OPINION

Explaining microbial population genomics through phage predation

Francisco Rodriguez-Valera, Ana-Belen Martin-Cuadrado, Beltran Rodriguez-Brito, Lejla Pašić, T. Frede Thingstad, Forest Rohwer and Alex Mira

Abstract | The remarkable differences that have been detected by metagenomics in the genomes of strains of the same bacterial species are difficult to reconcile with the widely accepted paradigm that periodic selection within bacterial populations will regularly purge genomic diversity by clonal replacement. We have found that many of the genes that differ between strains affect regions that are potential phage recognition targets. We therefore propose the constant-diversity dynamics model, in which the diversity of prokaryotic populations is preserved by phage predation. We provide supporting evidence for this model from metagenomics, mathematical analysis and computer simulations. Periodic selection and phage predation dynamics are not mutually exclusive; we compare their predictions to shed light on the ecological circumstances under which each type of dynamics could predominate.

Explaining patterns of diversity in microbial populations has been one of the great conundrums of microbiology. Historically, the approaches used for studying bacterial population genetics have been derived from eukaryotic models. They were based on distinguishing alleles and quantifying their presence in populations, as well as their degree of linkage, to infer mutation and recombination rates (for example, see REF. 1). However, with the advent of genomics, it has become apparent that the genomic diversity in prokaryotes arises more from having different sets of genes than from allelic differences at the same loci[2–4]. This is in contrast to genomic diversity in eukaryotic organisms, which preserve the gene content within the same species remarkably well, even across large phylogenetic distances[5]. The concept of the pan-genome has been developed to describe the increasing diversity of the gene pool that can be ascribed to one bacterial species as the number of sequenced strains increases[6]. A comparison of two strains, which may belong to highly related lineages within a single species, might reveal that approximately 10–35% of the genome content (typically in the range of 500–1,000 genes) is present in only one of the two strains[7]. The implications of these findings are still permeating the scientific community, but their importance for our way of thinking about prokaryotic microdiversity and evolution is paramount. Is this diversity important for the ecology and environmental adaptation of the different lineages, or is it just the result of junk DNA accumulation that has yet to be pruned by regular sweeps of natural selection? What are the evolutionary forces that preserve this degree of diversity within highly related populations?

The accepted models of bacterial population genetics state that a low phenotypic diversity is expected in asexual (or rarely sexual) microbial populations owing to purges involving fitter mutants, called periodic selection events[8]. This idea originated from classical laboratory experiments of mutational equilibrium in which populations were periodically replaced by new types[9,10]; it was later reinforced by epidemiological studies of pathogenic isolates, in which a rise and fall of clonal lineages was found to occur[11,12]. This process would purge diversity from the population, at least among the cells competing for the same resources. It has been claimed that the same kinds of dynamics are behind the genetic coherence of natural prokaryotic taxonomic units or 'ecotypes'[13,14]. Ecotypes are defined as "populations that are genetically cohesive and ecologically distinct"[15]. Cohesion is mostly ascribed to "periodic selection events that recurrently purge each ecotype of its genetic diversity"[15]. Divergence can become permanent when a mutation (or recombination) event places an organism into a new ecological niche and establishes a new ecotype[13]. Periodic selection will therefore keep the populations in each ecotype homogeneous and divergent from those in other ecotypes[1–5], making these population "the fundamental units of ecology and evolution"[16], in other words, a stable ecotype.

At present, the existence of natural diversity units (that is, discontinuities in a similarity gradient) in bacteria is widely accepted[14,16]. However, the origination of such units by regular periodic selection events seems difficult to reconcile with the wide gene content variability that is found among different strains with otherwise high sequence conservation among shared genes (that is, the pan-genome). We know that prokaryotes can easily acquire foreign DNA, and phages and plasmids can be easily transferred between bacteria and the DNA inserted into the chromosome of the new host. In fact, some (or most) periodic selection events could be due to the acquisition of new genes rather than mutation. However, it seems unlikely that the neutral accumulation of a large number of new genes could occur in a clone before it is purged during a clonal sweep, and such a phenomenon might have alternative explanations.

We propose that the main factor that generates diversity, which has been largely overlooked in many previous models, is the role of a crucial ecological factor: the presence of bacteriophages. In nature, prokaryotic cells must deal with a strong predation pressure that is mainly viral in origin, and therefore their fitness is measured not only by their adaptation to the available resources and the physical conditions (the niche) but also by their adaptation to the biotic environment[17]. Protozoan grazing might also contribute[18] but is less pervasive than viral attack[19].

In almost all ecosystems that have been investigated there are around ten phages for every microbial cell, making phages the most abundant biological entities on the planet[20]. By killing microorganisms, phages greatly influence global biogeochemical cycles, and because phages are species specific they have been predicted to help maintain microbial species diversity[21,22]. In fact, phages could have a fundamental role as guarantors of the microdiversity that is required to exploit ecological resources efficiently. We propose a new type of dynamics to explain bacterial microdiversity that incorporates phage predation. We begin by systematically describing the content of variable genomic regions among closely related bacterial strains, with the aim of identifying the force that drives microdiversity. We then introduce the constant-diversity (CD) dynamics model to explain the generation and maintenance of microbial diversity in natural ecosystems. We compare its predictions with those of the periodic-selection (PS) model and evaluate the evidence using mathematical modelling, metagenomic data and microbial ecology

studies. Finally, we reconcile the CD and PS dynamics models on the basis of the ecological features of different environments that would favour either one.

Single-species metagenomics

Metagenomic approaches have demonstrated the value of comparing the genomes of individual strains with the metagenomes from the environments in which these species are present. Such comparisons have shown that certain genomic regions are underrepresented within any given metagenome and are therefore predicted to be unique to individual isolates[23–29]. The metagenome represents all lineages within one sample and, if it has enough coverage, it will contain several representatives from the predominant populations. This virtual experiment has been carried out a number of times[24–30] and has shown that several genomic regions are underrepresented in all of the genomes of one species. These regions, which we refer to as metagenomic islands (MGIs), are predicted to be unique to individual cell types. It is important to note that these MGIs must not be confused with classical genomic islands — that is, regions of unusual DNA composition that are generally obtained by the bacterium through lateral gene transfer and that typically have the hallmarks of mobile DNA elements or recombinatorial hot spots. Nevertheless, some MGIs are obviously of a foreign nature and show extraneous compositional features.

An ideal system for this kind of comparison is an extreme environment that is heavily dominated by only a few species and, thus, has low genetic diversity — for example, a saturated brine that is largely dominated by the halophilic archaeon *Haloquadratum walsbyi*[24]. When the metagenome of the brine was compared with an already available genome of the strain isolated from the same pond, many genes were extremely divergent, highly rearranged or not found in the metagenome[27]. They included genes that encode cell surface components, such as glycoproteins, and factors involved in glycosylation of surface components[27] (FIG. 1). Many of the variable genes are substrate transporters or sensors of two-component regulatory systems. These two features were interpreted to indicate differential adaptation to phage sensitivity and organic carbon degradation. A similar pattern was found when a nearly monospecific metagenome that was dominated by *Candidatus* Accumulibacter phosphatis was sequenced in sludges from enhanced biological phosphorus removal reactors[30]. One could argue that these are special cases, because a single species has to perform many metabolic roles that would be shared among many different species in high-diversity environments. However, a strikingly similar pattern of inferred functions is found within metagenomic islands from high-diversity environments such as the open ocean (FIG. 1).

We have systematically compared MGIs larger than 10 kb in all sequenced marine bacterial species to the marine metagenome and found that specific functional categories were highly over-represented in MGIs (Supplementary information S1,S2 (figure, table)). Remarkably, all MGIs in the species that we studied contain genes encoding products that are exposed extracellularly (FIG. 1). Paramount among them was the variable O chain of the lipopolysaccharide (LPS), which has long been known to be highly variable and a target for phage receptors (for example, see REF. 31). Serotyping and phage typing are usually linked to changes in the O antigen that are determined by this gene cluster[32]. In pathogens, this variability has often been explained as a strategy to evade host immunity, but in free-living microbes other explanations are required.

Next most frequent were exopolysaccharide biosynthesis clusters and genes involved in sugar modifications of extracellular structures. Pili and flagellar components (particularly their extracellular components) were also found. Recently, an island was found that contains an alternative set of external flagellar proteins[33]. MGIs also frequently encode giant proteins that are probably extracellular[34]. All of those genes are potential phage recognition sites, suggesting a role in phage avoidance. When the functional classifications of genes encoded in MGIs are compared with the classifications found in the genome, the genes encoding potential phage recognition sites are heavily over-represented, as are genes involved in nutrient transport and environmental sensing (Supplementary information S3 (figure)). Thus, we conclude that, overall, the phage-interacting genes are often strain specific, and when they are shared between strains they tend to be more highly divergent.

Phages depend heavily on proper selection of the target cell, and for that they rely on a prominent structure for target recognition[35,36]; preventing this recognition is therefore the first line of defence for bacteria. It is important to note that the high divergence in potential phage recognition sites is found even in extremely compact genomes, such as that of *Candidatus* Pelagibacter ubique (FIG. 1), which has the smallest sequenced genome among marine prokaryotes. Despite its compact genome, the MGIs in the three available strains of *Candidatus* Pelagibacter ubique contain surface features (FIG. 1a) and transport and sensing genes. Thus, the presence of potential phage targets in the strain-specific areas of the genome is a pervasive phenomenon in the open ocean, suggesting that this feature may apply to other free-living prokaryotic microbes that are subject to phage predation pressure. The second line of phage defence is intracellular. Our MGI data show numerous examples of these types of signature, such as restriction–modification systems and the clustered, regularly interspaced short palindromic repeats (CRISPR) (Supplementary information S2 (table)) that are involved in phage interference[37–39]. However, they are probably involved in infection efficiency, providing a fine-tuning that would reinforce the primary control that is exerted by the receptor diversity.

CD dynamics

The CD dynamics model aims to explain microbial diversity in ecosystems in which bacterial populations can interact with each other; that is, the populations must compete with each other, and phage particles must have a similar chance of infecting any cell within the community. Let us consider an idealized aquatic habitat

Glossary

CRISPR
A widespread genetic system in bacteria and archaea that consists of multiple copies of palindromic repeats flanking short spacers of phage origin, which are believed to provide acquired resistance against viral infection.

Kill-the-winner dynamics
A model for the population dynamics of phage–bacteria interactions that postulates that an increase in a host population (the winner) is followed by an increase in its corresponding phage predator, resulting in an increase in the rate at which the winner is killed. It is analogous to classical Lotka–Volterra dynamics to explain predator–prey population dynamics.

Metagenome
The total genetic repertoire that exists in cells and viruses from a given environment. Metagenomes are often classified according to the predominant group of microorganisms that contribute to a sequence library, for example, viruses or bacteria.

Metagenomic island
A genomic region found in a bacterial genome that is absent or present at low frequency in the DNA pool of all microbes in the native environment of that bacterium.

Pan-genome
The total gene pool of a bacterial or archaeal taxon, a pan-genome is formed by the addition of all genes found in the different strains from a given species (species pan-genome) or from an ecologically distinct population (ecotype pan-genome).

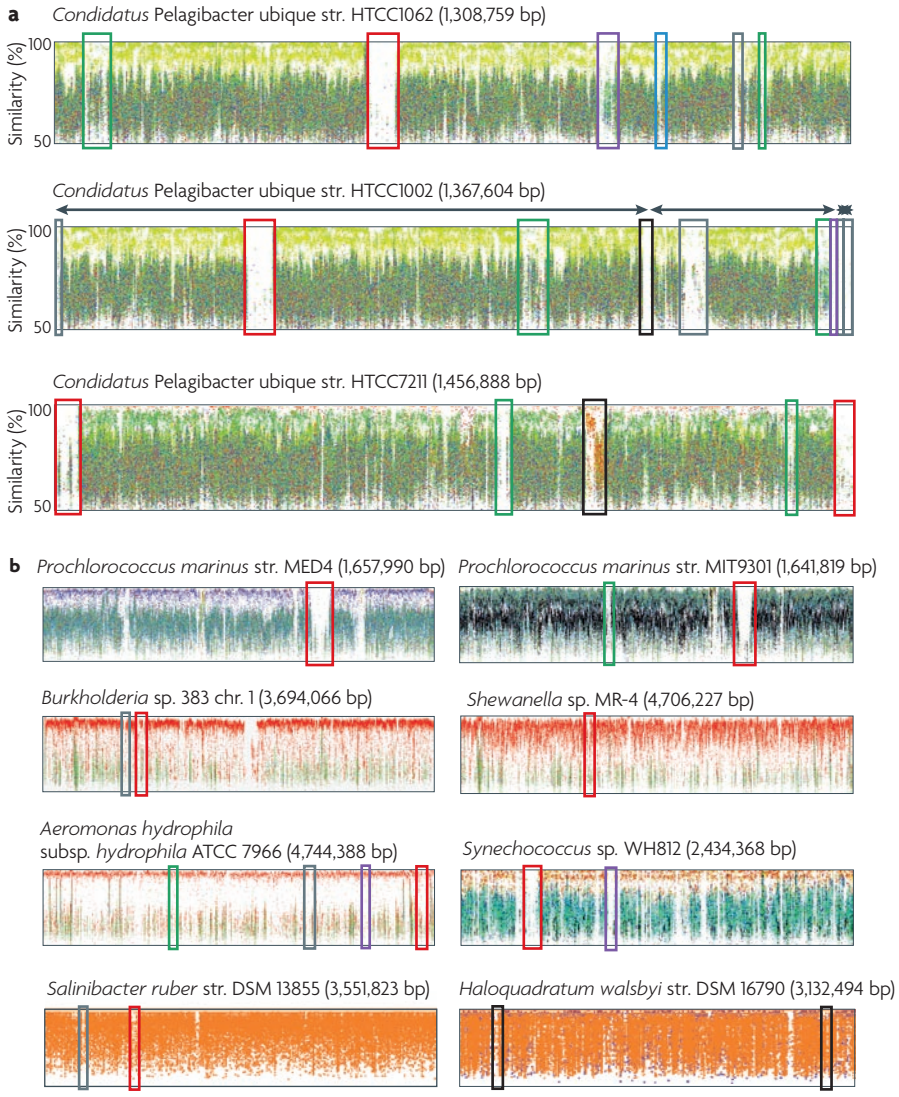


Figure 1 | **Metagenomic islands in selected bacteria identified by comparison with available metagenomes.** Representative species that showed >90% DNA sequence identity over 80% of the genome were selected to ensure that they are well represented in the environment. **a** | Comparison of the genomes of the marine alphaproteobacterium *Candidatus* Pelagibacter ubique strains HTCC1062, HTCC1002 and HTCC7211 with the Global Ocean Survey database (GOS; Phase I[64] and Phase II). Individual sequences were aligned to the genome of the sequenced strain, and the sequence conservation is visualized in the form of a percent identity plot. Regions larger than 10 kb with unusually low representation in the metagenome (metagenomic islands (MGIs)) are marked with a box and a brief description of their main features. Green boxes indicate pili, red boxes indicate O chain LPS, purple boxes indicate giant proteins, blue boxes indicate exopolysaccharide-related proteins, grey boxes indicate transmembrane and outer-membrane proteins and black boxes indicate proteins with other functions. **b** | Comparison of the genomes of *Prochlorococcus marinus* str. MED4, *P. marinus* str. MIT9301, *Burkholderia* sp. 383 chr. 1, *Shewanella* sp. MR-4, *Aeromonas hydrophila* subsp. *hydrophila* ATCC 7966 and *Synechococcus* sp. WH812 with the GOS dabatase and of *Salinibacter ruber* str. DSM 13855 and *Haloquadratum walsbyi* str. DSM 16790 with solar saltern metagenomes[24,65]. MGIs that are related to extracellular components are marked following the colour pattern indicated in part **a**. Data were collected from http://camera.calit2.net/, with the exceptions of those for *S. ruber* DSM 13855 and *H. walsbyi* DSM 16790. A full description of the MGIs' contents from all species with fragments of high similarity over 80% of their reference genomes can be found in Supplementary information S2 (table). LPS, lipopolysaccharide.

where organic nutrients are dissolved and in which a single prokaryotic population is present. A large diversity of phage sensitivity types is required to avoid catastrophic lysis of the population. Phage receptor types are each recognized by a different phage lineage. The phage receptor types differ in the gene clusters that control the synthesis of complex surface components such as the O chain or the pili, which are exposed extracellularly and are targets for phage recognition, analogous to a lock-and-key system. Periodic selection occurs when a new adaptive mutant (or recombinant) arises in the population and natural selection causes the mutant and its nearly clonal descendants to replace all competing variants in the population[16]. However, we predict that the increase in number of that fitter lineage would alter the predator–prey equilibrium, and the number of phages that target the receptor encoded by this lineage would also increase (FIG. 2a). This would select against the invasive clone, which would eventually be replaced by the original 'normal fitness' lineages. In this way, a constant high level of diversity of lineages would be maintained steadily.

Interestingly, the role of phage predation modifies the classical 'survival of the fittest' axiom such that metabolically superior microorganisms that are better adapted to a physical environment are selected against by the biological pressure that is imposed by density-dependent phage predation. These kill-the-winner dynamics have already been proposed for different species of marine bacteria[21,22], but here we propose that this process is responsible for maintaining diversity among closely related lineages. Under CD dynamics, no dominant lineage is found in a population, and phage dynamics maintain many concurrent cell types that are selected against ecological success. This success will be influenced by factors such as growth rate or the efficient use of different substrates, and therefore a corollary of this model is that, throughout the history of these clonal lineages, each lineage will acquire different but complementary capabilities for niche exploitation. As a consequence, a more efficient exploitation of the resources by the community is expected, and a better functioning of the ecosystem will be achieved. For example, one prokaryotic clone cannot contain even a fraction of the transporters required to internalize the chemical diversity of the organic compounds that are contained in a single eukaryotic cell. However, an ensemble of lineages carrying different sets of transporters could exploit every single one of them. A relationship between biodiversity and ecosystem efficiency has been demonstrated in plant ecosystems[40], and we predict that similar principles would apply for the biogeochemical cycles that are

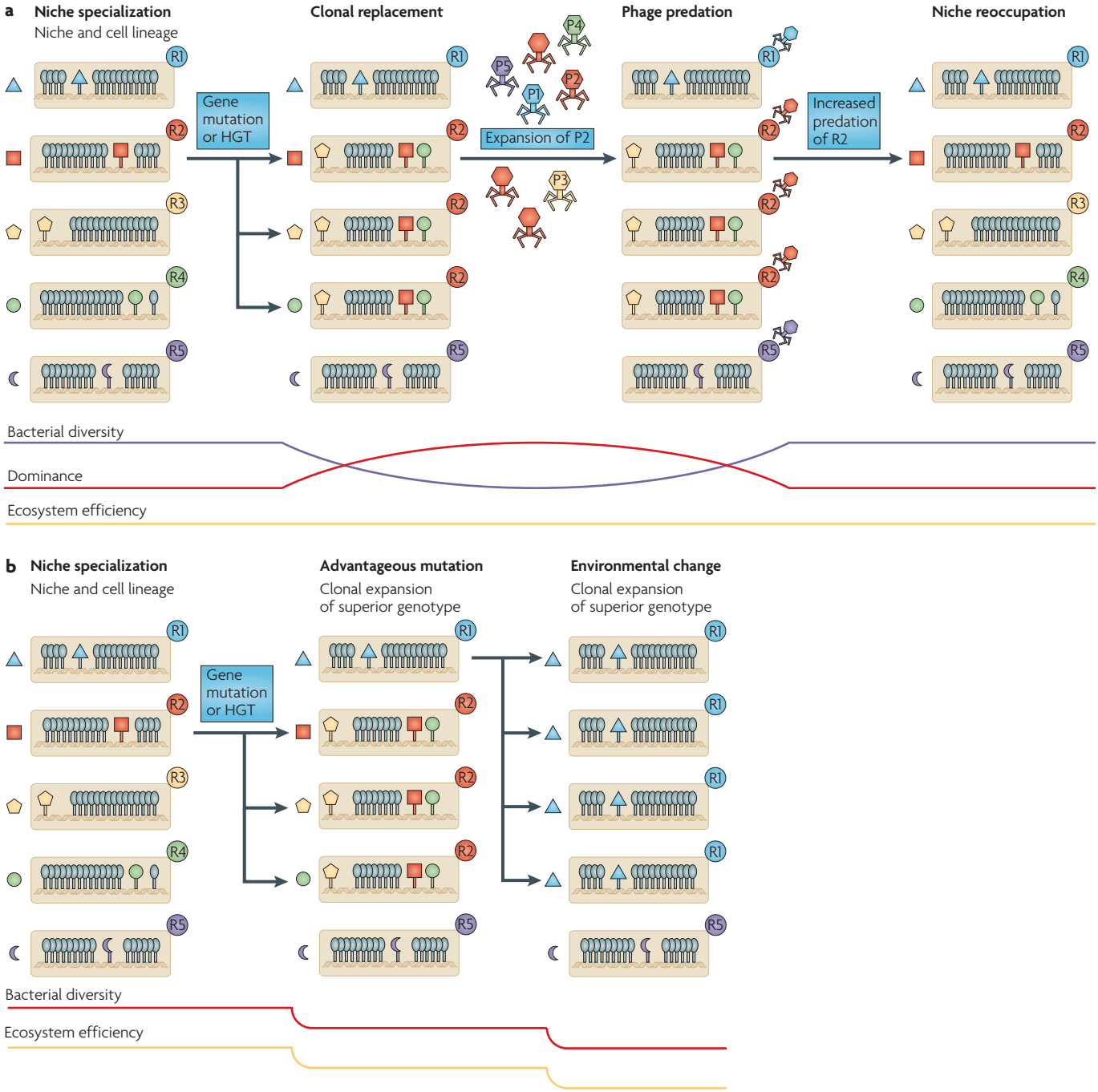


Figure 2 | **Population dynamics under constant-diversity and periodic-selection dynamics.** Several sympatric clones with diverse abilities to exploit environmental resources coexist within a population in an aquatic habitat. The different resources are indicated by geometric symbols, each representing a different niche. The adaptations required for the exploitation of the niche are represented by the same geometric figures on the genome of the lineage. These lineages also differ in the set of receptors (R1–R5) to phages (P1–P5) that they carry. Each strain contains only one major phage receptor, which is determined by genes located in a metagenomic island. One lineage can gain an increase in fitness — through mutations in the genome, through the introduction of new DNA by horizontal gene transfer (HGT) or through an environmental change that favours one particular genotype — and replace the other lineages. **a** | Under constant-diversity dynamics, this situation can occur only transiently, because an increase in the frequency of cells carrying the specific phage receptor that hitchhikes with the gene or genes that increase niche range would unsettle the predator–prey equilibrium and select against the invasive clone; thus, that clone would eventually be replaced by the original 'normal fitness' lineages. Kill-the-winner dynamics therefore give rise to a constant high diversity of lineages that would be maintained steadily. It is predicted that this equilibrium allows a more efficient exploitation of the resources by the community, leading to a more efficient ecosystem. **b** | The periodic-selection model predicts that lineages with higher fitness would expand and replace other cell types, giving rise to a clonal sweep. The process would be repeated after advantageous mutations or when environmental changes increase the fitness of a given strain, decreasing diversity. Ecosystem efficiency is also expected to be reduced, as the lineages exploiting substrates of low availability would severely decline.

controlled by microbial communities, in such a way that ecosystem functioning would be more efficient and stable under high-diversity situations than under periodic clonal sweeps, as the latter would cause fluctuations in nutrient and mineral recycling (FIG. 2b).

Computer simulations of bacterial strain replacement in the presence and absence of phages show that viral predation affects biodiversity (FIG. 3). In the absence of phages, microbial density for each lineage is dependent on the availability of the substrates and the efficiency of nutrient use. Thus, the lineage that uses the most abundant substrate at a given time is the most abundant (FIG. 3c). If an environmental change (for example, a variation in the concentration of nutrients) or a mutation (for example, one that changes the efficiency of nutrient use) occurs, the fitter lineage will be favoured, generating a clonal sweep. As a consequence, the diversity of the population will be low (FIG. 3b), as the better-adapted strain dominates over other strains that have a smaller representation in the consortium. In the presence of phages, however, a more successful lineage that uses the most abundant nutrient or that can metabolize more than one substrate will be preferentially attacked by phages, because phage–host interactions are density dependent[21]. In this way, the fitter bacterial strains are selected against, so the density of the different bacterial lineages fluctuates around stable levels (FIG. 3d). Therefore, bacterial diversity remains high, with all lineages present at similar levels regardless of the availability of the substrates that they use (FIG. 3c). The immediate consequence of this fine-tuning to different substrates would be the expansion of the gene pool in the population. Thus, contrary to what would be expected by clonal replacements, CD dynamics would predict a large pan-genome within populations (BOX 1).

Another interesting feature revealed by these simplified simulations relates to ecosystem functioning. In the presence of phage predation pressure and environmental changes, the lineage that feeds on the most abundant substrate (or that uses more than one substrate efficiently) undergoes substantial fluctuations in density. (FIG. 3d) By contrast, lineages using the least-abundant substrates show constant densities through time, undergoing only slight variations. Because the amount of density fluctuation is directly related to the probability of extinction, selection will favour less fit cells that use single, low-concentration substrates. This paradoxical selection against the fitter cells will give rise to lineages that feed on all accessible substrates, regardless of their availability. As a consequence, the ecosystem is expected to be more efficient. This high efficiency in substrate use at the ecosystem level will be sustained by the use of all available metabolites and minerals and by the presence of a high diversity, which is known to promote efficient ecosystem functioning[40,41]. Note that a high bacterial diversity cannot be the consequence of the availability of different resources alone: it is only in the presence of phages that all cell types reach a similar average density regardless of the substrate that they use; in the absence of phage predation, the exploitation of less favoured niches is selected against.

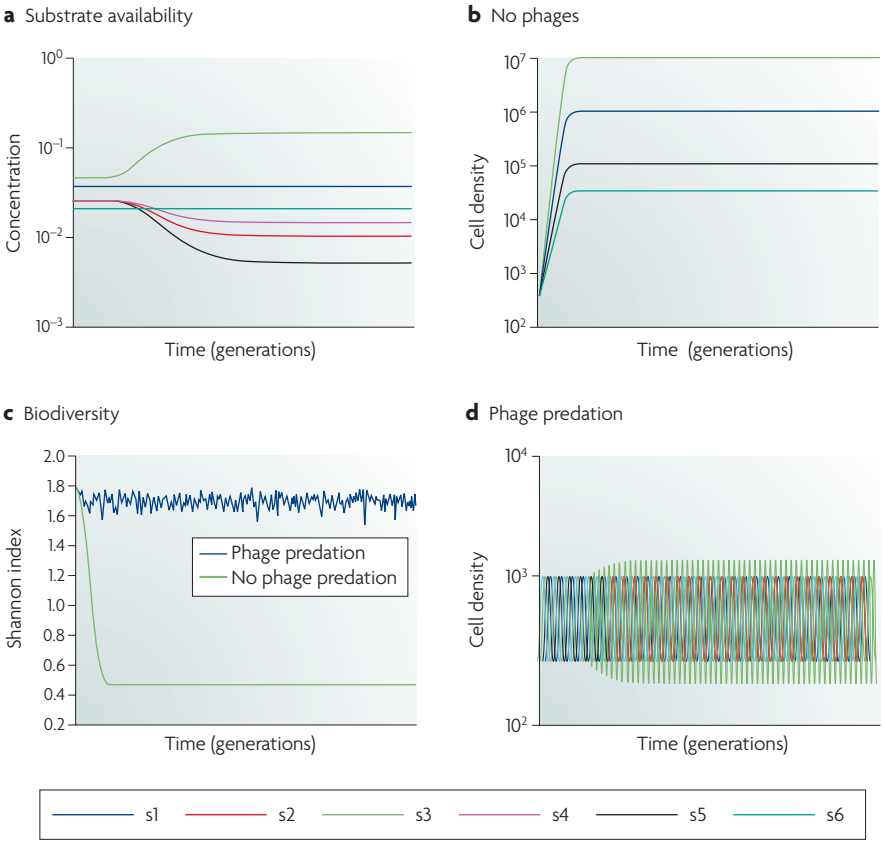


Figure 3 | **Cell density and microbial diversity under constant-diversity and periodic-selection dynamics.** The graphs show the outcome of computer simulations from a simplified model of an ecosystem that is inhabited by six heterotrophic prokaryotes (s1–s6). **a** | Each species optimally uses one specific substrate, the availability of which varies over time. **b** | In the absence of phages, microbial density is directly related to the availability of the substrates and the efficiency of nutrient use. **c** | After an environmental change (for example, a variation in nutrient concentration) or a mutation (for example, a gene transfer that improves nutrient utilization efficiency), the fitter lineage is favoured, generating a clonal sweep. As a consequence, the diversity of the system, as measured by the Shannon index, drops. **d** | In the presence of bacteriophages, the lineage of higher metabolic fitness is preferentially attacked by viruses, assuming density-dependent phage–host interactions. This type of kill-the-winner dynamics selects against the fitter bacterial strains, so the density of the different bacterial lineages fluctuates slightly (note the differences in scale) around stable levels, and bacterial diversity remains high, with all lineages present at similar values regardless of substrate availability. The model is a coupled set of differential equations and is composed of six pairs of equations, which are repeated for each of the six substrates. There is no direct interaction between different strains of viruses, and the interaction between strains of microbes comes in the form of a common carrying capacity for the whole system. We have considered a simplified case in which one phage infects one cell type; however, it must be noted that several experiments indicate that host–virus interactions in natural systems are more complex, with each phage infecting different bacterial strains with different efficiencies[50,66] (in other words, the binding probability between bacteriophages and cells is graded[67]). For the sake of simplicity, the decay rate was assumed to be the same for all of the viruses. The same was true for the mass-action constant and the burst size. This model was programmed in MATLAB (The MathWorks) and the code is available upon request.

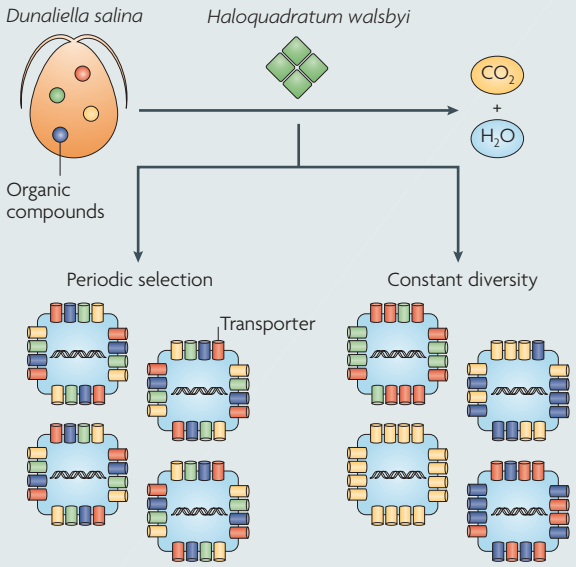
Predictions from CD dynamics

Despite several refinements, all of the ecotype-based models that have been proposed rely strongly on the importance of PS arising from competition for the resources among individual lineages (for example, see REF. 15). By contrast, CD predicts that PS of an individual lineage with larger capabilities to exploit resources would be prevented by phages, and a better exploitation of resources is achieved when many different lineages share the environment. These views make different predictions about microbial biodiversity and population dynamics (TABLE 1). Under PS, a dominant cell type would be expected to arise, and this dominant lineage would change periodically. Conversely, under CD dynamics, phages would keep dominant lineages in check, and therefore many coexisting cell types would be found at any one time. In addition, PS predicts that genomic variability among ecotypes is driven by niche exploitation, whereas CD predicts that it is driven by phage avoidance (TABLE 1).

What is the evidence for each of these models? The coexistence of multiple closely related strains is a commonly observed phenomenon in microbial communities and is apparent from studies of cultured isolates[42], marker gene surveys[43] and metagenomic data[27,44] (the ecological diversity of close relatives is discussed in REFS 15,45). In addition, mathematical modelling of bacterial dynamics in the presence of phages indicates that a dramatically high number of strains is expected, as indicated by metagenomic data (BOX 2). Although many strains have indeed been observed in free-living ecosystems, the maintenance of this high diversity through time remains to be explained.

The initial evidence for antagonistic dynamics between phages and bacteria came from mathematical modelling, as well as experimental studies[21,22,46,47]. Recently, work on marine samples has shown that phages influence their bacterial hosts in a density-dependent manner, mainly infecting a reduced number of phylotypes at any one time[48], as predicted by kill-the-winner dynamics. Similar results have been found in the horse gut, in a study of the relationship between coliphages and *Escherichia coli* strains[49]. In addition, detailed studies on marine flavobacteria have shown the degree of phage susceptibility in different cell types in the same environment, drawing a complex picture of phage–cell interactions[50]. Recent data on the temporal variation in phage and bacterial diversity also show oscillations that are consistent with a constant control of abundant genotypes by their infecting phages (B.R.-B., L.L. Li, L. Wegley, M. Furlan, F. Angly, M. Breitbart, J. Buchanan, C. Desnues, L. Dinsdale, R. Edwards, B. Felts, D. Willner, M. Haynes, H. Liu, D. Lipson, J. Mahaffy, A.-B.M.-C., A. Mira-Obrador, J. Nulton, L.P., S. Rayhawk, J. Rodriguez-Mueller, F.R.-V, P. Salamon, S. Srinagesh, T.F.T., T. Tran, R. Vega Thurber and F.R., unpublished observations): dominant genotypes are not found over time, and there is a dynamic equilibrium of functionally redundant microbial and viral strains that continuously replace each other in a kill-the-winner manner, thereby maintaining a stable metabolic potential and taxonomical signal. In terms of resource use, there is no experimental evidence to show that some bacterial cells in a population are more generalist or have superior fitness. Furthermore, mathematical modelling indicates that evolution of microbial systems does not lead to decreased biodiversity by the expansion of a variant that is ecologically superior[51]. On the contrary, it has been shown that some coexisting strains differentially specialize in micro-niches, thanks to their different gene content by, for example, using different sets of organic compounds heterotrophically[27] or exploiting different types of particulate matter[33]. This supports the idea that selection favours resource diversification rather than a dominant genotype that can use multiple substrates. It would nevertheless be interesting to measure experimentally the individual fitness of different bacterial strains isolated from a natural habitat to determine whether fitter variants are selected against in environments under phage pressure.

Box 1 | Constant diversity and the pan-genome



The gene pool of bacterial and archaeal species can be extremely large. Different lineages contain different genomes, enormously increasing the metabolic and ecological capabilities of one bacterial species. As such, bacterial species are more appropriately described nowadays by their 'pan-genome' (REFS 7,60), including a core genome, which contains genes that are present in all strains (the essence of this natural unit[61]), and an accessory genome, which consists of partially shared and strain-specific genes. The accessory (or adaptive) genome includes key genes that are required to survive under a specific lifestyle that is characteristic of the strain[62]. Ecotypes have been claimed to have the quintessential properties of species[15]. One key question concerns whether the pan-genome reflects the diversity of ecotypes in one species or whether ecotypes reflect real bacterial species and have a large pan-genome. Although individual ecotypes cannot be separated experimentally, simple environments can serve as a starting point to test whether there is a pan-genome structure in such ecotypes. One such environment is the saturated brine that is dominated by a single archaeon species (*Haloquadratum walsbyi*) (see the figure). In this extremely simplified and homogeneous ecosystem, the archaeon degrades the organic matter produced by the green algae *Dunaliella salina* and forms a coherent and distinct ecological unit that must approximate an ecotype (see the figure). When the metagenome of 5 litres of water from this single-species ecosystem was sequenced, a surprisingly large gene pool was found, and the inferred pan-genome was twice the size of the sequenced isolate[24]. A large portion of this pan-genome structure was composed of genes related to environmental sensing and nutrient transport[27]. This is consistent with the idea that a genomically diverse ecotype is needed to efficiently degrade all the organic matter, as the number of required metabolic and transport genes far exceeds those that can be packaged within a single genome. If periodic selection occurs, then the pan-genome of one ecotype would be narrow. Thus, single-species metagenomics supports the idea that a high diversity (or a large pan-genome) exists even within limited spatial and temporal frames and supports the model that constant-diversity dynamics is predominant over periodic-selection dynamics in homogeneous aquatic environments.

Table 1 | **Predictions from the clonal sweeps and constant-diversity models**

Periodic-selection or ecotype model	Constant-diversity model
Low diversity of cell types, with a dominant ecotype	Many concurrent cell types
Abrupt changes in dominant cell type	Stable, high diversity of cell types
Dominant lineage is more fit in exploiting resources	Each different lineage is suboptimal in exploiting resources
Dominant lineage is more generalist	Resource diversification among strains
Ecosystem is expected to be less efficient owing to lack of resource specialization	Ecosystem is expected to be more efficient owing to resource use by many ecotypes
Dominant lineage changes through time	Lineages are more stable
Absent or limited phage pressure	High phage pressure
Variable genomic regions are related to optimal niche exploitation	Variable genomic regions are related to different sensitivity to phage predation
Species coherence arises from regular clonal sweeps	Species coherence arises from phage specificity
Few niches or a homogeneous environment	Many niches (plankton paradox)
Physically constrained populations (i.e., hosts and biofilms)	Interacting populations (i.e., free living)
Selection is driven by the efficiency of resource use	Selection is driven by the evolutionarily stable strategy
Gaps between species arise from genotypes that exploit a niche suboptimally	Gaps between species arise from rare genotypes with a higher probability of phage infection

Other predictions are more difficult to evaluate with the available data. Ecosystem functioning, for example, has been thoroughly studied in terrestrial ecosystems, in which a high biodiversity has been related to efficiency of resource use[40,41]. This information is still limited for microbial ecosystems, but the data also relate microbial diversity to ecosystem efficiency and stability[52]. It would be interesting to quantify the efficiency and stability of ecosystem functioning in natural conditions and in the laboratory, where both bacterial diversity and the presence of phages can be manipulated in microcosm experiments[53,54]. We would predict that the presence of bacteriophages would promote resource specialization in their bacterial hosts, regardless of the respective availability of substrates, facilitating efficient nutrient recycling and system stability.

The CD dynamics counter-intuitively predicts that the strain of highest fitness would be selected against and that phage predation would select for cell types of lower fitness that use any substrate regardless of its availability (FIG. 3). By keeping a low profile in the population, the cells would be part of a selection unit that is formed by bacterial strains of suboptimal but nearly equal fitness, the predation risk for which is shared equally among the strains. Thus, the fitness concept under PS dynamics is dependent on the efficiency with which a resource is used, but the evolutionary success of a strain under CD dynamics is dependent on the fitness of the other strains that co-inhabit the environment and is therefore a relative concept. This is analogous to the outcome of game theory models for optimal market strategies: there is no optimal economic strategy *per se*; instead, the best market tactic depends on what the competitors are doing. This thinking has been successfully applied to evolutionary theory, giving rise to models that conclude that, once an optimal proportion of strategies is achieved in a population, those proportions are in equilibrium over time[55]. We believe that these evolutionarily stable strategies best define the situation in bacterial populations in the presence of phages, because under those conditions the fitness of bacteria would be lowered to an equilibrium point in which demographic success would be approximately equal among cell types. Any strain that increased its resource use evolutionarily would be quickly eliminated by phage predation, and an evolutionarily stable strategy would necessarily imply that growth rates would be adjusted to those of the rest of the population, unless the cell is immune to phage attack for some reason. The idea does not go against natural selection, but it does require that bacterial fitness is defined to include ecological aspects such as predation avoidance[17].

Microbial evolution under phage pressure

Under which circumstances should we therefore expect PS or CD to predominate? The influence of phage predation on bacterial diversity requires that bacterial populations interact with each other; therefore, host-associated niches can act as physical barriers that prevent direct cell competition and phage dispersal. Thus, pathogens are not expected to follow CD dynamics unless they become very numerous and persist in a host for a long time. CD dynamics are also not expected to occur in physically constrained microbial communities such as biofilms, in which populations cannot interact with each other or invade other niches, apart from their role in constraining phage attack and dispersal. By definition, CD dynamics require the presence of phages and are therefore expected to be less common in specific environments with limited viral presence or efficiency. Accordingly, large MGIs are not observed in biofilms, as work carried out in the acid mine drainage system shows[56]. Thus, although the presence of CD patterns mediated by phage predation is pervasive in natural environments that are inhabited by competing, free-living cells, the two types of population dynamics are not mutually exclusive, and PS selection is still expected to occur in circumstances of low predation pressure (for example, intracellular environments), physically constrained cells (for example, biofilms) or physically isolated environments that are subject to founder effects (for example, animal hosts).

The CD dynamics presented here may shed some light on different aspects of microbial evolution. One of them is the function of giant proteins, which are over-represented in MGIs and could therefore be involved in host–phage interactions. In addition, if predation pressure is so intense, we would predict that genes that provide phage resistance would be among the fastest evolving in free-living species in order to counteract the fast phage adaptability[57], giving rise to an evolutionary arms race[35] analogous to that which occurs between the immune system and surface antigens of bacterial pathogens. Another aspect that should be explored is the potential role of phages in defining the limits of bacterial populations or ecotypes. In this sense, phage infection may provide a link between the evolutionary process of generating diversity and the mechanistic explanation for it, which indicates that lack of recombination is responsible for population genetic divergence[58]. The phages themselves could facilitate homogenization of sequences by

Box 2 | **How much richness can lytic viruses produce in a prokaryote community?**

It has been proposed that the richness of heterotrophic prokaryotes in the photic zone of pelagic environments is regulated by lytic viruses[21]. A density-dependent loss mechanism such as this would be expected to act selectively on host populations that reach a sufficient size. By preventing the population from sequestering all of the limiting resources and filling up the total community size (typically 3×10^5 cells ml^{-1} to 3×10^6 cells ml^{-1}), viral lysis will leave room for other, potentially less competitive, host populations to fill in the remaining part of the prokaryote community. The richness that such a mechanism can produce is the ratio between the size of the total community and the average size of the host populations. Assuming a steady state, the balance between production and loss of viruses of type *i* is given by equation 1:

$$(m_i - 1)\beta_i B_i V_i = \delta_i V_i \qquad (1)$$

m_i, β_i and δ_i are the burst size, the effective adsorption constant and the specific decay rate, respectively, for the host–virus pair B_i,V_i. The "–1" reflects the loss of one virus at infection. Eliminating V_i and solving for B_i gives the size of the steady-state host population B_i (see equation 2):

$$B_i = \frac{\delta_i}{(m_i - 1)\beta_i} \qquad (2)$$

The highest richness allowable corresponds to the smallest possible values for all B_i and thus to low δ_i values but high m_i and β_i values. A classical determination of the value for the adsorption constant β (from diffusion estimates) is 0.24 10^{-8} ml min^{-1} $virus^{-1}$ (REF. 63). This is expected to be a high value because, among other factors, the defence mechanisms in the host could lower the effective value.

Combining this with, for natural systems, a high burst size of 100 viruses released per lysis and a slow decay rate of 1 per week, equation 2 gives a result of around 400 cells ml^{-1} as a low estimate for the host abundance that can sustain a population of free lytic viruses. For a community size of 10^6 ml^{-1}, viral lysis can therefore be argued to allow a richness of around 2,500 different hosts. Host groups in the sense used in this model could well be indistinguishable by methods such as 16S ribosomal RNA sequencing.

recombination, by providing a constant flux of homologous sequences among the population that is co-infected by the same phage type. With the sequencing of viral metagenomes[59], our understanding of the diversity and specificity of phages will be improved, and mathematical modelling of the virus–bacteria interactions should become possible under realistic assumptions. In addition, experimental evolution studies in the presence and absence of phages[17] and further genomic characterization of natural environments with different phage pressures should confirm whether the PS and CD dynamics complement each other and may explain the generation of intraspecies microbial diversity under different ecological circumstances.

Francisco Rodriguez-Valera and Ana-Belen Martin-Cuadrado are at the Departmento de Producción Vegetal y Microbiología, Universidad Miguel Hernandez, Apartado 18 San Juan de Alicante, 03550 Spain.

Beltran Rodriguez-Brito and Forest Rohwer are at the Department of Biology, San Diego State University, San Diego, California 92182, USA.

Lejla Pašić is at the Department of Biology, University of Ljubljana, Večna pot 111, 1000 Ljubljana, Slovenia.

T. Frede Thingstad is at the Department of Microbiology, University of Bergen, 5020 Bergen, Norway.

Alex Mira is at the Genomics and Health Department, Center for Advanced Research in Public Health, Avda. Cataluña 21 Valencia, 46020 Spain.

Correspondence to F.R.-V.
e-mail: frvalera@umh.es

doi:10.1038/nrmicro2235

1. Gupta, S. & Maiden, M. C. Exploring the evolution of diversity in pathogen populations. *Trends Microbiol.* **9**, 181–185 (2001).
2. Whitaker, R. J. & Banfield, J. F. Population genomics in natural microbial communities. *Trends Ecol. Evol.* **21**, 508–516 (2006).
3. Medini, D. *et al.* Microbiology in the post-genomic era. *Nature Rev. Microbiol.* **6**, 419–430 (2008).
4. Medini, D., Donati, C., Tettelin, H., Masignani, V. & Rappuoli, R. The microbial pan-genome. *Curr. Opin. Genet. Dev.* **15**, 589–594 (2005).
5. Liti, G. *et al.* Population genomics of domestic and wild yeasts. *Nature* **458**, 337–341 (2009).
6. Tettelin, H. *et al.* Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". *Proc. Natl Acad. Sci. USA* **102**, 13950–13955 (2005).
7. Konstantinidis, K. T. & Tiedje, J. M. Genomic insights that advance the species definition for prokaryotes. *Proc. Natl Acad. Sci. USA* **102**, 2567–2572 (2005).
8. Maharjan, R., Seeto, S., Notley-McRobb, L. & Ferenci, T. Clonal adaptive radiation in a constant environment. *Science* **313**, 514–517 (2006).
9. Atwood, K. C., Schneider, L. K. & Ryan, F. J. Periodic selection in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **37**, 146–155 (1951).
10. Atwood, K. C., Schneider, L. K. & Ryan, F. J. Selective mechanisms in bacteria. *Cold Spring Harb. Symp. Quant. Biol.* **16**, 345–355 (1951).
11. Feil, E. J. Small change: keeping pace with microevolution. *Nature Rev. Microbiol.* **2**, 483–495 (2004).
12. Hershberg, R. *et al.* high functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biol.* **6**, e311 (2008).
13. Cohan, F. M. & Koeppel, A. F. The origins of ecological diversity in prokaryotes. *Curr. Biol.* **18**, R1024–R1034 (2008).
14. Acinas, S. G. *et al.* Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* **430**, 551–554 (2004).
15. Gevers, D. *et al.* Opinion: Re-evaluating prokaryotic species. *Nature Rev. Microbiol.* **3**, 733–739 (2005).
16. Koeppel, A. *et al.* Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. *Proc. Natl Acad. Sci. USA* **105**, 2504–2509 (2008).
17. Kassen, R. & Rainey, P. B. The ecology and genetics of microbial diversity. *Annu. Rev. Microbiol.* **58**, 207–231 (2004).
18. Wildschutte, H. & Lawrence, J. G. Differential *Salmonella* survival against communities of intestinal amoebae. *Microbiology* **153**, 1781–1789 (2007).
19. Zubkov, M. V. & Tarran, G. A. High bacterivory by the smallest phytoplankton in the North Atlantic Ocean. *Nature* **455**, 224–226 (2008).
20. Suttle, C. A. Marine viruses — major players in the global ecosystem. *Nature Rev. Microbiol.* **5**, 801–812 (2007).
21. Thingstad, T. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic ecosystems. *Limnol. Oceanogr.* **45**, 1320–1328 (2000).
22. Thingstad, T. F. & Lignell, R. Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquat. Microb. Ecol.* **13**, 19–27 (1997).
23. Coleman, M. L. *et al.* Genomic islands and the ecology and evolution of *Prochlorococcus*. *Science* **311**, 1768–1770 (2006).
24. Legault, B. A. *et al.* Environmental genomics of "*Haloquadratum walsbyi*" in a saltern crystallizer indicates a large pool of accessory genes in an otherwise coherent species. *BMC Genomics* **7**, 171 (2006).
25. Coleman, M. L. & Chisholm, S. W. Code and context: *Prochlorococcus* as a model for cross-scale biology. *Trends Microbiol.* **15**, 398–407 (2007).
26. Wilhelm, L. J., Tripp, H. J., Givan, S. A., Smith, D. P. & Giovannoni, S. J. Natural variation in SAR11 marine bacterioplankton genomes inferred from metagenomic data. *Biol. Direct* **2**, 27 (2007).
27. Cuadros-Orellana, S. *et al.* Genomic plasticity in prokaryotes: the case of the square haloarchaeon. *ISME J.* **1**, 235–245 (2007).
28. Kettler, G. C. *et al.* Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus*. *PLoS Genet.* **3**, e231 (2007).
29. Frias-Lopez, J. *et al.* Microbial community gene expression in ocean surface waters. *Proc. Natl Acad. Sci. USA* **105**, 3805–3810 (2008).
30. Garcia Martin, H. *et al.* Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities. *Nature Biotechnol.* **24**, 1263–1269 (2006).
31. Sharma, R. S., Mishra, V., Mohmmed, A. & Babu, C. R. Phage specificity and lipopolysaccarides of stem- and root-nodulating bacteria (*Azorhizobium caulinodans*, *Sinorhizobium* spp., and *Rhizobium* spp.) of *Sesbania* spp. *Arch. Microbiol.* **189**, 411–418 (2008).
32. Fitzgerald, C., Sherwood, R., Gheesling, L. L., Brenner, F. W. & Fields, P. I. Molecular analysis of the *rfb O* antigen gene cluster of *Salmonella enterica* serogroup O:6,14 and development of a serogroup-specific PCR assay. *Appl. Environ. Microbiol.* **69**, 6099–6105 (2003).
33. Ivars-Martinez, E. *et al.* Comparative genomics of two ecotypes of the marine planktonic copiotroph *Alteromonas macleodii* suggests alternative lifestyles associated with different kinds of particulate organic matter. *ISME J.* **2**, 1194–1212 (2008).
34. Reva, O. & Tummler, B. Think big — giant genes in bacteria. *Environ. Microbiol.* **10**, 768–777 (2008).
35. Leiman, P. G. *et al.* The structures of bacteriophages K1E and K1–5 explain processive degradation of polysaccharide capsules and evolution of new host specificities. *J. Mol. Biol.* **371**, 836–849 (2007).
36. Mizoguchi, K. *et al.* Coevolution of bacteriophage PP01 and *Escherichia coli* O157:H7 in continuous culture. *Appl. Environ. Microbiol.* **69**, 170–176 (2003).
37. Wilmes, P., Simmons, S. L., Denef, V. J. & Banfield, J. F. The dynamic genetic repertoire of microbial communities. *FEMS Microbiol. Rev.* **33**, 109–132 (2009).

38. Sorek, R., Kunin, V. & Hugenholtz, P. CRISPR — a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature Rev. Microbiol.* **6**, 181–186 (2008).
39. Kunin, V. *et al.* A bacterial metapopulation adapts locally to phage predation despite global dispersal. *Genome Res.* **18**, 293–297 (2008).
40. Cardinale, B. J. *et al.* Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* **443**, 989–992 (2006).
41. Tylianakis, J. M., Didham, R. K., Bascompte, J. & Wardle, D. A. Global change and species interactions in terrestrial ecosystems. *Ecol. Lett.* **11**, 1351–1363 (2008).
42. Thompson, J. R. *et al.* Genotypic diversity within a natural coastal bacterioplankton population. *Science* **307**, 1311–1313 (2005).
43. Brown, M. V., Schwalbach, M. S., Hewson, I. & Fuhrman, J. A. Coupling 16S-ITS rDNA clone libraries and automated ribosomal intergenic spacer analysis to show marine microbial diversity: development and application to a time series. *Environ. Microbiol.* **7**, 1466–1479 (2005).
44. Allen, E. E. *et al.* Genome dynamics in a natural archaeal population. *Proc. Natl Acad. Sci. USA* **104**, 1883–1888 (2007).
45. Hunt, D. E. *et al.* Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* **320**, 1081–1085 (2008).
46. Hennes, K. P., Suttle, C. A. & Chan, A. M. Fluorescently labeled virus probes show that natural virus populations can control the structure of marine microbial communities. *Appl. Environ. Microbiol.* **61**, 3623–3627 (1995).
47. Wommack, K. E., Ravel, J., Hill, R. T. & Colwell, R. R. Hybridization analysis of Chesapeake Bay virioplankton. *Appl. Environ. Microbiol.* **65**, 241–250 (1999).
48. Parada, V., Baudoux, A. C., Sintes, E., Weinbauer, M. G. & Herndl, G. J. Dynamics and diversity of newly produced virioplankton in the North Sea. *ISME J.* **2**, 924–936 (2008).
49. Golomidova, A., Kulikov, E., Isaeva, A., Manykin, A. & Letarov, A. The diversity of coliphages and coliforms in horse feces reveals a complex pattern of ecological interactions. *Appl. Environ. Microbiol.* **73**, 5975–5981 (2007).
50. Holmfeldt, K., Middelboe, M., Nybroe, O. & Riemann, L. Large variabilities in host strain susceptibility and phage host range govern interactions between lytic marine phages and their *Flavobacterium* hosts. *Appl. Environ. Microbiol.* **73**, 6730–6739 (2007).
51. Shoresh, N., Hegreness, M. & Kishony, R. Evolution exacerbates the paradox of the plankton. *Proc. Natl Acad. Sci. USA* **105**, 12365–12369 (2008).
52. Wittebolle, L. *et al.* Initial community evenness favours functionality under selective stress. *Nature* **458**, 623–626 (2009).
53. Benton, T. G., Solan, M., Travis, J. M. & Sait, S. M. Microcosm experiments can inform global ecological problems. *Trends Ecol. Evol.* **22**, 516–521 (2007).
54. Brockhurst, M. A., Fenton, A., Roulston, B. & Rainey, P. B. The impact of phages on interspecific competition in experimental populations of bacteria. *BMC Ecol.* **6**, 19 (2006).
55. Maynard-Smith, J. *Evolution and the Theory of Games*. (Cambridge Univ. Press, Cambridge, 1982).
56. Simmons, S. L. *et al.* Population genomic analysis of strain variation in *Leptospirillum* group II bacteria involved in acid mine drainage formation. *PLoS Biol.* **6**, e177 (2008).
57. Strom, S. L. Microbial ecology of ocean biogeochemistry: a community perspective. *Science* **320**, 1043–1045 (2008).
58. Fraser, C., Hanage, W. P. & Spratt, B. G. Recombination and the nature of bacterial speciation. *Science* **315**, 476–480 (2007).
59. Schoenfeld, T. *et al.* Assembly of viral metagenomes from Yellowstone hot springs. *Appl. Environ. Microbiol.* **74**, 4164–4174 (2008).
60. Konstantinidis, K. T. & Tiedje, J. M. Towards a genome-based taxonomy for prokaryotes. *J. Bacteriol.* **187**, 6258–6264 (2005).
61. Ochman, H., Lerat, E. & Daubin, V. Examining bacterial species under the specter of gene transfer and exchange. *Proc. Natl Acad. Sci. USA* **102** (Suppl. 1), 6595–6599 (2005).
62. Read, T. D. & Ussery, D. W. Opening the pan-genomics box. *Curr. Opin. Microbiol.* **9**, 496–498 (2006).
63. Stent, G. S. *Molecular Biology of Bacterial Viruses*. (W. H. Freeman & Co, San Francisco,1963).
64. Rusch, D. B. *et al.* The Sorcerer II Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLoS Biol.* **5**, e77 (2007).
65. Dinsdale, E. A. *et al.* Functional metagenomic profiling of nine biomes. *Nature* **452**, 629–632 (2008).
66. Qimron, U., Marintcheva, B., Tabor, S. & Richardson, C. C. Genomewide screens for *Escherichia coli* genes affecting growth of T7 bacteriophage. *Proc. Natl Acad. Sci. USA* **103**, 19039–19044 (2006).
67. Forde, S. E., Thompson, J. N., Holt, R. D. & Bohannan, B. J. Coevolution drives temporal changes in fitness and diversity across environments in a bacteria–bacteriophage interaction. *Evolution* **62**, 1830–1839 (2008).

Acknowledgements

Work in the laboratory of F.R.-V. is supported by grant BIO2008-02444 from the Spanish Ministry of Science and Innovation. A.-B.M.-C. is supported by a Juan de la Cierva postdoctoral fellowship from the Spanish Ministry of Science and Innovation. A.M. is funded by the Centro Superior de Investigación en Salud Pública (CSISP) from Generalitat Valenciana. L.P. received a European Molecular Biology Laboratory (EMBO) short-term scholarship (ASTF366-2007) during her stay in the F.R.-V. laboratory. Funding for the F.R.-V. laboratory was received from the US National Science Foundation (NSF) (DEB-BE04-21955), the Gordon and Betty Moore Foundation, an Achievement Rewards for College Scientists Fellowship, NSF Postdoctoral Fellowship number DBI-0511948 and Advanced Technology Program/Kent Sea Tech.

DATABASES

Entrez Genome Project: http://www.ncbi.nlm.nih.gov/sites/entrez?db=genomeprj *Candidatus* Accumulibacter phosphatis | *Candidatus* Pelagibacter ubique | *Escherichia coli* | *Haloquadratum walsbyi*

FURTHER INFORMATION

Francisco Rodriguez-Valera's homepage: http://egg.umh.es/

The Mathworks: http://www.mathworks.com

SUPPLEMENTARY INFORMATION

See online article: S1 (figure) | S2 (table) | S3 (figure)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

ERRATUM

The biology and future prospects of antivirulence therapies

Lynette Cegelski, Garland R. Marshall, Gary R. Eldridge & Scott J. Hultgren

Nature Reviews Microbiology **6**, 17–27 (2008), doi:10.1038/nrmicro1818

In the above article, a mistake was introduced in figure 4. The correct figure is shown below. We wish to apologize to the authors, and to readers, for any confusion caused.

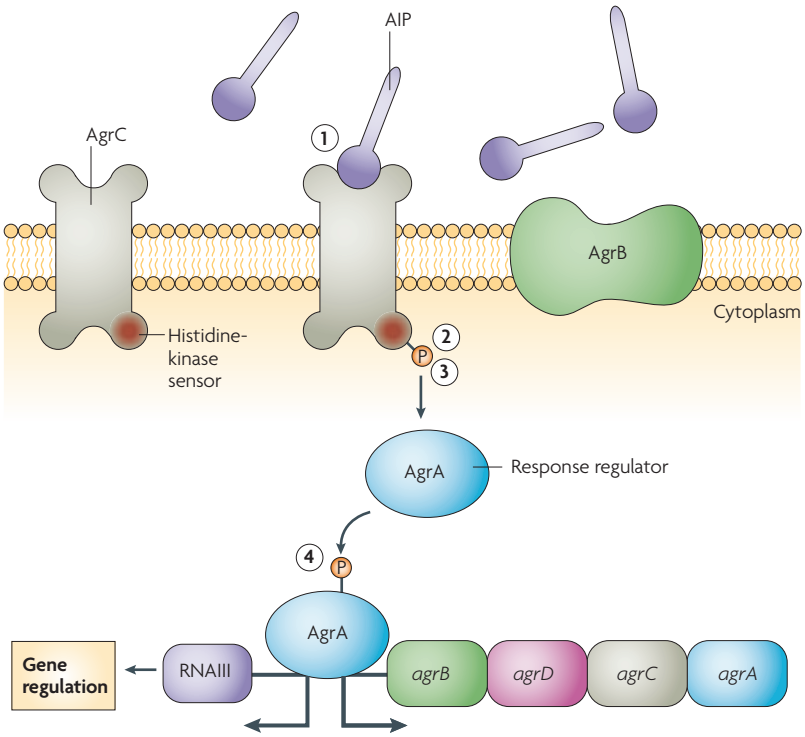


Figure 4 | **Pairing quorum sensing and two-component signalling in the staphylococcal *agr* system.** *Staphylococcus aureus* uses a two-component response system (TCRS) to mediate quorum sensing (QS). The regulation of QS involves the production of an autoinducer and an increase in its concentration, expression of RNAIII and the subsequent regulation of QS genes. *S. aureus* produces an autoinducing peptide (AIP) that accumulates extracellularly and activates the TCRS. The TCRS involves signal recognition by a histidine kinase (AgrC) (1), followed by histidine phosphorylation (2) and phosphotransfer to a response regulator (AgrA) (3), which then binds to the RNAIII transcript that encodes a small RNA that functions to modulate gene expression of *S. aureus* genes (4).