npş

www.nature.com/ismej

# **ORIGINAL ARTICLE**

# Resource limitation drives spatial organization in microbial groups

Sara Mitri<sup>1,2</sup>, Ellen Clarke<sup>3</sup> and Kevin R Foster<sup>2,4</sup>

<sup>1</sup>Department of Fundamental Microbiology, University of Lausanne, Lausanne, Switzerland; <sup>2</sup>Department of Zoology, University of Oxford, Oxford, UK; <sup>3</sup>All Souls College, Oxford, UK and <sup>4</sup>Oxford Centre for Integrative Systems Biology, University of Oxford, Oxford, UK

Dense microbial groups such as bacterial biofilms commonly contain a diversity of cell types that define their functioning. However, we have a limited understanding of what maintains, or purges, this diversity. Theory suggests that resource levels are key to understanding diversity and the spatial arrangement of genotypes in microbial groups, but we need empirical tests. Here we use theory and experiments to study the effects of nutrient level on spatio-genetic structuring and diversity in bacterial colonies. Well-fed colonies maintain larger well-mixed areas, but they also expand more rapidly compared with poorly-fed ones. Given enough space to expand, therefore, well-fed colonies lose diversity and separate in space over a similar timescale to poorly fed ones. In sum, as long as there is some degree of nutrient limitation, we observe the emergence of structured communities. We conclude that resource-driven structuring is central to understanding both pattern and process in diverse microbial communities.

The ISME Journal (2016) 10, 1471–1482; doi:10.1038/ismej.2015.208; published online 27 November 2015

### Introduction

A key determinant of whether a microbial cell survives and divides is the identity of surrounding cells. These neighboring cells determine the signaling molecules it will perceive, whether it will be infected by plasmids or viruses, whether it will be attacked by toxins, and, most fundamentally, its access to resources. On an evolutionary timescale, the pressures exerted by neighbors can lead to the competitive exclusion of a genotype or to new evolutionary adaptations (Kerr et al., 2002; Habets et al., 2006; Kim et al., 2014). These adaptations include both cooperative and competitive phenotypes that shape the productivity and functioning of the group, such as the secretion of polymers that allow cells to get better access to nutrients (Kim et al., 2014), or the evolution of lethal antibiotics (Kreft, 2004; Nadell et al., 2009; Mitri and Foster, 2013; Koch et al., 2014). In order to understand the population dynamics of microbial groups, it is therefore necessary to understand how and why genotypes and phenotypes organize in space (Shapiro, 1995; Johnson and Boerlijst, 2002; Korolev *et al.*, 2014; Murray *et al.*, 2014; Van Gestel *et al.*, 2014).

Theory and experiments have revealed that, in the absence of mutation and selection, initially diverse and well-mixed microbial populations will commonly lose their diversity when growing in dense groups such as bacterial colonies, leading to large patches of single genotypes (Ben-Jacob et al., 1994; Golding et al., 1999; Habets et al., 2006; Hallatschek et al., 2007; Xavier and Foster, 2007; Hallatschek and Nelson, 2010; Korolev et al., 2012). Theory suggests that nutrient limitation is a general factor that underlies this process (Nadell et al., 2010). Specifically, nutrient limitation ensures that only cells at the edge can grow, which leads to bottlenecks and genetic drift that drives a loss of diversity (Hallatschek et al., 2007; Nadell et al., 2010). However, we lack empirical tests that demonstrate this key link between nutrient levels and the genetic or phenotypic organization of microbial groups.

Here we use a combination of simulation modeling and mixed-genotype *Pseudomonas aeruginosa* bacterial colonies to evaluate the role of nutrient levels in the spatial structuring of microbial groups. We find that, as predicted by theory, colonies with abundant resources maintain large unstructured regions. High resource supply means that a large sub-population of cells will be dividing rapidly, allowing many different lineages to be maintained as the group expands. However, while high-nutrient groups remain well mixed for many more cell

Correspondence: S Mitri, Département de Microbiologie Fondamentale, Université de Lausanne, Bâtiment Biophore, UNIL-Sorge, Lausanne CH-1015, Switzerland.

E-mail: sara.mitri@unil.ch

or KR Foster, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

E-mail: kevin.foster@zoo.ox.ac.uk

Received 14 July 2015; revised 13 October 2015; accepted 16 October 2015; published online 27 November 2015

divisions than low-nutrient groups, they also expand more rapidly. Given sufficient space to expand, therefore, high- and low-nutrient groups experience a comparable loss of diversity over time. Our work suggests that the effects of resource limitation will be central to understanding the organization and evolution of many microbial communities.

### Materials and methods

### Experimental setup

To study the effects of varying nutrient abundance in a setup that can be easily compared with theoretical predictions (Hallatschek and Nelson, 2008; Hallatschek and Nelson, 2010; Korolev et al., 2010; Korolev et al., 2012), we use mixed-genotype colonies of Pseudomonas aeruginosa strain PAO1 and follow the patterning of two equally fit genotypes that differ only in the constitutive expression of either yellow or cyan fluorescent protein (YFP and CFP, respectively; Supplementary Figure S1). By restricting the duration of our experiments, we can study the physical mechanisms of diversity loss and assume the absence of genetic evolution through mutations (Supplementary Figure S2). This system allows us to follow the fate of different lineages present in an initially well-mixed droplet.

In preliminary experiments with a number of strains, we observed that in some isolates the two strains separated after 1 or 2 days (strain PA4, others remained mixed Figure 1a), while (strain PAO1, Figure 1b). Clean deletion of the *pilB* gene in PAO1 led to a much more defined spatial structure (Figure 1c), revealing that the lack of structuring was due to type IV pili-associated surface motility (Mattick, 2002; Burrows, 2012). To focus on the direct effects of nutrients on spatial structure, our detailed analysis was therefore conducted with PAO1  $\Delta$ pilB. This follows studies of patterning in Escherichia coli that also used non-motile strains (Hallatschek et al., 2007; Hallatschek and Nelson, 2010; Korolev et al., 2011). We focus throughout on images taken of the surface of colonies, but confocal microscopy reveals that the patterns observed at the surface are representative of growth throughout these relatively thick colonies (Figure 1d). To assess the localization of growth within the colonies, we used a strain with an rRNA transcriptional green fluorescent protein (GFP) fluorescent reporter (rrn), such that GFP is produced as cells grow.

We compared the spatial patterning of genotypes in colonies across a range of eight Luria Broth concentrations in agar plates. We selected this range by studying the growth rate of *P. aeruginosa*. This revealed inhibitory effects at high-nutrient concentrations (Heurlier *et al.*, 2005; Supplementary Figure S3). To avoid these effects, we focus on nutrient concentrations in the range where growth rate increased linearly  $(0.025-0.2 \times \text{ in steps})$ of  $0.025 \times$ , Supplementary Figure S3).

We initiated all colonies by mixing exponentially growing cells at a 1:1 ratio and spotting 2  $\mu$ l (or 1, 2 or 4  $\mu$ l in experiments where the drop volume was varied) onto the agar. Once the drops had dried, we overturned the plates and left them for 14 days at room temperature (22 °C). We imaged all  $\Delta$ pilB and rrn colonies using a stereoscope (Zeiss Lumar V.12, Jena, Germany) 1 h after the colonies had dried, every 24 h thereafter for the first 10 days and subsequently every 48 h. Additionally, we used confocal laser scanning microscopy to quantify rrn cell growth at the colony edge and to verify that patterns at the surface of  $\Delta$ pilB colonies were representative of spatial patterns deeper below the surface (Figure 1d).

### Image analysis

For all analyses involving colonies at a particular time point, we used images from day 12 (Supplementary Figure S2 justifies this choice). We assume that patterns along the circumference of the colony at a given distance from the inoculum can be mapped back to a specific time point. This assumption is based on two observations: (i) that the radial expansion velocity of the colony is approximately linear (see Figure 2b), and (ii) that patterns do not shift significantly within the colony over time: a generalized linear model showed no significant changes in demixing distances over time (from day 4 to day 14, P = 0.33, Supplementary Figure S4A), and



**Figure 1** *P. aeruginosa* colonies grown from a 1:1 mixture of YFP- and CFP-labeled cells. (a) Wild-type PA4, (b) wild-type PAO1 (c) PAO1  $\Delta$ pilB and (d) confocal image of green cells in a sector of a PAO1  $\Delta$ pilB colony; empty spaces represent blue cells. The small red square in the top right of the colony in panel (c) highlighted by the red arrow shows the scale of the confocal image in panel (d) relative to the whole colony (the position does not correspond).





**Figure 2** Growth of colonies of a 1:1 mixture of *P. aeruginosa* YFP- and CFP-labeled  $\Delta$ pilB mutants. (a) Images of the edges of colonies at different nutrient concentrations on day 12. White rings show the size of the initial drop; red rings show where our automated image analysis has defined that the two strains have demixed. (b) Colony radius (solid lines) over 14 days (three replicates for each treatment), and the distance between the inoculum and the demixing point (dots). Once segregation has occurred, the demixing distance does not change significantly (discussion in text, Supplementary Figure S4). (c) Heterozygosity *H* (see Materials and methods, Supplementary Figure S5) of colonies on day 12 as a function of distance from the inoculum. Inset: the demixing distance as a function of nutrient concentration. (d) *H* on day 12 as a function of time. This is obtained by dividing the *x* axis in panel (c) by the expansion velocity *v* of each colony (slope of lines in panel (b)). Inset: the demixing time as a function of nutrient concentration. (c-d) Data show that the demixing point in both time and space correlate with nutrient concentrations.

heterozygosity curves do not appear to change much over time (Supplementary Figure S4C).

We also used image analysis to estimate experimental parameters (details in Supplementary Text S2). Heterozygosity H was measured using algorithm visualized and described the in Supplementary Figure S5. In short, we computed H by estimating the diversity in individual pixels and averaging over pixels at a distance from the colony inoculum. Heterozygosity X H is 0.5—its maximum for two strains—if all pixels around the circumference contain equal concentrations of both colors and goes down as pixels become more likely to contain one or the other strain. The demixing distance was then determined by finding the distance at which the first derivative of H (dH/dx) was minimal. The logic behind this is that, after that point, the pattern begins to converge on its final sector number. Heterozygosity H is shown either as a function of distance (for example, Figure 2c), which is obtained as described above, or as a function of time (for example, Figure 2d), which is calculated by dividing the distance in x axis by each colony's radial expansion velocity v.

Colony heterozygosity (CH) was used to estimate the likelihood of a cell at an arbitrary place in the colony being in a clonal patch. Its measurement was similar to that of heterozygosity H, but pixels were sampled at equidistant points over the whole colony. For a time-controlled measure of CH, we took all points reaching from the inoculum all the way to the outside of the colony at that time point (for example, day 12). For the size-controlled measure, we determined the radius of the smallest colony at that time point (for example, day 12) and for each colony used all points between the inoculum and that radius to calculate CH.

The cellular diffusion rate  $D_s$  was determined by analyzing the boundaries between the sectors in the colonies. These boundaries were found automatically using an edge detection package developed by Kovesi (2000). We measured the distance  $R_i$  between the colony center and the closest point on the boundary, the furthest point  $R_f$ , and angle  $\alpha$  between each point on the boundary and a horizontal line drawn through the center of the colony. We then used these measurements to compute cellular diffusion rate  $D_s$  (Figure 6).

#### Estimating other experimental parameters

We estimated the growth rate  $\mu_{\rm N}$  by growing PAO1  $\Delta$ pilB YFP cells in triplicate in the eight Luria Broth concentrations N (0.025–0.2×) in a shaken 96-well plate and measuring OD<sub>600</sub> over 24 h (Supplementary Figure S6, Supplementary Text S2). We calculated colony expansion velocity v for each colony by fitting a line (linear regression) through the colony radii over the 14-day experiment and the initial colony radius  $R_0$ . Growing edge width w (the area at the colony edge in which cells were growing) was calculated as  $w = 2v/\mu_{\rm N}$ .

The number of sectors S was counted manually. Effective population size  $N_{\rm e}$  was calculated as  $N_{\rm e} = \frac{w h_{\rm e} w_{\rm c}}{2V_{\rm c}}$ , where w is the width of the growing edge,  $\overline{w_{c}}$  the width of a single cell in the edge (here  $w_{\rm c} = 1 \,\mu{\rm m}$ ),  $h_{\rm e}$  is the height of the colony at the inner point of the growing edge and  $V_{\rm c}$  is the volume of a single cell (here  $V_c = 1 \mu m^3$ ). Importantly, we verified that cell width  $w_c$  did not change significantly across nutrient concentrations (Mann–Whitney test, P=0.33, Supplementary Figure S7). The genetic diffusion constant  $D_{g}$  is inversely proportional to effective population size  $N_{\rm e}$  where  $\gamma$  is some constant:  $D_{g} = \frac{\gamma}{N_{e}}$  (Korolev *et al.*, 2010, 2011). Genetic diffusion  $D_{g}$  can also be estimated using Equation 2 in Supplementary text 2 (Supplementary Figure S8).

### Computational model

The computational model is described in detail in Nadell et al. (2010). It is a two-dimensional individual-based model, in which each cell has a radius and a position within a grid. Each grid element has a concentration of nutrients, which is determined by solving diffusion-reaction equations to steady state at each time step. A cell's growth rate is calculated using the nutrient concentration in the local grid element and its radius increased accordingly. Cell growth results in cells overlapping in space, which is resolved using a pushing algorithm. Once a cell increases beyond a given radius, it divides into two daughter cells of the same type. Apart from their color, green and blue cells were identical. Differences to previously published studies, as well as estimation of simulation parameters and measurements of growing edge width, colony expansion velocity and heterozygosity H, are detailed in Supplementary Text S2.

### Results

# Diversity loss occurs further in space when resources are abundant

In all experiments, the mixture separated into tens of sectors of apparently clonal patches, each composed of either YFP- or CFP-labeled cells (Figure 2a). This indicates that, of the initially well-mixed drop containing millions of cells, only a fraction formed clonal patches in the growing colony. To quantify the effect of nutrient limitation on diversity loss, we used a previously developed measure of 'heterozygosity' (Nei *et al.*, 1975; Hallatschek and Nelson, 2008; Korolev et al., 2010; Korolev et al., 2011; Materials and methods, Supplementary Figure S5), which refers to diversity at a particular position (pixel) in the colony, averaged over many such local diversity measures in a larger region. The region over which we measure is either an expanding concentric ring around the colony inoculum, which captures the change in diversity as the colony expands (Supplementary Figure S5), or the whole colony (below). At all nutrient concentrations, heterozygosity H was close to its maximum at the well-mixed center where the drop was initially placed. H then decreased further from the center as strains separated from each other, converging toward a value of approximately 0.25 at the edge of all colonies, where the sectors had formed (Figures 2c and d). As predicted by theoretical work (Nadell et al., 2010; Mitri et al., 2011), the concentration of nutrients had a clear effect on the change in heterozygosity H: it dropped more slowly in space (lower magnitude  $d\bar{H}/dx$ nutrients were of as increased (Supplementary Figure S9A). Furthermore. the distance at which the two strains demixed (see Materials and methods) showed a strong positive correlation with nutrient concentration (Figures 2a and c inset, Pearson's  $\rho = 0.92$ , P < 0.001).

This finding fits well with computational models, which predict that colonies with more limited access to nutrients should demix sooner because they have a narrower edge in which cells grow more slowly (Nadell *et al.*, 2010). This is thought to be because the nutrients do not diffuse as far into the colony, resulting in stronger nutrient gradients, such that only few cells at the edge of the colony can grow (Dockery and Klapper, 2002; Nadell *et al.*, 2010). A less intensely dividing growing edge not only translates into a slower group expansion rate (Pirt, 1967) but should also result in stronger bottlenecks, stronger genetic drift and the earlier formation of the observed sectors (Nadell *et al.*, 2010).

### Growing edge width increases with nutrients

We next tested whether, as predicted by theory, the growing edge width does indeed correlate with the rate of diversity loss. We quantified the growing edge using two methods. Taken together, both of these approaches strongly indicate that the more nutrients are available in the agar, the wider the zone at the edge of the colony in which cells are growing. We detail the two approaches next.

In our first approach, we calculated the growing edge width using the radial expansion velocity v (see Materials and methods, Supplementary Figure S10) and the maximum growth rate  $\mu_N$  of the strains which we estimated experimentally for each nutrient concentration N (see Materials and methods). Assuming that cells in the growing edge of a colony are growing at this maximal rate, the width of the growing edge approximates to  $w = 2v/\mu_N$  (Pirt, 1967), which correlates positively with nutrient concentrations (Spearman's  $\rho = 0.98$ , P < 0.001, Figure 3a).

We took a second approach to estimating the width of the growing edge, where we used a *P. aeruginosa* strain with a fluorescent marker linked to rRNA transcription (PAO1 rrn; Gourse *et al.*, 1996), such that GFP is synthesized whenever the



**Figure 3** Estimating the width of the growing edge. (a) Growing edge width w estimated from the radial expansion velocity v (or dr/dt) and the maximum growth rate  $\mu_N$ , measured at different nutrient concentrations in liquid (see Materials and methods, Supplementary Figure S6B):  $w = 2v/\mu_N$ , where the cross-section of the growing edge  $\Delta a$  is approximated to a triangle. Because v increases with nutrient concentration (see Figure 2b), so does w. (b) Image analysis of rrn colonies on day 7, showing the width of the radius of each colony that was above a given fluorescence threshold. Because at some thresholds many colonies were either never above or always above the threshold, multiple thresholds are shown. All data show an increase in growing edge width with nutrient concentration. (c) Fluorescence enhanced equally for all images to improve visualization.

cells are growing. By imaging all colonies at the same exposure to fluorescent light, we could detect where growth was occurring in the colony (see Materials and methods, Figures 3b and c). In agreement with the first method, the images show a significant positive correlation between the width of the growing edge and nutrient concentration (Spearman's  $0.75 < \rho < 0.96$ , all P < 0.001, Figure 3b). These data are further supported by confocal microscopy visualizing the cells at the very edge of these colonies (Supplementary Text S1, Supplementary Figure S11).

Computer simulations recapitulate empirical findings Although our data fit with predictions from published computational models, the previous models were based upon submerged biofilms rather than a colony model (Nadell *et al.*, 2010). We therefore modified the models to better approximate growth in a colony, as opposed to an attached biofilm. In each simulation, cells are placed in a circular cluster in the center of a 2D surface, with nutrients diffusing from around them. Starting from different initial nutrient concentrations, the cells in each of these colonies grew by consuming the nutrients available to them locally (see Supplementary Text S2). As observed in the laboratory experiments, the two simulated strains separated into single-color patches in all colonies, with demixing occurring at a greater radius in the colonies growing on higher initial nutrient concentrations (Figure 4a). Moreover, the model shows that higher nutrients do indeed slow the rate of diversity loss in space compared with lower-nutrient levels (Figure 4b). Finally, as expected, the growing edge width correlated positively with nutrients (Pearson's  $\rho = 0.98$ , P < 0.001, Supplementary Figure S12).

We can also use the simulations to perform a test that is not possible in our laboratory system: we can ask how the system behaves in the complete absence of nutrient limitation, such that cell division occurs at the same rate throughout the whole colony. Under these conditions, the two strains remained mixed independently of nutrient concentration (Figure 4c) and heterozygosity remained at its maximal value just below 0.5 (Figure 4d). This shows that, in these simulations, the strength of nutrient gradients at the edge of the colony dictates the rate at which two strains segregate. In the absence of such nutrient gradients, for example, when a cell population is small and nutrients can diffuse all the way through, diversity will be maintained throughout.



**Figure 4** Results of the modified computer simulations. (**a**, **b**) Cells take up nutrients as they grow, creating gradients at the edge of the colony. All colonies demixed into clonal sectors, with the demixing distance (see Materials and methods) increasing with nutrient concentration. (**a**, **c**) The background gray intensity represents the concentration of nutrients. Here we show images for four out of the eight nutrient concentrations. (**b**) Heterozygosity *H* at different distances from the initial inoculum for different nutrient concentrations. Inset: Demixing distance as a function of nutrients. Both plots show similar results to the experiments: demixing occurs further in space when nutrients are more abundant. (**c**, **d**) Nutrient gradients were removed by keeping the concentration of nutrients constant throughout the environment (cells grew without depleting nutrients). (**c**) With no gradients, the colonies did not form sectors, and (**d**) their heterozygosity remained high. Note that results from simulations are noisier than the experimental data owing to a much lower number of cells and should not be compared quantitatively.

#### Fitting data to a population genetics model

In order to further understand the behavior of our system, for the interested reader we next employ a previously developed population genetics model that captures diversity within microbial colonies in terms of a few key parameters (Korolev *et al.*, 2010; Korolev *et al.*, 2011). Applying this model to our data links our study to this seminal population genetics work and points to colony expansion velocity v as the main explanatory parameter of our data, which we follow-up on in the next section.

To adequately fit our data to the model, we ran an additional experiment where we not only varied nutrient concentrations but also the volume of the drop placed on the agar. Having this additional independent variable was important to better estimate the dependent parameters in the population genetic models (Korolev *et al.*, 2011). This resulted in a range of initial colony radii  $R_0$ , where colonies with greater  $R_0$  grew larger in size (Supplementary Figure S13A) and had more sectors *S* (Supplementary Figure S13B).

According to the population genetics model (Korolev *et al.*, 2010; Korolev *et al.*, 2011), equilibrium diversity is represented by the number of sectors *S* after demixing, which depends on four key parameters: the initial radius of the colony  $R_0$ ; its radial expansion velocity *v*; the genetic diffusion constant  $D_g$ , which characterizes the strength of genetic drift (Korolev *et al.*, 2010); and the cellular diffusion rate  $D_s$ , which describes

the amount of mixing of the two strains at the boundary between two sectors (that is, wandering of sector boundaries). In the published model (Korolev *et al.*, 2010), the number of sectors is then described as:

$$S = \frac{2\pi H_0 v}{D_{\rm g}} + H_0 \sqrt{\frac{2\pi R_0 v}{D_{\rm s}}}$$
(1)

where  $H_0$  is the initial heterozygosity (here  $H_0 = 0.5$ ). From our experiments, we already know expansion velocity v and initial radius  $R_0$  for each colony and can use the pooled data from both experiments to estimate  $D_g$  and  $D_s$  (see Materials and methods).

The genetic diffusion constant  $D_{\rm g}$  is inversely proportional to the effective population size  $N_{\rm e}$ . We estimated  $N_{\rm e}$  from the width of the growing edge together with colony height (Figure 5. Supplementary Figure S14; Korolev et al., 2010; Korolev et al., 2011) and found that it increases with nutrient concentrations (Figure 5). This suggests, as expected, that genetic drift  $D_{g}$  will decrease with resource abundance. We also attempted to estimate the magnitude of  $D_{\rm g}$  by fitting our data to the first term of Equation 1 (Supplementary Figure S8). This term is likely is to be extremely small independently of nutrients, suggesting that  $D_{g}$  is large. In our system then, although genetic drift decreases with increasing nutrients, it is likely to be quite high in all colonies, which is consistent with



**Figure 5** Effective population size  $N_e$  as a function of nutrient concentration at different inoculum sizes.  $N_e$  represents the number of cells in the purple sector of one-cell width and was calculated as indicated in the diagram on the right, where *w* is the width of the growing edge,  $h_e$  the height of the colony at the inside of the growing edge,  $w_c$  the width of a single cell (taken to be 1 µm) and  $V_c$  the volume of a cell, taken to be 1 µm<sup>3</sup>. For simplicity, we assume that the sector is triangular (hence the division by 2). The plot shows that  $N_e$  increases with nutrients. Because  $D_e$  is inversely proportional to  $N_e$ , we expect it to decrease with nutrient concentrations.



Figure 6 Quantifying the cellular diffusion constant  $D_{s}$ . (a) For each colony from both the original experiment and the experiment with varying inoculum size, we automatically traced the boundaries between sectors and determined the distance r and angle  $\alpha$  of each point along the boundary with respect to the center of the colony. For each boundary, we plotted  $1/R_i - 1/R_f$  on the x axis, where  $R_f$  and  $R_i$  are the distance of the furthest and closest point in a boundary to the center, respectively, and the variance in the angle  $\alpha$  for all points along the boundary on the y axis. We then binned boundaries into intervals of  $8 \times 10^{-5}$  on the x axis, took the mean of all variances within each bin (red dots) and fit a line through these means (green dots are >3 s.ds. from the mean and were discarded from the fit), in addition to a point at the origin (red line). The slope of this line was  $2D_s/v$ . By multiplying the slope by v/2, we obtained  $D_s$ . (b) Output of edge detection algorithm. Boundary colors were assigned randomly to improve visualization. These boundaries, together with the boundaries from 95 other colonies, are used to generate the data in panel (a). (c)  $D_s$  correlates positively with nutrient concentrations (Spearman's correlation test,  $\rho = 1$ , P < 0.001).  $D_s/v$  showed no significant correlation with nutrient concentrations ( $\rho = 0.024$ , P = 0.98), suggesting that the colony expansion velocity determines the cellular diffusion rate  $D_s$ . The error bars show s.e. of the fit.

our finding that genetic diversity decreases strongly across all nutrient concentrations.

Our estimate of cellular diffusion rate  $D_s$  increased with nutrient concentrations (Spearman's  $\rho = 1$ ,

P < 0.001) at a rate that was proportional to the increase in the radial expansion velocity v (Figure 6). Accordingly, the ratio of the two  $(D_s/v)$  did not change significantly across nutrient concentrations

 $(\rho = 0.024, P = 0.98, Figure 6)$ . In Equation 1 then,  $D_s$  and v cancel each other out. Importantly, this suggests that changes in the level of mixing  $(D_s)$  as a function of resource levels are due to differences in expansion velocity v.

Colony expansion velocity explains rate of diversity loss We next examined whether a single key parameter, expansion velocity v, could indeed explain the differences in spatial structure between colonies. Specifically, to what extent does v explain our finding that the demixing rate in space decreased with nutrients (Figure 2c)? Could it be that the colonies all demixed at approximately the same time but that the demixing distance differed because poorly-fed colonies were expanding more slowly? To answer this question, for each colony, we normalized the curve describing the change in heterozygosity over space and its demixing distance (Figure 2c) by the colony expansion velocity v(see Materials and methods). This normalization revealed that colonies demixed significantly later in time with increasing nutrient concentration (Pearson's  $\rho = 0.78$ , P < 0.001, Figure 2d, inset) and that the overall drop in heterozygosity H over time (dH/dt) was slower with more nutrients ( $\rho = 0.71$ , P < 0.001, Supplementary Figure S9B). However, the majority of the differences between the heterozygosity curves at different nutrient concentrations disappeared with this transformation, which is consistent with expansion velocity v being the key determinant of demixing distance (compare Figures 2c and d).

How robust is the effect of expansion velocity v in the face of other perturbations? If the radial expansion rate is indeed the main factor explaining the rate of diversity loss in space, then we expect to find little effect of the initial colony radius  $R_0$  on the rate of change in heterozygosity H, as colonies of different starting sizes but identical nutrient concentrations should have identical radial expansion rates. Consistent with this, a generalized linear model regression analysis revealed that nutrient concentration was the only significant predictor of the rate of diversity loss dH/dx ( $\beta = 1.76$ , P < 0.001), which was not significantly affected by starting radius  $R_0$  ( $\beta = 0.0003$ , P = 0.99, Supplementary Figure S13C).

Our work then suggests that expansion velocity explains why nutrients change the rate of diversity loss in space (dH/dx) but have little effect on the rate of loss in time (dH/dt). Another way to capture the latter effect is via the genetic diffusion constant  $D_g$ : although both  $D_g$  and dH/dt decrease with increasing nutrients, they are generally quite large in magnitude, such that diversity loss is similar throughout all colonies. In sum, when more nutrients are available, the boundaries between sectors fluctuate more strongly (larger  $D_s$ ) and the two strains separate into clonal patches further from the inoculum because there are more cells dividing per unit time at the growing edge. And, although nutrients also have a significant negative effect on the rate of diversity loss in time, and on the strength of genetic drift, the magnitude of this effect is small as all colonies demix strongly owing to the presence of nutrient gradients.

# Final genetic diversity is similar across all nutrient conditions

With our analysis, we have shown how nutrient abundance affects the rate of diversity loss over space and time. How though does it affect final genetic diversity throughout the cell group?

There are at least two ways to assess final diversity using our experiments, which correspond to two different ecological scenarios. The first is a system that is limited in space, such that cells grow to a fixed population size and then stop (size-controlled). The second scenario is where space is unconstrained and the final population size is driven purely by the amount of time that the system is allowed to grow (time-controlled). To calculate the final diversity in a size-controlled system, we compare colonies of the same size  $R_{\rm s}$  across the nutrient treatments and calculate the heterozygosity at the pixel scale averaged over each colony (CH, see Materials and methods). This analysis follows the patterns we have observed above for the rate of diversity loss in space: for size-controlled colonies, low-nutrient colonies show lower final heterozygosity than high-nutrient colonies (Pearson's  $\rho = 0.96$ , P < 0.001, Supplementary Figure S15).

We next ask how final diversity changes for time-controlled cell groups. Because our colonies are free to grow without constraints, this corresponds simply to assessing average heterozygosity over each colony (CH) at the end of the experiments. Pooling together data from all experiments, we find no significant correlation between final diversity and nutrients (Spearman's  $\rho = 0.06$ , P = 0.56). This suggests that, similar to genetic drift and the rate of diversity loss, colony expansion velocity has little effect on final colony diversity.

In summary, when comparing colonies over the same spatial scale, final diversity correlates positively with nutrient abundance, whereas observing colonies over similar timescales shows no variation owing to nutrients. Both of these cases contrast with the hypothetical case where there is never any nutrient limitation. Without nutrient limitation, there is no spatial structure and genetic diversity remains constantly high over time (Figures 4c and d). This final contrast emphasizes that, while variation in the degree of nutrient limitation has clear quantitative effects on genetic diversity, spatial structuring is inevitable in our experiments.



# Discussion

The spatial structure of genotypes within microbial communities is considered central to their form and function. Strong spatial structuring is generally associated with increased genetic drift and weakened natural selection (Hallatschek and Nelson, 2008; Hallatschek and Nelson, 2010; Korolev et al., 2011). Nevertheless, the emergence of structure can result in strong natural selection for cooperative traits that benefit the cells of the same genotype in the surroundings and promote productivity and general resilience (Nadell et al., 2010; Mitri et al., 2011; Mitri and Foster, 2013; Kim et al., 2014). In addition, spatial structure can have important ecological impacts, which includes making communities more stable by reducing the strength of interactions between species (Coyte et al., 2015; Kim et al., 2008). Understanding what drives spatial structuring in microbial groups, therefore, is a fundamental challenge for microbiology.

Our work suggests that nutrient limitation is a key mechanism underlying the emergence of spatial structure in cell groups. Unless nutrients are saturating through a cell group, there will be a limited number of cells dividing at the growing edge, which promotes the stochastic loss of cell lineages. The number of cells dividing at the growing edge increases with nutrient levels and determines the expansion velocity of the colony. Importantly, increases in this velocity are strongly associated with the maintenance of diversity. The power of this model to explain diversity loss can be summarized by plotting the rate of diversity loss in space as a function of colony expansion velocity (Figure 7). However, our work also emphasizes that any gradation in nutrient levels has the potential to lead to diversity loss, as long as there is sufficient space for expansion. Although high-nutrient conditions allow

**Figure 7** The rate at which colonies lose diversity (the absolute value of the average spatial change in heterozygosity |dH/dx| measured over the area from the inoculum to the demixing distance on day 12) correlates with the colony expansion velocity v. Data shown are from 96 colonies on day 12 at different nutrient concentrations and inoculum sizes.

cells to maintain large well-mixed areas, because the groups are growing rapidly, they too will lose diversity on similar timescales as groups under much lower-nutrient levels. Because of these differences in growth rate, then, colonies will end up with similar levels of genetic diversity independently of nutrient abundance, assuming that they are not limited by space. Therefore, in the absence of other mechanisms that strongly disrupt spatial structure such as the production of surfactants (Xavier *et al.*, 2011) or surface motility (Figure 1b)—nutrient gradients always have potential to generate strong spatial separation.

Consistent with this conclusion, evidence suggests that the demixing of diverse populations into clonal groups is common in microbes, including *P. aerugi*nosa (Korolev et al., 2011), E. coli (Korolev et al., 2010; Korolev et al., 2011), Saccharomyces cerevisiae (Hallatschek and Nelson, 2010; Korolev et al., 2010; Müller et al., 2014), Dictyostelium discoideum (Buttery et al., 2012) and Bacillus subtilis (Ben-Jacob et al., 1994; Golding et al., 1999), and under different growth conditions in the laboratory (on agar plates (Shapiro, 1995; Golding et al., 1999; Hallatschek et al., 2007; Freese et al., 2014) and in flow cells (Nielsen et al., 2000; Pamp and Tolker-Nielsen, 2007; Momeni et al., 2013a)). An exception can occur when the different strains and species depend on each other to grow, such as in cross-feeding interactions, where strains are immune to the demixing processes we describe here (Momeni et al., 2013a; Müller et al., 2014). In this case, however, the positively interacting species can effectively be viewed as a single ecological unit that will, in turn, segregate with respect to other competing strains or mixtures of strains (Mitri et al., 2011; Momeni et al., 2013b). It is important to note that the large-colony and high-nutrient conditions often studied in the laboratory setting are not reflective of many natural conditions where densities and growth rates can be much lower. Our study does, however, consider relatively low-nutrient conditions, and we observe that spatial structure can emerge on very small spatial scales (tens of microns, Figure 1d).

Our observation of fine-scale structuring contrasts with the emphasis on sequencing studies that sample over large scales and detect extremely high species diversity (Gans et al., 2005; Roesch et al., 2007). However, there is a growing body of work that applies fluorescent in situ hybridization to natural communities, which also highlight a significant potential for fine-scale spatial structuring of genotypes (Stacy *et al.*, in press), such as in burn wounds (Fazli et al., 2009; Malic et al., 2009), in the gut microbiota (Dejea *et al.*, 2014; Millet *et al.*, 2014; Engel et al., 2015; Earle et al., 2015), in bladder infections (Kikuchi et al., 2009) or on leaf surfaces (Monier and Lindow, 2005). There is also variability in the degree of structure, however, and significant genotypic mixing is also seen in some contexts, such as human dental plaque (Palmer et al., 2003; Zijnge et al., 2010) or infections of the middle ear

(Weimer *et al.*, 2010). The observation of natural spatial structure does not, of course, demonstrate that this structure was driven purely by nutrient-limited growth. Spatial separation can result from other factors, including sparse seeding of a population or a patchy environment of microniches (Kikuchi *et al.*, 2009; Kubo *et al.*, 2011).

Although our experiments were conducted with microbes, we hypothesize that our findings will be common to any initially diverse population containing multiple heritable cell types with limited access to resources. Resource gradients are expected to be common in all cellular systems, occurring not only owing to nutrient limitations but also low-nutrient diffusivity, high cell growth rates or low yields of converting nutrients to biomass (Nadell *et al.*, 2010). The logic also applies to any diffusible substrate that is necessary for growth, such as signaling molecules.

Outside of microbial communities, our conclusions are likely to be particularly relevant to the organization and evolution of cancerous tumors. Viewed through the lens of evolutionary ecology, tumors are populations of cells that acquire mutations and diversify as they grow and expand in space, followed by migrations to different parts of the body (metastases; Korolev et al., 2014; Gundem et al., 2015). These parallels to other evolving populations may help predict tumor evolution, leading to novel therapies (González-García et al., 2002; Gatenby et al., 2013; Korolev et al., 2014). Indeed, preliminary data using computer simulations together with 'radiogenomics'-where the spatio-genetic structures resulting from these processes are revealed by linking radiology to the genetic characterization of cancer cells (Gatenby et al., 2013)—suggest that competition between neighboring cells is likely to lead to the spatial separation of genotypes, as seen in our colonies (González-García et al., 2002; Carmona-Fontaine et al., 2013; Gatenby et al., 2013). Similarly, cell competition over resources has been suggested to be important during multicellular development. Experiments with neutral markers to distinguish lineages of developing cells in the Drosophila melanogaster wing have shown spatial separation of cells reminiscent of that observed in microbial colonies, which is driven by gradients in signaling molecules rather than nutrient gradients (Johnston, 2009).

In sum, our analyses show that resource gradients can have a pivotal role in purging diversity and shaping the spatial structure of microbial groups. It is clear that these effects will act alongside a number of other processes that affect spatial structure, ecology and evolution. Nevertheless, our findings predict that the process of spatial structuring and the concomitant loss in diversity will be common due to the commonness of nutrient gradients (Stewart *et al.*, 2008). Although genomics often documents extremely high species diversities, the near-universality of resource limitation suggests that local diversity may be much lower than these estimates. A gradual reduction in diversity as populations grow may prove fundamental to the structure and function of microbial communities and other cellular groups.

# **Conflict of Interest**

The authors declare no conflict of interest.

## Acknowledgements

We are indebted to Marina Caldara, who helped with the construction of the  $\Delta$ pilB mutant; Mack Durham and Kirill Korolev, who have been extremely helpful regarding the fitting of our data to the theoretical model; Flora Sánchez who helped with imaging; and Kirill Korolev, Carey Nadell, William Smith, Clément Vulin and three anonymous reviewers for very constructive comments on the final manuscript. SM is supported by an Ambizione grant from the Swiss National Science Foundation, EC is supported by a Post-Doctoral Research Fellowship from All Souls College and KRF is supported by European Research Council Grant 242670. This research was funded by the European Research Council, the Swiss National Science Foundation and All Souls College, Oxford, UK.

## References

- Ben-Jacob E, Schochet O, Tenenbaum A, Cohen I, Czirók A, Vicsek T. (1994). Generic modelling of cooperative growth patterns in bacterial colonies. *Nature* **368**: 46–49.
- Burrows LL. (2012). *Pseudomonas aeruginosa* twitching motility: type iv pili in action. *Annu Rev Microbiol* **66**: 493–520.
- Buttery NJ, Jack CN, Adu-Oppong B, Snyder KT, Thompson CRL, Queller DC *et al.* (2012). Structured growth and genetic drift raise relatedness in the social amoeba *Dictyostelium discoideum*. *Biol Lett* **8**: 794–797.
- Carmona-Fontaine C, Bucci V, Akkari L, Deforet M, Joyce JA, Xavier JB. (2013). Emergence of spatial structure in the tumor microenvironment due to the warburg effect. *Proc Natl Acad Sci USA* **110**: 19402–19407.
- Coyte K, Schluter J, Foster KR. (2015). The ecology of the microbiome: networks, competition, and stability. *Science* **350**: 663–666.
- Dejea CM, Wick EC, Hechenbleikner EM, White JR, Mark Welch JL, Rossetti BJ *et al.* (2014). Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc Natl Acad Sci USA* **111**: 18321–18326.
- Dockery J, Klapper I. (2002). Finger formation in biofilm layers. *SIAM J Appl Math* **62**: 853.
- Earle KA, Billings G, Sigal M, Lichtman JS, Hansson GC, Elias JE. (2015). Quantitative imaging of gut microbiota spatial organization. *Cell Host Microbe* **18**: 478–488.
- Engel P, Bartlett KD, Moran NA. (2015). The bacterium *Frischella perrara* causes scab formation in the gut of its honeybee host. *mBio* **6**: e00193–15.
- Fazli M, Bjarnsholt T, Kirketerp-Møller K, Jørgensen B, Andersen AS, Krogfelt KA *et al.* (2009). Nonrandom

distribution of *Pseudomonas aeruginosa* and *Staphylococcus aureus* in chronic wounds. *J Clin Microbiol* **47**: 4084–4089.

- Freese PD, Korolev KS, Jiménez JI, Chen IA. (2014). Genetic drift suppresses bacterial conjugation in spatially structured populations. *Biophys J* 106: 944–954.
- Gans J, Wolinsky M, Dunbar J. (2005). Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science* **309**: 1387–1390.
- Gatenby RA, Grove O, Gillies RJ. (2013). Quantitative imaging in cancer evolution and ecology. *Radiology* **269**: 8–15.
- Golding I, Cohen I, Ben-Jacob E. (1999). Studies of sector formation in expanding bacterial colonies. *Europhys Lett (EPL)* 48: 587–593.
- González-García I, Solé RV, Costa J. (2002). Metapopulation dynamics and spatial heterogeneity in cancer. *Proc Natl Acad Sci USA* **99**: 13085–13089.
- Gourse RL, Gaal T, Bartlett MS, Appleman JA, Ross W. (1996). Rrna transcription and growth rate-dependent regulation of ribosome synthesis in *Escherichia coli*. *Annu Rev Microbiol* **50**: 645–677.
- Gundem G, Van Loo P, Kremeyer B, Alexandrov LB, Tubio JM, Papaemmanuil E *et al.* (2015). The evolutionary history of lethal metastatic prostate cancer. *Nature* **520**: 353–357.
- Habets MGJL, Rozen DE, Hoekstra RF, de Visser JA. (2006). The effect of population structure on the adaptive radiation of microbial populations evolving in spatially structured environments. *Ecol Lett* **9**: 1041–1048.
- Hallatschek O, Hersen P, Ramanathan S, Nelson DR. (2007). Genetic drift at expanding frontiers promotes gene segregation. *PNAS* **104**: 19926–19930.
- Hallatschek Ö, Nelson DR. (2008). Gene surfing in expanding populations. *Theor Popul Biol* **73**: 158–170.
- Hallatschek O, Nelson DR. (2010). Life at the front of an expanding population. *Evolution* **64**: 193–206.
- Heurlier K, Dénervaud V, Haenni M, Guy L, Krishnapillai V, Haas D. (2005). Quorum-sensing-negative (lasR) mutants of Pseudomonas aeruginosa avoid cell lysis and death. *J Bacteriol* 187: 4875–4883.
- Johnson CR, Boerlijst MC. (2002). Selection at the level of the community: the importance of spatial structure. *Trends Ecol Evol* **17**: 83–90.
- Johnston LA. (2009). Competitive interactions between cells: death, growth, and geography. *Science* **324**: 1679–1682.
- Kerr B, Riley MA, Feldman MW, Bohannan BJM. (2002). Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* 418: 171–174.
- Kikuchi Y, Bomar L, Graf J. (2009). Stratefied bacterial community in the bladder of the medicinal leech. *Environ Microbiol* **71**: 5484–5493.
- Kim HJ, Boedicker JQ, Choi JW, Ismagilov RF. (2008). Defined spatial structure stabilizes a synthetic multispecies bacterial community. PNAS 105: 18188–18193.
- Kim W, Racimo F, Schluter J, Levy SB, Foster KR. (2014). Importance of positioning for microbial evolution. *PNAS* 111: E1639–E1647.
- Koch G, Yepes A, Förstner Ku, Wermser C, Stengel ST, Modamio J et al. (2014). Evolution of resistance to a last-resort antibiotic in *Staphylococcus aureus* via bacterial competition. *Cell* **158**: 1060–1071.

- Korolev KS, Avlund M, Hallatschek O, Nelson DR. (2010). Genetic demixing and evolution in linear stepping stone models. *Rev Mod Phys* **82**: 1691–1718.
- Korolev KS, Müller MJI, Karahan N, Murray AW, Hallatschek O, Nelson DR. (2012). Selective sweeps in growing microbial colonies. *Phys Biol* **9**: 026008.
- Korolev KS, Xavier JB, Gore J. (2014). Turning ecology and evolution against cancer. *Nat Rev Cancer* **14**: 371–380.
- Korolev KS, Xavier JB, Nelson DR, Foster KR. (2011). A quantitative test of population genetics using spatiogenetic patterns in bacterial colonies. Am Nat 178: 538–552.
- Kovesi PD. (2000). Matlab and Octave Functions for Computer Vision and Image Processing. Centre for Exploration Targeting, School of Earth and Envrionment, The University of Western Australia: Western Australia, Australia.
- Kreft J-U. (2004). Biofilms promote altruism. *Microbiology* **150**: 2751–2760.
- Kubo K, Knittel K, Amann R, Fukui M, Matsuura K. (2011). Sulfur-metabolizing bacterial populations in microbial mats of the Nakabusa hot spring, Japan. *Syst Appl Microbiol* **34**: 293–302.
- Malic S, Hill KE, Hayes A, Percival SL, Thomas DW, Williams DW. (2009). Detection and identification of specific bacteria in wound biofilms using peptide nucleic acid fluorescent in situ hybridization (PNA FISH). *Microbiol* **155**: 2603–2611.
- Mattick JS. (2002). Type iv pili and twitching motility. Annu Rev Microbiol 56: 289–314.
- Millet YA, Alvarez D, Ringgaard S, Von Andrian UH, Davis BM, Waldor MK. (2014). Insights into *Vibrio cholerae* intestinal colonization from monitoring fluorescently labeled bacteria. *PLoS Pathogens* **10**: e1004405.
- Mitri S, Foster KR. (2013). The genotypic view of social interactions in microbial communities. *Annu Rev Genet* **47**: 247–273.
- Mitri S, Xavier J, Foster K. (2011). Social evolution in multispecies biofilms. *PNAS* **108**: 10839–10846.
- Momeni B, Brileya KA, Fields MW, Shou W. (2013a). Strong inter-population cooperation leads to partner intermixing in microbial communities. *eLife* **2**: e00230.
- Momeni B, Waite AJ, Shou W. (2013b). Spatial selforganization favors heterotypic cooperation over cheating. *eLife* **2013**: 1–18.
- Monier J-M, Lindow SE. (2005). Spatial organization of dual-species bacterial aggregates on leaf surfaces. *Appl Environ Microbiol* **71**: 5484–5493.
- Müller MJI, Neugeboren BI, Nelson DR, Murray AW. (2014). Genetic drift opposes mutualism during spatial population expansion. *Proc Natl Acad Sci USA* **111**: 1037–1042.
- Murray JL, Connell JL, Stacy A, Turner KH, Whiteley M. (2014). Mechanisms of synergy in polymicrobial infections. *J Microbiol* **52**: 188–199.
- Nadell CD, Foster KR, Xavier JB. (2010). Emergence of spatial structure in cell groups and the evolution of cooperation. *PLoS Comput Biol* **6**: e1000716.
- Nadell CD, Xavier JB, Foster KR. (2009). The sociobiology of biofilms. *FEMS Microbiol Rev* **33**: 206–224.
- Nei M, Maruyama T, Chakraborty R. (1975). The bottleneck effect and genetic variability in populations. *Evolution* **29**: 1–10.
- Nielsen AT, Tolker-Nielsen T, Barken KB, Molin S. (2000). Role of commensal relationships on the spatial structure of a surface-attached microbial consortium. *Environ Microbiol* **2**: 59–68.

- Palmer RJ, Gordon SM, Cisar JO, Kolenbrander PE. (2003). Coaggregation-mediated interactions of streptococci and actinomyces detected in initial human dental plaque. *J Bacteriol* **185**: 3400–3409.
- Pamp SJ, Tolker-Nielsen T. (2007). Multiple roles of biosurfactants in structural biofilm development by *Pseudomonas aeruginosa*. J Bacteriol 189: 2531–2539.
- Pirt S. (1967). A kinetic study of the mode of growth of surface colonies of bacteria and fungi. J Gen Microbiol 47: 181–197.
- Roesch LFW, Fulthorpe RR, Riva A, Casella G, Hadwin AKM, Kent AD *et al.* (2007). Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J* **1**: 283–290.
- Shapiro JA. (1995). The significances of bacterial colony patterns. *BioEssays* 17: 597–607.
- Stacy A, Mcnally S, Darch SE, Brown SP, Whiteley M. (in press). The biogeography of infection. Nat Rev Microbiol.
- Stewart PS, Franklin MJ. (2008). Physiological heterogeneity in biofilms. *Nat Rev Microbiol* **6**: 199–210.
- Van Gestel J, Weissing FJ, Kuipers OP, Kovács AT. (2014). Density of founder cells affects spatial pattern formation and cooperation in bacillus subtilis biofilms. *ISMEJ* 8: 2069–2079.
- Weimer KED, Armbruster CE, Juneau RA, Hong W, Pang B, Swords WE. (2010). Coinfection with *Haemophilus*

*influenzae* promotes pneumococcal biofilm formation during experimental otitis media and impedes the progression of pneumococcal disease. *J Infect Dis* **202**: 1068–1075.

- Xavier JB, Foster KR. (2007). Cooperation and conflict in microbial biofilms. *PNAS* **104**: 876–881.
- Xavier JB, Kim W, Foster KR. (2011). A molecular mechanism that stabilizes cooperative secretions in Pseudomonas aeruginosa. *Mol Microbiol* **79**: 116–179.
- Zijnge V, Van Leeuwen MBM, Degener JE, Abbas F, Thurnheer T, Gmür R *et al.* (2010). Oral biofilm architecture on natural teeth. *PLoS One* **5**: e9321.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/ licenses/by-nc-sa/4.0/

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)