Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Microbiology

How clonal are bacteria over time?

B Jesse Shapiro

Bacteria and archaea reproduce clonally (vertical descent), but exchange genes by recombination (horizontal transfer). Recombination allows adaptive mutations or genes to spread rapidly within (or even between) species, and reduces the burden of deleterious mutations. Clonality — defined here as the balance between vertical and horizontal inheritance — is therefore a key microbial trait, determining how quickly a population can adapt and the size of its gene pool. Here, I discuss whether clonality varies over time and if it can be considered a stable trait of a given population. I show that, in some cases, clonality is clearly not static. For example, non-clonal (highly recombining) populations can give rise to clonal expansions, often of pathogens. However, an analysis of time-course metagenomic data from a lake suggests that a bacterial population's past clonality (as measured by its genetic diversity) is a good predictor of its future clonality. Clonality therefore appears to be relatively — but not completely — stable over evolutionary time.

Address
Département de sciences biologiques, Université de Montréal, Montréal, QC H3C 3J7, Canada

Corresponding author: Shapiro, B Jesse (jesse.shapiro@umontreal.ca)

Current Opinion in Microbiology 2016, **31**:116–123

This review comes from a themed issue on **Environmental microbiology**

Edited by **Monica Vasquez** and **Steven Hallam**

http://dx.doi.org/10.1016/j.mib.2016.03.013

Introduction

Here, I revisit the question posed in the title of a classic paper by John Maynard Smith and colleagues [1]: How clonal are bacteria, and more specifically how does clonality vary among different microbial populations and over time? First, what do we mean by clonality? Perfectly clonal bacteria replicate by cell division (vertical descent) and evolve by random mutations that occur during DNA replication. In this theoretical population, there is negligible horizontal transfer of DNA by recombination across the resulting tree of vertical descent. Very few (if any) natural bacterial populations fit this idealized, theoretical definition of clonality. Or, as discussed below, they might only fit it for a short amount of time. However, knowing where a bacterial population of interest happens to fall along a spectrum of clonality can help us understand its biology, and even make predictions about its evolution.

The opposite of clonality is panmixis — a situation in which the rate of horizontal transfer is higher than the rate of vertical cell division, resulting in random association (linkage equilibrium) among loci in the genome [1,2]. However, rates of horizontal transfer (recombination) vary widely across the genome, such that a population can be mostly clonal, except for a few loci in the genome [3]. These loci came to be termed genomic islands — a metaphor I will build upon below. Some of the first islands identified were called pathogenicity islands because they contained virulence factors [4]. However, non-pathogenic environmental bacteria also contain islands, conferring adaptation to different ecological niches. For example, genes in *Prochlorococcus* genomic islands confer adaptation to light and nutrient conditions [5,6]. But islands need not confer niche adaptation to their host genome; they can be neutral to host fitness or even detrimental, selfish parasites. Here, I define genomic islands broadly as any piece of DNA that is transferred horizontally (by either homologous or nonhomologous recombination) from cell to cell and therefore evolves independently (i.e. is unlinked) from the rest of the genome.

I will begin by extending the use of island analogies to include continents, peninsulas and archipelagos (Table 1). I will then use these analogies to discuss to what extent microbial populations are clonal or panmictic, and how often they transition between the two regimes.

Are some islands really peninsulas?

In the classic analogy, an island is totally disconnected from the mainland, meaning that genes in the island evolve independently of the genome (Table 1). Examples of islands that fit this strict independence might include integrated phages and other 'selfish' elements, or genes that reside in a particular niche but not in a particular genome (e.g. a gene ecology model [7•]). Peninsulas provide an analogy that might better describe how islands are related to microbial genomes. A peninsula (or 'presque-ile,' from the French for 'almost island') is a geographic term for a very narrow strip of land connected to (but distinct from) the mainland. In my analogy, an island is evolutionarily independent of the mainland genome, but their fates may become linked, forming a peninsula. For example, a bacterium may acquire a gene from a vast microbial gene pool. This gene allows the bacterium to invade a new ecological niche, triggering a

Table 1

Extended island metaphors of microbial genome evolution

Geographic metaphor	Genetic unit to which the metaphor applies	Type of selective sweep experience by the unit	Dominant mode of genetic transmission	Example
Island	Gene	Gene-specific	Horizontal	Genes in the *V. cholerae* integron [22•,23]
Peninsula	Gene	Genome-wide	Vertical (clonal)	The cholera toxin gene, acquired horizontally, then linked to a clonal *V. cholerae* genome [9,21]
Continent	Genome	Genome-wide	Vertical (clonal)	Clonal expansions of *S. aureus* [28], *M. tuberculosis* [31,43]
Archipelago	Genome	Gene-specific	Horizontal	Hotspring cyanobacteria [11•], ocean vibrios [13•,14], pneumococcus [44,46•]

clonal expansion in which the fate of the gene and its new host genome are linked, at least for the duration of the clonal expansion. One such example could be *Yersinia pestis*, which acquired a single gene allowing flea-borne transmission and triggering a clonal expansion in the form of Plague pandemics [8]. Another peninsula, the prophage-encoded cholera toxin, and its links to the mainland *Vibrio cholerae* genome [9,10], is discussed below.

Are some genomes archipelagos?

The very concept of one or a few islands implies a contrast with the large, clonal genomic mainland or continent. But some microbial genomes may contain so many islands that there is no mainland, only a vast archipelago (Table 1). A striking recent example is a population of hotspring cyanobacteria in which virtually every gene in the genome evolved independently due to frequent recombination [11•], leading the authors to call the population 'quasi-sexual' (in other words, panmictic). Frequent recombination was confirmed by another group, using different methods to study the same cyanobacteria [12]. This group also found that despite a history of panmixis over long time scales, populations are more clonal over shorter time scales. Similarly, the Asian ocean population of *Vibrio parahaemolyticus* also forms a panmictic gene pool, with each recombination block of ~1.8 kbp evolving independently [13•]. However, the panmictic gene pool occasionally gives rise to pandemic clones. In another example, we found that almost every gene in a population of *Vibrio cyclitrophicus* genomes showed signs of recombination over relatively recent time scales [14]. Such apparently high rates of recombination in natural populations were mysterious at first, contradicting recombination rates measured in the lab [15,16] and predicted by theory [7•,17]. However, theoretical models (discussed below) suggest mechanisms capable of explaining how genes can spread through populations more rapidly by recombination than by clonal expansion [18,19•,20].

Clonal expansions from panmictic pools

I propose that archipelagos are not necessarily static over time, and that archipelagos can sometimes coalesce into continents. Given the right ecological opportunity, a genome from a panmictic gene pool can escape the 'gravitational pull' of recombination and take off into a clonal expansion. An example mentioned earlier is *V. cholerae*, a genetically diverse group of coastal marine bacteria, some of which cause cholera. Virulence is mainly determined by two loci in the genome: the cholera toxin and the toxin-coregulated pilus. Both genes are frequently gained and lost by recombination [21,22•], but are always found in one lineage of *V. cholerae* — the lineage causing severe disease with pandemic potential, known as the phylocore genome (PG) group [10]. It remains a mystery why the PG lineage evolved once, and only once. If PG *V. cholerae* really did evolve just once, this would be surprising because *V. cholerae* draws on a diverse, global gene pool and can be considered panmictic [23]. Therefore multiple different lineages would be expected to acquire the two (or perhaps a handful of) genetic elements required for pandemic disease. This leads to the hypothesis that pandemic cholera emergence is *selection limited* rather than *diversity limited*. In other words, benign *V. cholerae* strains constantly acquire virulence genes. However, these strains rarely encounter the right ecological niche to flourish, for example, a human population consuming brackish water. 'The right niche' has appeared a few times in human history: for example in India in the 1800s, when the Classical lineage evolved, and again in Indonesia in the 1950s, when the El Tor lineage evolved [24]. When the right conditions appear, the PG lineage, along with its virulence factors, takes off in a clonal expansions which continue to wreak havoc today (e.g. cholera pandemics from the 1800s to today, all caused by the PG clonal group). The virulence factors, previously islands in an archipelago, became a peninsula connected to the PG mainland. The linkage between virulence factors and PG remains imperfect because different variants of the cholera toxin continue to flow in and out of the PG continent [10,21]; hence the toxin remains a peninsula, not firmly part of the mainland.

V. cholerae is a particularly well-characterized example of a panmictic gene pool giving rise to a clonal expansion, but similar evolutionary dynamics are seen in other pathogens as well (e.g. *V. parahaemolyticus* described above [13•]).