bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Enchained growth and cluster dislocation : a possible mechanism for microbiota homeostasis

Florence Bansept[1], Kathrin Schumann-Moor[2,3], Médéric Diard[2], Wolf-Dietrich Hardt[2], Emma Slack[2], and Claude Loverdo[1,*]

[1]Laboratoire Jean Perrin, Sorbonne Université / CNRS, Paris, France.
[2]Institute of Microbiology, ETH Zürich, Zürich, Switzerland.
[3]Present address: Center of Dental Medicine / Oral and Maxillofacial Surgery, University of Zürich, Switzerland
[*]Corresponding author

Abstract

Immunoglobulin A is a class of antibodies produced by the adaptive immune system and secreted into the gut lumen to fight pathogenic bacteria. We recently demonstrated that the main physical effect of these antibodies is to enchain daughter bacteria, i.e. to cross-link bacteria into clusters as they divide, preventing them from interacting with epithelial cells. These links between bacteria may break over time. Using analytical and numerical calculations on several models to check the results robustness, we study the rate of increase in the number of free bacteria as a function of the replication rate of bacteria, and the resulting distribution of chain sizes. At higher replication rates, the bacteria replicate before the link between daughter bacteria breaks, leading to growing cluster sizes. However at low growth rates two daughter cells have a high probability to break apart. Thus the gut could produce IgA against all the bacteria it has encountered, but the most affected bacteria would be the fast replicating ones, which could destabilize the microbiota.

Introduction

The digestive system has a large surface area[1][2], covered by a single layer of epithelial cells, essential for nutrient absorption, but also a gateway for many pathogens. Contrary to the inside of the body, where the presence of any bacteria is abnormal, the lumen of the digestive system is home to a very important microbiota. These microbiota bacteria are present in extremely high densities[3] and have profound effects on the host's health and nutrition. Bacteria are necessary to break down and absorb certain nutrients, and can compete against potentially pathogenic intruders[4]. Inside the organism, the immune system can fight generically against any bacteria. However, in the digestive system, the host has to find alternative ways to fight dangerous bacteria while sparing beneficial ones. As closely related bacteria (e.g. *Salmonella* spp. and commensal

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

E. coli) can show highly variable behaviors in the intestine, identifying which bacteria are good or bad is challenging. Besides, overgrowth of any type of bacteria, even those that do not cause acute pathology, can impair the functionality of the microbiota. Thus the host needs mechanisms to maintain homeostasis of the gut microbiota.

The adaptive response is the only strong handle that the host has on directly controlling microbiota composition at the species level[5][6]. The main effector of the adaptive immune response in the digestive system is secretory IgA, an antibody. sIgA specifically bind to targets that the organism has already encountered via infection or vaccination. It was observed more than 40 years ago that this prevents infection[7]. Many studies have focused on the complex molecular and cellular pathways that trigger an immune response on the host side of the digestive surface[8]. However, we are only just beginning to understand how induced immunity really acts after the secretion of immune effectors into the intestinal lumen.

We have shown that mice vaccinated with inactivated *Salmonella* Typhimurium do produce specific sIgA which bind to *S.*Typhimurium, but this neither kills them nor prevents them from reproducing[9][10]. The initial colonization of the intestinal lumen by *S.*Typhimurium is in fact unchanged in either kinetics or magnitude in vaccinated animals. These mice are nevertheless protected against pathogen spread from the gut lumen to systemic sites like lymph nodes, liver or spleen. A classic idea in immunology is that, as one antibody has several binding sites, antibodies aggregate bacteria when they collide into each other. But this effect would be negligible at realistic densities of a given bacterium in the digestive system, simply due to a low probability of bacteria recognized by the same sIgA encountering one another (see section 1 in appendix). We have shown that actually, the main effect is that upon replication, daughter bacteria remain attached to one another by sIgA, driving the formation of clusters derived from a single infecting bacterium[10]. This "enchained growth" is effective at any bacterial density. Clustering has physical consequences: the produced clusters do not come physically close to the epithelial cells. And as interaction with the epithelial cells is essential for *S.*Typhimurium virulence, this is sufficient to explain the observed protective effect.

If sIgA was perfectly sticky, we would expect all bacteria to be in clusters of ever increasing size. In these experiments, despite observing *S.*Typhimurium clusters in the presence of sIgA, there are still free *S.*Typhimurium, and small clusters. One possibility would be that not all bacteria are coated with sIgA. But in these experiments, it has been demonstrated that they are (extended figure 2c of[10]). Indeed, a gram of digestive content contains at most 10^{11} bacteria, and typically 50 micrograms or more of sIgA[11], of molecular mass of about 385kD. This leads to about 800 sIgA per bacteria. sIgA may not be all bound to bacteria, and sIgA for different specific antigens may be produced in proportions not matching the proportions of antigens present in the digestive system, so that not all bacteria are coated with 800 sIgA. Nevertheless, most bacteria already encountered by the organism will be coated with many sIgA, and thus the cluster size is not limited by the number of available sIgA. Another possibility is that the sIgA-mediated links break. Such breaking has been demonstrated to be dependent on the applied forces in related systems[12][13]. As there is shear in the digestive system, because mixing is needed for efficient nutrients absorption, it is plausible that links break over time.

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Small clusters are linear chains of bacteria, bound by sIgA, with these links being broken over time by the forces induced by the flow. As bacteria are similar to each other, it is, at another scale, analogous to other physical systems[14], such as polymers breaking under flow[15]. The main difference is that these chains grow by bacterial replication. Growth and fragmentation are competing effects, and the modelling of these clusters can be viewed as statistical physics, to predict their size distribution, whether there is a typical cluster size, or if large clusters of ever-increasing size dominate the distribution, and how the growth in number of free bacteria depends on the bacterial replication rate.

This could have very important biological consequences. To illustrate this point, let us consider a simplified model: bacteria remain enchained by sIgA when they grow (replication time τ_{div}), and this link between 2 bacteria breaks at a specific time τ_{break} (although this latter hypothesis is not realistic, we make it for now for the sake of simplicity). If $\tau_{div} > \tau_{break}$, then when a bacteria divides, it forms a 2-bacteria cluster, which dislocates into 2 free bacteria before the next replication steps, so the bacteria remain in the state of free or 2-bacteria clusters and there are no larger clusters. If $\tau_{div} < \tau_{break}$, when a bacteria divides, it forms a 2-bacteria cluster, which becomes a 4 bacteria cluster before the first link breaks, so there cannot be free bacteria. In this model, the fast-growing bacteria are selectively targeted by the action of the immune system. The immune system does not need to sense which bacteria are growing faster, it only has to produce sIgA targeted to all the bacteria it has encountered, and bacteria with $\tau_{div} > \tau_{break}$ are unaffected, whereas bacteria with $\tau_{div} < \tau_{break}$ are trapped in clusters. That could be a simple physical mechanism to target the action of the immune system to the fast-growing bacteria which are destabilizing the microbiota, and thus to preserve microbiota homeostasis.

In the following, we present different plausible models of bacteria clusters dynamics, and the methods to study them. Then we give, for each model, the resulting dynamics and cluster size distribution, before putting these results in perspective with experimental data. Eventually, we discuss the results.

Models and methods

General methods

We consider low bacterial densities, so encounters between unrelated bacteria are negligible. Thus, we consider each free bacteria and each cluster of bacteria independently of the others. *Salmonella* are rod shaped bacteria, which divide at the middle of the longitudinal axis. Thus if the daughter bacteria remain enchained, they are linked to each other by their poles. With further bacterial replications, the cluster will then be a linear chain. This is consistent with experimental observations, in which clusters are either linear chains, with bacteria attached to one or two neighbors by their poles, or larger clusters which seem to be formed as bundles of such linear clusters (pannel A figure 1). Our aim is to model the dynamics of these clusters.

A first element is the bacterial replication (see figure 1 C). One way to model it is to assume that bacteria replicate every τ_{div}. Another way, that we will generally use, less realistic but easier for calculations, is to assume that there is a fixed replication rate r.

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A second element is that when bacteria replicate, they may be able to escape enchainment (see figure 1 B), but likely with low probability (see discussion in section 2 in appendix). In general, we will take the limit with perfect enchainment upon replication.

A crucial element is the possibility for the links between bacteria to break. We usually assume that the breaking rate α is the same for all links and over time. We will also explore the case when the link breaking rate is force-dependent, in which case not all the links have the same breaking rate.

Another crucial element, is to model what happens when the chain breaks (see figure 1 D). If the subparts come in contact again at the same poles and get linked again, then this could simply be modeled by an effectively lower breaking rate. More likely, if the subparts come in contact again, they do so laterally, forming larger clusters of more complex shapes. Because in these clusters, most bacteria have more than two neighbors, and more contact surface, they are much less likely to escape. To simplify, we will consider that these clusters do not contribute anymore to releasing either free bacteria or linear chains. Thus when a link breaks, either the two subparts move sufficiently away and become two independent chains (probability q); or collide and become a more complex cluster which does not contribute anymore to both free bacteria and linear chains (probability $1-q$). For simplicity, we consider that when an outermost link breaks, the single bacteria, more mobile, always escapes ($q_{outermost}=1$), but that else q is size independent. We will take $q=0$ for the base model.

As digestive content leaves the digestive system, or the part of the digestive system under consideration, due to flow, we define c the loss rate of free bacteria, and c' the loss rate of clusters. We assume no death (which could break clusters). As free bacteria have more autonomous motility, enabling them to swim towards the epithelial cells, it is likely that $c' \geq c$. We will usually take $c=c'$. Crucially, in this latter case, free bacteria, and all clusters are lost at the same rate. The c value has a complex effect on stochastic quantities, such as the probability to have at least one cluster of a given size. However, here we study the mean numbers of free bacteria and clusters of different sizes, then the case with $c=c'$ is equivalent to $c=c'=0$, with all numbers of bacteria and clusters multiplied by $\exp(-ct)$.

We start with the most basic model, with a replication rate r, bacteria perfectly bound upon replication, a fixed breaking rate per link α, and bacterial chains always binding into a more complex cluster when a link breaks (except for the outermost links) ($q=0$). We then study variations of the model to test the robustness of the results: with an non-zero escape probability upon replication and $c \neq c'$; with a replication time τ instead of a replication rate r; with the possibility for chains to escape when an inner link breaks ($q>0$); with a force-dependent breaking rate.

We consider the beginning of the process, early enough so that the carrying capacity is far from reached, and thus the replication rate is constant. We do not consider generation of escape mutants which are not bound by IgA. We consider only the average numbers of free bacteria and linear clusters of different sizes, and we do not count more complex clusters, as they do not contribute to free bacteria dynamics in our model. For each model, we write the equations for the derivative of these numbers with respect to time. With N the vector of the mean number of free bacteria, linear clusters of size 2, 3, etc., these equations give the coefficients of the matrix M, such that $dN/dt = MN$. The results are obtained in

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

part via analytical derivations and in part via numerical studies. The latter are obtained in Mathematica by numerically solving the eigensystem written for clusters up to size n_{max} (chosen large enough not to impact the results), and looking at the largest eigenvalue λ and the corresponding eigenvector. For each model, we study how the growth of free bacteria – the ones which are capable of causing systemic infection[10] – which is λ in the steady state, depends on the bacterial replication rate. Besides, we obtain distributions of the cluster sizes, which could be compared to experimentally observed distributions.

Base model : replication rate, no bacteria escape upon replication, fixed breaking rate, $q = 0$

Equations

In the base model, bacteria have a replication rate r, daugthers are perfectly bound upon replication, each link has a breaking rate α, and when a link which is not at a tip breaks, the resulting two chains of bacteria always bind into more complex clusters and thus do not contribute to free bacteria dynamics anymore ($q = 0$). With $n_i(t)$ the number of linear clusters of size i as a function of time, (n_1 is the number of free bacteria),

$$\frac{dn_1}{dt} = -rn_1 + \sum_{i=2}^{\infty} 2\alpha n_i \tag{1}$$

and for $i \geq 2$,

$$\frac{dn_i}{dt} = rn_{i-1}(i-1) - irn_i - (i-1)n_i\alpha + 2\alpha n_{i+1} \tag{2}$$

Free bacteria growth rate as a function of the bacterial replication rate

Even for this simple version, the system of equations is hard to solve in the general case. We start by studying numerically the growth rate in the long term (the maximum eigenvalue λ of the matrix of coefficients $m_{i,j} = r(i-1)\delta_{i-1,j} - ir\delta_{i,j} - (i-1)\alpha\delta_{i,j} + 2\alpha(\delta_{i+1,j} + \delta_{i,1}(1-\delta j, 1-\delta j, 2)))$, as a function of the replication rate (see figure 2A). The growth rate has a maximum for a finite replication rate, of the order of α (the link breaking rate): the higher the replication rate, the higher the potential for growth in the number of free bacteria, but when the replication rate becomes too large compared to the breaking rate, the bacteria get trapped in clusters, which break and re-attach in more complex clusters from which independent bacteria cannot escape.

Chain length distribution

In the long time limit, the number of clusters of size i is of the order of $b_i \exp(\lambda t)$, with λ the largest eigenvalue. Equation (2) simplifies to:

$$\lambda b_i = -irb_i + rb_{i-1}(i-1) - (i-1)b_i\alpha + 2\alpha b_{i+1} \tag{3}$$

Assuming that i is large,

$$b_i \simeq \frac{r}{r+\alpha} b_{i-1} \tag{4}$$

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

is required. Using this approximation for all i, the probability that a randomly chosen chain is of size k is:

$$p_k = \left(1 - \frac{r}{r+\alpha}\right)\left(\frac{r}{r+\alpha}\right)^{k-1} \tag{5}$$

This approximation works relatively well, especially for smaller r/α values (see figure 2B). Part of the discrepancy is that equation (4) is an approximation for large i, and thus does not hold at small clusters sizes.

Model with bacteria escape

Equations

This is similar to the base model presented before, except that we take into account that upon replication, bacteria may not be perfectly bound, and may escape (pannel B of figure 1). We note δ the probability for the two daughter bacteria to become free bacteria upon replication of a free bacteria. We note δ' the probability that when a bacteria at the tip of a cluster replicates, the daughter cell on the outside of the cluster escapes the enchainment. We note δ'' the probability that when a bacteria at the interior of the cluster divides, the daughter cells will not be enchained, effectively clipping the cluster in two. As free bacteria are more motile than clusters, then $\delta \geq \delta' \geq \delta''$. We also add here the possibility that the loss rate c for free bacteria and c' for clusters are different. Then the base equations are:

$$\frac{dn_1}{dt} = r(-1+2\delta)n_1 + \sum_{i=2}^{\infty} 2r\delta' n_i + \sum_{i=2}^{\infty} 2\alpha n_i - cn_1 \tag{6}$$

$$\frac{dn_2}{dt} = r(1-\delta)n_1 - 2r(1-\delta')n_2 - \alpha n_2 + 2n_3\alpha - c'n_2 \tag{7}$$

and for $i \geq 3$,

$$\frac{dn_i}{dt} = r(2\delta'-i)n_i + rn_{i-1}(i-1-2\delta'+3\delta''-i\delta'') - (i-1)n_i\alpha + 2\alpha n_{i+1} - c'n_i. \tag{8}$$

Free bacteria growth rate as a function of the bacterial replication rate

Similarly to the base model, we study numerically the growth rate as a function of the replication rate (see figure 2C). The larger the replication rate, the more the deviation between the growth rate and the replication rate, which would be its value in the absence of clusters. If δ, δ', δ'' are small enough, the qualitative behavior is similar to the base model. But for larger δ, δ' and δ'', the growth rate continues to increase monotonously with the replication rate. The same is true when δ, δ' and δ'' are different (see supplementary figure 3). If $c = c'$, the growth rate is simply offset by minus the loss rate (see supplementary figure 3), and if $c \neq c'$, the effect is more complex, but for small r/α values it corresponds to an offset of $-c$.

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Chain length distribution

We can reason similarly to the base model (more details in section 3 in appendix), and find:

$$p_k = \left(1 - (1-\delta'')\frac{r}{r+\alpha}\right)\left((1-\delta'')\frac{r}{r+\alpha}\right)^{k-1} \tag{9}$$

This approximation works relatively well (figure 2D, and supplementary figure 4). The approximation (9) depends on δ'', but neither on δ nor δ', but δ and δ' could actually matter when i is small, and indeed we observe (see supplementary figures 5 and 6) that the approximation (9) works slightly less well when δ'' is different from δ or δ'. If $c = c'$, the distribution does not change, and if $c \neq c'$, the distribution changes very little (see supplementary figure 7).

Model with fixed replication time

In this variant of the base model, bacteria divide every τ. The effective growth rate is r_{eff} such that $\exp(r_{eff}t) = 2^{t/\tau}$, thus $r_{eff} = \log(2)/\tau$.

Equations

Let us start by considering a chain of n bacteria at $t = 0$. In the absence of replication, and with $l(n,i,t)$ the probability that at t, the chain has lost i bacteria in total on the extremities, and consequently is of size $n-i$ at t (since we assume $q = 0$ as in the base model, if the chain breaks somewhere else, the subparts form a more complex cluster and thus are "lost" for the system):

$$\frac{dl(n,i,t)}{dt} = -\alpha(n-1-i)l(n,i,t) + 2\alpha l(n,i-1,t). \tag{10}$$

At $t = 0$, $l(n,0,0) = 1$ and for $0 < i < n-1$, $l(n,i,0) = 0$. The solution for any $0 \leq i \leq n-2$ is:

$$l(n,i,t) = \frac{2^i}{i!}\exp(-\alpha t(n-1))(\exp(\alpha t)-1)^i \tag{11}$$

For any chain of size > 2, there are two outermost links, each breaking at rate α, liberating one free bacteria; and a chain of size 2 breaks at rate α, but liberates two free bacteria. Consequently, the average number of free bacteria generated during τ by this chain of n bacteria is:

$$l(n,free,\tau) = 2\alpha\sum_{i=0}^{n-2}\int_0^\tau l(n,i,t)dt = 2\alpha\sum_{i=0}^{n-2}\int_0^\tau \frac{2^i}{i!}\exp(-\alpha t(n-1))(\exp(\alpha t)-1)^i dt. \tag{12}$$

Upon replication, a chain of length n will become a chain of length $2n$, and will contribute to chains of size k by $l(2n, 2n-k, \tau)$, and to the free bacteria by $l(2n, free, \tau)$. We write the corresponding matrix, cut to size $n_{max} \times n_{max}$, and numerically solve the eigensystem.

Free bacteria growth rate as a function of the bacterial replication rate

The shape of the relation between free bacteria growth rate and (effective) replication rate (figure 2E) is very similar in the fixed replication time vs.

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

fixed replication rate models, with a maximum of the growth rate for a finite value of the (effective) replication rate, at close values ($r_{eff} = 1.15\alpha$ vs. $r = 1.09\alpha$). When the replication is at fixed time intervals instead of a fixed replication rate, the maximum growth rate is higher, and it dips faster at increasing effective replication rate. Indeed, in the case of fixed replication rate, the distribution of durations between two replications is exponential, thus more spread. Close to the maximum, the presence of short replication intervalls makes that there can be more cluster formation, and conversely, at higher replication rates, the presence of longer replication intervalls results in more production of free bacteria.

Chain length distribution

We show here the main steps to calculate analytically an approximation for the chain size distribution, and more details are given in section 4 in appendix. We define $u(N,t)$ the number of chains of size N at t. Assuming N even,

$$u(N,t+\tau) = \sum_{i=0}^{\infty} q\left(\frac{N}{2}+i,t\right) l(N+2i,2i,\tau). \tag{13}$$

In the long time, $u(N,t) = f(N)\exp(\lambda t)$, with λ the long term growth rate, that is such that $\exp(\lambda\tau) = \mathcal{N}$, with $\mathcal{N}$ the largest eigenvalue of the matrix. Then previous equation leads to:

$$\mathcal{N} f(N) = \sum_{i=0}^{\infty} f(\frac{N}{2}+i) \exp\left(-\alpha\tau\left(N-1+2i\right)\right) \left(\exp(\alpha\tau)-1\right)^{2i} \frac{2^{2i}}{(2i)!}. \tag{14}$$

We make the assumption that the first term of the sum is large compared to the rest of the sum (assumption discussed in appendix, section 4). Then,

$$f(N) \simeq \frac{1}{\mathcal{N}} f\left(\frac{N}{2}\right) \exp\left(-\alpha\tau\left(N-1\right)\right) \tag{15}$$

and recursively,

$$f(N) \simeq f(1) N^{\frac{\alpha\tau-\log(\mathcal{N})}{\log(2)}} \exp\left(-2\alpha\tau N\right). \tag{16}$$

When $\alpha\tau \gg 1$, links typically break before the next replication, thus there is little impact of the clustering on the growth, and thus the growth will be close to its value in the abscence of clustering, i.e. doubling every τ, thus in this limit $\mathcal{N} = 2$:

$$f(N) \simeq f(1) N^{\frac{\alpha\tau}{\log(2)}-1} \exp\left(-2\alpha\tau N\right). \tag{17}$$

This rough approximation allows to explain the core of the observed distribution (figure 2F). There are bumps, due to the replication every τ (which in the absence of link breaking would results in clusters of size 2^k only), which makes that clusters of power-of-two length are overrepresented. Compared to the case with fixed replication rate, the distribution is much narrower.

Model with linear chains independent after breaking ($q > 0$)

Limit case : subchains always remain independent linear chains after breaking ($q = 1$)

In this model, when a cluster breaks, the two resulting clusters remain independent and can thus continue to participate in the dynamics of the

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

system:

$$\frac{dn_i}{dt}(t) = (i-1)r\ n_{i-1}(t) + (-\alpha(i-1) - ir)\ n_i(t) + 2\alpha \sum_{j=i+1}^{\infty} n_j(t) \quad (18)$$

We recognize here the equation studied in[16], where they described chains of growing unicellular algae. As it has been shown, the steady state solution of the system is:

$$n_i(t) = C \exp(rt) \left(\frac{r}{\alpha + r}\right)^i. \quad (19)$$

In the steady state, the growth rate is equal to the replication rate. The average cluster size is $1 + \frac{r}{\alpha}$, which shows that, as expected, if the link breaking rate is high compared to the replication rate ($r/\alpha \ll 1$), the average length is close to one as no cluster has the time to form: all the bacteria remain free.

Intermediate case : chains can either be independent or trapped after breaking

More realistically, after breaking, chains will have some probability to either encounter each other and remain trapped in more complex clusters, or to escape and become independent. We will assume in the following that if a chain of size N breaks at a link at the extremity, releasing a cluster of size $N-1$ and a free bacteria, then the free bacteria, smaller and likely more mobile, will escape in all cases; but that if the link that breaks is elsewhere, the probability for the new clusters of sizes $N-k$ and k ($k > 1$) to escape and continue as two independent linear clusters will be q, and the probability that they bind and form a more complex cluster will be $1-q$, with q independent of k. We write the equations for the number $n_i(t)$ of cluster of i bacteria:

$$\frac{dn_1}{dt} = -rn_1 + 2\alpha \sum_{j=2}^{\infty} n_j \quad (20)$$

$$\frac{dn_i}{dt} = -rin_i + r(i-1)n_{i-1} - \alpha(i-1)n_i + 2\alpha n_{i+1} + 2\alpha q \sum_{j=2}^{\infty} n_{i+j}. \quad (21)$$

In the long time, $n_i \to f_i \exp(\lambda t)$ with λ the largest eigenvalue.

$$\lambda f_i = -rif_i + r(i-1)f_{i-1} - \alpha(i-1)f_i + 2\alpha f_{i+1} + \sum_{j=i+2}^{\infty} 2\alpha q f_j \quad (22)$$

This is valid for any i. We assume that f_i decreases fast enough with i such that the sum from $i+2$ to ∞ of the f_i is an order of magnitude less than if_i. Then, the largest elements of equation (22) when i is large enough are the terms multiplied by i, and consequently:

$$0 \simeq -rf_i + rf_{i-1} - \alpha f_i \quad (23)$$

Leading to:

$$f_i \simeq \frac{r}{\alpha + r} f_{i-1} \quad (24)$$

and then by recursion,

$$f_i \simeq C \left(\frac{r}{\alpha + r}\right)^i. \quad (25)$$

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

If this is valid for any i, the probability that a cluster taken at random is of size i is:

$$p_i = \frac{\alpha}{\alpha + r}\left(\frac{r}{\alpha + r}\right)^{i-1}. \tag{26}$$

We compare this approximation with the numerical results and they are in good agreement (figure 2H), except when both q is small and r/α is large, and even in this case it gives a reasonable approximation.

Replacing f_i by $C\left(\frac{r}{\alpha+r}\right)^i$, equation (22) simplifies to:

$$\lambda \simeq -r\left(1 + \frac{\alpha}{r}\right) + \alpha + 2\alpha\left(\frac{r}{\alpha + r}\right) + \sum_{j=2}^{\infty} 2\alpha q\left(\frac{r}{\alpha + r}\right)^j \tag{27}$$

which after simplifications leads to:

$$\lambda \simeq r\frac{\alpha + (2q-1)r}{\alpha + r}. \tag{28}$$

This approximation does not work for $q < 0.5$, but it works well for q close to 1, and gives the right dependence for r/α large for $q > 0.5$ (figure 2G). Intuitively, if $q > 0.5$, when a cluster breaks it leads to more than one independent linear cluster, thus the population of linear clusters and thus free bacteria may continue to increase with $r\alpha$, whereas if $q < 0.5$, clusters that break lead to less than one independent cluster on average, and thus, as the breaking rate increases with the size, the growth of the population is stunted when r/α increases.

Model with force-dependent breaking rate

Equations

What drives link breakage? The links could break if there was some process degrading the sIgA, but the sIgA are thought to be very stable[17]. Another possible explanation for link breaking is that the bound antigen can be extracted from the bacterial membrane, which may vary exponentially with the force[18][13]. The forces applied on the links are likely mostly due to the hydrodynamic forces exerted by the digesta flow on the bacterial chain. Taking the linear chain as a string of beads, as done for polymer chains, and in a flow with a constant shear rate, the force is predicted to be larger as the chain grows longer, and the largest at the center of the chain[15]. A more detailed discussion and the calculations can be found in section 5.1 in the appendix. Taking α as the breaking rate in the absence of shear, and β a constant expressing the strength of the coupling between hydrodynamic forces and link breaking, the resulting equations for this minimal model taking into account the forces are:

$$\frac{dn_1}{dt} = -rn_1 + 2\sum_{i=2}^{\infty} \alpha n_i \exp\left(\beta\frac{i-1}{2}\right) \tag{29}$$

and for i even,

$$\frac{dn_i}{dt} = -rin_i - \alpha n_i e^{\beta i^2/8}\left(1 + 2\sum_{j=2}^{i/2} e^{-(j-1)^2\beta/2}\right) + r(i-1)n_{i-1} + 2\alpha n_{i+1}e^{\beta i/2} \tag{30}$$

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and for $i > 1$ odd,

$$\frac{dn_i}{dt} = -rin_i - 2\alpha n_i e^{\beta i^2/8} \sum_{j=1}^{(i-1)/2} e^{-(j-1/2)^2\beta/2} + r(i-1)n_{i-1} + 2\alpha n_{i+1} e^{\beta i/2} \tag{31}$$

Free bacteria growth rate as a function of the bacterial replication rate

The growth rate as a function of the replication rate has a qualitatively similar shape as for the base model (figure 2I), with a finite replication rate maximizing the growth rate. The limit $\beta \to 0$ corresponds well to the base model, as expected. When β increases, the replication rate maximizing the growth rate increases, as the effective breaking rate is higher. Numerically, we find (see supplementary figure 9) that the replication rate maximizing the growth rate scales as $\alpha \exp(0.8\beta)$.

Chain length distribution

Similarly to the other models, for t long enough, $n_i \simeq p_i \exp(\lambda t)$ (with λ the largest eigenvalue), and assessing which terms in equations (30) and (31) will be dominant, we ultimately obtain (details in supplementary section 5.3):

$$p_i \simeq \left(\frac{r}{\alpha}\right)^{i-1} \frac{(i-1)!}{Y^{floor(i/2)} Z^{floor((i-1)/2)}} \exp\left(-\frac{\beta}{8}\left(-1 + \frac{i + 3i^2 + 2i^3}{6}\right)\right), \tag{32}$$

with $Y = 1 + 2\sum_{j=1}^{\infty} \exp(-\beta j^2/2) = \theta_3(0, \exp(-\beta/2))$ and $Z = 2\sum_{j=1}^{\infty} \exp(-\beta(j-1/2)^2/2) = \theta_2(0, \exp(-\beta/2))$. This approximation works well, except for small β (figure 2J, and supplementary figure 10). Compared to the base model, the number of clusters decreases much faster with their size. Indeed, the breaking rates for each link increase importantly with the cluster size, thus larger clusters are much less stable than in the base model.

Comparison with experimental data

We analyzed (see section 6 in the appendix) experimental data from[10], of vaccinated mice infected with *S.*Typhimurium. Most clusters are large, and of complex shape. But smaller clusters are linear, and we obtained the following distribution : clusters of size 2 (106), 3 (40), 4 (94), 5 (11), 6 (15), 7 (19), 8 (19), 9(1), 10 (2), 11 (1), 12 (2), 13 (2), 14 (1). The data may be biased, as longer chains may not be fully in the focal plane. As there are not enough data points, we cannot quantitatively fit the data, in particular for larger chain lengths. We can nevertheless give some qualitative points. The larger value at 4 is in line with a fixed time between divisions. Clusters of uneven size could be evidence that linear chains do break. The distribution is relatively narrow, which could be compatible with force-dependent breaking rates.

Summary of results and discussion

We started from the recent finding[10] that the protection effect of sIgA, the main effector of the adaptive immune system in the gut, can be explained by enchained growth. Because sIgA are multivalent, they can

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

stick identical bacteria together if they encounter each other. Early in infection, bacteria of the same type are at low density, thus typical encounter times are very long, but when a bacteria replicates, the daughter bacteria are in contact and thus can remain enchained to each other by IgA. Bacteria in clusters are less motile than individual bacteria, and in particular, are not observed close to the epithelial cells. In the case of wild type *S.*Typhimurium, only free bacteria which can interact with the epithelial cells contribute to the next steps of the infection process. Despite the presence of sIgA, some free bacteria are observed. It could be that they escape at the moment of replication. But, along with the observation that clusters do not grow indefinitely, it could also be a sign that the links between bacteria break. It is also physically expected that the links have some finite breaking rate. If the typical time between two bacterial divisions is much larger than the typical time for the link to break, then there would be no cluster. Conversely, in the inverse case, bacteria will be very likely to be trapped in large clusters. Then, even if sIgA are produced against all bacterial types, the bacteria dividing faster will be disproportionately affected.

We investigated if this qualitative idea holds with more realistic models. We started from a base model in which: bacteria replicate at a fixed rate; remain enchained upon replication; until the link between them breaks at a given fixed breaking rate, identical for all links; and considering that, because of the way bacteria such as *Salmonella* or *E.coli* divide, the early clusters are linear chains of bacteria; when the chain breaks at an outermost link, we assumed the free bacteria will escape; but if the chain breaks elsewhere, we assumed that the two resulting sub-chains encounter each other quickly and form clusters of more complex shapes from which individual bacteria do not escape. We studied this base model with a combination of analytical and numerical approaches. We also tested the robustness of our findings by studying separately several variations of the base model: a probability of escaping upon replication, loss rates, fixed replication time, non-zero probability for the subchains to escape, and force-dependent breaking-rates. For each model, we studied how the growth rate of the free bacteria varies with the replication rate (which would be equal if there were no clusters), and the distribution of cluster sizes. Clusters seem unable to come close to the epithelial cells[10], thus only free bacteria interact directly with epithelial cells and may lead to systemic infections.

We find that, except in the very specific case in which subchains always escape upon link breaking, the growth rate of the number of free bacteria is lower than the replication rate. And more spectacularly, in most of the models studied (but not if more than half the subchains escape upon link breaking, or if there is a significant probability for bacteria to escape enchainement upon replication), the growth rate of the number of free bacteria is non-monotonous with the replication rate : there is a finite replication rate which maximizes the growth rate of non-clustered bacteria. At very high replication rates, bacteria get trapped in more complex clusters and cannot contribute anymore to the free bacteria dynamics and thus to the next steps of the infection process. The replication rate maximizing the growth rate is of the order of the breaking rate, though its specific value depends on the details of the model.

The cluster size distribution is dependent on the model. In most cases, the probability for a linear cluster to be of size k decreases as γ^k, with γ some constant sm+aller than 1. When replication occurs at fixed time,

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and when breaking rates are force-dependent, the probability of larger clusters decreases faster. There are models with different cluster size distributions but qualitatively similar dependence of the growth rate on the replication rate, and the opposite is true too. This shows that large clusters have little importance for free bacteria production, what matters most is the small clusters dynamics. It is reassuring, as we did not consider buckling, which would make long linear chains fold on themselves and produce more complex clusters, and may bias the linear cluster distribution for very large sizes. It should also be noted that with fixed division time, not only the distribution is bumpy, as clusters comprising a power of two number of bacteria are more frequent than others, but the distribution is also narrower. Bacteria divide at approximately fixed division times, while replication is most often taken as occuring at fixed rates, because this makes calculations easier. Sometimes this modelling choice can lead to significant differences.

We analyzed experimental data on clusters of *S.*Typhimurium in the cecum of vaccinated mice. We have not enough data to quantitatively fit the cluster size distribution, but the distribution is qualitatively plausible with the fixed division time model (which is indeed more realistic for bacteria), and with force-dependent breaking rates. With more data, the shape of the distribution could be fitted to compare which model is the most plausible. To test the dependence of the growth rate with the replication rate, an ideal experiment would be to compare similar bacterial strains, but with differing replication rates, and compete them in the same individual. It is however very challenging to obtain bacteria that differ only by their replication rate, particularly in vivo.

sIgA-enchained bacterial clusters could be studied in vitro to measure how they break. However, using in vitro results to draw conclusions on in vivo systems is limited. First, there could be chemical or enzymatic components of the lumen that could facilitate or hinder link breaking, and the non-Newtonian viscosity of the digesta could play a role in the mechanic forces felt by the links, thus a simple buffer may not mimic well the real conditions. More crucially, the exact forces felt by particles of the size of bacterial clusters are not well characterized. Most studies of the flow characteristics in the digestive system rely either on external observations of the peristaltic muscles[19] or indirect measures of times for a marker to exit some section of the digestive track[20]. More quantitative study of the digestive flow at small scales is just beginning[21, 22, 23, 24, 25] and in the future it may give more clues to assess to which forces bacteria are subjected to in the digestive track.

The mechanism we propose is nevertheless plausible. The observation in vaccinated mice of the existence of single bacteria and small clusters, and particularly small linear clusters with an odd number of bacteria, are pieces of evidence that clusters do break in these in vivo conditions. An alternative explanation could be that some bacteria escape enchainement upon replication. However, at higher bacterial densities, we have evidence of independent bacteria binding when they encounter, thus sIgA coated bacteria are adhesive. When two daughter bacteria divide, they are in contact, thus if sIgA is adhesive, escape is unlikely (see appendix section 2). Importantly, even though our results show that specific conditions are needed for the growth rate to decrease with high replication rates, we almost always find that the higher the replication rate, the higher the proportion of bacteria trapped in clusters. Thus, even when it does not reverse the relationship between the growth rate of the free bacteria

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and the replication rate, it is at least dampening this relationship, and can be a tool both to control pathogenic bacteria, but also to maintain homeostasis of the gut microbiota. It is also interesting that there are other host effectors besides sIgA that bind bacteria together: neutrophil extracellular traps for instance[26], and there could also be an interplay between replication rates and the breaking of the links mediated by these other effectors, as the mechanism we propose here is generic.

As for any mechanism to fight against bacteria, how easily resistance can be evolved is crucial. On the one hand, the replication rate could evolve. But a bacteria replicating slower would be less competitive with other bacteria in the absence of sIgA, and a slower growth leaves more time for further host response. On the other hand the typical link breaking time could evolve. On the host side, sIgA is thought to be mechanically very stable, and experiments about the bonding of cells by sIgA seem to point to the link failing because of the extraction of the antigen rather than because of sIgA breaking, and rather than the sIgA/antigen bond detaching[18][12]. In the case of IgA defficiency, there is more secretion of IgM, and microbiota is disturbed[27]: we may speculate that IgM being less powerful for microbiota homeostasis is related to these immunoglobulins being more protease-sensitive than IgA and thus cleaved on shorter time scales[28]. On the other side, bacteria could evolve surface antigens. It could be interesting to think that bacteria could produce decoy antigens with no functional value, but against which the immune system will mount an immune response, and that are more easily released from the bacteria, thus disabling the main sIgA mode of action (being easily evolvable would also be a benefit). Such decoys would however be a metabolic cost for the bacteria, and when breaking, may unmask other antigens corresponding to crucial functions of the bacteria. It could be argued that the capsule around bacteria such as *Salmonella* spp., and also common in pathogenic *E.coli*, may behave as a decoy, though it has also other functions. Along the same lines, we may speculate whether mechanical aspects could be a reason why sIgA against some antigens are not efficient for protection. For instance, while anti-flagella sIgA aggregate very well *Salmonella* Enteriditis together, they are not efficient for protection[29]. A main reason could be that as *Salmonella* can switch flagella production on and off, then some *Salmonella* will always escape these sIgA, and seed the infection[30]. An additional possibility could be that flagella may more easily break, especially as distance between bacteria bound by flagella (long) is likely larger than for bacteria bound by O-antigens (on chains shorter than flagellas), and thus the shear forces would be larger. Further, the mechanical properties of the outer sugar layer of the gram negative bacteria could vary, and thus could be used to tune interactions. However, it would add another constraint on bacteria, and the general result that the growth rate compared to the replication rate is at least dampened by the cluster formation would remain.

In the crowded environment of the gut, it is hard for the host to identify the good and the bad bacteria. That vaccination with dead bacteria is sufficient to produce sIgA and protection, shows that the host does not discriminate well against which bacteria they produce sIgA, as these dead bacteria do not harm. Linking the effect (here the clustering) of the immune effectors with a property directly relevant to the potential bacterial pathogeneicity (here the replication rate) avoids to make complex decisions about which bacteria to produce effectors against.

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

[1] Helander, H. F. and Fändriks, L. (2014) Surface area of the digestive tract–revisited. *Scandinavian journal of gastroenterology*, **49**, 681–689.

[2] Casteleyn, C., Rekecki, A., Van der Aa, A., Simoens, P., and Van den Broeck, W. (2010) Surface area assessment of the murine intestinal tract as a prerequisite for oral dose translation from mouse to man. *Laboratory animals*, **44**, 176–183.

[3] Sender, R., Fuchs, S., and Milo, R. (2016) Revised estimates for the number of human and bacteria cells in the body. *PLoS biology*, **14**, e1002533.

[4] Stecher, B. and Hardt, W.-D. (2008) The role of microbiota in infectious disease. *Trends in microbiology*, **16**, 107–114.

[5] Dishaw, L. J., Cannon, J. P., Litman, G. W., and Parker, W. (2014) Immune-directed support of rich microbial communities in the gut has ancient roots. *Developmental & Comparative Immunology*, **47**, 36–51.

[6] Donaldson, G. P., Lee, S. M., and Mazmanian, S. K. (2016) Gut biogeography of the bacterial microbiota. *Nature Reviews Microbiology*, **14**, 20.

[7] Williams, R. and Gibbons, R. (1972) Inhibition of bacterial adherence by secretory immunoglobulin a: a mechanism of antigen disposal. *Science*, **177**, 697–699.

[8] Strugnell, R. A. and Wijburg, O. L. (2010) The role of secretory antibodies in infection immunity. *Nature Reviews Microbiology*, **8**, 656.

[9] Endt, K., et al. (2010) The microbiota mediates pathogen clearance from the gut lumen after non-typhoidal salmonella diarrhea. *PLoS pathogens*, **6**, e1001097.

[10] Moor, K., et al. (2017) High-avidity IgA protects the intestine by enchaining growing bacteria. *Nature*, **544**, 498–502.

[11] Kudoh, K., Shimizu, J., Ishiyama, A., Wada, M., Takita, T., Kanke, Y., and Innami, S. (1999) Secretion and excretion of immunoglobulin a to cecum and feces differ with type of indigestible saccharides. *Journal of nutritional science and vitaminology*, **45**, 173–181.

[12] Tees, D. F., Coenen, O., and Goldsmith, H. L. (1993) Interaction forces between red cells agglutinated by antibody. IV. Time and force dependence of break-up. *Biophysical Journal*, **65**, 1318–1334.

[13] Evans, E. A. and Calderwood, D. A. (2007) Forces and bond dynamics in cell adhesion. *Science*, **316**, 1148–1153.

[14] McGrady, E. and Ziff, R. M. (1987) "shattering"transition in fragmentation. *Physical review letters*, **58**, 892.

[15] Odell, J. and Keller, A. (1986) Flow-induced chain fracture of isolated linear macromolecules in solution. *Journal of Polymer Science Part B: Polymer Physics*, **24**, 1889–1916.

[16] Gherardi, M., Amato, A., Bouly, J.-P., Cheminant, S., Ferrante, M. I., d'Alcalá, M. R., Iudicone, D., Falciatore, A., and Lagomarsino, M. C. (2016) Regulation of chain length in two diatoms as a growth-fragmentation process. *Physical Review E*, **94**, 022418.

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[17] Brandtzaeg, P. (2003) Role of secretory antibodies in the defence against infections. *International Journal of Medical Microbiology*, **293**, 3–15.

[18] Evans, E., Berk, D., and Leung, A. (1991) Detachment of agglutinin-bonded red blood cells. I. Forces to rupture molecular-point attachments. *Biophysical Journal*, **59**, 838–848.

[19] Hulls, C., Lentle, R. G., de Loubens, C., Janssen, P. W., Chambers, P., and Stafford, K. J. (2012) Spatiotemporal mapping of ex vivo motility in the caecum of the rabbit. *Journal of Comparative Physiology B*, **182**, 287–297.

[20] Janssen, P. W. M., Lentle, R. G., Asvarujanon, P., Chambers, P., Stafford, K. J., and Hemar, Y. (2007) Characterization of flow and mixing regimes within the ileum of the brushtail possum using residence time distribution analysis with simultaneous spatio-temporal mapping. *The Journal of Physiology*, **582**, 1239–1248.

[21] Lentle, R. and Janssen, P. (2008) Physical characteristics of digesta and their influence on flow and mixing in the mammalian intestine: a review. *Journal of Comparative Physiology B*, **178**, 673–690.

[22] Kim, H. J., Huh, D., Hamilton, G., and Ingber, D. E. (2012) Human gut-on-a-chip inhabited by microbial flora that experiences intestinal peristalsis-like motions and flow. *Lab on a Chip*, **12**, 2165.

[23] Lentle, R. G. and Loubens, C. (2015) A review of mixing and propulsion of chyme in the small intestine: fresh insights from new methods. *Journal of Comparative Physiology B*, pp. 369–387.

[24] Cremer, J., Segota, I., Yang, C.-y., Arnoldini, M., Sauls, J. T., Zhang, Z., Gutierrez, E., Groisman, A., and Hwa, T. (2016) Effect of flow and peristaltic mixing on bacterial growth in a gut-like channel. *Proceedings of the National Academy of Sciences*, **113**, 11414–11419.

[25] Cremer, J., Arnoldini, M., and Hwa, T. (2017) Effect of water flow and chemical environment on microbiota growth and composition in the human colon. *Proceedings of the National Academy of Sciences*, **114**, 6438–6443.

[26] Fournier, B. and Parkos, C. (2012) The role of neutrophils during intestinal inflammation. *Mucosal immunology*, **5**, 354.

[27] Mirpuri, J., Raetz, M., Sturge, C. R., Wilhelm, C. L., Benson, A., Savani, R. C., Hooper, L. V., and Yarovinsky, F. (2014) Proteobacteria-specific iga regulates maturation of the intestinal microbiota. *Gut microbes*, **5**, 28–39.

[28] Richman, L. K. and Brown, W. R. (1977) Immunochemical characterization of igm in human intestinal fluids. *The Journal of Immunology*, **119**, 1515–1519.

[29] Iankov, I. D., Petrov, D. P., Mladenov, I. V., Haralambieva, I. H., and Mitov, I. G. (2002) Lipopolysaccharide-specific but not anti-flagellar immunoglobulin a monoclonal antibodies prevent salmonella enterica serotype enteritidis invasion and replication within hep-2 cell monolayers. *Infection and immunity*, **70**, 1615–1618.

[30] Diard, M., Garcia, V., Maier, L., Remus-Emsermann, M. N., Regoes, R. R., Ackermann, M., and Hardt, W.-D. (2013) Stabilization of cooperative virulence by the expression of an avirulent phenotype. *Nature*, **494**, 353.

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

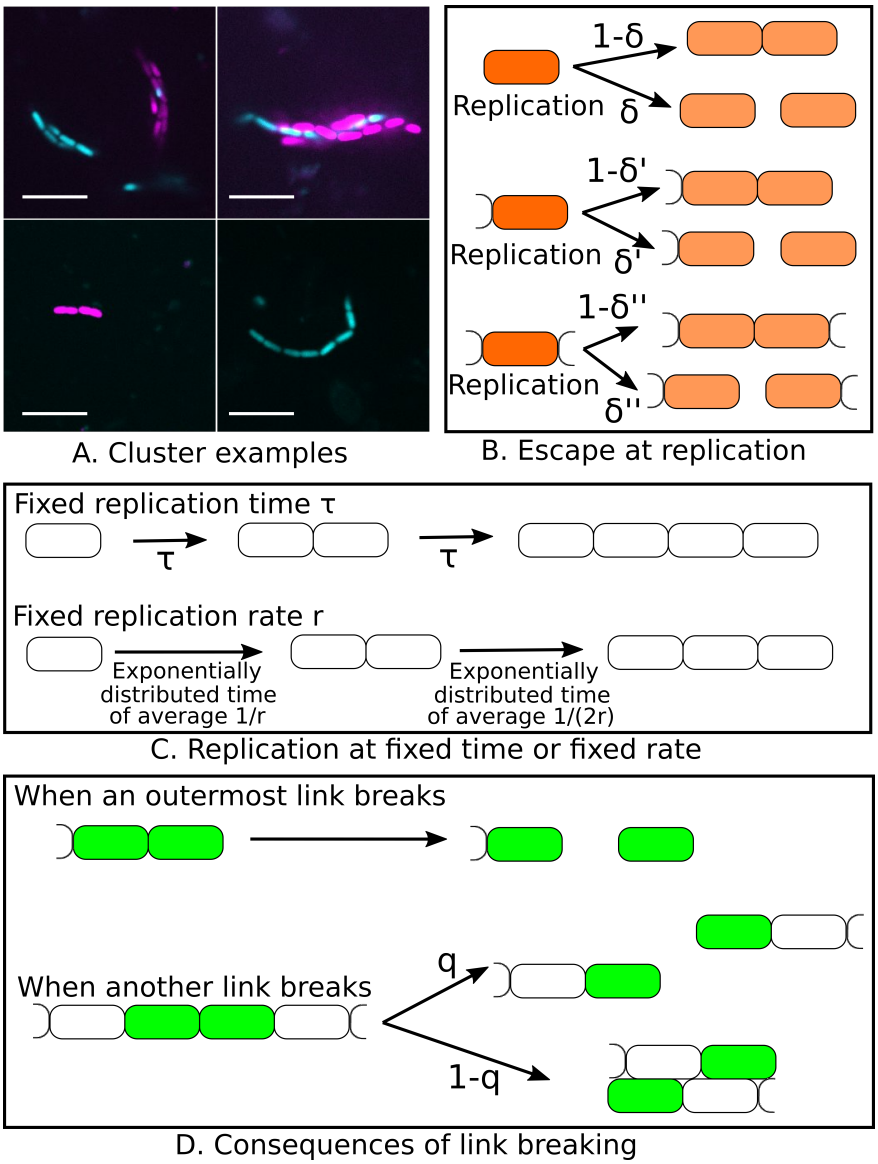

Figure 1: Bacterial cluster modeling. A. Representative experimental images of bacterial clusters in cecal content of vaccinated mouse at 5h post infection with isogenic GFP and mCherry expressing S.typhimurium. The scale bar is $10\mu m$. Top images: complex clusters made from bundles of linear clusters, which could be re-linked single chains (left) or formed from at least two independent clones (indicated by fluorescence, right). Bottom images: linear clusters which dynamics we aim to model. B. Potential bacterial escape at replication (in the base model, $\delta = \delta' = \delta''$). C. Fixed replication time or fixed replication rate (the latter is chosen for the base model). D. Consequences of link breaking. In the base model, $q = 0$.

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

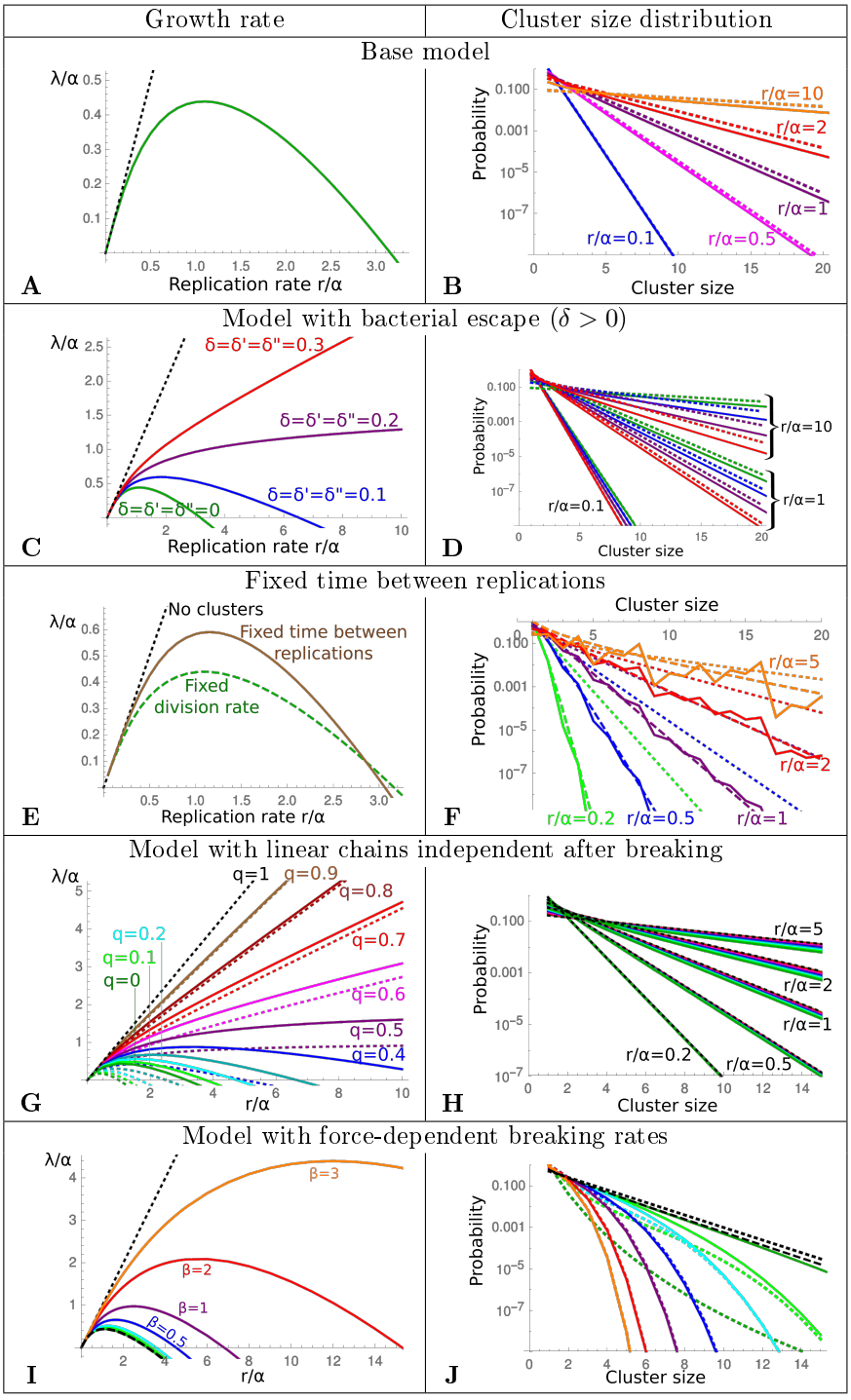

Figure 2:

Legend figure 2 :
A,C,E,G,I: Growth rate λ of the free bacteria as a function of the bac-

bioRxiv preprint doi: https://doi.org/10.1101/298059; this version posted April 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

teria replication rate r, both in units of α. Numerical results (solid colored lines), and limit with no clusters ($\lambda = r$) (black dotted line). B,D,F,H,J: Cluster size distribution. Solid lines: numerical results. A,B: Base model, $n_{max} = 40$. B. dotted lines: approximation (5) (almost overlaid with the numerical results for $r/\alpha = 0.1$). C,D: Model with bacterial escape. $\delta = \delta' = \delta'' = 0$, 0.1, 0.2, 0.3. $c = c' = 0$, $n_{max} = 40$. D. dotted lines: approximation (9). E,F: Fixed time between replications. $r_{eff} = \log(2)/\tau$. $n_{max} = 32$. F. approximation (17) (dashed lines), numerical result in the base model (dotted lines). $r/\alpha = 0.2$, 0.5, 1, 2, 5. G,H: Model with linear chains independent after breaking. $n_{max} = 100$. G. The dotted black line is the case $q = 1$, for which $\lambda = r$, like in the absence of clusters. The colored dotted lines are the analytical approximation (28). H. The dotted black lines are the approximate distribution (26) for each r/α, which is the exact distribution for $q = 1$. The colours represent the same q values than for the left panel. All curves are almost overlaid for small r. I,J: Model with force-dependent breaking rates. Each color represents a different β: $\beta = 0.01$ ($n_{max} = 20$), $\beta = 0.1$ ($n_{max} = 15$), $\beta = 0.2$ ($n_{max} = 15$), $\beta = 0.5$ ($n_{max} = 15$), $\beta = 1$ ($n_{max} = 15$), $\beta = 2$ ($n_{max} = 10$), $\beta = 3$ ($n_{max} = 10$). The black dashed lines are the numerical results for the base model, equivalent to $\beta = 0$. The curves for $\beta = 0.01$ (dark green) are almost overlaid with the curves for $\beta = 0$. J. Distribution of the cluster sizes for $r/\alpha = 1$. The colored dotted lines the analytical approximation (32), and the black dotted line the approximation for the base model (5).