Detecting rare gene transfer events in bacterial populations

Kaare M. Nielsen[1,2]*, Thomas Bøhn[1,2] and Jeffrey P. Townsend[3,4,5]

[1] Department of Pharmacy, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway
[2] GenØk-Centre for Biosafety, The Science Park, Tromsø, Norway
[3] Department of Biostatistics, Yale University, New Haven, CT, USA
[4] Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT, USA
[5] Program in Microbiology, Yale University, New Haven, CT, USA

*Correspondence:

Kaare M. Nielsen PhD (1997) Norwegian University of Science and Technology. Postdoctoral studies at DLO, NL, and Harvard University, USA. Sabbatical stay at Yale Univ. in 2011. Tenured faculty in the Department of Pharmacy, the University of Tromsø, Norway and full professor since 2005. Leader of the MMPE research group since 2009 (www.uit.no/forskning/mmpe). We aim to improve the knowledge of how bacteria utilize horizontally transferred DNA and the factors that regulate genetic recombination processes in natural environments. Advisor to Genok-Center for Biosafety, and member of the EFSA GMO panel since 2009.
kaare.nielsen@uit.no

Horizontal gene transfer (HGT) enables bacteria to access, share, and recombine genetic variation, resulting in genetic diversity that cannot be obtained through mutational processes alone. In most cases, the observation of evolutionary successful HGT events relies on the outcome of initially rare events that lead to novel functions in the new host, and that exhibit a positive effect on host fitness. Conversely, the large majority of HGT events occurring in bacterial populations will go undetected due to lack of replication success of transformants. Moreover, other HGT events that would be highly beneficial to new hosts can fail to ensue due to lack of physical proximity to the donor organism, lack of a suitable gene transfer mechanism, genetic compatibility, and stochasticity in tempo-spatial occurrence. Experimental attempts to detect HGT events in bacterial populations have typically focused on the transformed cells or their immediate offspring. However, rare HGT events occurring in large and structured populations are unlikely to reach relative population sizes that will allow their immediate identification; the exception being the unusually strong positive selection conferred by antibiotics. Most HGT events are not expected to alter the likelihood of host survival to such an extreme extent, and will confer only minor changes in host fitness. Due to the large population sizes of bacteria and the time scales involved, the process and outcome of HGT are often not amenable to experimental investigation. Population genetic modeling of the growth dynamics of bacteria with differing HGT rates and resulting fitness changes is therefore necessary to guide sampling design and predict realistic time frames for detection of HGT, as it occurs in laboratory or natural settings. Here we review the key population genetic parameters, consider their complexity and highlight knowledge gaps for further research.

Keywords: lateral or horizontal gene transfer, DNA uptake, modeling, monitoring, sampling, antibiotic resistance, GMO, biosafety

INTRODUCTION TO HGT IN BACTERIAL POPULATIONS

Bacteria in natural populations are known to import and integrate exogenous genetic material of diverse, often unidentified, origins (Eisen, 2000; Lawrence, 2002; Nakamura et al., 2004; Chen et al., 2005; Didelot and Maiden, 2010). Bacterial genomes can be exposed not only to the multitude of sources of exogenous DNA present in their natural environments (Levy-Booth et al., 2007;

Nielsen et al., 2007; Pontiroli et al., 2007; Pietramellara et al., 2009; Rizzi et al., 2012), but also to introduced sources of novel DNA such as the fraction of recombinant DNA present in genetically modified organisms (GMOs) (Nielsen et al., 2005, 2007). DNA exposure can potentially lead to horizontal gene transfer (HGT) events dependent on the multitude of parameters that govern HGT processes in various environments (Dubnau, 1999; Bensasson et al., 2004; Thomas and Nielsen, 2005; Popa and Dagan, 2011; Domingues et al., 2012a; Seitz and Blokesch, 2013).

For long-term persistence of infrequently acquired genetic material in new bacterial hosts, a conferred selective advantage is considered necessary (Feil and Spratt, 2001; Berg and Kurland, 2002; Pettersen et al., 2005; Johnsen et al., 2009; Kuo and Ochman, 2010). Experimental investigations have shown that most HGT events that integrate into the bacterial chromosome are deleterious (Elena et al., 1998; Remold and Lenski, 2004; Starikova et al., 2012). Thus, in terms of the persistence of its signature and its effects on fitness, HGT processes resemble routine mutational processes that take place at similarly low frequencies in bacteria and that are eventually lost from the population due to lack of a conferred advantage (Kimura and Ohta, 1969; Jorgensen and Kurland, 1987; Lawrence et al., 2001; Mira et al., 2001; Koonin and Wolf, 2008; Johnsen et al., 2011). However, the larger size range of DNA transferable through single HGT events increases the potential for rapid acquisition of functional traits in bacteria when compared to single mutational events (Overballe-Petersen et al., 2013).

Both HGT and mutation should be seen as processes that occur continuously in bacterial populations. For HGT, the combinatorial possibilities are nearly unlimited as the substrates are diverse evolving DNA sequences. Due to a number of barriers and limitations to HGT, only a few of these combinatorial possibilities will materialize at a given time. When individual cells with rare HGT events (and or mutations) are positively selected under particular conditions, they may become important sources of bacterial population adaptation and evolution (Imhof and Schlötterer, 2001; Townsend et al., 2003; Orr, 2005; Barrett et al., 2006; Sousa and Hey, 2013). Different bacterial species and strains are likely to experience variable contribution of HGT to their evolutionary trajectories (Spratt et al., 2001). For instance, positive selection of bacteria after HGT of drug resistance determinants plays a central role in the evolution of resistance to antibacterial agents (Bergstrom et al., 2000; Heinemann and Traavik, 2004; Aminov and Mackie, 2007; Aminov, 2010, 2011).

The detection of HGT events in a given bacterial genome can be performed retrospectively through bioinformatics-based comparative analyses (Spratt et al., 2001; Nakamura et al., 2004; Didelot and Maiden, 2010; Didelot et al., 2010). Alternatively, events may be detected via focused experimental efforts on defined bacterial populations under controlled conditions in the laboratory (Nielsen et al., 1997, 2000), or monitoring efforts on subsamples taken from bacterial populations present in various environments, e.g., from soil, water, wounds, or gastrointestinal tracts (Nielsen and Townsend, 2004; Pontiroli et al., 2009; Aminov, 2011). The latter monitoring approach has limitations, but may enable the identification of HGT events as they occur in the context of complex interactions in and between diverse bacterial communities.

Representative analysis of HGT events in bacterial communities depends on knowledge of the structure and population dynamics of the study population and the sequence of the DNA transferred. Detection strategies nevertheless frequently rely on hidden or implicit assumptions regarding the distribution and proportion of the individual cells in the sampled larger bacterial population that would carry the transferred DNA sequences (Nielsen and Townsend, 2004; Heinemann et al., 2011).

Risk assessments of genetically modified (GM) organisms also consider the potential for HGT of recombinant DNA inserts (Nielsen et al., 2005; EFSA, 2009). For instance, the large-scale cultivation of GM-plants, i.e., on about 170 million hectares worldwide (James, 2012), results in multitudinous opportunities for bacterial exposure to **recombinant DNA** and therefore, opportunities for unintended horizontal dissemination of **transgenes** (EFSA, 2004, 2009; Nielsen et al., 2005; Levy-Booth et al., 2007; Wögerbauer, 2007; Pietramellara et al., 2009; Brigulla and Wackernagel, 2010). In laboratory settings, experimental studies have demonstrated that single bacterial species can take up and recombine with DNA fragments from GM-plants under optimized conditions (e.g., Gebhard and Smalla, 1998; de Vries et al., 2001; Kay et al., 2002; Ceccherini et al., 2003). In natural settings, negative or inconclusive evidence for HGT has been reported from most sampling-based studies of agricultural soils, run-off water and gastrointestinal tract contents (Gebhard and Smalla, 1999; Netherwood et al., 2004; Mohr and Tebbe, 2007; Demanèche et al., 2008; Douville et al., 2009).

KEY CONCEPT 1 | Recombinant DNA
DNA that has been recombined in the laboratory using *in vitro* methods, and then transferred into the genome of transgenic organism via various gene transfer techniques. Such DNA usually originate from several different species.

KEY CONCEPT 2 | Transgene
Functional unit of recombined DNA present in the genome of a recombinant organism (also called genetically modified or transgenic organism). The word transgene is sometimes used synonymously with "insert" or "inserted DNA."

Between idealized laboratory conditions and investigations of complex ecosystems, HGT research suffers from significant methodological limitations, model uncertainty, and knowledge gaps. Most research on HGT from GM-plants to bacteria has been performed via bacterial screening after a limited time period following transgene exposure, perhaps in part because only limited explicit considerations of the population dynamics of HGT events have been available to guide sampling design and data analysis (Heinemann and Traavik, 2004; Nielsen and Townsend, 2004; Nielsen et al., 2005; Townsend et al., 2012).

Given the generally low mechanistic probability of horizontal transfer of non-mobile DNA in complex environments such as soil or the gastrointestinal tract, HGT events will initially be present at an exceedingly low frequency in the overall bacterial population. It may therefore take months, years, or even

longer for the few initially transformed cells to divide and numerically out-compete non-transformed members of the population to reach proportions that can be efficiently detected by a particular sampling design (Nielsen and Townsend, 2004). The generation time of bacterial populations that are potential recipients for HGT events is therefore of high importance for the determination of sample size and choice of detection methods.

A time lag between initial occurrence of rare HGTs and the opportunity of detection will therefore be present in most environments even though the relevant HGT events lead to positive selection of **transformant** bacteria (Nielsen and Townsend, 2001, 2004). Quantifying this time lag and determining the relationship between HGT frequencies and probability of detection requires the application of mathematical models with dependency on several key parameters: HGT frequencies, changes in relative fitness of the transformants, bacterial population sizes, and generation times in nature (**Figure 1** and **Box 1**). A few studies have accordingly begun to characterize the effects of **natural selection** and the probability of **fixation** of HGT events in bacterial populations (Landis et al., 2000; Nielsen and Townsend, 2001, 2004; Johnson and Gerrish, 2002; Pettersen et al., 2005; Townsend et al., 2012). The multiple levels of, and importance of population genetic considerations in understanding HGT have recently also been reviewed by Baquero and Coque (2011 and references within) and Zur Wiesch et al. (2011).

KEY CONCEPT 3 | Transformant
The individual cells in a larger bacterial population that have acquired DNA through horizontal gene transfer. The transformation frequency is often calculated as the number of transformant cells per the total number of recipient cells.

KEY CONCEPT 4 | Natural selection
The process by which traits become either more or less common in a population as a function of their heritable effect on reproductive success.

KEY CONCEPT 5 | Fixation
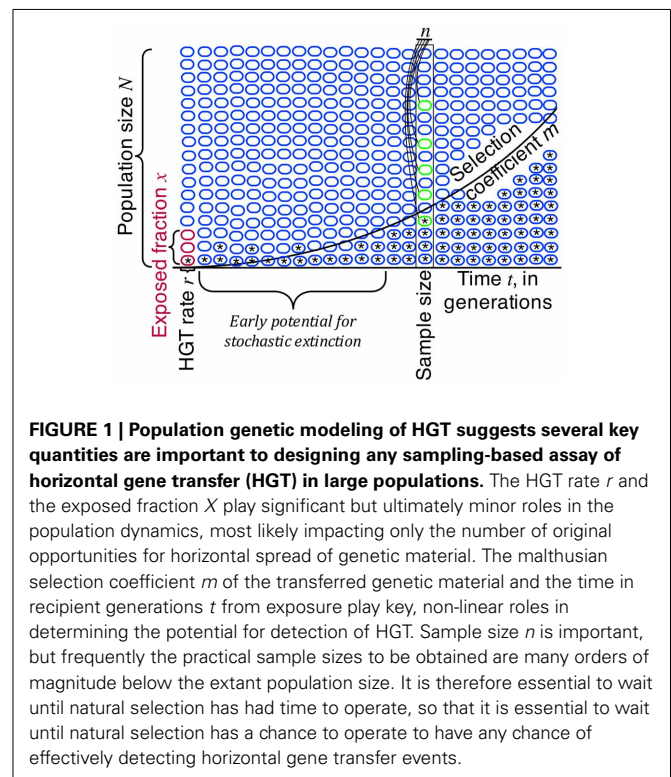A particular allele/gene is present in all members of the population.

In the recent Frontiers publication, Townsend et al. (2012), we integrated previous theory into a cohesive probabilistic framework to address current methodological shortcomings in the detection of HGT events. We assessed the key parameters and how they interact in a 5-dimensional graphical space. We also probabilistically modeled the time lag between HGT occurrence and detectability, accounting for the stochastic timing of rare HGT events; exploring also scenarios where bacteria are exposed to relevant DNA sources for a relatively short time. Our analysis yielded a simple formulation for the probability of detection given that a HGT actually occurred. This formulation facilitates computation of the statistical power of an experimental sampling design. Here we review key parameters determining the fate of HGT events in bacterial populations, consider complexity and uncertainty in parameter estimation, and identify knowledge gaps for further research.

POPULATION GENETIC PARAMETERS DETERMINING THE OUTCOME OF HGT EVENTS

Within the population genetic framework, the key parameters determining the fate of rare HGT events in larger bacterial populations are

- (i)   The rate of HGT (r),
- (iia) The bacterial population size (N),
- (iiib) The fraction (x) of the bacterial population that is exposed to the relevant DNA,
- (iii) The strength of selection on transformants (m), and
- (iv)   The bacterial generation time (t).

All these parameters are subject to considerable measurement uncertainty. Moreover, they are impacted by natural variability in biological systems and tempo-spatial interdependencies. Therefore, quantitative estimates of such parameters are best presented by ranges or probabilistic distributions, rather than point values. Measurement uncertainty can be reduced by proper statistical analysis and methodological design. Accounting for uncertainty in parameter estimation due to natural variability requires a much larger amount of data collection. Accounting for structural uncertainty due to insufficient knowledge of the broader biological system including implicit assumptions made on the nature of the HGT models applied can be acknowledged but requires extensive intense investigation of the system. In the absence of long-term basic research programs is challenging to address quantitatively. Below we consider the basis for obtaining quantitative parameter ranges that are informative in modeling

FIGURE 1 | Population genetic modeling of HGT suggests several key quantities are important to designing any sampling-based assay of horizontal gene transfer (HGT) in large populations. The HGT rate r and the exposed fraction X play significant but ultimately minor roles in the population dynamics, most likely impacting only the number of original opportunities for horizontal spread of genetic material. The malthusian selection coefficient m of the transferred genetic material and the time in recipient generations t from exposure play key, non-linear roles in determining the potential for detection of HGT. Sample size n is important, but frequently the practical sample sizes to be obtained are many orders of magnitude below the extant population size. It is therefore essential to wait until natural selection has had time to operate, so that it is essential to wait until natural selection has a chance to operate to have any chance of effectively detecting horizontal gene transfer events.

Box 1 | Key Model Equations for the Population Dynamics of HGT Events.

Under fairly general assumptions, population genetic theory facilitates quantitative estimation of probabilities of population fixation of HGT events, as well as probabilities of detection of transfers that are en route to fixation. The first concept to undertake in consideration of these probabilities is the number of HGT events that occur over time t. Assuming that transfers are independent from one another and that the product of the exposed population size n and rate of transfer r is low,

$$\Pr\{k \text{ actual transfers}\} = \frac{(xrt)^k e^{-nrt}}{k!}. \tag{1}$$

Nielsen and Townsend (2001), where $\Pr\{\}$ denotes "the probability of." Assuming a standard decay rate λ for recombinogenic DNA material (a transgene), and a decay function for the population exposed over time of of $x(t) = x_0 e^{-\lambda t}$, the

$$\Pr\{\text{neutral transgene fixation}\} = \frac{x_0 r}{N\lambda}. \tag{2}$$

Assuming large population size N, small selection coefficient s, and no interference between multiple positively selected alleles in the population,

$$\Pr\{\text{selected transgene fixation}\} = \frac{2x_0 r m}{\lambda} \tag{3}$$

(Nielsen and Townsend, 2001). In most practical contexts, HGT events will not have fixed throughout a population during the exposure period and time and time scale studied. Bacterial growth and competition may be modeled as in Hartl and Clark (1997) to yield the recombined DNA frequency p over time t after transfer based on a Malthusian selection coefficient m:

$$p = \frac{e^{mt}}{N - 1 + e^{mt}} \tag{4}$$

(Nielsen and Townsend, 2004).
Applying Equation 1 and Kimura's diffusion equation result for the spread of a positively selected allele, it is further possible to integrate over all possible timings of potential HGT, from time t of zero, through time t_s over which selection occurs, to time t_x of sampling, to calculate in a given scenario that the

$$\Pr\{\text{transgene detected}\} = \left(\frac{1 - e^{-2m}}{1 - e^{-2Nm}}\right) rx \int_0^{t_x} \left(1 - \left(1 - \frac{e^{m(t_s - t)}}{N - 1 + e^{m(t_s - t)}}\right)^n\right) e^{-\left(\frac{1 - e^{-2m}}{1 - e^{-2Nm}}\right)rxt} dt \tag{5}$$

(Townsend et al., 2012). The primary experimental design criterion, however, would be the restricted, higher probability of detection given that a successful HGT has occurred; and frequently $t_x = t_s$. Provided sufficiently large N compared to m, then, a simpler result obtains:

$$\Pr\{\text{HGT detected given that it occurred}\} = 1 - e^{-(1 - e^{-2m})rxt_x}. \tag{6}$$

---

approaches, and discuss their robustness and dependency on the model system in which they were generated.

RATES OF HGT

Several methodological approaches have been used to estimate or quantify HGT rates in bacterial populations. These methods have different limitations and advantages as discussed below.

Bioinformatics-based quantification

Bacterial genome sequencing has lead to major changes in the appreciation of the importance of HGT in bacterial evolution and adaptation (Ochman et al., 2000; Posada et al., 2002; Welch et al., 2002; Pal et al., 2005); although there is no well-established standard of proof, such studies frequently identify HGT events. Bacterial populations may also differ in their degree of accumulating HGT events (Boucher and Bapteste, 2009). For instance, bioinformatics-based genome comparisons by Dagan and Martin (2007), Dagan et al. (2008) suggest that up to 80% of the genes in

proteobacterial genomes have been involved in lateral gene transfer. Due to constraints on representative sampling and sample size, genomics-based studies are often limited to identifying HGT events that have disseminated and persisted in the larger proportion of bacterial populations over relatively long time scales. Recently, the reduction in genome sequencing costs now enables the identification of HGT events taking place in metapopulations (Whitaker and Banfield, 2006; Forsberg et al., 2012) and sampled populations of related bacteria over shorter time scales (Hiller et al., 2010; Harris et al., 2012; Sousa and Hey, 2013). Despite their clear potential to illustrate the role of HGT in evolution of bacteria, these comparative genome analyses provide fundamentally imprecise quantifications of HGT rates. This imprecision arises because such analyses *de facto* examine the evolutionary successful progeny of the bacteria in which the HGT events occurred. Any direct numeric quantification of the HGT rate from comparative genomics data is therefore challenged by a lack of distinction between the rate that HGT events

occur vs. the frequency that population dynamics spreads those events throughout species. Equivalently, comparative genomic analyses alone provide no information on the exposure rate of the sequenced genome to the same DNA fragment prior to the successful integration event, or the relative number of bacteria carrying the same HGT event that escape detection or have gone extinct during population growth (Pettersen et al., 2005). Finally, comparative genomic identification of HGT events rarely provides a precise identification of the time and location where the HGT event(s) occurred, the dependency on the specific donor DNA source, and the HGT vectors involved. These weaknesses arise inevitably from any approach relying on the analysis of single individuals sampled from an evolved and disseminated population of descendants of the primary transformants.

In summary, comparative genomic analyses, including metagenomics-based approaches, have been essential to elucidating the qualitative impact and evolutionary importance of HGT in bacterial populations. Such analyses do not yet, however, provide a quantitative understanding of HGT rates as they occur on a contemporary time scale.

Experimental-based quantification

The determination of HGT rates in a contemporary perspective is usually based on data obtained from environmental and laboratory-based experimental studies. These experimental studies can quantify HGT rates of DNA fragments of known composition in the model system used, as well as explore their dependency on experimental variables (Nordgård et al., 2007). Limitations to such studies arise from the biological model system considered, the choice and growth characteristics of the bacterial populations studied, from scaling issues and the limited capacity for analysis of individual cells in larger bacterial sample sizes. For instance, both experimentally-based laboratory studies as well as environmental sampling-based studies typically have a detection limit of 1 HGT event per 10^9–10^{10} bacteria exposed (Nielsen and Townsend, 2004). Moreover, the precision of laboratory estimates for purposes of understanding processes in natural environments is compromised because the relevant environmental variables influencing HGT rates could easily be missed or misrepresented in many, most, or all model systems.

A frequent limitation to quantitative measurement of HGT rates is the narrow focus on the transferability of a particular DNA sequence into an introduced single recipient species over a limited time frame. In particular, studies of recombination with chromosomal DNA at predefined loci in bacteria suggest a log-linear relationship with increasing sequence divergence. Drawing general conclusions on HGT rates based on experimental HGT studies that have characterized the transferability and rate of a particular, predefined gene/locus (Zawadzki et al., 1995; Vulic et al., 1997; Majewski et al., 2000; Costechareyre et al., 2009; Ray et al., 2009) is, however, problematic. For instance, HGT rates established for a particular locus and donor/recipient pair may not necessarily represent the relative gene transfer potential for that locus into a broader set of bacterial recipients. Retchless and Lawrence (2007) showed by bioinformatics-based analyses that different parts of a bacterial genome have a

non-uniform likelihood of recombination. Accordingly, Ray et al. (2009) demonstrated experimentally a high variability in gene transfer frequencies for the same donor/recipient pair depending on the genomic location of the transferred gene. The initial gene transfer frequencies that are determined between species will quantify only one aspect of the broader dynamics of the gene transfer process, as subsequent spread of the acquired DNA within a bacterial population can occur at several order of magnitude higher frequencies within species (Domingues et al., 2012a). Moreover, different members of the same bacterial population can also, due to random single mutations, have different HGT rates with divergent DNA (Rayssiguier et al., 1989; Matic et al., 1997; Townsend et al., 2003). The presence of mobile genetic elements may also increase local DNA sequence similarity between otherwise unrelated bacterial species, thereby providing opportunities for gene exchange through homologous recombination (Bensasson et al., 2004; Domingues et al., 2012a,b). Thus, the genomic location in the donor and the nature of the recombination site in the recipient cell, as well as the recombinogenic characteristics of single cells within the larger bacterial population will determine the transfer rates for a particular locus. This one example of the complexity of HGT illustrates the challenges in obtaining robust estimates of the HGT rates for a particular gene within or into a bacterial population.

To address this issue (in part), genome-wide studies of the recipients in HGT studies in laboratory populations are now emerging (Mell et al., 2011; Sauerbier et al., 2012). These latter studies suggest that multiple gene transfer events into the recipient genome often occur simultaneously after exposure to homologous genomic DNA. The outcome typically appears to be bacterial transformants with a variable proportion of between 1 and 3% of their genome carrying recombined regions. A recent study by Overballe-Petersen et al. (2013) suggest bacterial uptake of highly fragmented DNA can result in single HGT events that are producing as few as a single nucleotide change, an outcome rarely tested in previous experimental studies. Taken together, there is experimental evidence that a broad size range of DNA fragments can be acquired through HGT and that experimental data are available mostly for a subset of these size-ranges.

In summary, the quantification of HGT rates between bacterial genomes will vary with the locus examined and the model system used including the choice of donor/recipient pair as well as experimental variables. Considerations of HGT rates in bacterial populations should take into account that initially rare interspecies HGT events may subsequently transfer at high frequencies among members of the same population (Novozhilov et al., 2005; Domingues et al., 2012a). Theory (Jain et al., 1999) as well as recent, locus-context dependent or locus independent transfer models suggest more complex patterns of HGT. Such studies also suggest that robust estimates of HGT rates cannot currently be obtained in experimental models that will reflect transfer rates as they occur in complex natural systems. Nevertheless, laboratory studies are frequently viewed as representing upper levels of HGT rates for a given donor-recipient combination, and can therefore be useful for limiting the parameter ranges considered.

Sampling-based detection

Taking into account the complexity of the HGT process, experimental models will only rarely reproduce environmental conditions that can reflect HGT rates as they occur naturally. In the face of this complexity, more precise quantitative estimates of HGT rates can conceivably be obtained by the empirical approach of sampling native bacterial populations exposed to defined DNA sources. Such sampling can circumvent several of the methodological limitations of laboratory-based gene transfer models. Sampling and characterization of environmental populations of bacterial cells and populations for HGT, however, presents their own challenges (Nielsen and Townsend, 2004). A restricted sampling capacity of large populations results in an inevitably low power. In heterogeneous environments, the ability to examine relevant recipient species populations for specific HGT events is limited due to methodological constraints related to cultivation and genetic analysis. Current methodologies are mostly limited to the genetic analysis of approx. 10^3–10^5 single bacterial cells. Recent HGT events occurring at low frequencies may be nearly impossible to detect in populations that are much larger than this.

In summary, quantification of HGT rates by sampling natural bacterial populations is feasible but restricted by practical methodological constraints. The quantification of HGT events as they occur under natural conditions undoubtedly provides the better estimate of HGT rates used for modeling approaches. Only a few studies are however available on HGT rates based on sampling based approaches. Both experimental and natural sampling-based approaches carry important assumptions on HGT rates, and in both cases the relevant rates can be measured only within a limited sampling period that undoubtedly cannot be comprehensive of all potentially relevant environmental conditions.

BACTERIAL POPULATION SIZE

Bacterial populations often fluctuate in size reflecting highly opportunistic population dynamics. Sizes depend on habitat and resource availability. In the gastrointestinal tract of mammals, population sizes are typically in the range of 10^6–10^9, and in soil, 10^7–10^9 per gram of material. Quantification of bacterial population sizes can be challenging as often only a small fraction of bacteria is amendable to cultivation (e.g., in soil or the gastrointestinal tract; Janssen et al., 2002; McNamara et al., 2002; Zoetendal et al., 2004). Moreover, the ability to culture bacteria (including those in a viable but non-culturable state) will vary with environmental conditions, resulting in major methodological challenges to the culture-based determination of the population sizes for bacterial species in their communities and environments.

Additionally, fluctuating populations of bacteria are usually not evenly distributed in a given environment (Bulgarelli et al., 2012). Bacterial populations form local aggregates, clusters, and biofilms that can be isolated, partly isolated, or in contact with one another. For instance bacteria on leaf surface forms spatially structured aggregates (Kinkel et al., 1995; Morris et al., 1997; Monier and Lindow, 2004). Incorporating both the spatially structured dispersion and population fluctuations, bacterial

species may best be described as **metapopulations** (Walters et al., 2012; Fren et al., 2013). Metapopulations in close proximity are more likely to share genetic material than more distant populations (Cadillo-Quiroz et al., 2012; Polz et al., 2013). Moreover, strongly positively selected traits (and phenotypes) are likely to survive and expand locally before further dissemination to more distant populations. Such dynamics call for careful consideration in terms of constructing a proper sampling design.

KEY CONCEPT 6 | Metapopulation
A group of spatially separated populations of the same species that nonetheless have interactions, maintaining a species identity.

For bacterial populations, the spatial aspects engendered by bacterial growth dynamics are relevant on diverse scales. For example, in a soil environment, some bacteria will be found in the immediate proximity of the plant roots, and are likely to be exposed to root exudates. Other bacteria, just a few millimeters away will not accrue benefit from the same nutrient source, must compete for sparser resources, and hence exhibit smaller population sizes and slower growth rates. On a larger scale, the root systems of individual plants may be seen as discrete patches, islands or sources, but with a potential for microbial dispersal between. Topographically or ecologically demarcated fields or plots represent another level of scale, with different modes and vectors for potential transfer (dispersal) in between. Spatially structured population models have not yet been incorporated in modeling of HGT (Townsend et al., 2012). Incorporation of spatial structuring and heterogeneity is likely to be an important next step for quantitative modeling, as it will affect the population dynamics of rare bacterial phenotypes in larger populations (Fren et al., 2013). An accurate determination of N is also challenged by the lack of a precise understanding of coherent bacterial populations and the species definition in prokaryotes (Cohan, 2002, 2005; Fraser et al., 2007, 2009; Achtman and Wagner, 2008).

In summary, bacterial populations vary in size and structure within the environment. This variation will have implications for the likelihood of occurrence as well as the subsequent population dynamics of HGT events. While quantitative estimates of bacterial population sizes (N) can usually be obtained for the biological system of interest, it is challenging for such estimates to accurately account for the local fluctuations and patterns of bacterial distribution, and differences between relevant bacterial phenotypes (species).

FRACTION OF BACTERIAL POPULATION EXPOSED TO DNA

Given a spatially structured population, DNA exposure will be uneven across the recipient population; the distribution of recipients is expected to be dissimilar from the distribution of the DNA source. These differences in exposure make accurate modeling challenging as well as muddling empirical efforts to parameterize a model with the fraction of bacteria in a population that may undergo HGT as a result of exposure to a given DNA source. The presence of physical and biotic barriers to DNA exposure in most environments suggests that the fraction of potential recipients exposed would inevitably be considerably less than one. Empirical data or experimental models that permit quantification of actual DNA exposure levels in microbial communities are

typically not available. DNA exposure will also depend on the biological and physical properties of the DNA vector or DNA source. An upper bound for exposure can be determined if the absolute concentration (or copy number) and the dynamics or decay rate of the particular DNA source in question is known. Nevertheless, parameter estimates of the fraction of the bacterial population (n) exposed to the DNA source will be based on theoretical considerations relying on the known properties of the DNA source and recipient population.

STRENGTH OF SELECTION

The detection of rare HGT events in larger populations is typically feasible only if the few initial transformants have a growth advantage, so that they increase their relative proportion in the overall population. Selection in microbial populations is typically described by the Malthusian fitness parameter m, which works out to be equal to the selection coefficient in haploids (Hartl and Clark, 1997). The parameter m represents the relative cost or advantage conferred by the HGT event to the transformed bacterium compared to untransformed members of the same population ($m = 0$ represents no positive or negative selection). In nature, values of m that are relevant to HGT success would range from very weak positive selection (the reciprocal of population size, perhaps $m = 10^{-12}$) to strong positive selection (perhaps at most $m = 1$, representing a doubling of the rate of reproduction).

In certain circumstances, m could be greater than one in relative terms: stronger positive selection would, for instance, arise under intense antibiotic treatment when the acquisition of a resistance gene is exceptionally advantageous, such as when 100% of the susceptible population is likely to die. Thus, the relative growth advantage can be immense, leading to rapid population expansion and replacement of the transformant population in the absence of competitors. Clinical antibiotic usage produces strong fluctuations in the selection for a given resistance trait over time. Such exceptionally strong periodic selection is not considered further here and would require different modeling approaches such as **dynamic epidemiological modeling**. In opposite cases, horizontally acquired DNA may be costly (Baltrus, 2013) or outright lethal to the recipient cell (Sorek et al., 2007). In most cases, these two extremes are unexpected. For most relevant HGT events and acquired traits, values of m of low magnitude are expected, and constant values of m are assumed over time (Townsend et al., 2012). These typical, low values of m will not correspond to levels of selection that are easily discriminated in the laboratory populations of bacteria over limited time periods, but will affect relative growth in nature over time.

> **KEY CONCEPT 7 | Dynamic epidemiological modeling**
> The use of mathematical models to project how infectious diseases progress and to predict the outcomes of interventions.

Quantification of the strength of selection is affected by model inconstancy. The selection coefficient of a given trait is typically *not* constant over evolutionary time but will fluctuate over space and time, responding to environmental variables, varying among bacterial genotypes, and showing gene-by-environment interactions (Kimura, 1954; Barker and Butcher, 1966; Via and

Lande, 1985, 1987; Hodgins-Davis and Townsend, 2009). For instance, an antibiotic resistance trait can be highly advantageous in the presence of antibiotics but confer a fatal fitness cost in the absence of antibiotics (Johnsen et al., 2009, 2011). Thus, a numerically constant selection coefficient will by nature be an inexact approximation and careful consideration is required for quantification over variable environments. In addition to environmental fluctuations, the genome of a given bacterial transformant is not constant. Further HGT events and continual mutational processes may rapidly change the initial effects on host fitness of a given HGT event (Lenski et al., 1991; Gerrish, 2001; Heffernan and Wahl, 2002; Rozen et al., 2002; Barrett et al., 2006; Starikova et al., 2012). For instance, biological constraint to functional HGT events as represented by unconstrained gene expression or differences in codon usage patterns may be ameliorated over time through spontaneous mutations (Lawrence and Ochman, 1997; Tuller et al., 2011; Starikova et al., 2012).

In summary, the strength of selection for a given trait depends on host genetics and environmental conditions. Natural variability and complexity in such systems obstruct precise quantification of selection. Similar to HGT rates, selection is better expressed with a value range and modeled as multiple scenarios using long-term average values. Experimental models, in particular, will rarely provide a precise estimate of the selection coefficient under natural conditions, and have potential only to be able to identify relatively strong selection coefficients. Acknowledging the limitations to quantification, mathematical models should incorporate uncertainty in the estimation of the strength of selection and examine a broad parameter range of m.

BACTERIAL GENERATION TIME

Bacterial cell division time varies with species and environments (Powell, 1956; Kovarova-Kovar and Egli, 1998). Division time can be as short as <1 h in nutrient-rich environments with stable temperatures such as the gastrointestinal tract, and can be as long as several weeks in nutrient-limited and temperature-fluctuating environments such as soil. Bacterial populations with spore-forming capacity or with "persister" stages (Dawson et al., 2011) lead to non-uniform duplication rates among cells present in a population. Variable nutrient access over time and space also lead to variation in generation time, even in homogeneous bacterial culture conditions (Kovarova-Kovar and Egli, 1998). Estimates of bacterial generation time should therefore be interpreted as average population measurements with potentially very large cell-to-cell differences. Quantitative values for bacterial generation time can be obtained for cultivable species though experimental measurements in the laboratory. Actual generation times of bacterial species, as measured in their natural environments, are only rarely available.

Bacterial generation time will also be determined by the overall bacterial population and community structure; particularly if opportunities for population expansion exist. The infectious lifestyle of some pathogens leads to exceptionally rapid changes in their population sizes and generation time (during

infections) followed by strong **bottlenecks** (e.g., during trans- mission). Thus, depending on the pathogen in question, the transformed cells may or may not be competing for resources with non-transformed members of their populations. With the exception of some chronic bacterial infections, most changes in relative population sizes of pathogens are expected to take place through population replacement. The speed of replace- ment will be based on competitive advantage of the transfor- mant population as expressed through the selection coefficient and observed through a more rapid generation time. Thus, the materialized growth advantage is observed in the presence of a non-transformed population.

KEY CONCEPT 8 | Bottleneck
An event in which a population is reduced to a very small size, eliminating most genetic variation.

In summary, quantitative estimates of bacterial generation time (t) can be obtained for cultivable bacterial species. Large tempospatial variation is expected within individual members of a population and different lifestyles will determine the character- istics and limits to bacterial growth rates.

General considerations

Above we have described some key features of the parameters defining the fate of rare HGT events occurring in larger bacterial populations. These parameters have in common that they can- not be precisely quantified as a defined single numeric value for a given bacterial population and environment. Complex bacte- rial population dynamics and interactions lead to large temporal and spatial variability in parameter values in structured environ- ments. This natural variability can nevertheless be captured in modeling approaches by determining the effect of a range or likely distribution of parameter values for a given environmental HGT scenario. Multiple outcomes/scenarios can be quantified for their probability and critical parameter ranges can be identified that can structure further hypothesis formulation, guide experimental design and contribute to theory development.

The utility of a quantitative approach to understand HGT processes as presented in Townsend et al. (2012) is, although dependent on some knowledge of the rates of the relevant pro- cesses, independent of the specific bacterial mechanism of HGT (e.g., transduction, conjugation, transformation). Quantitative adjustments to the DNA exposure and HGT rate can in our model accommodate diverse mechanisms of transfer of non-mobile DNA. Other models have been developed to better understand how mobile genetic elements can move within and between bac- terial populations (e.g., Bergstrom et al., 2000; Novozhilov et al., 2005; Ponciano et al., 2005). Although out of the scope of this review, information from both types of models may be needed to generate a more complete picture of HGT in bacteria. **Figure 1** illustrates the key parameters of a quantitative approach.

An outcome of the quantitative analysis is that measured HGT frequencies are typically highly insufficient as predictors of the short and long term evolutionary impact of HGT events. As long as such events occur repeatedly, other factors will deter- mine the biological impact of these events (Pettersen et al., 2005). Directional selection typically dominates determination of the probability of detection of HGT events (Townsend et al.,

2012). Strong experimental sampling designs would therefore avoid overly focusing on the possibility of detection of the ini- tial HGT *recipients* (and associated HGT frequencies), but rather attempt to capture the population dynamics of positively selected *descendants* of the primary transformants. Generally, experi- mental design in this context will incorporate a delay between exposure and assessment, facilitating the action of selection to bring descendants of recipients to higher, detectable frequency. Experimental designs should explicitly address the intensity of selection that they should be able to detect given the design, just as (for example) clinical trial designs specify an effect size for a treatment that they should be able to detect.

Accordingly, a key consideration is when the transformant proportion rises to the point where subsequent evolution is largely deterministic based on the current level of directional selection (Rouzine et al., 2001). Understanding threshold lev- els of transformant populations also have practical implications. For instance, the prevalence of a pathogenic strain carrying an HGT event encoding antibiotic resistance for first line antibiotic therapy is of highest interest when its relative proportion among sensitive strains has reached <0.1–0.3, as such levels will call for changes in clinical prescription guidance (Daneman et al., 2008).

Even when bacteria experience positive selection, stochastic processes contribute to variability in the predicted fate of nearly neutral and weakly selected transformants. **Genetic drift**, uneven survival rates due to bottlenecks and **selective sweeps** in struc- tured bacterial populations can also play important roles in determining the fate of individual genotypes in larger popula- tions (Majewski and Cohan, 1999; Heffernan and Wahl, 2002; Pettersen et al., 2005). Random or seasonal variations in local population sizes may also cause particular genotypes (e.g., trans- formants and their descendants) to fluctuate at low frequencies above or below detection for long periods of time in spatially structured populations (Gerrish and Lenski, 1998).

KEY CONCEPT 9 | Genetic drift
The counterpart of natural selection, genetic drift is change in the frequency of a gene variant within a population due to "random sampling"–the chance association of the variant with individuals of differing reproductive success.

KEY CONCEPT 10 | Selective sweep
An event in which a highly selected genetic variant confers a large reproduc- tive advantage and carries with it only the linked genetic variants, sweeping other variation out of a population.

Spatially structured models have not yet been incorporated in modeling of HGT (Townsend et al., 2012) and are likely to be an important next step for quantitative assessment and mod- eling. Experimental designs should accommodate the potential for rare HGT events to be unevenly distributed in large, struc- tured bacterial populations. Bacterial species are expected to have non-uniform distribution of genes at spatial and temporal scales (Reno et al., 2009). For instance, antibiotic resistance genes or transgenes are likely to be initially present only in a limited number of patches (e.g., patients/hospitals, or soil sites/fields), representing sub- or meta-populations of the larger global pop- ulation (Maynard Smith, 2000). The fate of initial HGT events established in some metapopulations will depend on migration

or dispersal for further dissemination in the larger bacterial population. Bacterial dispersal relies on a multitude of factors and is expected to be of variable intensity and directionality. Bacterial dispersal can be dependent on or at least enhanced by vectors: for antibiotic-resistant bacteria, humans have been excellent vectors, carrying bacteria and their transferable traits between hospitals across continents. In soil, numerous small and large invertebrates may carry soil and bacteria around. At larger scales, water, wind, animals, birds, food and feed products, and human activities carry and disseminate bacteria both locally and globally.

Understanding the relationship between the exposed and total population size, HGT rates, bacterial generation time, selective advantage, and delayed sampling, is not an easy task. Nevertheless, it is essential to the fate of HGT events occurring in various environments. The model published by Townsend et al. (2012) examines how these parameters define the outcome of the fate of HGT occurring in various environmental scenarios. It was concluded that under some conditions HGT is likely to occur over temporal and spatial scales that are not amenable to direct experimental observation; emphasizing that the probability of detection can only correspond to a calculable level of selection, and that a powerful experimental design requires a delayed sampling strategy (i.e., not close to the initial exposure and the HGT event itself). Greater informed quantitative analysis of the population genetics of the bacterial system(s) investigated will contribute to make the assumptions behind hypothesis formulation less arbitrary and more explicit; and therefore, improve the robustness of future experimental designs.

ACKNOWLEDGMENTS

REFERENCES

Achtman, M., and Wagner, M. (2008). Microbial diversity and the genetic nature of microbial species. *Nat. Rev. Microbiol.* 6, 431–440. doi: 10.1038/nrmicro1872

Aminov, R. I. (2010). A brief history of the antibiotic era: lessons learned and challenges for the future. *Front. Microbiol.* 1:134. doi: 10.3389/fmicb.2010.00134

Aminov, R. I. (2011). Horizontal gene exchange in environmental microbiota. *Front. Microbiol.* 2:158. doi: 10.3389/fmicb.2011.00158

Aminov, R. I., and Mackie, R. I. (2007). Evolution and ecology of antibiotic resistance genes. *FEMS Microbiol. Lett.* 271, 147–161. doi: 10.1111/j.1574-6968.2007.00757.x

Baltrus, D. A. (2013). Exploring the costs of horizontal gene transfer. *Trends Ecol. Evol.* 28, 489–495. doi: 10.1016/j.tree.2013.04.002

Baquero, F., and Coque, T. M. (2011). Multilevel population genetics in antibiotic resistance. *FEMS Microbiol. Rev.* 35, 705–706. doi: 10.1111/j.1574-6976.2011.00293.x

Barker, J. S. F., and Butcher, J. C. (1966). A simulation study of quasi-fixation of genes due to random fluctuation of selection intensities. *Genetics* 53, 261–268.

Barrett, R. D. H., M'Gonigle, L. K., and Otto, S. P. (2006). The distribution of beneficial mutant effects under strong selection. *Genetics* 174, 2071–2079. doi: 10.1534/genetics.106.062406

Bensasson, D., Boore, J. L., and Nielsen, K. M. (2004). Genes without frontiers. *Heredity* 92, 483–489. doi: 10.1038/sj.hdy.6800451

Berg, O. C., and Kurland, G. C. (2002). Evolution of microbial genomes: sequence acquisition and loss. *Mol. Biol. Evol.* 19, 2265–2276. doi: 10.1093/oxfordjournals.molbev.a004050

Bergstrom, C. T., Lipsitch, M., and Levin, B. R. (2000). Natural selection, infectious transfer and the existence conditions of bacterial plasmids. *Genetics* 155, 505–519. Available online at: http://www.genetics.org/content/155/4/1505.full.pdf+html

Boucher, Y., and Bapteste, E. (2009). Revisiting the concept of lineage in prokaryotes: a phylogenetic perspective. *Bioessays* 31, 526–536. doi: 10.1002/bies.200800216

Brigulla, M., and Wackernagel, W. (2010). Molecular aspects of gene transfer and foreign DNA acquisitions in prokaryotes with regard to safety issues. *Appl. Microbiol. Biotechnol.* 86, 1027–1041. doi: 10.1007/s00253-010-2489-3

Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F., et al. (2012). Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* 488, 91–95. doi: 10.1038/nature11336

Cadillo-Quiroz, H., Didelot, X., Held, N. L., Herrera, A., Darling, A., Reno, M. L., et al. (2012). Patterns of gene flow define species of Thermophilic Archaea. *PLoS Biol.* 10:e1001265. doi: 10.1371/journal.pbio.1001265

Ceccherini, M. T., Poté, J., Kay, E., Tran Van, V., Maréchal, J., Pietramellara, G., et al. (2003). Degradation and transformability of DNA from transgenic plants. *Appl. Environ. Microbiol.* 69, 673–678. doi: 10.1128/AEM.69.1.673-678.2003

Chen, I., Christie, P. J., and Dubnau, D. (2005). The ins and outs of DNA transfer in bacteria. *Science* 310, 1456–1460. doi: 10.1126/science.1114021

Cohan, F. M. (2002). What are bacterial species? *Annu. Rev. Microbiol.* 56, 457–487. doi: 10.1146/annurev.micro.56.012302.160634

Cohan, F. M. (2005). "Periodic selection and ecological diversity in bacteria," in *Selective Sweeps*, ed D. Nurminsky (Georgetown, Texas: Landes Bioscience), 78–93.

Costechareyre, D., Bertolla, F., and Nesme, F. (2009). Homologous recombination in *Agrobacterium*: potential implications for the genomic species concept in bacteria. *Mol. Biol. Evol.* 26, 167–176. doi: 10.1093/molbev/msn236

Dagan, T., and Martin, W. (2007). Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution. *Proc. Natl. Acad. Sci. U.S.A.* 104, 870–875. doi: 10.1073/pnas.0606318104

Dagan, T., Artz-Randrup, Y., and Martin, W. (2008). Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. *Proc. Natl. Acad. Sci. U.S.A.* 105, 10039–10044. doi: 10.1073/pnas.0800679105

Daneman, N., Low, D. E., McGeer, A., Green, K. A., and Fisman, D. N. (2008). At the threshold, defining clinically meaningful resistance thresholds for antibiotic choice in community-acquired pneumonia. *Clin. Infect. Dis.* 46, 1131–1138. doi: 10.1086/529440

Dawson, C. C., Intapa, C., Jabra-Rizk, M. A. (2011). "Persisters" survival at the cellular level. *PLoS Pathog.* 7:e1002121. doi: 10.1371/journal.ppat.1002121

Demanèche, S., Sanguin, H., Poté J., Navarro, E., Bernillon, D., Mavingui, P., et al. (2008). Antibiotic resistant soil bacteria in transgenic plant fields. *Proc. Natl. Acad. Sci. U.S.A.* 105, 3957–3962. doi: 10.1073/pnas.0800072105

de Vries, J., Meier, P., and Wackernagel, W. (2001). The natural transformation of the soil bacteria Pseudomonas stutzeri and Acinetobacter sp. by transgenic plant DNA strictly depends on homologous sequences in the recipient cells. *FEMS Microbiol. Lett.* 195, 211–215. doi: 10.1016/S0378-1097(01)00012-X

Didelot, X., Lawson, D., Darling, A., and Falush, D. (2010). Inference of homologous recombination in bacteria using whole genome sequencing. *Genetics* 186, 1435–1449. doi: 10.1534/genetics.110.120121

Didelot, X., and Maiden, M. C. (2010). Impact of recombination on bacterial evolution. *Trends Microbiol.* 18, 315–322. doi: 10.1016/j.tim.2010.04.002

Domingues, S., Harms, K., Fricke F. W., Johnsen, P. J., da Silva, G., and Nielsen, K. M. (2012a). Natural transformation facilitates transfer of transposons, integrons and gene cassettes between bacterial species. *PLoS Pathog.* 8:e1002837. doi: 10.1371/journal.ppat.1002837

Domingues, S., da Silva, G. J., and K. M. Nielsen. (2012b). Integrons: vehicles and pathways for horizontal dissemination in bacteria. *Mob. Gen. Elements* 2, 211–223. doi: 10.4161/mge.22967

Douville, M., Gagné, F., Andre, C., and Blaise, C. (2009). Occurrence of the transgenic corn cry1Ab gene in freshwater mussels (*Elliptio complanata*) near corn fields: evidence of exposure by bacterial ingestion. *Ecotoxicol. Environ. Saf.* 72, 17–25. doi: 10.1016/j.ecoenv.2008.02.006

Dubnau, D. (1999). DNA uptake in bacteria. *Annu. Rev. Microbiol.* 53, 217–244. doi: 10.1146/annurev.micro.53.1.217

EFSA. (2004). European food safety authority: opinion of the scientific panel on genetically modified organisms on the use of antibiotic resistance genes as marker genes in genetically modified plants. *EFSA J.* 48, 1–18. doi: 10.2903/j.efsa.2004.48

EFSA. (2009). Use of antibiotic resistance genes as marker genes in genetically modified plants. Scientific opinion of the panel on genetically modified organisms (GMO) and the panel on biological hazards (BIOHAZ). *EFSA J.* 1034, 1–82. doi: 10.2903/j.efsa.2009.1108

Eisen, J. (2000). Horizontal gene transfer among microbial genomes: new insights from complete genome analysis. *Curr. Opin. Genet. Dev.* 10, 606–611. doi: 10.1016/S0959-437X(00)00143-X

Elena, S. F., Ekunwe, L., Hajela, N., Oden, S. A., and Lenski, R. E. (1998). Distribution of fitness effects caused by random insertion mutations in *Escherichia coli. Genetica* 102–103, 349–358. doi: 10.1023/A:1017031008316

Feil, E. J., and Spratt, B. G. (2001). Recombination and the population structure of bacterial pathogens. *Annu. Rev. Microbiol.* 55, 561–590. doi: 10.1146/annurev.micro.55.1.561

Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M., Sommer, M. O., and Dantas, G. (2012). The shared antibiotic resistome of soil bacteria and human pathogens. *Science* 337, 1107–1111. doi: 10.1126/science.1220761

Fraser, C., Hanage, W. P., and Spratt, B. G. (2007). Recombination and the nature of bacterial speciation. *Science* 315, 476–480. doi: 10.1126/science.1127573

Fraser, C., Alm, E. J., Polz, M. F., Spratt, B. G., and Hanage, W. P. (2009). The bacterial species challenge: making sense of genetic and ecological diversity. *Science* 323, 746. doi: 10.1126/science.1159388

Fren, M., Rainey, P. B., and Traulsen, A. (2013). The effect of population structure on the rate of evolution. *Proc. R. Soc. B Biol. Sci.* 280, 20130211. doi: 10.1098/rspb.2013.0211

Gebhard, F., and Smalla, K. (1998). Transformation of *Acinetobacter* sp. strain BD413 by transgenic sugar beet DNA. *Appl. Environ. Microbiol.* 64, 1550–1554.

Gebhard, F., and Smalla, K. (1999). Monitoring field releases of genetically modified sugar beets for persistence of transgenic plant DNA and horizontal gene transfer. *FEMS Microbiol. Ecol.* 28, 261–272. doi: 10.1111/j.1574-6941.1999.tb00581.x

Gerrish, P. (2001). The rhythm of microbial adaptation. *Nature* 413, 299–302. doi: 10.1038/35095046

Gerrish, P., and Lenski, R. E. (1998). The fate of competing beneficial mutations in an asexual population. *Genetica* 102–103, 127–144. doi: 10.1023/A:1017067816551

Harris, S. R., Clarke, I. N., Seth-Smith, H. M. B., Solomon, A. W., Cutcliffe, L. T., Marsh, P., et al. (2012) Whole-genome sequencing of diverse *Chlamydia trachomatis* strains identifies phylogenetic relationships masked by current clinical typing. *Nat. Genet.* 44, 413–419. doi: 10.1038/ng.2214

Hartl, D. L., and Clark, A. G. (1997). *Principles of Population Genetics*. 3rd Edn. Sunderland: Sinauer Associates Inc.

Heffernan, J. M., and Wahl, L. M. (2002). The effects of genetic drift in experimental evolution. *Theor. Pop. Genet.* 62, 349–356. doi: 10.1016/S0040-5809(02)00002-3

Heinemann, J. A., and Traavik, T. (2004). Problems in monitoring horizontal gene transfer in field trials of transgenic plants. *Nat. Biotechnol.* 22, 1105–1109. doi: 10.1038/nbt1009

Heinemann, J. A., Kurenbach, B., and Bleyendaal, N. (2011). Evaluation of horizontal gene transfer monitoring experiments conducted in New Zealand between 2004 and 2009. *J. Org. Syst.* 6, 3–19. Available online at: http://www.organic-systems.org/journal/Vol_6(1)/abstracts/3-19_Heinemann-et-al.html

Hiller, N. L., Ahmed, A., Powell, E., Martin, D. P., Eutsey, R., et al. (2010) Generation of genic diversity among *Streptococcus pneumoniae* strains via horizontal gene transfer during a chronic polyclonal pediatric infection. *PLoS Pathog.* 6:e1001108. doi: 10.1371/journal.ppat.1001108

Hodgins-Davis, A., and Townsend, J. P. (2009). Evolving gene expression: from G to E to G x E. *Trends Ecol. Evol.* 24, 649–658. doi: 10.1016/j.tree.2009.06.011

Imhof, M., and Schlötterer, C. (2001). Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. *Proc. Natl. Acad. Sci. U.S.A.* 98, 1113–1117. doi: 10.1073/pnas.98.3.1113

Jain, R., Rivera, M. C., and Lake, J. A. (1999). Horizontal gene transfer among genomes: the complexity hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 96, 3801–3806. doi: 10.1073/pnas.96.7.3801

James, C. (2012). *Global Status of Commercialized Biotech/GM Crops: 2012*. Ithaca, NY: ISAAA. ISAAA Brief No. 44. 2012.

Janssen, P. H., Yates, P. S., Grinton, B. E., Taylor, P. M., and Sait, M. (2002). Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions Acidobacteria, Acinobacteria, Proteobacteria and Verrucomicrobia. *Appl. Environ. Microbiol.* 68, 2391–2396. doi: 10.1128/AEM.68.5.2391-2396.2002

Johnsen, P. J., Townsend, J., Bøhn, T., Simonsen, G. S., Sundsfjord, A., and Nielsen, K. M. (2009). Factors affecting the reversal of antimicrobial resistance. *Lancet Infect. Dis.* 9, 357–364. doi: 10.1016/S1473-3099(09)70105-7

Johnsen, P. J., Townsend, J. P., Bøhn, T., Simonsen, G. S., Sundsfjord, A., and Nielsen, K. M. (2011). Retrospective evidence for a biological cost of vancomycin resistance in the absence of glycopeptide selective pressures. *J. Antimicrob. Chemother.* 66, 608–610. doi: 10.1093/jac/dkq512

Johnson, T., and Gerrish, P. J. (2002). The fixation probability of a beneficial allele in a population dividing by binary fission. *Genetica* 115, 283–287. doi: 10.1023/A:1020687416478

Jorgensen, F., and Kurland, C. G. (1987). Death rates of bacterial mutants. *FEMS Microbiol. Lett.* 40, 43–46. doi: 10.1111/j.1574-6968.1987.tb01979.x

Kay, E., Vogel, T. M., Bertolla, F., Nalin, R., and Simonet, P. (2002). In situ transfer of antibiotic resistance genes from transgenic (transplastomic) tobacco plants to bacteria. *Appl. Environ. Microbiol.* 68, 3345–3351. doi: 10.1128/AEM.68.7.3345-3351.2002

Kimura, M. (1954). Process leading to quasi-fixation of genes in natural populations due to random fluctuation of selection intensities. *Genetics* 39, 280–295.

Kimura, M., and Ohta, T. (1969). Average number of generations until extinction of an individual mutant gene in a finite population. *Genetics* 61, 763–771.

Kinkel, L. L., Wilson, M., and Lindow, S. E. (1995). Effects of scales on estimates of epiphytic bacterial populations. *Microb. Ecol.* 29, 283–297. doi: 10.1007/BF00164891

Koonin, E. V., and Wolf, Y. I. (2008). Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. *Nucl. Acids Res.* 36, 6688–6719. doi: 10.1093/nar/gkn668

Kovarova-Kovar, K., and Egli, T. (1998). Growth kinetics of suspended microbial cells: from single-substrate-controlled growth to mixed-substrate kinetics. *Microbiol. Mol. Biol. Rev.* 62, 646–666.

Kuo, C. H., and Ochman, H. (2010). The extinction dynamics of bacterial pseudogenes. *PLoS Genet* 6:e1001050. doi: 10.1371/journal.pgen.1001050

Landis, W. G., Lenart, L. A., and Spromberg, J. A. (2000). Dynamics of horizontal gene transfer and the ecological risk assessment of genetically engineered organisms. *Hum. Ecol. Risk Assess.* 6, 875–899. doi: 10.1080/10807030091124248

Lawrence, J. G. (2002). Gene transfer in bacteria: speciation without species? *Theor. Pop. Biol.* 61, 449–460. doi: 10.1006/tpbi.2002.1587

Lawrence, J. G., and Ochman, H. (1997). Amelioration of bacterial genomes: rates of change and exchange. *J. Mol. Evol.* 44, 383–397. doi: 10.1007/PL00006158

Lawrence, J. G., Hendrix, R. W., and Casjens, S. (2001). Where are the pseudogenes in bacterial genomes? *Trends Microbiol.* 9, 535–540. doi: 10.1016/s0966-842x(01)02198-9

Lenski, R. E., Rose, M. R., Simpson, S. C., and Tadler, S. C. (1991). Long-term experimental evolution in *Escherichia coli*. 1. Adaptation and divergence during 2 000 generations. *Am. Nat.* 138, 1315-1341. doi: 10.1086/285289

Levy-Booth, D. J., Campbell, R. G., Gulden, R. H., Hart, M. M., Powell, J. R., Klironomos, J. N., et al. (2007). Cycling of extracellular DNA in the soil environment. *Soil Biol. Biochem.* 39, 2977–2991. doi: 10.1016/j.soilbio.2007.06.020

Majewski, J., and Cohan, F. M. (1999). Adapt globally, act locally: the effect of selective sweeps on bacterial sequence diversity. *Genetics* 152, 1459–1474.

Majewski, J., Zawadzki, P., Pickerill, P., Cohan, F. M., and Dowson, C. G. (2000). Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation. *J. Bacteriol.* 182, 1016–1023. doi: 10.1128/JB.182.4.1016-1023.2000

Matic, I., Radman, M., Taddei, F., Picard, B., Doit, C., Bingen, E., et al. (1997). Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science* 277, 1833–1834. doi: 10.1126/science.277.5333.1833

Maynard Smith, J., Feil, E. J., and Smith, N. H. (2000). Population structure and evolutionary dynamics of pathogenic bacteria. *Bioessays* 22, 1115–1122. doi: 10.1002/1521-1878(200012)22:12<1115::AID-BIES9>3.0.CO;2-R

McNamara, C. J., Lemke, M. J., and Leff, L. G. (2002). Culturable and non-culturable fractions of bacterial populations in sediments of a South Carolina stream. *Hydrobiologia* 482, 151–159. doi: 10.1023/A:1021268516231

Mell, J. C., Shumilina, S., Hall, I. M., and Redfield, R. J. (2011). Transformation of natural genetic variation into *Haemophilus influenzae* genomes. *PLoS Pathog.* 7:e1002151. doi: 10.1371/journal.ppat.1002151

Mira, A., Ochman, H., and Moran, N. A. (2001). Deletion bias and the evolution of bacterial genomes. *Trends Genet.* 17, 589–596. doi: 10.1016/S0168-9525(01)02447-7

Mohr, K., and Tebbe, C. (2007). Field study results on the probability and risk of a horizontal gene transfer from transgenic herbicide-resistant oilseed rape pollen to gut bacteria of bees. *Appl. Microbiol. Biotechnol.* 75, 573–582. doi: 10.1007/s00253-007-0846-7

Monier, J. M., and Lindow, S. E. (2004). Frequency, size, localization of bacterial aggregates on bean leaf surfaces. *Appl. Environ. Microbiol.* 70, 346–355. doi: 10.1128/AEM.70.1.346-355.2004

Morris, C. E., Monier, J. M., and Jacques, M. A. (1997). Methods for observing microbial biofilms directly on leaf surfaces and recovering them for isolation of culturable microorganisms. *Appl. Environ. Microbiol.* 63, 1570–1576.

Nakamura, Y., Itoh, T., Matsuda, H., and Gojobori, T. (2004). Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nat. Genet.* 36, 760–766. doi: 10.1038/ng1381

Netherwood, T., Martín-Orúe, S. M., O'Donnell, A. G., Gockling, S., Graham, J., Mathers, J. C., et al. (2004). Assessing the survival of transgenic plant DNA in the human gastrointestinal tract. *Nat. Biotechnol.* 22, 204–209. doi: 10.1038/nbt934

Nielsen, K. M., Bones, A. M., and van Elsas, J. D. (1997). Induced natural transformation of *Acinetobacter calcoaceticus* in soil microcosms. *Appl. Environ. Microbiol.* 63, 3972–77.

Nielsen, K. M., van Elsas, J. D., and Smalla, K. (2000). Transformation of *Acinetobacter* sp. BD413(pFG4ΔnptII) with transgenic plant DNA in soil microcosms and effects of kanamycin on selection of transformants. *Appl. Environ. Microbiol.* 66, 1237–1242. doi: 10.1128/AEM.66.3.1237-1242.2000

Nielsen, K. M., and Townsend, J. P. (2004). Monitoring and modeling horizontal gene transfer. *Nat. Biotechnol.* 22, 1110–1114. doi: 10.1038/nbt1006

Nielsen, K. M., and Townsend, J. P. (2001). "Environmental exposure, horizontal transfer and selection of transgenes in bacterial populations," in *Enhancing Biocontrol Agents and Handling Risks*, Vol. 339, eds M. Vurro, J. Gressel, T. Butt, G. E. Harman, A. Pilgream, R. J. St. Leger, et al. (Amsterdam: NATO Science Series, IOS Press), 145–158.

Nielsen, K. M., Berdal, K. G., Kruse, H., Sundsfjord, A., Mikalsen, A., Yazdankhah, S., et al. (2005). *An Assessment of Potential Long-term Health Effects Caused by Antibiotic Resistance Marker Genes in Genetically Modified Organisms Based on Antibiotic Usage and Resistance Patterns in Norway*. Oslo: VKM-Report, Norwegian Scientific Committee for Food Safety.

Nielsen, K. M., Johnsen, P. J., Bensasson, D., and Daffonchio, D. (2007). Release and persistence of extracellular DNA in the open environment. *Environ. Biosafety Res.* 6, 37–53. doi: 10.1051/ebr:2007031

Nordgård, L., Nguyen, T., Midtvedt, T., Benno, Y., Traavik, T., and Nielsen, K. M. (2007). Lack of detectable uptake of DNA by bacterial gut isolates grown *in vitro* and by *Acinetobacter baylyi* colonizing rodents *in situ*. *Environ. Biosafety Res.* 6, 149–160. doi: 10.1051/ebr:2007029

Novozhilov, A. S., Karev, G. P., and Koonin, E. V. (2005). Mathematical modeling of evolution of horizontally transferred genes. *Mol. Biol. Evol.* 22, 1721–1732. doi: 10.1093/molbev/msi167

Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405, 299–304. doi: 10.1038/35012500

Orr, H. A. (2005). The genetic theory of adaptation. *Nat. Genet.* 6, 119–127. doi: 10.1038/nrg1523

Overballe-Petersen, S., Harms, K., Orlando, L. A. A., Moreno Mayar, J. V., Rasmussen, S., Dahl, T. W., et al. (2013). Natural transformation by degraded DNA allows for bacterial genetic exchange across geological time *Proc. Natl. Acad. Sci. U.S.A.* 110, 19860–19865. doi: 10.1073/pnas.1315278110

Pal, C., Papp, B., and Lercher, M. J. (2005). Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat. Genet.* 37, 1372–1375. doi: 10.1038/ng1686

Rayssiguier, C., Thaler, D. S., and Radman, M. (1989). The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature* 342, 396–401. doi: 10.1038/342396a0

Pettersen, A. K., Primicero, R., Bøhn, T., and Nielsen, K. M. (2005). Modeling suggests frequency estimates are not informative for predicting the long-term effect of horizontal gene transfer in bacteria. *Environ. Biosafety Res.* 4, 222–233. doi: 10.1051/ebr:2006008

Pietramellara, G., Ascher, J., Borgogni, F., Ceccherini, M.-T., Guerri, G., and Nannipieri, P. (2009). Extracellular DNA in soil and sediment: fate and ecological relevance. *Biol. Fertil. Soils.* 45, 219–235. doi: 10.1007/s00374-008-0345-8

Polz, M. F., Alm, E. J., and Hanage, W. P. (2013). Horizontal gene transfer and the evolution of bacterial and archaeal population structure. *Trends Genet.* 29, 170–175. doi: 10.1016/j.tig.2012.12.006

Ponciano, J. M., De Gelder, L., Top, E. M., and Joyce, P. (2005). The population biology of plasmids: a hidden Markov model approach. *Genetics* 176, 957–968. doi: 10.1534/genetics.106.061937

Pontiroli, A., Simonet, P., Frostegard, A., Vogel, T. M., and Monier, J. M. (2007). Fate of transgenic plant DNA in the environment. *Environ. Biosafety Res.* 6, 15–35. doi: 10.1051/ebr:2007037

Pontiroli, A., Rizzi, A., Simonet, P., Daffonchio, D., Vogel, T. M., and Monier, J. M. (2009). Visual evidence of horizontal gene transfer between plant and bacteria in the phytosphere of transplastomic tobacco. *Appl. Environ. Microbiol.* 75, 3314–3322. doi: 10.1128/AEM.02632-08

Popa, O., and Dagan, T. (2011). Trends and barriers to lateral gene transfer in prokaryotes. *Curr. Opin. Microbiol.* 14, 1–9. doi: 10.1016/j.mib.2011.07.027

Posada, D., Crandall, K. A., and Holmes, E. C. (2002). Recombination in evolutionary genomics. *Annu. Rev. Genet.* 36, 75–97. doi: 10.1146/annurev.genet.36.040202.111115

Powell, E. O. (1956). Growth rate and generation time of bacteria, with special reference to continous culture. *J. Gen. Microbiol.* 15, 492–511. doi: 10.1099/00221287-15-3-492

Ray, J. L., Harms, K., Wikmark, O. G., Starikova, I., Johnsen, P., and Nielsen, K. M. (2009). Sexual isolation in *Acinetobacter baylyi* is locus-specific and varies 10,000-fold over the genome. *Genetics* 182, 1165–1181. doi: 10.1534/genetics.109.103127

Remold, S. K., and Lenski, R. E. (2004). Pervasive joint influence of epistasis and plasticity on mutational effects in *Escherichia coli*. *Nat. Genet.* 36, 423–426. doi: 10.1038/ng1324

Reno, M. L., Held, N. L., Fields, C. J., Burke, P. V., and Withaker, R. J. (2009). Biogeography of the *Sulfolobus islandicus* pan-genome. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8605–8610. doi: 10.1073/pnas.0808945106

Retchless, A., and Lawrence, J. (2007). Temporal fragmentation of speciation in bacteria. *Science* 317, 1093–1096. doi: 10.1126/science.1144876

Rizzi, A., Raddadi, N., Sorlini, C., Nordgård, L., Nielsen, K. M., and Daffonchio, D. (2012). The stability and degradation of dietary DNA in the gastrointestinal tract of mammals - implications for horizontal gene transfer and the biosafety of GMOs. *Crit. Rev. Food Sci. Nutr.* 52, 142–161. doi: 10.1080/10408398.2010.499480

Rouzine, I. M., Rodrigo, A., and Coffin, J. M. (2001). Transition between stochastic evolution and deterministic evolution in the presence of selection: general theory and application to virology. *Microbiol. Mol. Biol. Rev.* 65, 151–185. doi: 10.1128/MMBR.65.1.151-185.2001

Rozen, D. E., de Visser, J. A., and Gerrish, P. J. (2002). Fitness effects of fixed beneficial mutations in microbial populations. *Curr. Biol.* 12, 1040–1045. doi: 10.1016/S0960-9822(02)00896-5

Sauerbier, J., Maurer, P., Rieger, M., and Hakenbeck, R. (2012). *Streptococcus pneumoniae* R6 interspecies transformation: genetic analysis of penicillin resistance determinants and genome-wide recombination events. *Mol. Microbiol.* 86, 692–706. doi: 10.1111/mmi.12009

Seitz, P., and Blokesch, M. (2013). Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. *FEMS Microbiol. Rev.* 37, 336–363. doi: 10.1111/j.1574-6976.2012.00353.x

Sorek, R., Zhu, Y., Creevey, C. J., Francino, M. P., Bork, P., and Rubin, E. M. (2007). Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* 318, 1449–1452. doi: 10.1126/science.1147112

Sousa, V., and Hey, J. (2013). Understanding the origin of species with genome-scale data: modeling gene flow. *Nat. Rev. Genet.* 14, 404–414. doi: 10.1038/nrg3446

Spratt, B. G., Hanage, W. P., and Feil, E. J. (2001). The relative contribution of recombination and point mutation to the diversification of bacterial clones. *Curr. Opin. Microbiol.* 4, 602–606. doi: 10.1016/S1369-5274(00)00257-5

Starikova, I., Harms, K., Lunde, T. T. M., Haugen, P., Primicerio, R., Samuelsen, Ø., et al. (2012). A fitness trade-off between gene cassette capture and stability of integrons. *PLoS Pathog.* 8:e1003043. doi: 10.1371/journal.ppat.1003043

Thomas, C. M., and Nielsen, K. M. (2005). Mechanisms and barriers to horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* 3, 711–721. doi: 10.1038/nrmicro1234

Townsend, J. P., Nielsen, K. M., Fisher, D. S., and Hartl, D. L. (2003). Horizontal acquisition of divergent chromosomal DNA: consequences of mutator phenotypes. *Genetics* 164, 13–21. Available online at: http://www.genetics.org/content/164/1/13.full.pdf+html

Townsend, J. P., Bøhn, T., and Nielsen, K. M. (2012). Probability of detecting horizontal gene transfer in bacterial populations. *Front. Microbiol.* 3:27. doi: 10.3389/fmicb.2012.00027

Tuller, T., Girshovvich, Y., Sella, Y., Kreimer, A., Freilich, S., Kupiec, M., et al. (2011). Association between translation efficiency and horizontal gene transfer within microbial communities. *Nucl. Acids Res.* 39, 4743–4755. doi: 10.1093/nar/gkr054

Via, S., and Lande, R. (1985). Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* 39, 505–522. doi: 10.2307/2408649

Via, S., and Lande, R. (1987). Evolution of genetic variability in a spatially heterogeneous environment: effects of genotype–environment interaction. *Genet. Res.* 49, 147–156. doi: 10.1017/S001667230002694X

Vulic, M., Dionisio, F., Taddei, F., and Radman, M. (1997). Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 94, 9763–9767. doi: 10.1073/pnas.94.18.9763

Walters, M. S., Lane, C., Vigil, P. D., Smith, S. N., Walk, S. T., and Mobley, H. L. T. (2012). Kinetics of uropathogenic *Escherichia coli* metapopulation movement during urinary tract infection. *MBio* 3, e00303–e00311. doi: 10.1128/mBio.00303-11

Welch, R. A., Burland, V., Plunkett, G. 3rd., Redford, P., Roesch, P., Rasko, D., et al. (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *E. coli. Proc. Natl. Acad. Sci. U.S.A.* 99, 17020–17024. doi: 10.1073/pnas.252529799

Whitaker, R., and Banfield, J. (2006). Population genomics in natural microbial communities. *Trends Ecol. Evol.* 21, 508–516 doi: 10.1016/j.tree.2006.07.001

Wögerbauer, M. (2007). "Risk assessment of antibiotic resistance marker genes in genetically modified organisms," in *BMGFJ – Forschungsberichte der Sektion IV, Band 5/2007* (Vienna).

Zawadzki, P., Roberts, M. S., and Cohan, F. M. (1995). The log-linear relationship between sexual isolation and sequence divergence in Bacillus transformation in robust. *Genetics* 140, 917–932.

Zoetendal, E. G., Cheng, B., Koike, S., and Mackie, R. I. (2004). Molecular microbial ecology of the gastrointestinal tract: from phylogeny to function. *Curr. Issues Intest. Microbiol.* 5, 31–48.

Zur Wiesch, P. A., Kouyos, R., Engelstadter, J., Regoes, R. R., and Bonhoeffer, S. (2011). Population biological principles of drug-resistance evolution in infectious diseases. *Lancet Inf. Dis.* 11, 236–247. doi: 10.1016/S1473-3099(10)70264-4