Journal of Theoretical Biology 373 (2015) 111–116

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

How probable is common ancestry according to different evolutionary processes?

Elliott Sober [a], Mike Steel [b,*]

[a] Philosophy Department, University of Wisconsin, Madison 53706, WI, USA
[b] School of Mathematics and Statistics, Allan Wilson Centre, University of Canterbury, Christchurch 8140, New Zealand

HIGHLIGHTS

- Darwin and contemporary biologists argue that all present-day life traces back to one or a few common ancestors.
- We investigate the relationship of different evolutionary processes to the hypothesis of common ancestry.
- We describe how different evolutionary processes confer different probabilities on the common ancestry thesis.

ARTICLE INFO

Article history:
Received 17 November 2014
Received in revised form
9 March 2015
Accepted 10 March 2015
Available online 18 March 2015

Keywords:
Common ancestry
Darwin
Divergence
Extinction
Natural selection

ABSTRACT

Darwin and contemporary biologists argue that all present-day life traces back to one or a few common ancestors. Here we investigate the relationship of different evolutionary processes to this hypothesis of common ancestry. We identify the property of an evolutionary process that determines what its probabilistic impact on the common ancestry thesis will be. The point of this exercise is to understand how the parts of Darwin's powerful theory fit together, not to call into question common ancestry or natural selection, since these two pillars of Darwin's theory enjoy strong support.

1. Introduction

Darwin's views about natural selection (including his claim that selection is the main but not the exclusive cause of evolution) and his thesis about common ancestry (that everything now alive on earth traces back to one or a few original progenitors) are logically independent of each other (Mayr, 2007), but that did not prevent Darwin from illustrating them both in the single figure that appears in the *Origin of Species*. That diagram (which occurs on a fold-out page immediately following p. 116 of Darwin, 1859) is redrawn below in Fig. 1; there are 11 ancestors (labeled A, B, C, …, L) at the bottom of the page and 15 descendants at the top. Those 15 trace back to only 3 of the original 11. Why did the lineages stemming from 8 of the original 11 go extinct? Darwin's answer is natural selection. The horizontal axis in Darwin's figure represents a quantitative phenotype. Notice that when an ancestor produces several offspring in Darwin's drawing, the offspring with extreme phenotypes are usually the ones that go on to have offspring of their own; those with intermediate phenotypes usually do less well. This is Darwin's principle of divergence. This selection process not only causes organisms in the present to trace back to a small number of ancient ancestors; it also serves to increase life's diversity. The 15 descendants at the top of the diagram have more horizontal spread than the 11 at the bottom. Darwin says that his principle of divergence describes what “tends” to happen in processes of natural selection, not what happens invariably. This is why he includes in his figure the lineage stemming from F. F and its present-day descendant have the same phenotype.

Darwin (1859, pp. 111–126) gives different characterizations of his principle of divergence (Kohn, 2009). In addition to the idea that selection favors extreme phenotypes, he says that selection favors parents that diversify their offspring and that it favors organisms that diversify their own internal structure. These three formulations are logically independent of each other. Current biology views two of them with reserve. Selection often favors extreme phenotypes, but it also often favors intermediate phenotypes; a classic example of the latter is birth weight in humans (Bell, 1997). Natural selection

---

* Corresponding author. Fax: +64 33642587.
E-mail address: mike.steel@canterbury.ac.nz (M. Steel).

http://dx.doi.org/10.1016/j.jtbi.2015.03.012

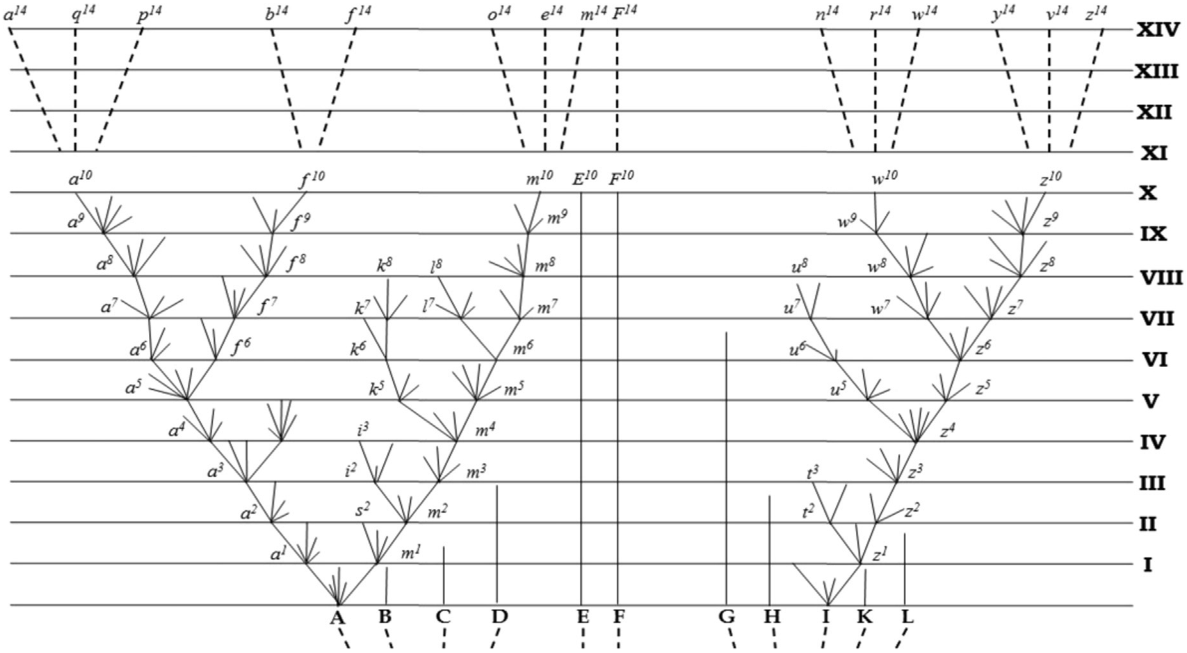

Fig. 1. The only illustration in Darwin's *Origin* represents both common ancestry and the principle of divergence.

sometimes is disruptive and sometimes it is stabilizing; there is no a priori reason to think that one is common and the other rare. A similar reservation can be lodged against the idea that selection favors organisms that exhibit greater internal diversity. Selection sometimes promotes increased division of labor (specialization of parts), but often it does not. Darwin (1859, p. 148) notes in his discussion of parasites that selection sometimes favors simplification; parasites often lose structures that were present in their free-living ancestors. The idea in Darwin's trifecta that corresponds most closely to current biology is his suggestion that parents that diversify their offspring will be favored over parents that do not. This idea is alive and well in discussions of the evolution of sexual reproduction (Burt, 2000); organisms that reproduce sexually will do better than organisms that reproduce asexually when the environment is sufficiently unpredictable.

The question we will investigate here—how different evolutionary processes (including different forms of natural selection) confer probabilities on the thesis of common ancestry–is distinct from the question of which observations lend strong support to natural selection and which do so for common ancestry. Darwin addresses this last question in the following passage from the *Origin*:

> … adaptive characters, although of the utmost importance to the welfare of the being, are almost valueless to the systematist. For animals belonging to two most distinct lines of descent, may readily become adapted to similar conditions, and thus assume a close external resemblance; but such resemblances will not reveal–will rather tend to conceal their blood-relationship to their proper lines of descent (Darwin, 1859, p. 427).

Darwin's idea is that adaptive similarities provide scant evidence for common ancestry, whereas similarities that are neutral or deleterious provide evidence that is more weighty. As noted by Sober and Steel (2014), a simple likelihood comparison supports Darwin's claim about adaptive similarities under a rather general model of evolution. Supposing two extant taxa A and B share a trait x, let us consider the likelihood ratio $LR_{CA/SA}$ of the following two hypotheses:

(CA) Taxa A and B have a most recent common ancestor that existed t units of time in the past, trait x was present in this ancestor with some probability p, and the trait's evolution down each of the two lineages leading from that ancestor to A and to B followed continuous-time stochastic processes.

(SA) Taxa A and B do not trace back to a common ancestor, state x was independently present in these two taxa with probability p at t time units in the past, and the trait's subsequent evolution down these two lineages to the present followed continuous-time stochastic processes.

Sober and Steel point out that if t is small, then:

$$LR_{CA/SA} = \frac{\Pr(A, B \text{ share trait } x|CA)}{\Pr(A, B \text{ share trait } x|SA)} \approx p/p^2 = 1/p.$$

If an adaptive trait has a higher value for p than a neutral trait does, and if a neutral trait has a higher value for p than a deleterious trait possesses, then the value of $LR_{CA/SA}$ is higher for neutral and deleterious traits than it is for traits that are adaptive, thus vindicating Darwin's statement about adaptive similarities. This argument has two limitations: it requires t to be small and it considers only two taxa. While the value of $LR_{CA/SA}$ is not the main subject of the present paper, we expand on the $1/p$ argument by providing an exact expression for the likelihood ratio for two taxa when the evolving trait has two states; we also provide a bound on the ratio that applies for any number of taxa when the underlying continuous-time stochastic process is a stationary Markov process. Proofs for both are given in the Appendix.

Proposition 1. *For the evolution of a trait under a stationary continuous-time Markov process on two states*:

$$LR_{CA/SA} = 1 + \left(\frac{1}{p} - 1\right)e^{-2r\cdot t},$$

where r is a rate parameter associated with the model. Moreover, if n present taxa have their most recent common ancestor at t time units