as little relevant. Take for example complex concepts such as "adaptive evolution", "cryptic information" and "energy dissipative systems". The combination "adaptive evolution" is richer in meaning than the term evolution. Evolution and variability exist in both alive and non-alive systems, yet living systems have an edge for survival by being built for "adaptive evolution". Also, in all life forms that we recognize on earth today the genetic information is encoded. This is a very important attribute of life because it allows living entities to accumulate a collection of virtual realities (dissimilar states) that are all possible but do not interfere with each other, because when one is expressed all others are hidden. Obviously, the simple term information is too poor in meaning to describe the complexity of what life does. Lastly, if we discard the connection between the energy dissipative properties of living systems and their capacity for self-control we loose the very reason for the origin of life (3, 4).

Can life be reduced to a collection of dimensionless qualities, or it has some properties that require reaching specific size before a system can become alive? The obvious example is the complexity-level analysis made by Stuard Kauffman (5), which a vocabulary analysis will simply ignore. The lesson learned from Dr. Trifonov's approach is that we can circumvent such limitations by analyzing major consequences of being alive rather than analytically dissect and list all life's properties and achievements.

At the syntactic level a definition cannot be constructed as a popular saying that leaves grammatical parts in limbo. In order to be useful, the construction of a definition has to obey a specific set of minimal rules. The expression, "*Life is self-reproduction with variations*" albeit inspirational, does not inform whether life is a physical system or the property of a physical system. This "definition" does not clarify whether these features are restricted to life or if they may also occur in other systems that are not alive. Lastly, is this "definition" sufficient to explain life or is it just one important aspect of it? The analysis from (1) correctly replaces this truncated definition with a grammatically correct one developed by Oparin (6): "Any system capable of replication and mutation is alive". This is not however sufficient reason to accept Oparin's definition of life as complete, because it still lacks many important properties of life.

Last but not least, can life exist that is not RNA-based? If the answer is YES, then no need exists to pound on the RNA-world drum in order to explain what life is. If NOT, then the author (1) has to state that no alien life can exist unless it is based on RNA-like molecules, which everybody will probably doubt. To summarize, one recognizes in a vocabulary method the need to resolve present ambiguities about defining life though, in my opinion, not the best avenue for reaching this goal. This aside, the reader interested in the subject of defining life and explaining its early evolution will find sufficient substance in (1) to make this article worth reading and instructive. Ultimately, "*We may never agree on a definition of life, which will remain forever subject to a personal perspective. The measure of one's scientific maturity may actually be his/her latest definition of life and the acceptance that it cannot be ultimate*" (7).

References

1. E. N. Trifonov. *J Biomol Struct Dyn 29*, 259-266 (2011).
2. V. Kunin. *Orig Live Evol Biosph 30*, 459-466 (2000).
3. P. Bak. How nature works. Copernicus, New York (1996).
4. E. Chaisson. *Int J Astrobiol 2*, 91-101 (2002).
5. S. A. Kauffman. The origins of order, Oxford University Press, New York (1993).
6. A. I. Oparin. Life: its nature, origin and development, Academic Press, New York (1961).
7. R. Popa. *Orig Live Evol Biosph 40*, 183-190 (2010).