Journal of Theoretical Biology 420 (2017) 200–212

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

Reconstructing the genotype-to-fitness map for the bacterial chemotaxis network and its emergent behavioural phenotypes

Alberto Nakauma*, G. Sander van Doorn

Groningen Institute for Evolutionary Life Sciences, University of Groningen, 9700 CC Groningen, The Netherlands

ARTICLE INFO

Keywords:
Fitness landscape
Complex traits
Network evolution
Evolutionary systems biology
Compensatory mutations

ABSTRACT

The signal-transduction network responsible for chemotaxis in *Escherichia coli* has been characterised in extraordinary detail. Yet, relatively little is known about eco-evolutionary aspects of chemotaxis, such as how the network has been shaped by selection and to what extent natural populations may fine-tune their chemotactic behaviour to the ecological conditions. To address these questions, we here develop an evolutionary-systems-biology model of the chemotaxis network of *E. coli*, which we apply to estimate the resource accumulation rate (here used as a proxy for fitness) of wildtype and a large number of potential mutant genotypes. Mutant genotypes differ from the wildtype in the concentrations of one or more constituent proteins of the chemotaxis signalling network or in one or more of its kinetic parameters. To guarantee model consistency across the genotype space, we explicitly incorporated biochemical constraints that underly observed phenotypic trade-offs. The model was validated by reconstructing the phenotypic properties of several known mutant genotypes. We also characterised differences in the fitness distribution between genotypes, and reconstructed adaptive walks in genotype space for populations exposed to different environmental conditions. We found that the local fitness landscape is rugged, due to non-additive interactions between mutations. When selection has a consistent direction, just a few adaptive mutations are required to reach a local peak, and different local peaks can be reached by adaptive walks starting from the same initial genotype. However, when the direction of selection is fluctuating, evolutionary paths are much longer and genotype space is explored further. Longer adaptive walks were also observed when evolution was started from a low-fitness genotype such as a *CheZ* knockout mutant. In line with empirical observations, the initial Δ*cheZ* mutant did not respond to a step-down stimulus, but a dynamic response similar to the wildtype was recovered following the fixation of compensatory mutations.

1. Introduction

The movement of bacteria in response to environmental gradients of chemicals (bacterial chemotaxis), is governed by a small protein-protein interaction network that has evolved from a classical two-component bacterial signalling system (Wadhams and Armitage, 2004). The network has been extensively studied in *Escherichia coli* (Hazelbauer, 2012; Tindall et al., 2008), where detailed information is available on the receptors involved in the detection of stimuli, the proteins responsible for transmission of the signal and the swimming behaviour of cells (Wadhams and Armitage, 2004; Minamino et al., 2008; Porter et al., 2011; Hamer et al., 2010; Szurmant and Ordal, 2004; Berg, 2003; Thormann and Paulick, 2010; Kirby, 2009; Brown et al., 2011; Sourjik and Armitage, 2010). Chemotaxis of *E. coli* relies on a temporal comparison of chemical concentrations along the path of movement, which is used to bias transitions between phases of straight, smooth swimming (runs) and rotational motions (tumbles), that induce a random reorientation of the direction of movement (Segall et al., 1986).

In *E. coli*, attractants in the extracellular environment suppress the activity of receptors clustered at the cell poles (Sourjik, 2004). Active receptors trigger a phosphorylation cascade of cytoplasmic Che proteins. The first protein in the cascade, CheA, is coupled to the receptor. Phosphorylated CheA (CheAp) transfers its phosphate group to the methylesterase CheB or to the effector protein CheY. Phosphorylated CheY (CheYp) in turn binds to the flagellar motor complex, where it induces a change in the rotational state of the flagellum. The CheYp signal is deactivated by CheZ, a dedicated CheY-phosphatase. When the flagella are rotating counterclockwise (CCW), they form a single bundle that allows for smooth swimming. If one or several flagella rotate clockwise (CW), the bundle disintegrates, causing the cell to start tumbling.

A feature of *E. coli*'s chemotactic response that has received considerable attention is that it exhibits sensory adaptation (Segall et al., 1986; Alon

* Corresponding author.
E-mail address: j.a.nakauma.gonzalez@rug.nl (A. Nakauma).

http://dx.doi.org/10.1016/j.jtbi.2017.03.016
Received 22 November 2016; Received in revised form 10 February 2017; Accepted 16 March 2017
Available online 18 March 2017
0022-5193/

et al., 1999; Mello and Tu, 2003; Clausznitzer et al., 2010) (though the accuracy of adaptation is observed to vary between receptor types Wong-Ng et al., 2016). When a change in attractant concentration persists, the network will at first respond but then return to a state similar to the one it had before the stimulus was presented. This property allows bacteria to sense gradients and to stay in the vicinity of an attractant peak irrespective of the background concentration of attractant (Mello and Tu, 2007; Celani and Vergassola, 2010). The molecular mechanism responsible for sensory adaptation is receptor methylation, which is affected by a balance between the activities of two proteins: CheR, a methyltransferase that methylates certain residues of the receptor, and phosphorylated CheB (CheBp) (Kehry et al., 1985), which demethylates those same residues. Demethylated receptors exhibit a lower baseline activity level. Hence, the dependence of the demethylation rate on the activity of the receptor establishes a negative feedback loop resulting in sensory adaptation.

Most molecular studies of the chemotaxis network are conducted using well-documented *E. coli* lab strains and, as a result, little is known about the extent of phenotypic variation in the chemotaxis network that exists across natural populations. Variation in the concentrations of constituents of the network, resulting in measurable differences in swimming behaviour, has been observed within populations of isogenic bacteria and between genetically distinct isolates from natural populations (Dzinic et al., 2008; Davidson and Surette, 2008). At present, it is not clear how much of this variation is caused by genetic factors, whether it is a target of selection, and how it is related to ecological conditions (Davidson and Surette, 2008). Comparative phylogenomic studies, however, indicate that the chemotaxis network has been subject to extensive evolution, potentially in response to ecological selection pressures (Hamer et al., 2010; Szurmant and Ordal, 2004; Briegel et al., 2015). For example, in comparison to *E. coli*, *Bacillus subtilis* has several additional Che proteins involved in sensory adaptation and features a qualitatively different network topology, even though core constituents such as CheA, CheB and CheY have been conserved. Another species, *Rhodobacter sphaeroides* has evolved two parallel chemosensory pathways (Porter et al., 2011), which has been hypothesised to reflect the complex ecology of its habitat.

An additional motivation for analysing the evolution of the chemotaxis network is that chemotaxis provides a minimal model for investigating the adaptation of a complex trait. Selection on such traits acts on phenotypic properties that emerge from molecular interactions between genes, rather than on individual genes directly, posing a challenge to the classical gene-centred view of evolution. Chemotactic performance has a straightforward link to fitness and relies on emergent properties of the chemotaxis network including the precision of sensory adaptation, adaptation time and CW bias (Yi et al., 2000; Frankel et al., 2014), which are not simply associated with the characteristics of a single chemotaxis protein. At the same time, the underlying signalling network is characterised in detail and small enough to allow for an accurate computational reconstruction of the mapping from molecular mechanism to fitness (Loewe, 2009). In this way, bacterial chemotaxis presents a unique model system for studying the molecular evolution of emergent phenotypes and for developing novel methods for the functional analysis of molecular data.

As a case in point, we here develop an evolutionary-systems-biology model of the bacterial chemotaxis network of *E. coli*. We use this model to quantify the molecular and phenotypic signatures of selection on chemotactic performance and study how these are related to each other. Starting from a mechanistic quantitative description of molecular interactions, our model predicts fitness (i.e., chemotactic efficiency, measured as energy accumulation rate), based on the simulated swimming behaviour of bacteria. Finally, the model is applied to predict how bacteria may adapt evolutionarily to different environmental conditions by fine-tuning interactions in their chemotaxis network (López-Maury et al., 2008; de Vos et al., 2013).

2. Materials and methods

2.1. Overview and specific features of the model

Our goal is to build a mechanistic model of the chemotaxis network of *E. coli* and be able to quantify the fitness effects of mutations. To this end, we integrate several earlier models of chemotaxis, which concentrated on the activity of chemoreceptors (Shimizu and Bray, 2002; Lan et al., 2012), the phosphorylation cascade (Bray et al., 1993; Rao et al., 2004), its interaction with the flagellar motor (Duke et al., 2001) and the dependence of the swimming pattern on the rotational state of the flagella (Sneddon et al., 2012; Saragosti et al., 2012). Two aspects deserve special attention in light of the evolutionary focus of our analysis: (1) the specification of fitness as a function of chemotactic performance and (2), the effect of mutations on components of the chemotaxis network.

One approach to quantify fitness is to rely on observed statistical relationships between fitness and key phenotypes (Frankel et al., 2014; Soyer et al., 2006). However, this statistical approach may fail to capture non-additive interactions between phenotypic effects, is vulnerable to extrapolation errors when applied to novel genotypes and is difficult to generalise across different environments. The second approach, which we use here, is to derive fitness from a mechanistic model of swimming behaviour, in which the local attractant concentration is used as a proxy for the momentary rate at which the cell accumulates resources needed for growth (Soyer et al., 2006). A time-average of the attractant concentration can thus be used to estimate the average growth rate (taking into account potential costs of signal transduction in the chemotaxis pathway), which is a suitable measure for the evolutionary fitness of bacteria. A major drawback of this method is that it is computationally demanding.

A challenge to modelling the effects of mutation is that not enough data are available to formulate a sequence-level model of protein evolution from which the effects of specific mutations could be predicted. Therefore, we model mutations as heritable changes that modify the rate of interactions between proteins in the chemotactic signalling cascade. One class of mutations that we consider are those that affect only the concentration of a protein (including knockout mutations). To allow for such mutations, the concentration of a protein is considered as an evolutionary parameter that can change as a result of 'regulatory' mutations with no other phenotypic effects. A second class of mutations contribute to variability in protein interaction rates by affecting kinetic parameters, as a result of changes in the coding sequence, post-translational modification or other molecular genetic mechanisms. However, unlike protein concentrations, reaction rate parameters are subject to thermodynamic constraints, such as energy-conservation conditions. Without taking these into account (i.e, when mutations are allowed to act independently on the kinetic parameters of the model), mutation is likely to create thermodynamically inconsistent protein-interaction networks. In order to maintain consistency, we therefore model the effect of mutations based on their effect on the free energy of reactants in a reaction. The mutational effects on kinetic model parameters are calculated subsequently from the modified free energy values. A worked-out example is provided in Fig S1.

2.2. Model description

Below, we provide a description of our model, structured according to the different components of the chemosensory cascade. The model has 42 parameters, 15 of which were allowed to evolve (5 are associated with the total concentration of Che proteins and 10 are associated with energy parameters). For each of the evolving parameters (indicated in bold face below), we allowed for 16 different genetic variants (alleles), each linked to a specific value of the evolving parameter (Table S1). The other parameters, which were held fixed during the simulations, were given values based on previous models and estimates from the literature (Table S2 and Text S1).