Journal of Theoretical Biology 253 (2008) 355–362

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

The evolution of group-level pathogenic traits

Dominique Chu *

Computing Laboratory, University of Kent, Cornwallis, Canterbury CT2 7NF, UK

ARTICLE INFO

Article history:
Received 12 December 2007
Received in revised form
16 March 2008
Accepted 17 March 2008
Available online 19 March 2008

Keywords:
E. coli
Group-selection
Type-I fimbriae
Virulence factors

ABSTRACT

A group-selection model for the evolutionary origin of phase-variation in *E. coli* is proposed. Populations of commensal strains of *E. coli* populating mammalian hosts modulate its immune defenses through population-level control of the expression of fimbriae. At any time only a proportion of the population expresses these cell-surface adhesins. Collectively they elicit a host-based nutrient release if the fimbriae expression is low. Too high levels of fimbriation would provoke an inflammatory response and thus intolerable conditions for the cells. The optimal level of fimbriation is a group property and its evolution is difficult to explain by naive individual selection scenarios. This article presents a computational model to simulate the evolution of fimbriae. The two main conclusions of this contribution are: (i) the evolution of this group property requires the population to be partitioned into weakly interacting sub-populations. (ii) Given certain scenarios evolution consistently under-performs, in the sense that it does not find the optimal level of fimbriation.

1. Introduction

1.1. Fimbriation and its role

Fimbriae are adhesins that allow bacteria to attach to host cells, and are a virulence factor in urinary tract infections and possibly in meningitis, yet are also produced by many commensal (i.e. non disease causing) strains (Teng, 2005; Bahrani-Mougeot et al., 2002; Connell et al., 1996). Typically colonies of *E.coli* are a mix of fimbriate (i.e. expressing fimbriae) and afimbriate cells. In the urinary tract, high levels of fimbriation in the population colonizing the host triggers an inflammatory host response, with the risk of elimination by host defenses (Godaly et al., 1998; Hedlund et al., 2001; Fischer et al., 2006). If none of the parasites are fimbriate there will be little or no host response; as the fimbriation levels increase, there will be a point at which the host response rapidly reaches full levels (Fischer et al., 2006; Gunther et al., 2002). In terms of biological function, there is mounting evidence that the *fim* system has evolved to elicit the release of nutrients (such as *N*-acetylneuraminic acid or GlcNAc) by activating host defenses while avoiding a full scale (and for the bacteria lethal) host-based inflammation response (Sohanpal et al., 2005; El-Labany et al., 2003; Sohanpal et al., 2004; Chu and Blomfield, 2006); this interpretation is corroborated by the observation that both *N*-acetylneuraminic acid and GlcNAc are (i) indicators of an inflammatory host response, (ii) readily metabolized by *E. coli* (Chu et al., 2008) and (iii) down-regulate the fimbriation probability. We will henceforth refer to the hypothesis that fimbriation is regulated so as to maximize nutrient release while avoiding a full-scale inflammatory reaction as the *milking model*.

1.2. Phase-variation in fim is a group property

The main mechanism by which population level control of fimbriation is achieved is *phase variation*: A gene (or set of genes) is phase-variable if even a clonal population with identical environmental conditions is heterogeneous with respect to the activation of this gene/set of genes. At the level of the individual cell the transition between the fimbriate ("on") and the afimbriate ("off") state (and vice versa) is stochastic (van der Woude, 2006; van der Woude and Bäumler, 2004), although the switching probability is modulated by environmental conditions. In this sense, each cell can be seen as a molecular random bit generator, with the probability for a particular outcome depending on the environmental conditions (as sensed via the cytoplasmic *N*-acetylneuraminic acid and GlcNAc concentration). The above mentioned observation that even commensal *E. coli* populations are a mix of fimbriate and afimbriate cells is precisely due to the random switching between these two states. Phase variation of fimbriation is often cited as a means to evade the host's immune response (van der Woude, 2006). This interpretation of the adaptive significance of phase-variation is well corroborated by theoretical and experimental evidence (although it might not be the whole story). Yet, the mechanism by which it could evolve is somewhat unclear because the fimbriae-mediated parasite–host

* Tel.: +44 122 7827690.
E-mail address: D.F.Chu@kent.ac.uk

0022-5193/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2008.03.017

interaction is a group property of the colony in the following sense:

- The activation of host-defenses and the emission of nutrients by the host are controlled by the number of fimbriate cells. Hence, the state of any individual cell only matters in that it contributes to the larger statistical ensemble.
- Consequences from host reactions (both beneficial and detrimental) equally apply to all cells. In the case of a host-based immune response, all cells will be killed; on the other hand, all cells in the group will get the benefit of an increased nutrient availability when the overall fimbriation increases moderately.

In the spirit of notational economy the words "phase variation" and "fimbriation" will henceforth be overloaded to refer to these group properties *as well as* to their proper meanings. The context will disambiguate the usage of these terms.

1.3. The evolution of group properties

The evolution of group properties is a much debated issue in evolutionary biology (Dugatkin et al., 2003; Nowak and Sigmund, 2005; Hofbauer and Sigmund, 1998). One important (if very simplified) model to study aspects of group-selection is the so-called *prisoner's dilemma* (Worden and Levin, 2007; Fletcher and Zwick, 2007); here two cooperating agents would get a high pay-off each; but if one agent defects, the other one also has to defect in order to prevent prohibitively low pay-off for herself. Altogether in the prisoner's dilemma there is a tendency to defect leading to a globally (i.e. taking both agents together) sub-optimal pay-off.

A recent (general) model (Traulsen and Nowak, 2006) considered multi-level selection in a population of compartments each of which contains a population of individuals. These individuals interact with one another in games, which also determines their fitness. Once a population within a compartment reaches a certain size, the compartment splits into two. Each individual has a strategy (either defection or co-operation); compartments with more co-operators will grow faster, but defectors will reproduce faster when they are interacting with cooperators, which creates the conflict between the individual- and the compartment-level of selection. Another model of group selection in a different context is Szathmáry and Demeter's (1987)"stochastic corrector model." It consists of compartments containing a mixture of two different templates that can replicate. The replication rate of the templates depends on the composition of the compartment, in that there is an optimal mix of templates to optimize compartment replication. At the same time, the model is set up such that the template-types compete with one another, which creates the tension between individual- and compartment-level selection. Similar to the Traulsen model, the compartments split once the number of templates reaches a certain limit. The contents of the original compartment is then allocated randomly to the two newly created offspring-compartments.

In the context of bacterial evolution group selection models are often focused around "common goods" produced by cells at a metabolic cost. In this context the analogue of a non-cooperating agent in the prisoner's dilemma are cheater cells, i.e. cells that benefit from the common good, but avoid the metabolic cost of producing it. An often cited example of this scenario is the production of pathogens, such as for example siderophores (West et al., 2007; Kreft and Bonhoeffer, 2005; Harrison et al., 2006; Griffin et al., 2004; Buckling et al., 2007). In this case free-riding "cheater" cells would benefit from the released siderophores (through enhanced growth) while avoiding the costs of producing them.

In the case of the evolution of fimbriation cheating (in the above sense) might be relevant given that the expression of fimbriae requires resources that could otherwise be invested into growth. However, there is another aspect to it: Other than in the case of, for example, siderophores the underlying idea of fimbriation is that the group modulates the host response to achieve an "optimal" nutrient release, rather than maximizing its virulence. It thus balances the benefit of increased nutrient release as the proportion of fimbriate cells increases with the risk of triggering a full-blown inflammatory reaction—with lethal results. Ignoring the (potentially relevant but confounding) issue of cheating, the focus of this contribution will be to investigate under which conditions individual-based evolutionary processes can effectively tune the fimbriation probability of *individual* cells so as to optimize the interaction of the *colony* with the specific host. Individual-based evolution means the following:

- Any genotypic change can only originate at the level of the individual cell. Hence, the behaviour of the colony as a whole is strictly emergent on the behaviour of its constituent cells.
- Cells do not have, at any time, any information about the state of the group as a whole nor about the shape of the host response.

Standard evolutionary narratives based on differential fitness of mutants would fail to account for the evolution of group properties such as fimbriation even if there is no cheating involved. To see this, assume a population of phase-varying cells that is not at some optimal point (i.e. maximal exploitation of nutrients with minimal risk of causing an infection); for the sake of the argument assume that the proportion of fimbriate cells is below its optimal value (for a given set of circumstances). In this situation, if a mutation in a particular cell increases its probability of switching into the fimbriate state, then this also brings the population closer to the optimal point (albeit by a possibly minute amount).

The problem of this scenario is that any positive feedback is not only provided to the mutant cell, but equally to all other cells, i.e. the mutant does not have any intra-group differential benefit and will thus not spread (except by random drift). Furthermore, a similar argument applies to maladaptive mutants (here a mutant that has a lower probability of switching its fimbriation-state). Such mutants will leave the group as a whole worse off, but will themselves not have a reduced fitness relative to other members of the same group. Yet, even if the lack of specific feed-back was not a problem, another issue remains: In a stochastic system such as fimbriation, the effects of a single mutant would be barely detectable against the background of random switching events, leaving the efficiency of selection somewhat doubtful.

Since phase-variation in *E. coli* exists and is very likely also an adaptation to balance host-nutrient release and inflammatory responses, there ought to be a mechanism for group properties such as fimbriation to evolve. In this contribution an individual-based evolutionary computer model to investigate possible scenarios for the evolution of group properties will be proposed. Nominally the model—to be described below in Section 2—will be about the particular problem of fimbriation and its evolution. The details of how cells switch their fimbriation states is modeled on the genetic mechanism by which fimbriation is controlled in *E. coli* (often called "orientational control", Chu and Blomfield, 2006). Yet, most other biological detail has by necessity, been ignored. The core of the assumption, namely that phase-variation of fimbriation serves to "milk" the host for nutrients, takes center stage in this model.

The focus of this contribution is mainly on the evolution of fimbriation. However, due to the non-specific nature of the model, any conclusions drawn from it will remain valid for any system

that fulfills the same essential characteristics. These characteristics are:

- There is intra-group differential pay-off.
- Pay-off is directly related to the prevalence of a trait via a unimodal reward curve.
- There is no (relevant) communication between the individuals.

The next section will describe an (agent-based) computational model of the evolution of phase-variation implementing the "milking" model (i.e. the idea that fimbriation evolved to optimize the release of nutrient by the host). The main insights derived from the model are that (i) group effects such as those described by the milking model can only evolve in a spatially structured population, i.e. if there are sub-groups that only have limited gene-flow from outside. (ii) In variable environments, cells evolve regulation mechanisms that allow them to adjust to changing conditions and thus better exploit resources than cells with fixed fimbriation functions. (iii) At least in the current scenario where the onset of a lethal inflammation reaction is sudden, individual-based evolution is not capable of fine-tuning but only achieves an approximate adaptation (with a population size well below the possible maximum).

2. Description of the model

This section describes the computational model on which the simulations in this article are based. In the model, every cell is explicitly represented as a data structure. Each cell has the following key-properties:

- Age.
- Energy holdings.
- Fimbriation state, i.e. either on or off.
- Internal fimbriation function (see below).

The value of these properties is (simultaneously for all cells) updated in discrete time-steps according to the following rules: At every time step the age of a cell increases by one. Once its age surpasses a (user defined) life span it will (with a user defined probability) be removed from the system. At every time-step cells receive an amount of energy (i.e. nutrient) from their environment (according to rules described below), which they accumulate. If the energy holdings of a cell are greater than a certain (user defined) threshold value, it will (with a certain probability per time step) produce offspring. Following reproduction the energy holdings of both parent and offspring are set to 0. Offspring is genetically identical to parents, however, at birth offspring may be subject to a mutation that changes its genome. The genome consists of five real valued numbers, p_1, p_2, p_3, p_4, p_5. The "gene" $p_3 \in (0, 10]$ whereas all other p_x are constrained to values between 0 and 1. Initially all p_x take random values within their respective range. Mutations consist of small adjustments of the values of the p_x. The genome specifies the fimbriation function which in turn determines the probability for a cell to switch its fimbriation state. If the cell is in the fimbriate state then at every time step it will switch to the afimbriate state with probability:

$$P(\text{on} \rightarrow \text{off}) = p_5 \tag{1}$$

If at time-step t the agent receives the amount v_t of nutrient, then independently of the fimbriation state it will switch with the probability:

$$P(\text{switch}) = p_1\left(1 - \frac{(p_2 v_t)^{p_3}}{p_4 + (p_2 v_t)^{p_3}}\right) \tag{2}$$

This function will make the transition from its maximum to its minimum value as v_t increases. The parameter p_1 determines the maximum switching probability; increasing p_3 makes the function more step-like.

The environment is partitioned into 625 compartments. Cells do not move between compartments and offspring is placed into the same compartment as the parent. There are two alternative methods to bring new cells into a compartment:

(A.I) At every time-step a randomly chosen cell from a randomly chosen compartment is moved to another randomly chosen compartment. This mechanism provides (limited) contact between compartments.
(A.II) At every time-step a randomly initialized (newly created) cell is placed into a randomly chosen compartment.

In each of the simulations reported here only one of these two mechanisms was used. Note that under the alternative (A.II) all compartments are screened off from one another.

At every time step compartments release "nutrients," representing the host-response. The amount of released nutrient depends on the number of fimbriate cells in the respective compartment and is determined as follows:

$$F = 50\frac{n_F}{(600 + n_F)} + g \tag{3}$$

Here, n_F is the absolute number of fimbriate agents in the environment and g is a basic (but low) nutrient allowance that is released into the cell independently of the number of fimbriate cells. The amount F is equally divided between all cells in the compartment (so each cell receives F/N). The nutrient is taken up by the cells. In the model it was assumed that once the amount of emitted energy crosses a threshold, then all cells in this compartment are removed; this "tolerance threshold" represents the transition from a bearable to a lethal host-response.

The main purpose of the model is to create an environment that realizes an adaptive pressure on simulated evolving cells relating to the scenario described in the introduction (the milking model). Individual-based selection is implicit in the model in the sense that the reproductive success (and hence fitness) of the cells depends on their ability to collect nutrients (i.e. there is no explicit fitness function); note that since all cells in a compartment receive the same amount of food, their fitness is equal as well; this realizes the group property. In addition to this implicit individual-based selection a (naive) explicit group selection mechanism is implemented (group selection can be turned on or off by the user). Group selection works as follows: First the compartments with the highest and lowest population are determined. Then the population of the latter will be deleted and replaced with a copy of the population of the former. In the simulations reported below, this naive group-selection mechanism is turned off unless explicitly stated otherwise.

2.1. Analysis of a simplified model

A colony of bacteria triggers an inflammatory response, if it has more than a certain threshold number (N_r) of fimbriate cells per compartment.[1] In order to simplify the analysis it will be assumed that each compartment is independent of all other compartments. In this case, the probability of a particular colony causing an inflammation at a particular time will be given by a binomial

[1] As described above, it is actually the amount of emitted nutrients that determines whether or not a sub-population will go extinct, but this is uniquely determined by the number of fimbriate cells in the compartment.

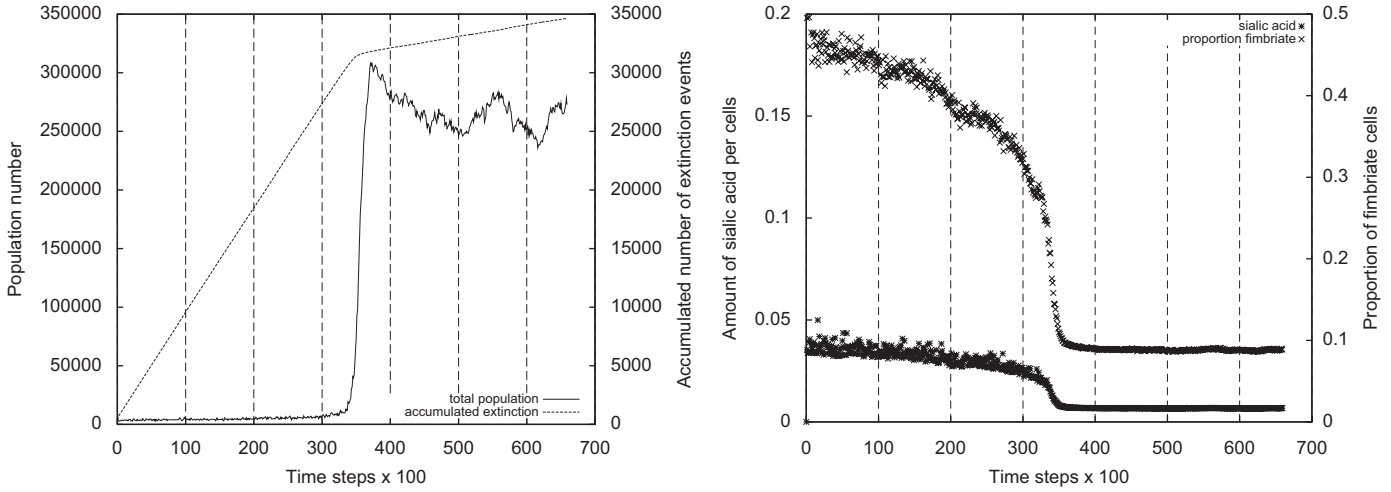

Fig. 1. The dashed curve shows the number of accumulated extinction events whereas the solid curve is the population number. A sudden increase of the population accompanied by a reduction of the slope of the extinction curve is clearly visible between time 30 000 and 40 000. The bottom curve shows the corresponding transition of the fimbriation levels to under 0.1. The parameters of this simulation were as follows: Initial population size: 10 000 cells; life-time of a cell: 40 times steps; energy threshold for reproduction: 0.2; mutation probability/birth event: 0.1; probability of transferring a randomly chosen agent to a new random site: 1; basic allowance per compartment: 0.5; inflammation tolerance threshold: 10 energy units; size of the environment: 625 compartments.

distribution.

$$P(\text{no response}) = \sum_{i=0}^{i=N_r} \binom{N(p)}{i} p^i q^{N(p)-i} \tag{4}$$

where p is the probability for a particular cell to be fimbriate and $q = 1 - p$. In this equation the population size N is unknown. However, an approximate solution can be obtained by considering that in equilibrium[2] each bacterium will reproduce exactly once; this is so because otherwise the population would either grow or shrink (hence there would be no equilibrium). We get:

$$\frac{F}{N}\frac{L}{f_r} = 1 \tag{5}$$

Here F is the total amount of nutrient that is shared among the members of the population of size N. The variable L is the lifetime of the agents and f_r is the minimum amount of nutrient required for reproduction. The amount of nutrient F in the model is given by

$$F = \frac{50pN}{600 + pN} + g \tag{6}$$

The first term on the right-hand side is essentially the amount of nutrient a single compartment releases in response to a certain number of fimbriate agents. The second term is the basic amount of nutrient released by each compartments independently of the number of fimbriate bacteria (i.e. the basic provision). Substituting Eq. (6) into Eq. (5) gives two possible solutions for N depending on the fimbriation probability.

$$N = \frac{-600f_r + 50Lmp + Lgp}{2pf_r} \pm \sqrt{(50Lmp - 600f_r)^2 + 1200f_r pLg + 100L^2mp^2g + L^2g^2p^2} \tag{7}$$

Clearly only the "+" solution is physically realistic; hence N is uniquely determined by the model. Note however, that this solution might overestimate real population numbers by a factor of up to 2, depending on the combination of parameters. This can be seen as follows: Assume that (once equilibrium is reached) the cells reproduce (on average) k time steps before the end of their "natural" life span:

$$k = L - T_r = L - \frac{F}{Nf_r} \tag{8}$$

Here the time required to reproduce is denoted by T_r. The two extreme cases to consider are that $T_r = L$ in which case $k = 0$ and the above formula in Eq. (6) is correct. The other extreme is the case where $T_r = (L + 1)/2$ in which case $k = (L - 1)/2$. Given that $L = T_r + k$ the range of possible population sizes can be obtained. If N^{real} refers to the actually observed population number, then

$$N^{real} = \frac{F}{f_r(L-k)}$$

$$\frac{F}{f_r(L+1)} \leqslant N^{real} \leqslant \frac{2F}{f_r L} \tag{9}$$

3. Results

3.1. Simulation of the basic system

Fig. 1 shows an example of a typical simulation of the system. The population was initialized with 10 000 cells each with a random genome. Initially, the total population remains well below 10 000 cells; only after about time step 35 000 it suddenly increases to a very high value of about 300 000; after a short period the population then falls back to approximately 250 000 where it remains. The qualitative characteristics of this example run (such as the sudden transition from a low population to a high population, the approximate value of the population size and so on) can be observed in all runs with these parameters although quantitative details such as the time of the transition vary significantly for different random seeds. We also performed a number of simulations with the system seeded with a single ancestor only; these simulations also show the transition, although the waiting time tends to be longer than in the simulations with high initial diversity.

[2] The assumption of an equilibrium point is not strictly correct; it can be shown that the long term attractor of the population is oscillatory, rather than a single fixed point.

In the case of a vanishing mutation rate individual-based evolution only rarely leads to the transition (approximately 1 in 10 simulations). If group selection is turned on combined with a vanishing mutation rate and alternative (A.I), then the transition occurs but at lower and widely varying population levels; group selection with vanishing mutation and alternative (A.II) gives somewhat better results than individual-based evolution (see below and Fig. 2). In individual-based selection, alternative (A.II) practically never supports the transition (data not shown).

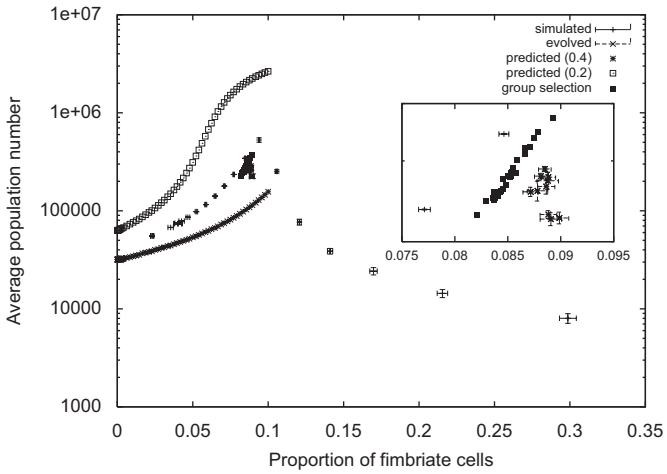

Fig. 2. The population number as a function of the proportion of fimbriate agents; the graph shows both simulated and predicted values. Each data point labeled "simulated," or "evolved" or "group selection" corresponds to the average over 3000 (500 in the case of "group selection") time steps. In order to obtain the points labeled "simulated" mutation was turned off and the population was seeded with identical cells. The error-bars of both the x and the y axis represent the standard deviation of the sample mean. The points labeled "simulated" were obtained from a simulation with identical agents and no mutation. The points labeled "evolved" represent the results of a set of simulations with non-uniform evolving agents. The points labeled "predicted" correspond to results obtained from the mathematical model for assumed minimum energy values for reproduction of 0.2 and 0.4, respectively. The points labeled "group selection" represent a simulation with group selection turned on, a vanishing mutation probability but injection of agents from outside (i.e. alternative (A.II)). The inset shows a detail of the cluster of "predicted" and "evolved" points.

Fig. 2 shows the "optimality curve" indicating results obtained by simulating genetically homogeneous and non-evolving populations (see caption of Fig. 2 for more details). The different points on the optimality curve were obtained by varying the values of the genes $p_1 - p_5$ and recording the steady state population numbers.

The optimality curve shows that there is an optimum fimbriation probability before the population starts to decrease with increasing fimbriation. This optimum is close to the onset of extinction; for the sampling frequency chosen here it does in fact co-incide with the onset of extinctions (data not shown). Models with non-homogeneous populations, i.e. runs with individual selection and group selection are below the optimality curve. The long-term population size in individual selection is below that of group selection simulations. The simulated data is enveloped by the curves predicted by the mathematical model equation (7); the top curve assumes a reproduction threshold of 0.2 whereas the lower one a value of 0.4. The mathematical model does not take into account extinction hence becomes irrelevant once significant proportions of the population go extinct.

Fig. 3 compares the group selection results in Fig. 2 with simulations that have been performed with the same parameters except for a reduced probability of injecting cells from outside; concretely the injection probability per time step was reduced from 1 to 0.1. Reducing the injection probability leads to higher population sizes for a given fimbriation probability.

3.2. Variable environments

Variable environments have been implemented by varying the amount of nutrient released by the compartment over time. In practice, this has been done by multiplying the host-emission function (Eq. (3)) by a fixed factor. Note that this also changes the number of fimbriate cells required to provoke an extinction event (because extinction in the model is coupled to nutrient release). In the specific case of Fig. 4 every 1000 time steps the amount of nutrient released was reduced by a factor of 10 for 500 time steps;

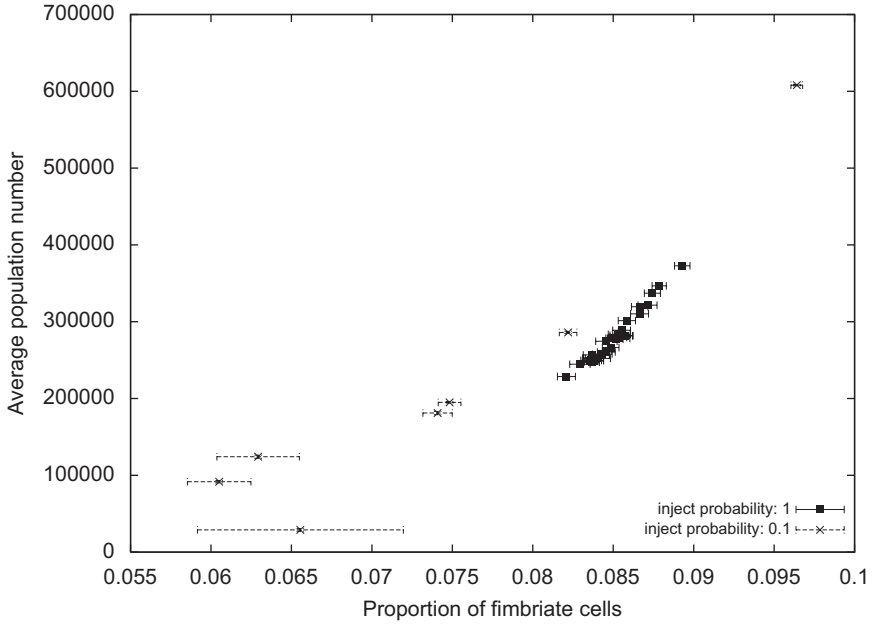

Fig. 3. This graph shows the proportion of fimbriate cells versus the average population number. Each point corresponds to the time average of the respective measures taken over 2000 time steps (as in Fig. 2). The data labeled "injectprobability: 1" is the same data as the group-selection data in Fig. 2; the points labeled "inject probability: 0.1" are evolved using the same parameters but with a lower inject probability.

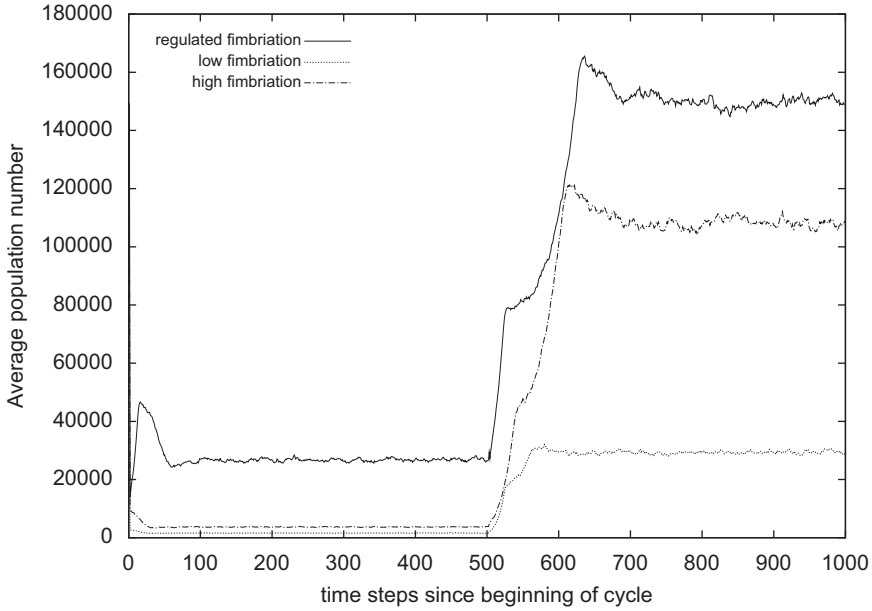

Fig. 4. This graph shows the average size of three homogeneous population over 84 000 cycles of the same simulation run. The graphs labeled high/low have a fixed fimbriation probability, whereas the population labeled "regulated fimbriation" has its fimbriation function regulated by external nutrient according to the values determined by the evolutionary run. The averages were obtained as follows: The first 85 000 time steps were considered; of those the first 1000 were discarded (because they contain transient behaviour). Then the average of the corresponding points in the phase were taken (so, the first point in the graph is the average of the points at times $1001, 2001, \ldots, 85\,001$).

then it was brought back to original levels. Under these conditions the cells evolve regulation mechanisms that give them higher fitness (measured in population number) than cells that have fixed fimbriation probabilities. Fig. 4 shows that regulated fimbriation has a consistently higher population across all environments than the cells with fixed probability.

Fig. 4 was obtained as follows: A population was evolved using standard parameters but with a variable environment. After the transition had taken place and the population stabilized, the simulation was stopped and the average parameters settings from the run were used to seed three new simulation runs; in all those three simulations mutation was turned off, but the environmental variability was the same as in the previous model. The difference between the three simulation runs was as follows: Two of the models were seeded with an ancestor with fixed fimbriation probability (corresponding to the high/low nutrient level in the variable environment of the original model). The cells in the third model were equal to those, but regulated their fimbriation probability in response to the host's nutrient release (e.g. according to Eq. (2)).

4. Discussion

4.1. Mode of evolution

The simulation results indicate that group properties such as fimbriation can evolve in individual-based evolution, at least as long as the population is divided into different sub-populations. From the simulation results it is not immediately clear by which mechanism the transition is driven. One possibility is that it is a classical evolutionary scenario (with a mutation driven discovery of novel/fitter individuals); another possibility is that it is a mere sieving effect.

Sieving means the following: If the initial population is diverse, then there will be considerable variability among the subpopulations in the compartments. The extinction probability of a particular sub-population will depend on the genetic composition of its members. In particular, if it contains many members that have a high probability of being fimbriate, then it is more likely to go extinct; by the same token, if it consists of a high number of cells that have a low probability of being fimbriate, then this sub-population tends to survive longer. Sub-populations that survive longer have a higher chance to contribute founder cells to re-populate empty compartments. Hence over time the fimbriation probability will tend to approach levels that minimize the chance of going extinct. If sieving is the only mechanism guiding the evolution then one would expect that:

- Transitions always happen shortly after the start of the simulation or not at all.
- The system converges to a regime with no (or a very low intensity of) extinction events.
- Reducing the mutation rate to zero will have only a minor or no effect on the frequency of transitions.
- Simulations initialized with a single ancestor never show the transition.

In simulations presented here none of these propositions holds true: Individual-based evolution with vanishing mutation probability only rarely leads to the transition and even then the population number is lower than in the case of a finite mutation rate (data not shown). Furthermore, when the mutation rate is turned on, then transitions are not limited to early stages of the simulation (see for example Fig. 1 where the transition happens after 30 000 time steps). Finally, experiments with a single seeded cell also show the transitions (data not shown).

While this suggests that the creative force of mutation is important, sieving does play an essential role in the evolution of the system. This is evidenced by the observation that experiments with no mobility between compartments (i.e. essentially turning off sieving) will not lead to a transition; furthermore, even in the absence of mutation the transition sometimes (if rarely) occurs as long as the initial population is sufficiently diverse; in these cases the transition can only be explained by sieving.

The partitioning of the population into only weakly interacting sub-populations is crucial for the evolution of the system (i.e. alternative (A.II) does not lead to a transition in individual-based evolution). Evolution within individual compartments is not efficient, which is not surprising (see the relevant argument in the introduction section). The evolutionary dynamics in the model relies on the antagonistic "forces" of mutations and sieving. The former creates novelty and the latter is a selection mechanism that is effective at the group level: Re-colonization of empty compartments (due to previous extinction events) creates genetically relatively homogeneous sub-populations founded by a single seed cell (randomly chosen from across the entire population). The fate of any such newly formed sub-population will crucially depend on the genome of the founder cell. The higher the population number of a given compartment (and the longer it is surviving without going extinct) the higher the chance that one of its members will be the founder of yet another new sub-population. Any such member is very likely to be from the lineage of the original founder of the original compartment (and thus its genome is a variant of this founder). If the chosen variant is fitter, i.e. leads to larger and more longveous populations, then the chance that one of its offspring will in turn be a founder-cell will be higher. Altogether, this provides a scenario for how the regulation of fimbriation as required by the milking model (and other group level properties) can evolve.

This scenario can also be interpreted biologically. The assumption of semi-separated populations is plausible if one assumes that there is only little exchange of cells between hosts. Inflammatory conditions (or bouts of anti-biotic treatment) on the other hand could sufficiently weaken a resident population to make it vulnerable to influx from outside. In real systems, the situation will be, of course, far more complex than in the model. For one, fimbriation is by no means the only virulence factor of *E. coli*. Secondly, the dynamics of extinction and re-colonization will be more fine-graded in that an inflammation will not remove the entire colony but only large parts of it. Bacteria are also capable of horizontal gene transfer; this potentially changes the evolutionary dynamics of the system.

4.2. Evolution leads to sub-optimal exploitation of resources

The optimality curve defines the achievable population size of a homogeneous, non-evolving population with a given fimbriation probability and the (fixed) host response function of Eq. (3); in particular the optimality curve indicates that there is a fimbriation probability (≈ 0.1) that maximizes the population size. Fig. 2 shows that the evolved solutions (both individual and group selection) remain below the optimality curve, and are also far away from the optimum fimbriation for homogeneous populations. In particular, the (average) fimbriation levels of the evolved solutions are below the optimum fimbriation levels for homogeneous populations. The homogeneous population does not suffer any extinction events at these fimbriation levels, but at the same *average* fimbriation level the evolving populations do. These extinction events also cause the difference in population size. If success is measured in terms of overall population size, then under the conditions of this model evolution leads to sub-optimal outcomes (even in the absence of cheating *sensu* prisoner's dilemma).

On closer inspection this sub-optimality of evolution is not surprising: Assume for the moment that there is an effective mechanism to push sub-populations along fitness gradients (as discussed above such a mechanism involves migration between compartments). The gradient experienced by the evolving populations would effectively be shaped like the optimality curve in Fig. 2: As long as the population remains below the extinction threshold, there will be adaptive pressure to increase fimbriation, thus favoring sub-populations with higher fimbriation probabilities. Yet, once the threshold is reached the population will simply go extinct. The information about where the point of extinction is, is not embedded into the fitness landscape, except at the point where it actually happens; in this sense, the evolving population cannot "know" at what point there will be a sudden increase in the inflammation reaction leading to its extinction. Seen from this perspective, it is not surprising that individual-based evolution leads to populations under the optimality curve.

A similar explanation can be given for the group-selection results. These are also below the optimality curve. The relevant antagonistic force to sieving in the group selection results is the injection of (random) cells that constantly disturb any point found by evolution. Reducing the injection probability should therefore lead to populations closer to the optimality curve; this is indeed the case (see Fig. 3).

4.3. Adaptation to changing environments

One aspect of the genetic network of *E. coli* has so far not been explained: If the host response function is constant then one would expect that there is no need for the fimbriation probability of cells to be modulated by environmental factors (i.e. *N*-acetylneuraminic acid, GlcNAc). Given that in real *E. coli* the fimbriation probability is modulated (see Section 1.1), one would expect it to be an adaptation to changing environments (as has been hypothesized previously Chu and Blomfield, 2006). Fig. 4 shows that cells with evolved regulation outperform cells with a fixed fimbriation probability in changing environments.

The ability to evolve such a regulation mechanism requires the lowest level entities in the model (i.e. the cells) to be subject to evolutionary change. Similar group-selection models, such as for example Szathmáry and Demeter's (1987) stochastic corrector model (described in Section 1.3), often do not have this feature. In both models there is a group-level selection pressure to regulate the composition of compartments, but the mechanisms by which the composition of the compartments is controlled are very different in the two models. In the stochastic corrector model there is no genetic information determining the behaviour of the most basic units (i.e. the templates), or alternatively there are only two possible genotypes. In a sense the genetic information of the protocells is their template composition. The templates have a fixed behaviour that cannot be modified by mutation. The evolutionary dynamics of the system is driven by random re-allocation of templates from parent to offspring compartments, and a process of elimination of those compartments that are not reproducing (or not reproducing fast enough). A similar type of process (i.e. sieving) is also important in the fimbriation model, but as discussed above, mutations changing the behavior of the individual components are essential for the efficiency of the evolutionary dynamics of this system. Adaptation to changing environments is a particular case in point that by necessity requires mutations of the individual in order to be efficient.

5. Conclusion

Despite being a group-property, fimbriation can emerge in individual-based evolution if the evolving population is partitioned into only weakly interacting sub-populations. Furthermore, the process by which evolution proceeds in this model goes beyond mere sieving, but involves the evolution of a genuine control mechanism that allows populations of bacteria to cope with changing environments. Altogether, these conclusions

support the milking model as a functional explanation for the regulation of fimbriation in *E. coli*.

Another finding of this model—that evolution is sub-optimal—could also have a biological interpretation in terms of the host–parasite interactions. Commonly virulence is interpreted as a strategy that by itself entails an adaptive benefit. In the case of the milking model the situation is somewhat different in that virulence is more an undesired by-product of inefficient evolution rather than a desired end in itself. In the simulations presented here, the bacteria could maximize their overall population if they avoided host responses altogether; in individual-based evolution, however, there was always a residual virulence resulting from the fact that the point at which the host reaction sets in is not "knowable" before it is encountered. There is circumstantial evidence that the inefficiency of evolution, as observed in the model is relevant for real populations as well. *E. coli* exists in commensal strains, however, even in those cases it leads to episodic inflammatory reactions. This residual virulence might be the symptom of evolution testing its boundaries. At this point this is speculation only. It is unclear to what degree the effect of inefficient evolution is preserved as more continuous extinction curves are used.

One simplification that has been made here is that the transmission rate of infections, i.e. the rate with which compartments get infected is independent of the virulence of the parasite population; changing this assumption could have some impact given that the transmission rate has been identified as an important factor determining the evolution of virulence (see for example Lenski and May, 1994). In particular, if the transmission probability increases with the virulence, then this could lead to different evolutionary pathways. Future models will need to take this into account. The present model has shown that phase-variation can be evolved to regulate the collective action of bacterial colonies in response to incipient host-inflammations. This process will only be effective if the overall population is spatially structured into semi-independent sub-populations. In terms of real bacteria and their hosts, this assumption is justifiable: Each host can be thought of as a semi-separated population with only limited influx of new strains from the outside.

Acknowledgments

This work was partially funded by the EPSRC (Grant number EP/F035152/1). I thank Ian Blomfield for discussions on fimbriation and its importance for the host–parasite relation. I also thank Jon Rowe and Boris Mitavskiy for discussions on evolutionary and mathematical aspects of fimbriation.

References

Bahrani-Mougeot, F., Buckles, E., Lockatell, C., Hebel, J., Johnson, D., Tang, C., Donnenberg, M., 2002. Type 1 fimbriae and extracellular polysaccharides are preeminent uropathogenic *Escherichia coli* virulence determinants in the murine urinary tract. Mol. Microbiol. 45, 1079–1093.

Buckling, A., Harrison, F., Vos, M., Brockhurst, M., Gardner, A., West, S., Griffin, A., 2007. Siderophore-mediated cooperation and virulence in pseudomonas aeruginosa. FEMS Microbiol. Ecol. 62 (2), 135–141 URL: ⟨http://dx.doi.org/10.1111/j.1574-6941.2007.00388.x⟩.

Chu, D., Blomfield, I., 2006. Orientational control is an efficient control mechanism for phase switching in the *E. coli* fim system. J. Theor. Biol. 244 (3), 541–551.

Chu, D., Roobol, J., Blomfield, I., 2008. A theoretical interpretation of the transient sialic acid toxicity of a *nanR* mutant of *Escherichia coli*. J. Mol. Biol. 375, 875–889.

Connell, I., Agace, W., Klemm, P., Schembri, M., Marild, S., Svanborg, C., 1996. Type 1 fimbrial expression enhances *Escherichia coli* virulence for the urinary tract. Proc. Natl Acad. Sci. USA 93, 9827–9832.

Dugatkin, L., Perlin, M., Atlas, R., 2003. The evolution of group-beneficial traits in the absence of between-group selection. J. Theor. Biol. 220 (1), 67–74.

El-Labany, S., Sohanpal, B., Lahooti, M., Akerman, R., Blomfield, I.C., 2003. Distant cis-active sequences and sialic acid control the expression of *fimB* in *Escherichia coli* K-12. Mol. Microbiol. 49, 1109–1118.

Fischer, H., Akira, M.Y.S., Beutler, B., Svanborg, C., 2006. Mechanism of pathogen-specific tlr4 activation in the mucosa: fimbriae, recognition receptors and adaptor protein selection. European J. Immunol. 36, 267–277.

Fletcher, J., Zwick, M., 2007. The evolution of altruism: game theory in multilevel selection and inclusive fitness. J. Theor. Biol. 245 (1), 26–36 URL: ⟨http://dx.doi.org/10.1016/j.jtbi.2006.09.030⟩.

Godaly, G., Frendeus, B., Proudfoot, A., Svensson, M., Klemm, P., Svanborg, C., 1998. Role of fimbriae-mediated adherence for neutrophil migration across *Escherichia coli*-infected epithelial cell layers. Mol. Microbiol. 30, 725–735.

Griffin, A., West, S., Buckling, A., 2004. Cooperation and competition in pathogenic bacteria. Nature 430 (7003), 1024–1027 URL: ⟨http://dx.doi.org/10.1038/nature02744⟩.

Gunther, I., Snyder, J., Lockatell, V., Blomfield, I., Johnson, D., Mobley, H., 2002. Assessment of virulence of uropathogenic *Escherichia coli* type 1 fimbrial mutants in which the invertible element is phase-locked on or off. Infect. Immunity 70, 3344–3354.

Harrison, F., Browning, L., Vos, M., Buckling, A., 2006. Cooperation and virulence in acute pseudomonas aeruginosa infections. BMC Biol. 4, 21 URL: ⟨http://dx.doi.org/10.1186/1741-7007-4-21⟩.

Hedlund, M., Frendéus, B., Wachtler, C., Hang, L., Fischer, H., Svanborg, C., 2001. Type 1 fimbriae deliver an LPS- and TLR4-dependent activation signal to CD14-negative cells. Mol. Microbiol. 39, 542–552.

Hofbauer, J., Sigmund, K., 1998. Evolutionary Games and Population Dynamics. Cambridge University Press, Cambridge.

Kreft, J., Bonhoeffer, S., 2005. The evolution of groups of cooperating bacteria and the growth rate versus yield trade-off. Microbiology 151 (Pt. 3), 637–641 URL: ⟨http://dx.doi.org/10.1099/mic.0.27415-0⟩.

Lenski, R.E., May, R.M., 1994. The evolution of virulence in parasites and pathogens: reconciliation between two competing hypotheses. J. Theor. Biol. 169 (3), 253–265 URL: ⟨http://dx.doi.org/10.1006/jtbi.1994.1146⟩.

Nowak, M., Sigmund, K., 2005. Evolution of indirect reciprocity. Nature 437 (7063), 1291–1298 URL: ⟨http://dx.doi.org/10.1038/nature04131⟩.

Sohanpal, B., El-Labany, S., Plumbridge, J., Blomfield, I., 2004. Integrated regulatory responses of fimb to *N*-acetylneuraminic (sialic) acid and glcnac in *Escherichia coli* K-12. Proc. Natl Acad. Sci. USA 101, 16322–16327.

Sohanpal, B., El-Labany, S., Lahooti, M., Plumbridge, J., Blomfield, I., 2005. Integrated regulatory responses of *fimB* to *N*-acetylneuraminic (sialic) acid and GlcNAc in *Escherichia coli* K-12. Proc. Natl Acad. Sci. USA 101, 16322–16327.

Szathmáry, E., Demeter, L., 1987. Group selection of early replicators and the origin of life. J. Theor. Biol. 128 (4), 463–486.

Teng, C., 2005. *Escherichia coli* K1 Rs218 interacts with human brain microvascular endothelial cells via type 1 fimbria bacteria in the fimbriated state. Infect. Immunity 73, 2923–2931.

Traulsen, A., Nowak, M., 2006. Evolution of cooperation by multilevel selection. Proc. Natl Acad. Sci. USA 103 (29), 10952–10955 URL: ⟨http://dx.doi.org/10.1073/pnas.0602530103⟩.

West, S., Diggle, S., Buckling, A., Gardner, A., Griffin, A., 2007. The social lives of microbes. Annu. Rev. Ecol. Evol. Systematics 38, 53–77.

Worden, L., Levin, S., 2007. Evolutionary escape from the prisoner's dilemma. J. Theor. Biol. 245 (3), 411–422 URL: ⟨http://dx.doi.org/10.1016/j.jtbi.2006.10.011⟩.

van der Woude, M., 2006. Re-examining the role and random nature of phase variation. FEMS Microbiol. Lett. 254 (2), 190–197 URL: ⟨http://dx.doi.org/10.1111/j.1574-6968.2005.00038.x⟩.

van der Woude, M., Bäumler, A., 2004. Phase and antigenic variation in bacteria. Clin. Microbiol. Rev. 17 (3), 581–611.