Journal of Theoretical Biology 371 (2015) 127–136

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

Phenotypic plasticity, the Baldwin effect, and the speeding up of evolution: The computational roots of an illusion

Mauro Santos [a,*], Eörs Szathmáry [b,c,d], José F. Fontanari [e]

[a] Departament de Genètica i de Microbiologia, Grup de Genòmica, Bioinformàtica i Biologia Evolutiva (GGBE), Universitat Autonòma de Barcelona, 08193 Bellaterra (Barcelona), Spain
[b] Parmenides Center for the Conceptual Foundations of Science, Kirchplatz 1 Pullach, D-82049 Munich, Germany
[c] Biological Insititute and Research MTA-ELTE Theoretical Biology and Evolutionary Ecology Research Group, Eötvös University, Budapest, Hungary
[d] Faculty of Biology, Ludwig Maximilians University Munich, Munich, Germany
[e] Instituto de Física de São Carlos, Universidade de São Paulo, Caixa Postal 369, 13560-970 São Carlos, SP, Brazil

HIGHLIGHTS

- Some current debates revolve around the role of genes as leaders or followers in the evolutionary process.
- The tenet that phenotypes are leaders and genes are followers is an old one and relates to the Baldwin effect.
- This effect was not part of evolutionary thinking during the Modern Synthesis in the 1930s and 1940s.
- People working in evolutionary computation revitalized the Baldwin effect by showing how and why it could work.
- A standard population genetics treatment can achieve what many people thought was unsolvable without the Baldwin effect.

ARTICLE INFO

Article history:
Received 25 November 2014
Received in revised form
4 February 2015
Accepted 9 February 2015
Available online 19 February 2015

Keywords:
Evolutionary search
Genetic algorithm
Learning
The Baldwin effect
Speed of evolution

ABSTRACT

An increasing number of dissident voices claim that the standard neo-Darwinian view of genes as 'leaders' and phenotypes as 'followers' during the process of adaptive evolution should be turned on its head. This idea is older than the rediscovery of Mendel's laws of inheritance, with the turn-of-the-twentieth-century notion eventually labeled as the 'Baldwin effect' as one of the many ways in which the standard neo-Darwinian view can be turned around. A condition for this effect is that environmentally induced variation such as phenotypic plasticity or learning is crucial for the initial establishment of a trait. This gives the additional time for natural selection to act on genetic variation and the adaptive trait can be eventually encoded in the genotype. An influential paper published in the late 1980s claimed the Baldwin effect to happen in computer simulations, and avowed that it was crucial to solve a difficult adaptive task. This generated much excitement among scholars in various disciplines that regard neo-Darwinian accounts to explain the evolutionary emergence of high-order phenotypic traits such as consciousness or language almost hopeless. Here, we use analytical and computational approaches to show that a standard population genetics treatment can easily crack what the scientific community has granted as an unsolvable adaptive problem without learning. Evolutionary psychologists and linguists have invoked the (claimed) Baldwin effect to make wild assertions that should not be taken seriously. What the Baldwin effect needs are plausible case-histories.

1. Introduction

What role does the Baldwin effect play in evolution? By Baldwin effect —a term coined by Simpson (1953)— we refer to a turn-of-the-twentieth-century idea (Baldwin, 1896; Morgan, 1896; Osborn, 1896) cogently described by Maynard Smith (1987, p. 761) as follows: "If individuals vary genetically in their capacity to learn, or to adapt developmentally, then those most able to adapt will leave most descendants, and the genes responsible will increase in frequency. In a fixed environment, when the best thing to learn remains constant, this can lead to the genetic determination of a character that, in earlier generations, had to be acquired afresh in each generation". The Baldwin effect

* Corresponding author.
E-mail addresses: mauro.santos@uab.es (M. Santos), szathmary.eors@gmail.com (E. Szathmáry), fontanar@ifsc.usp.br (J.F. Fontanari).

http://dx.doi.org/10.1016/j.jtbi.2015.02.012
0022-5193/

involves two transitions (Turney et al., 1997; Godfrey-Smith, 2003): the first has to do with the evolutionary value of phenotypic plasticity, or some particular form of plasticity such as learning; the second with the 'genetic accommodation' [i.e., evolution in response to both genetically based and environmentally induced novel traits (Griffiths, 2003; West-Eberhard, 2003; Crispo, 2007)] of the learned trait. We use genetic accommodation instead of the more familiar term 'genetic assimilation' coined by Waddington (1953) because this last term should not be equated to the Baldwin effect (Crispo, 2007; see also West-Eberhard, 2003, pp. 153–154). Genetic assimilation can work with pre-existing genetic variation; the Baldwin effect (as originally posited) requires a new gene or genes.

Some towering figures in the Modern Synthesis —an expression borrowed from the title of Julian Huxley's (1942) book— were either indulgent with the theoretical plausibility of the Baldwin effect (Simpson,1953) or utterly hostile towards it, recommending to discard this concept altogether (Mayr, 1963; Dobzhansky, 1970). This advice is followed by several influential textbooks in evolutionary biology (Ridley, 2004; Futuyma, 2005; Barton et al., 2007) that do not even mention Baldwin at all. However, although at present there appears to be no clear empirical evidence for the Baldwin effect, several authors have called for a radical revision of the consensus view and argued that much evolution involves genetic accommodation (Schlichting and Pigliucci, 1998; Avital and Jablonka, 2001; West-Eberhard, 2003; Schlichting and Wund, 2014; but see Braendle and Flatt, 2006). The current tension among evolutionary biologists (Laland et al., 2014) is unmatched by scholars working in evolutionary computation (Mitchell, 1996; Back et al., 1997; Turney et al., 1997) and in others disciplines (typically evolutionary psychologists and cognitive scientists; Weber and Depew, 2003), who invoke the Baldwin effect as a major evolutionary force that could have led to the emergence of mind (Dennett, 1995; Deacon, 1997; Pinker, 1997) and to modern language (Pinker and Bloom, 1990; Briscoe, 1997; Calvin and Bickerton, 2000; Dor and Jablonka, 2001). As Yamauchi (2004, p. 3) put it, "the Baldwin effect is particularly appealing because … It may provide a natural Darwinian account for language evolution: It is an especially popular idea among linguists that language evolution is somehow saltational. This leads them to conclude that *neo*-Darwinian theories are 'incompetent' for accounting for language evolution" (our addition in italics). (Neo-Darwinism is used here to describe the Modern Synthesis version of Darwinism.)

Much of the recent 'excitement' about the Baldwin effect stems from a seminal paper published by computer scientists Geoffrey Hinton and Steven Nowlan in the late 1980s (Hinton and Nowlan, 1987), which has been cited 1,101 times (Google Scholar) to date. They developed a computational model combining a genetic algorithm with learning by trial and error in a sexual population of chromosomes (the 'organisms') that were initially segregating at $L=20$ loci with three alleles each: 1, 0, and ? This chromosome determines the connectivity of a neural network: allele 1 at a given locus indicates that a particular connection exists whereas allele 0 at that locus indicates that it does not. The question marks are plastic alleles that allow the organism to set (or not) the connection at the end of a learning period. The neural network has only one correct configuration of connections and the task the organisms had to solve was to find this configuration out of the $2^L \approx 10^6$ possible configurations. We can assume without loss of generality that the right answer is the chromosome with all alleles1; i.e., a fully connected neural network. The catch is that any other configuration provides no information whatsoever about where the correct answer might be. In such problems, there is no better way to search than by exhaustively sampling the entire combinatorial space; a situation termed a 'needle-in-the-haystack' problem. In other words, there is no efficient algorithm that can find the fitness maximum unless we introduce some 'trick'; namely, to somehow smooth the spiked fitness landscape through phenotypic plasticity (Frank, 2011). Hinton and Nowlan (1987) assumed that each organism could try up to a maximum of $G=1000$ random guesses for the settings of the ? states; these alleles define the 'plastic genome'. The organisms were also given the ability to recognize whether they have found the correct settings after $g < G$ learning trials and, in such a case, stop guessing (see below for details). Therefore, those organisms that were relatively fast at learning the correct configuration of alleles enjoyed a fitness advantage and produced more offspring. In the long run —well before 50 generations in the simulation performed by Hinton and Nowlan (1987)—, natural selection redesigned the genotypes in the population and the correct alleles 1 increased in frequency. Nonetheless, they did not take over and undecided alleles ? remained segregating at relatively high frequency because in the end organisms were able to learn quickly and, therefore, there was not much selective pressure to fix the 'innately correct' fitter alleles. (Note, therefore, that the model that worked was the one in which what really mattered was how close an organism got to the correct configuration.)

The scenario in Hinton and Nowlan (1987) showed (i) that the Baldwin effect can be observed *in silico*, and (ii) that once the wrong 0 alleles are being eliminated by selection learning can dramatically accelerate adaptive evolution in a flat fitness landscape with a single isolated peak; what Ancel (2000) characterized as the 'Baldwin expediting effect'. Maynard Smith (1987, p. 762) explained this effect by making a simple contrast with a population where organisms do not learn: "In a sexual population of 1000 with initial allele frequencies of 0.5, a fit individual would arise about once in 1000 generations … Mating would disrupt the optimum genotype, however, and its offspring would have lost the adaptation. In effect, a sexual population would never evolve the correct settings… (or does so excessively slowly)". Actually, "the problem was never solved by an evolutionary search without learning" (Hinton and Nowlan, 1987, p. 497). Conversely, Maynard Smith (1987) claimed that in the absence of learning a large asexual population would include optimal individuals and the correct settings would soon be established by selection.

The first claim about non-learning sexual organisms has been taken for granted, whereas the second claim concerning asexual organisms was analytically investigated by Fontanari and Meir (1990) to answer the question: how soon is 'soon'? Using their recursion equation (3.1) to analyze the evolution of correct alleles, the answer is that it would take more than 3000 generations for the population to evolve the correct settings with initial allele frequencies 0.5 and no mutation. Therefore, the conclusion seems to be fairly clear: in the single-peaked fitness landscape assumed by Hinton and Nowlan (1987) learning has a drastic effect on evolution.

Here, we show that this conclusion is generally incorrect and requires careful considerations. The heart of the problem was also pointed out by Maynard Smith (1987) and relates to the strong positive epistasis in Hinton and Nowlan's (1987) scenario. This epistasis generates, in turn, strong positive associations between the correct alleles in the non-learning organisms that can greatly accelerate evolution (Appendix A). The former solution of more than 3000 generations for the asexual population to evolve the correct settings is likely to be a gross overestimate as Fontanari and Meir (1990) ignored the generation of linkage disequilibrium due to directional selection. The remainder of the paper is organized as follows. First, we discuss Hinton and Nowlan (1987) model in more detail as it will make the reason for our skepticism about what they have really demonstrated very clear. Second, we derive the exact recursion equations for the asexual case and show that evolution is indeed quite fast in this case. Third, challenging the conventional wisdom we show that a finite population of sexual organisms that do not learn does evolve the correct settings, and estimate the probability of fixation and mean time to fixation of the correct genotype as a function of population size N and chromosome length