BioSystems 123 (2014) 106–113

Contents lists available at ScienceDirect

BioSystems

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

Punctuated evolution and robustness in morphogenesis

D. Grigoriev[a], J. Reinitz[b,c,d,e], S. Vakulenko[f,g,*], A. Weber[h]

[a] CNRS, Mathématiques, Université de Lille, Villeneuve d'Ascq 59655, France
[b] Department of Statistics, University of Chicago, Chicago, IL 60637, United States
[c] Department of Ecology and Evolution, University of Chicago, United States
[d] Department of Molecular Genetics and Cell Biology, University of Chicago, United States
[e] Institute for Genomics and Systems Biology, University of Chicago, United States
[f] Institute for Mechanical Engineering Problems, Bolshoy pr. V. O.61, Sankt Petersburg, Russia
[g] ITMO University, Sankt Petersburg, Russia
[h] Computer Science Department, University of Bonn, 53113 Bonn, Germany

ARTICLE INFO

Article history:
Received 8 April 2014
Received in revised form 25 June 2014
Accepted 29 June 2014
Available online 1 July 2014

Keywords:
Genetic networks
Evolution
Hard combinatorial problems
Robustness
Canalization

ABSTRACT

This paper presents an analytic approach to the pattern stability and evolution problem in morphogenesis. The approach used here is based on the ideas from the gene and neural network theory. We assume that gene networks contain a number of small groups of genes (called hubs) controlling morphogenesis process. Hub genes represent an important element of gene network architecture and their existence is empirically confirmed. We show that hubs can stabilize morphogenetic pattern and accelerate the morphogenesis. The hub activity exhibits an abrupt change depending on the mutation frequency. When the mutation frequency is small, these hubs suppress all mutations and gene product concentrations do not change, thus, the pattern is stable. When the environmental pressure increases and the population needs new genotypes, the genetic drift and other effects increase the mutation frequency. For the frequencies that are larger than a critical amount the hubs turn off; and as a result, many mutations can affect phenotype. This effect can serve as an engine for evolution. We show that this engine is very effective: the evolution acceleration is an exponential function of gene redundancy. Finally, we show that the Eldredge–Gould concept of punctuated evolution results from the network architecture, which provides fast evolution, control of evolvability, and pattern robustness. To describe analytically the effect of exponential acceleration, we use mathematical methods developed recently for hard combinatorial problems, in particular, for so-called *k*-SAT problem, and numerical simulations.

1. Introduction

Robustness is an important property of biological systems. Wild-type organisms are buffered, or "canalized", against environmental or genetic variation during the course of both development (Manu et al., 2009b) and evolution (Rendel, 1959). This term was coined by Waddington (1942), who stated that "developmental reactions, as they occur in organisms submitted to natural selection. . .are adjusted so as to bring about one definite end-result regardless of minor variations in conditions during the course of the reaction". More recent work has shed light on the mechanistic origins of canalization behavior. In some cases a specific gene, called a "genetic capacitor" is responsible for canalizing behavior (Bergman and Siegal, 2003; Levy and Siegal, 2008; Moczek, 2007) while in other cases the canalization behavior arises from a small network of genes (Manu et al., 2009a). These genes or networks of genes buffer environmental or genetic variations, thus canalizing pattern formation and evolution. This situation implies an apparent paradox, because canalized systems are nevertheless able to evolve successfully to adapt to environmental changes. Different mechanisms of canalization have been discussed, for example, in Gunji and Ono (2012), Gursky et al. (2012), Gunji et al. (2014). In the paper by Gunji and Ono (2012), a cellular automata-based model is proposed to generate a French flag pattern that arises as a model of canalization due to agents equipped with sociality. In this model cell can be considered to be an agent that transports morphogens. A pattern occurs as a result of interaction of neighboring agents. The paper by Gursky et al. (2012) contains a review of the canalization problem, describes different canalization mechanisms and

* Corresponding author at: Institute for Mechanical Engineering Problems, Bolshoy pr. V. O.61, Sankt Petersburg, Russia. Tel.: +7 9215842914.
E-mail addresses: dmitry.grigoryev@math.univ-lille1.fr (D. Grigoriev), reinitz@galton.uchicago.edu (J. Reinitz), vakulenfr@mail.ru (S. Vakulenko), weber@cs.uni-bonn.de (A. Weber).

http://dx.doi.org/10.1016/j.biosystems.2014.06.013
0303-2647/

considers canalization as a result of gene interactions. Here, we develop these ideas and suggest that the basis of robustness and canalization is a network architecture in which the genes or small networks responsible for canalization are hubs in larger gene networks. Recall that the hubs are strongly connected nodes in networks (Albert and Barabási, 2002), which is a key element of network architecture. In systems level studies of metabolic networks, this architecture has been described as "bow-tie" connectivity (Zhao et al., 2006). Here, we use a variant of this idea, called an "empire structure" by Vakulenko (2013), wherein highly connected hubs play the role of organizing centers, and each center interacts with many weakly connected nodes, which are called satellites (see Fig. 2). Note that hubs are a universal feature of scale-free networks, and have been identified in a wide variety of natural and human-generated networks, including the genetic, metabolic and economic networks, as well as the internet (Albert and Barabási, 2002; Lesne, 2006). In particular, centralized networks have been empirically identified in molecular biology, where the centers can be, for example, transcription factors, while the satellite regulators can be small regulatory molecules such as microRNAs (Li et al., 2010) or the target genes of the *Drosophila* segmentation network.

In this work, we address the apparent paradox of evolutionary change in the face of canalizing stability by means of an analytically tractable mathematical model of the canalization effect and its abrogation. We show that populations that experience an increase in the mutation rate or pass through a bottleneck, possibly because of environmental stress, release hidden genetic information. Specifically, hubs in the network can create an abrupt change or transition. In normal conditions, the hubs stabilize the gene expression pattern against all mutations. When the mutation rate increases or the population size N decreases, the mutation probability p becomes higher (specifically, this effect can result from the genetic drift effect because the drift induces noise in the intensity that is proportional to $N^{-1/2}$). More specifically, there is a threshold effect in which—when the probability p of a mutation becomes more than some critical value p_c, i.e., $p > p_c$—the hub stabilization fails. This effect produces an abrupt change and it is sharper for large gene networks. At this point, hidden genetic information is manifested in the phenotype in such a way that hidden mutations that are captured during a long stability period start begin to play a role in creating new phenotypes that are better adapted to new ecological conditions. This evolutionary mechanism corresponds to the punctuated evolution ideas of Eldredge and Gould (1972). Here we show that punctuated evolution is a natural consequence of an empire-type network organization.

In our approach to this problem, we also address a serious difficulty in evolutionary theory, namely that classical population genetics does not consider the hierarchical organization of multicellular organisms into differentiated cell types that in turn constitute tissues and organs. Classical population genetics instead treats the organism as a unitary entity that possesses organismal fitness. This formulation avoids the problem that a condition that could be selectively advantageous for a specific cell type with a specific gene expression state is selectively disadvantageous for another cell type that has a different gene expression state, where both cell types contain the same genetic material. In this work, we represent the selective effects of mutations on a hierarchically organized multicellular organism by using an idea from theoretical computer science (Valiant, 2009). Specifically, in this paper Valiant (2009) has found a connection between the evolution problems and the fundamental computer science problem P $\neq$ NP (Cook, 1971); this connection helps to formulate the problem in a more rigorous mathematical way.

Our approach to the problem of understanding punctuated equilibrium in the evolution of multicellular organisms in the face of canalization is based on an analogy between these evolution processes and hard-combinatorial problems. In recent decades these problems have received a large amount of attention from mathematicians and theoretical physicists (Friedgut and Bourgain, 1999; Deroulers and Monasson, 2006; Achlioptas, 2001; Mertens et al., 2006; Mézard and Zecchina, 2002). This analogy enables us to obtain an analytical relation for the evolution speed as a function of the gene redundancy.

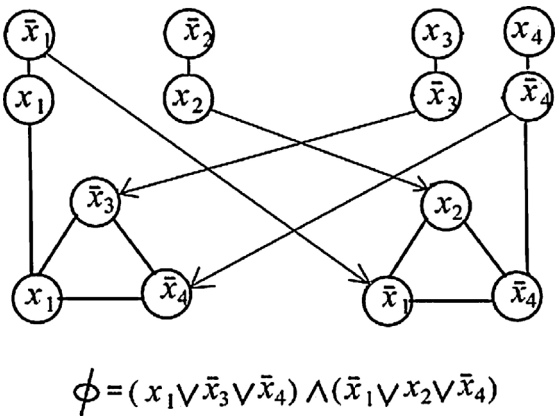

Fig. 1. This image illustrates a toy k-SAT problem for $n = 4$ logical variables x_1, x_2, x_3, x_4 for $k = 3$ and $m = 2$. The clauses are triangles, here we have the two clauses of length 3. An assignment of the logical variables is shown by arrows. This assignment is correct and gives a solution of the problem if all the clauses are true. In this case we can take, for example, $x_1 = 0$, $x_2 = 0$, $x_3 = 1$ and $x_4 = 1$. This toy example is easy to resolve, however, the problem becomes difficult for n, $m \gg 1$. In our biological model, we interpret x_i as genes and clauses as gene expression patterns for different cell types.

The connection of evolution with hard combinatorial problems allows us to formulate a precise statement of the meaning of *feasible* evolution (Valiant, 2009). In the framework of this model, evolution is feasible if one can find a local search algorithm, for example, a greedy one, that resolves the problem in Poly(n) elementary steps. These steps can be thought of as single mutations. This idea, which connects the P $\neq$ NP problem with evolutionary biology was first formulated by Valiant (2009). Note that from this paper we have used only idea about connection between evolution feasibility and algorithm feasibility, and also a connection with problem P $\neq$ NP, but that our model significantly differs from the model used in Valiant (2009).

Note that if P = NP, an equality that most do not believe to be correct, then evolution is always feasible.

The most fundamental hard-combinatorial problem is the famous k-SAT one. We state some important facts about k-SAT in the next section.

2. *k*-SAT problem

2.1. Formulation of the problem

Cook (1971) have shown that k-SAT problem is NP-complete, moreover, this problem has important applications for bioinformatics.

The k-SAT problem can be formulated as follows. Let us consider the set $V_n = \{x_1, \ldots, x_n\}$ of Boolean variables $x_i \in \{0, 1\}$ and a set C_m of m clauses. The clauses C_j are disjunctions (logical ORs) involving k literals $y_{i_1}, y_{i_2}, \ldots, y_{i_k}$, where each y_i is either x_i or the negation $\bar{x}_i$ of x_i. The problem is to test whether one can satisfy all of the clauses by an assignment of Boolean variables.

It can be illustrated by the following picture (see Fig. 1).

A biological interpretation of k-SAT is quite transparent and can be formulated as follows. The number n is the gene number. Each gene is involved in the formation of many differentiated cell types,