Many bacteria are highly sexual, but the reasons for their promiscuity remain obscure. Did bacterial sex evolve to maximize diversity and facilitate adaptation in a changing world, or does it instead help to retain the bacterial functions that work right now? In other words, is bacterial sex innovative or conservative? Our aim in this review is to integrate experimental, bioinformatic and theoretical studies to critically evaluate these alternatives, with a main focus on natural genetic transformation, the bacterial equivalent of eukaryotic sexual reproduction. First, we provide a general overview of several hypotheses that have been put forward to explain the evolution of transformation. Next, we synthesize a large body of evidence highlighting the numerous passive and active barriers to transformation that have evolved to protect bacteria from foreign DNA, thereby increasing the likelihood that transformation takes place among clonemates. Our critical review of the existing literature provides support for the view that bacterial transformation is maintained as a means of genomic conservation that provides direct benefits to both individual bacterial cells and to transformable bacterial populations. We examine the generality of this view across bacteria and contrast this explanation with the different evolutionary roles proposed to maintain sex in eukaryotes.

This article is part of the themed issue 'Weird sex: the underappreciated diversity of sexual reproduction'.

1. Introduction

Bacteria have long been appreciated as genomic shape-shifters that gain and lose genes with great regularity [1]. This fluidity is intuitively encapsulated by the idea of the core genome, or the fraction of genes shared by all or most strains of a given species. By this measure only around 50% of the genomic content of many bacterial strains is shared by other strains of the same species, while the rest, the accessory genome, is either unique to a particular genome or shared sporadically across the species [2]. In other estimates, and depending on the species, up to 25% of the genome is the result of horizontal gene transfer (HGT) [3–5]. Although there are wide margins of error on these estimates, owing to difficulties of discerning recent from distant gene-transfer events or HGT occurring within or across species, it is clear that HGT plays a key role in shaping bacterial genomes. Equally, the genes subject to HGT seem to have an outsized role on bacterial ecology and evolution, influencing where bacteria are found, what they can consume or degrade, their susceptibility to antibiotics and their virulence as pathogens, among others [3].
Because of these conspicuous benefits, it is easy to draw the conclusion that HGT is uniformly positive, and that bacteria engage in promiscuous sex in order to enhance their adaptability, much like meiotic sex enhances eukaryotic adaptation. However, this conclusion is too narrow and potentially misguided. First, it is important to bear in mind that there is a perception bias with respect to the benefits of HGT because the only instances of HGT that are observed in bacterial genomes are those that have passed the filter of natural selection [6]. Second, just like in eukaryotes, recombination in prokaryotes is associated with several potential costs that limit when, where and how recombination can occur [7,8]. Finally, while there are superficial similarities between eukaryotic sex and recombination in bacteria, the mechanisms underlying these processes are vastly diverged, as are their potential costs and benefits [6]. Thus, although it is tempting to look to eukaryotes to help understand the benefits of sex in prokaryotes, this should be done with caution. The regulation and mechanisms underlying bacterial sex appear to have evolved independently across prokaryotic groups, although some of the core genes associated with recombination are broadly conserved. Also, the frequency of recombination can be highly variable even within a species [6,9,10]. In spite of this, we believe a unifying benefit to bacterial sex can still be found. Here, we argue that the benefits of bacterial sex, and transformation in particular, lie in its genomic conservatism and not in the opportunities transformation can provide for genomic innovation. So rather than the ‘weird sex’ observed in some of the other species considered in this special issue, we argue that bacteria use ‘safe sex’ to repair DNA or purge deleterious mutations, to overcome stress, and to combat genomic parasites.

Bacteria utilize three primary mechanisms to acquire exogenous DNA as a substrate for recombination: conjugation, which is plasmid mediated; transduction, which is caused by bacterial viruses called phages; and transformation, the regulated uptake and incorporation via homologous recombination (HR) of exogenous DNA. Although all three of these mechanisms of recombination contribute to HGT, only natural transformation is exclusively encoded by genes present on the bacterial chromosome. For that reason, transformation is the only process that may have evolved as a form of bacterial sex [6,11]. Accordingly, and as opposed to genes for conjugation and transduction that are predominantly carried by accessory infectious elements, the costs and benefits of transformation are borne solely and directly by the competent bacteria themselves. In addition to the three classic mechanisms of DNA exchange, it has become clear in recent years that DNA may also be transferred between bacteria by other mechanisms, including nanotubes [12], micro-vesicles [13] or gene-transfer agents [14]. However, the prevalence and impact of these agents still remains to be established and we therefore focus on transformation for the remainder of this review.

We first discuss the diverse evolutionary costs and benefits of natural competence. Next, we outline the numerous strategies bacteria use to increase the likelihood that transformed DNA is derived from within the same species. Finally, we consider experimental and theoretical support for the idea of conservative sex and conclude with suggestions for further study. In addition to the discussion below, we also refer interested readers to several excellent reviews that provide more mechanistic or species-specific details of competence induction and transformation [15–17].

## 2. Costs and benefits of natural transformation

### (a) Physiological costs of transformation

Natural transformation is coordinated by a large and complex molecular machinery dedicated to the uptake of DNA from the environment, its intracellular processing and, potentially, its genetic incorporation through recombination. Although there are similarities in the mechanisms of DNA binding, uptake and incorporation across species, distinct patterns of competence regulation together with the sporadic distribution of competence across bacteria imply that transformation has had multiple independent origins [18,19]. Equally, the rates of natural transformation can vary markedly within a single species [9,10], suggesting that transformation is evolutionarily labile and that bacteria face a trade-off between its costs and benefits. Potential costs are numerous and diverse. The very act of transformation is energetically costly, sometimes involving the transcription of more than 100 genes, only a fraction of which are required for recombination [18]. Moreover, DNA binding, uptake and incorporation may require a ‘handling’ time that reduces rates of vegetative growth; indeed, in some species competent cells temporarily arrest growth, thereby reducing fitness when in competition with non-competent cells [20]. This cost of entering this persister-like state was proposed to explain why only a fraction of cells in *Bacillus subtilis* become competent [21]. On the other hand, under conditions where rapidly dividing bacteria experience high levels of mortality (e.g. because of antibiotics that target cell wall biosynthesis), the persister-state induced by competence may actually be favoured and thus indirectly select for the maintenance of competence [20]. Another cost arises through recombination itself because recombination is by default a DNA damaging process; evidence suggests that chromosomal integration of DNA that is successfully taken up by the cell can ultimately lead to double-stranded breaks that are lethal unless repaired [22,23]. Finally, there are the potentially considerable costs associated with competence-induced cell lysis in, for example, *Streptococcus pneumoniae*, where competent cells actively lyse non-competent members of the same population [24]. Although this last cost is undoubtedly harmful to non-competent cells, it remains possible that competent ‘killers’ benefit by liberating DNA from competing cells, which then serves as a substrate for transformation.

### (b) DNA as food

DNA is metabolically costly and one of the simplest and most intuitive benefits for transformation lies in the possibility that bacteria could use DNA as a resource in the form of nucleotides and nucleotide precursors [25]. Were cells induced to become competent by starvation, DNA could in principle provide sufficient energy for continued replication or repair. Consistent with this possibility, some species, like *Haemophilus influenzae*, require nutritional down-shifts to induce competence and purine depletion activates the competence activator *sxy* [26]. However, several factors argue against this hypothesis as a general explanation for the maintenance of transformation: (i) as yet, there is no clear evidence that the integration of nucleotides taken up by transformation become routed into DNA metabolism, (ii) the presence of exogenous DNA does not appear to induce competence in any transformable species, (iii) competence in streptococci, like *S. pneumoniae*, is induced for only a short time period during exponential growth when...
other resources are highly abundant [15], (iv) transported DNA is heavily protected against nuclease digestion within the cell, potentially enabling transported fragments to remain intact as a substrate for recombination [27], and (v) the hypothesis does not explain why several species that become competent for natural transformation only take up DNA from close relatives due to conserved DNA uptake sequences (DUS) despite the fact that non-homologous DNA could be used as a source of nucleotides for direct use or degradation [19,28]. In addition to these concerns, it remains uncertain, on energetic grounds, whether the costs of DNA transport are sufficiently offset by any metabolic savings provided by exogenous DNA [11]. Thus, despite the intuitive appeal of this idea, the evidence in its favour is currently limited.

(c) DNA repair

An old idea for a potential benefit of transformation is that acquired DNA is used as a substrate for genome repair [29,30]. Early experimental evidence indicated an immediate benefit of DNA uptake on transformant survival relative to the remainder of the population in B. subtilis [30–32]. However, these earlier results were countered in the same species by evidence that genotoxic stress did not induce competence [33], as predicted by the original idea. A more recent extension of this ‘DNA repair hypothesis’ proposes that transformation is a general stress response [34]. This idea is supported by the fact that some, but not all [35], naturally transformable human pathogens such as S. pneumoniae, Legionella pneumophila and Helicobacter pylori lack a general SOS response and that DNA damaging agents, including some antibiotics, induce competence in these species [34,36,37]. Current evidence suggests, however, that the benefits of competence induction in these species may be unlinked from the effects of transformation per se (i.e. DNA uptake and integration) [38,39]. Thus, although these reports reveal that at least some forms of stress can induce competence, they do not always provide evidence that this response is adaptive, nor provide clear indications of the mechanisms underlying these benefits. Furthermore, they do not take into account the abundant evidence showing that competence in many species is induced by quorum sensing, irrespective of exogenous stress, and that other unambiguous forms of stress, like temperature and pH, can even repress natural transformation [40–43]. Taken together, the available data do not clearly support the classical ‘DNA for repair’ hypothesis; instead, they are more consistent with the more general hypothesis that competence development in itself is beneficial, albeit under restricted conditions.

(d) Homologous recombination

Natural transformation has the potential to shuffle alleles at different loci within a population. This gives rise to a suite of potential benefits and costs of transformation associated with genetic recombination that are largely identical to those studied for meiotic sex in eukaryotes. Below, we briefly review two of the principles—epistasis and Hill–Robertson interference—through which transformation may be favoured, in both cases through either helping to purge deleterious or fix beneficial mutations. For more in-depth reviews of this large field, we refer to Otto [44], Hartfield & Keightley [45] and, in the context of bacteria, Vos [46].

Selection with epistasis, i.e. non-independent fitness effects of alleles at different loci, produces non-independent gene associations within a population (linkage disequilibrium). At the most basic level, recombination can have a detrimental effect as it breaks up co-adapted gene complexes (e.g. [47]). Nevertheless, if deleterious mutations that interact with negative epistasis continually arise within a population, recombination can be favoured because it can increase genetic variance and therefore, the efficacy by which natural selection purges deleterious mutations (the ‘deterministic mutation hypothesis’) [48]. This was shown theoretically to also provide a benefit to natural transformation in bacterial populations [49,50]. However, empirical work has shown that while negative epistasis sometimes exists, it is far from pervasive and there are many systems in which positive or no epistasis was reported (reviewed in [51,52]). For this reason, the deterministic mutation hypothesis has been broadly disregarded as a main contender to explain the ubiquity of sex, including natural transformation.

A second class of explanations for why recombination can be beneficial relies on the Hill–Robertson effect [53,54]. Here, an interaction between natural selection and stochastic effects in finite populations (through random genetic drift or mutation) produces genetic associations (negative linkage disequilibria) that reduce genetic variance for fitness. By breaking up these associations, recombination can increase the efficacy of natural selection and genes that increase the recombination rate can be indirectly selected for. Advantages of recombination stemming from the Hill–Robertson effect come in different forms, and may involve both beneficial and deleterious mutations (e.g. [55,56]). Extreme manifestations are the Fisher–Muller model [57,58], in which recombination brings together beneficial mutations that would in asexual populations compete with each other (‘clonal interference’), and Muller’s ratchet [59], the perpetual loss of mutation-free individuals in asexual populations subject to deleterious mutations. A number of mathematical and simulation models have been developed to specifically investigate variants of the Hill–Robertson effect in the context of bacterial sex, confirming that natural transformation can be favoured both in populations subject to recurring deleterious mutations [21,60] and in adapting bacterial populations [21,61,62].

Experimental work testing these ideas has focused on the Fisher–Muller model, the central prediction of which is that recombination accelerates the adaptation rate of sexual relative to asexual variants [57,58]. Evolution experiments with the Gram-negative bacterium Escherichia coli [63] undergoing plasmid-mediated recombination and the yeast Saccharomyces cerevisiae [64] support this hypothesis, while both reports strongly suggest the crucial role of recombination in relieving the effects of clonal interference. Reduced clonal interference has also been proposed as a key genetic mechanism underlying the evolutionary maintenance of natural transformation [46]. For example, transformable H. pylori adapted more rapidly than non-transformable, otherwise isogenic, populations [65]. By contrast, a study using the highly transformable species Acinetobacter baylyi failed to detect any consistent evolutionary advantage of natural transformation during 1000 generations of experimental evolution [66]. It was later demonstrated that the benefits of natural transformation in A. baylyi are growth-phase dependent, whereby a positive effect was demonstrated during active growth/early stationary phase which was offset by reduced adaptation to the experimental conditions during
Stationary/death phase [67]. Similar context-dependent benefits of transformation were shown in the human pathogen *S. pneumoniae*. Here, competence for natural transformation was disadvantageous when populations evolved in benign experimental conditions, but not when they evolved in the presence of periodic mild stress (sub-inhibitory concentrations of kanamycin) [68]. In addition, evolving competent *S. pneumoniae* populations fixed significantly fewer mutations than isogenic non-competent lineages, while non-competent lineages were more likely to evolve mutator genotypes [68]. Taken together, the few experimental tests of Fisher–Müller advantages to bacterial transformation provide a mixed picture: while some studies support the idea that natural transformation can accelerate adaptation, other results highlight that these responses are either context dependent or absent altogether (at least under the conditions examined).

(e) Biased allelic replacement

Because of the symmetry of meiotic crossovers, the central influence of eukaryotic recombination at the population level is to reduce linkage disequilibria. By contrast, natural transformation can also affect allele frequencies within a population because of its inherent asymmetry: genetic material is taken from a donor gene pool of free DNA molecules in the environment (‘eDNA’) and replaces material within a recipient gene pool of living cells, and those two gene pools are not necessarily identical in terms of allelic composition. This was first recognized by Redfield [49], who showed with a mathematical model that if bacteria carrying deleterious alleles are more likely to die and thereby release their DNA into the environment than wild-type bacteria, this creates a bias towards taking up deleterious alleles and thus a distinctive cost of natural transformation that is independent from any effects derived from gene shuffling. More recently, models of well-mixed and spatially structured populations were developed that explicitly incorporated a pool of free eDNA subject to decay [62,69]. These models recovered a similar detrimental effect of transformation in populations adapting to new environmental conditions because the eDNA may build up an over-representation of old, non-beneficial alleles. Interestingly, however, taking up and incorporating ‘old alleles’ may also turn out to be beneficial in situations where environmental conditions change frequently. According to this idea of ‘genetic time travel’, natural transformation may allow the bacteria to make use of a reservoir of genetic material preserved in the form of eDNA [70].

(f) Defence against genomic parasites

Recently, yet another hypothesis for the raison d’être of natural transformation has been proposed [71], according to which transformation helps bacteria to cure their genomes of costly integrated mobile genetic elements (MGEs) such as phages or conjugative elements. This hypothesis is based on the observation that successful transformation only requires homology between short stretches of DNA at the ends of the transformed fragment and the host chromosome, while the regions in the middle can freely vary [72,73]. As a consequence, transformation can lead to either the incorporation of new genetic material (as is usually emphasized) or its expulsion, if, for example the transforming DNA lacks a MGE that is present in the transformed recipient. But under what conditions would this work? Croucher et al. [71] motivate their hypothesis with the observation that in both *S. pneumoniae* and *H. influenzae*, the length of newly incorporated DNA fragments approximately follows a geometric distribution [74,75]. They argue that as a consequence of this bias transformed cells will be more likely to incorporate short rather than long DNA stretches, in turn preferentially leading to the loss of MGE. However, generalizing their idea, we suspect that such a bias may not be strictly necessary. Although transformation with long DNA fragments can lead to MGE acquisition, it can also—and potentially at the same rate—lead to their loss if transformed fragments cover the insertion site of the MGE. Conversely, transformation with DNA fragments shorter than the MGE can only ever lead to MGE loss. Thus, as long as there is any natural transformation with DNA fragments that are shorter than the MGE in question, there should be an automatic bias towards exclusion of the element. The existence of this bias should then be independent of the size distribution of incorporated DNA, even though its magnitude may still be affected by the size distribution. This more general version of the defence hypothesis remains to be specifically tested.

In addition to corroborating their hypothesis by means of a mathematical model, Croucher et al. [71] also showed through whole-genome analyses that transformable pneumococci harboured fewer prophages than non-transformable ones and that phages often insert into genes involved in the DNA uptake machinery, thus disrupting the putative cellular defence mechanism against MGEs. In a small preliminary analysis for this review, we screened the whole genomes of 10 strains of *S. pneumoniae* for insertion sequences, transposons and phages (putative and confirmed) [76] from which we have precisely quantified transformation rates [10]; while five of the strains are unable to become competent, the other five show variable rates of transformation across several orders of magnitude. Consistent with the predictions of the Croucher et al. model, we observed clear but non-significant negative correlations between transformation rate and the numbers of putative and confirmed phage. Although no strong conclusions can be drawn from these limited tests, we believe this approach is likely to be very powerful for further tests of these ideas, as it can tie experimentally validated differences in transformation rates to the number of mobile elements carried in a given genome.

3. The potential for consensus

Despite the numerous studies of competence across a broad swath of competent bacterial species, there still remains considerable uncertainty about its function because no single explanation seems to capture the particulars of different species growing in different contexts. And this may in fact be the answer: as with eukaryotic sex, a plurality of explanations must be considered [77]. While *Haemophilus* may become competent when it is hungry or to repair DNA damage, streptococci, *H. pylori* [36] and *L. pneumophila* [37,78] may only do so when they are damaged, while *A. baylyi* [42] and *Neisseria meningitidis* [79] are constitutively competent (although not necessarily all cells in the population at the same time). However, a consensus option may lie in the fact that not all transformable DNA is of equal appeal or value to each bacterial species. Most obviously, non-homologous DNA is a poorer substrate for recombination than homologous DNA uptake machinery, thus disrupting the putative cellular defence mechanism against MGEs. In a small preliminary analysis for this review, we screened the whole genomes of 10 strains of *S. pneumoniae* for insertion sequences, transposons and phages (putative and confirmed) [76] from which we have precisely quantified transformation rates [10]; while five of the strains are unable to become competent, the other five show variable rates of transformation across several orders of magnitude. Consistent with the predictions of the Croucher et al. model, we observed clear but non-significant negative correlations between transformation rate and the numbers of putative and confirmed phage. Although no strong conclusions can be drawn from these limited tests, we believe this approach is likely to be very powerful for further tests of these ideas, as it can tie experimentally validated differences in transformation rates to the number of mobile elements carried in a given genome.
DNA. But it goes further than this. As we detail below, all competent bacterial species go to extreme lengths to ensure that the DNA they take up or recombine is either from the same species or from the same strain. This source of DNA is a safe and conservative way to either repair DNA or purge deleterious mutations (including mutator alleles [68]) by restoring the wild-type allele, to re-establish functions from related strains that have already been tested in a highly similar genetic background or to purge genomes of potentially harmful parasitic elements. This conservative explanation builds on the multi-layered barriers in place to reduce outcrossing, and offers a general and testable hypothesis for the role of transformation in bacteria.

4. Barriers to transformation outside and within the cell

Barriers to transformation are mediated by passive and active processes that serve to dramatically increase the likelihood that competent cells will be transformed with DNA from clones. While passive barriers offer approximate routes to restrict transformation, they are imprecise. By contrast, active restrictions to DNA recognition, uptake and recombination are significantly more limiting, and together these redundant processes strongly bias which DNA becomes available for transformation, relative to the broader pool of eDNA.

(a) Barriers acting outside the cell

Ecological barriers to recombination require no special mechanisms of exclusion and arise as a simple consequence of the fact that persistence of eDNA is limited and because bacterial growth is spatially structured. For these passive reasons, HGT will tend to take place between cells that are in close proximity to one another and thus often clonal. At a broad scale, this leads to networks of gene exchange that are ecologically confined. For example, by examining routes of HGT via a network approach, Popa & Dagan [7] estimated that 74% of identified HGT events occurred among bacteria residing in the same habitat. Similarly, Smillie et al. [80] found that bacteria in the human microbiome residing in the same body sites were more likely to exchange genes than those from different sites. Although compelling, these examples deal with HGT broadly and may only be partly due to transformation. However, even at this finer scale, proximity, together with other processes examined below, is likely to increase the probability of clonal transfer. *Vibrio cholerae*, for example, becomes competent in the presence of chitin, a substance these bacteria encounter during growth on the surface of zooplankton [81]. Adding to this, *Vibrio* competence is induced by a quorum-sensing system that is activated by high concentrations of a secreted chemical signal that serves as a proxy measure for high densities of cells of the same species that share this signal [82].

Several other naturally competent species rely on similar modes of species or clone-level recognition to regulate induction. The Gram-positive species *B. subtilis* and *S. pneumoniae* become competent following detection of a threshold concentration of a secreted peptide signal that is species specific [16,34]. Because the signal is produced constitutively, its environmental concentration approximately scales with the density of producing cells [15,83]. While quorum sensing is generally sufficient to ensure induction is coordinated among cells of the same species, it would not allow finer coordination within species. A possible solution to this problem lies in the fact that *B. subtilis* and *S. pneumoniae* are polymorphic for the peptide signals they produce (called pherotypes) [84,85], and there is no apparent cross-talk between a given peptide and non-cognate receptors [86]. In *S. pneumoniae*, several groups have argued that this type of polymorphism could permit transformation to occur preferentially among cells expressing the same phenotype—much like bacterial mating types, although current results on this issue are mixed [84,87].

By using ecological proximity and quorum-sensing systems to estimate cell density, bacteria can regulate competence so that it is only turned on when cells are surrounded by individuals of the same species. More strikingly, some naturally competent bacterial species can also strongly influence which DNA becomes available as substrate for transformation by coordinately regulating competence induction with diverse modes of cell killing. Some streptococci, for example, use a process dubbed ‘fratricide’ whereby a sub-population of competent cells kill and lyse non-competent members of the same population [24,88]. In *S. pneumoniae*, competence coincides with the production of both hydrolytic enzymes that degrade the cell wall and also the secretion of multiple classes of bacteriocins, including the highly diverse blp (bacteriocin-like peptide) gene-cluster [89–92]. Importantly, killer cells are protected from the action of these secreted weapons, which leads to a one-way routing of DNA from target cells to transformed recipients. Another recent discovery confirms this coincident expression of competence and killing, but via an entirely distinct mechanism. Many Gram-negative bacteria express a needle-like machine called a Type VI secretion system (T6SS) that injects toxins into non-immune target cells, resulting in their death and lysis [93]. Remarkably, *V. cholerae* co-regulates competence induction with T6SS-mediated predation after which the liberated DNA is immediately available for uptake and incorporation [94]. As with fratricide, this form of predation is not a fail-safe means by which cells can acquire clonal DNA; however, together with ecological proximity and quorum sensing-dependent regulation of competence, clonal uptake is likely to be the most common outcome.

While external barriers to transformation may be reliable, they are also susceptible to failure in multi-species bacterial communities where eDNA will inevitably be present due to the continuous lysis of dead cells. To offset these risks, bacteria have evolved a suite of redundant mechanisms that can limit transformation by excluding cells on the basis of specific sequence tags, called uptake sequences, or by restricting incorporation or expression following DNA update.

(b) Sequence tags restricting DNA uptake

Several species in two phylogenetically distinct bacterial families, the Pasteurellaceae and the Neisseriaceae, have evolved sequence-specific uptake of DNA [19,28,95]. In short, these sequences act like small barcodes that identify that the DNA a competent cell binds comes from the correct species. Discrimination between ‘self’ and ‘other’ DNA occurs at the bacterial surface. In the Neisseriaceae, the small competence protein ComP binds a small signature in DNA and initiates uptake [96]. A similar factor has not yet been identified in the Pasteurellaceae and the exact molecular interactions responsible for sequence specificity await characterization in this
family. Nonetheless, the genomes of these competent bacteria carry hundreds to several thousands of these small specific signatures in DNA termed DUS in the Neisseriaceae [97] and uptake signal sequences (USS) in the Pasteurellaceae [28]. Typically, DUS extend for 12 nucleotides with the consensus sequence 5'-ATGCCGCTGGAA-3' and USS for 9 nt as 5'-AAGTGCGGT-3', with a defined core of 3–4 essential nucleotides, although different dialects of both DUS and USS exist in different species [98,99]. USS and DUS are not homologous, suggesting that the uptake specificities in the Neisseriaceae and the Pasteurellaceae are the results of convergent evolution. The DUS/USS are spread throughout their respective chromosomes so that most parts of the genomes are probably substrates for transformation. There is a close concordance between DUS distribution and conversion fragments identified from whole-genome alignments [100]. In some species, the number of DUS is so high that they occupy more than 3% of the entire genome, reflecting the substantial genetic investment in maintaining specific transformation [98]. How high is the DUS/USS barrier against non-DUS/USS DNA? The dependency for DUS in transformation has been shown to vary from absolute to less pronounced between strains [95]. The location of DUS relative to homologous and heterologous stretches of DNA, together with DUS/USS integrity and strandedness also influence transformation efficacy [95,98,101–103]. However, DUS/USS always biases transformation to involve homologous DNA. DUS/USS are remarkably conserved and single deviations from the signal are rare in their respective genomes and experimentally reduce or eliminate transformability [98,99]. Interestingly, DUS are clustered in the core genome [100] and both DUS and USS are biased towards genes associated with genome maintenance [104]. Because these important and often essential genes of the core genome carry evolved and costly signatures for transformation, often inside coding regions, an association between their conserved status and transformation may therefore exist. By contrast, DUS are underrepresented in genes encoding surface-exposed epitopes that typically are hypervariable and in the accessory genome. DUS and transformation in these species are therefore not apparently associated with variability (e.g. for immune evasion) and support a conservative rationale for transformation.

(c) Restriction modification systems

Once in the cytoplasm, and in species that do not rely on DUS/USS for discrimination, DNA faces another threat to its integrity: restriction modification systems (RMSs). Typically, RMSs are constituted by a restriction endonuclease that cuts double-stranded DNA (dsDNA) at non-methylated specific recognition sites and a corresponding methylase that can protect the same recognition site. RMSs are nearly ubiquitous in bacteria [105] where they raise barriers against the stable integration of alien DNA from transduction, conjugation and transformation [8]. More than 5000 RMSs are biochemically or genetically characterized and even more are added to the comprehensive RMS database REBASE daily [106]. At which exact stage during transformation the restriction endonuclease cuts DNA has been the matter of some debate because the endonucleolytic activity is clearly directed towards dsDNA, whereas it is ssDNA that enters the cytoplasm according to current models for transformation [16]. A post-replication and hence post-recombination model has been proposed where transformed chromosomes are subject to endonucleolytic attack once double-strandedness has been restored by chromosomal replication [73]. Using commercially available methylases, it has become possible to specifically study the extent of individual barriers to transformation due to RMSs in many different bacteria such as Pseudomonas putida (e.g. [107]). Specific knock-outs of restriction endonucleases or specific pre-methylation of transforming DNA has also revealed several further species that can become competent for transformation, such as the human pathogen Staphylococcus aureus [108,109] and the thermophilic cellulytic Caldusiruptor bescii [110]. In this latter organism, the restriction barrier is apparently absolute when using transforming homologous DNA propagated in another species (E. coli). Many other bacteria display very high and some absolute RMS barriers to transformation that depend on the DNA source, sequence and homology [101,111]. As discussed above, competence may be regulated and induced in some bacteria when conditions for transformation are favourable. About half of the sequenced S. pneumoniae strains harbour an unusual ssDNA methylase, DpnA, that protects incoming DNA from restriction at 5'-GATC-3' sites by DpnII during transformation. Notably, DpnA forms part of the competence regulon and is expressed from a competence-inducible promoter [27,73]. Streptococcus pneumoniae can thereby transiently lower the transformation barrier in periods when competence is induced.

By limiting stable integration of heterologous DNA and favouring homologous DNA of equal modification status, RMSs may therefore be important drivers of speciation in bacteria [112]. Genomic studies of the naturally competent N. meningitidis have indeed shown that divergent RMS profiles in different lineages of this species have contributed to the sexual isolation of phylogenetic clades [113]. In addition to RMSs, the bacterial immune system based on clustered, regularly interspaced, short palindromic repeats (CRISPRs) found in 40% of all bacteria [114] have been shown to prevent natural transformation in a streptococcal transformation/infection model [115]. A genomics study of the naturally competent Aggregatibacter actinomycetemcomitans showed that the loss of competence was followed by the loss of CRISPRs in an evolutionary time frame [116] linking the two phenomena. However, the evolutionary consequence of CRISPRs inhibiting HGT remains a matter of debate [117].

(d)Mismatch repair

HR is the last step in transformation where incoming DNA is exchanged with a similar or identical allele in the chromosome. An HR event that involves two identical segments of DNA leaves no genetic signature thereby hiding it from retrospective detection. By contrast, HR involving similar but not identical alleles is traceable and has hence been more studied. This in turn may have fostered the common view that transformation evolved to innovate and not to conserve. HR is ubiquitous in nature and in E. coli more than 25 genes are involved in the process [118]. During transformation, incoming ssDNA is coated with single stranded binding protein (SSB) and sequentially the recombination mediator DprA [119] or RecO facilitates loading of the recombinase RecA that ultimately leads to strand exchange [120,121]. RecA allows the ssDNA to match up with a homologous segment in the chromosome with high fidelity and to initiate the allelic
exchange [122]. RecA can very rapidly identify its relatively small homologous target in a suspension where heterologous DNA targets are in a 200,000-fold excess [123]. Homology must not be complete for strand exchange to take place and variable degrees of RecA fidelity have been documented. However, a strong and persistent negative exponential correlation between sequence divergence and transformation has been shown in different competent bacteria and archaea [124–127]. Furthermore, the recombination potential between particular loci of divergent, but related, Acinetobacter species has been shown to be dependent on the composition of DNA sequences up to several kilobases away from the loci under consideration [128]. Also, mismatch repair (MMR) systems may control the tolerance for heterology in HR. In the competent B. subtilis, S. pneumoniae and P. stutzeri, MMR plays a lesser role in controlling mismatches than in the non- or less-competent species such as E. coli. In these competent species, RecA is the main discriminator against sequence heterology during HR [72,129].

(e) Protein H-NS associated gene silencing

Although the barriers described above are efficient, successful transfers of DNA between species and genera regularly occur and are important drivers of microbial evolution. DNA can be mobilized by means of plasmids, transposons and bacteriophages [130]. Regions in bacterial chromosomes of foreign and phylogenetically distant (xenogeneic) origin can be detected by searching for particular signatures such as unusual base composition (GC-content). The base composition in such regions ameliorate over time and ‘tune’ into the genomic signature of its new host as they experience the same mutational processes affecting all genes in the recipient genome [3]. Notably, both Gram-positive and Gram-negative bacteria have evolved repressors that target such non-ameliorated xenogeneic DNA [131,132]. In effect, these repressors silence the expression of genes that could cause harmful effects in the host organism. However, the longer term fate of such silenced regions is not well understood, and it could even be that such gene silencing mitigates costs of carrying novel genes thereby facilitating their longer term persistence. Alternatively, silenced regions may be more likely to acquire nonsense mutations, as there is little selection for the maintenance of these unexpressed genes.

(f) Fitness costs

In the event that passive and active barriers to transformation do not prevent transformation of non-homologous sequences, transformed cells may still fail to persist and spread. Focusing on HGT generally, Baltrus [133] outlined a series of cellular costs associated with recombination of foreign DNA. Among these, new sequences can cause toxicity due to protein misfolding, misregulation or inefficient translation or transcription due to different GC content (if not adequately silenced by H-NS) or misregulation. More simply, transformed DNA can potentially displace functional alleles that are highly coevolved with the host genome, a history that will not exist for diverged sequences. Although these diverse fitness costs are difficult to quantify, several studies have examined the issue by transforming recipient cells (not necessarily natural competent ones) with donor DNA of different size and origins. Knoppel et al. [134] found that of 90 such transfers, surprisingly few introduced significant costs while most were neutral. A clear caveat of this work, however, is that fitness was measured during short-term assays with a detection limit that may fail to identify costs that are relevant over longer periods. At the same time, over the longer term and like antibiotic resistance, fitness costs of transformation can potentially be compensated by second-site mutations or by amelioration of the transformed sequence itself [135].

5. Conclusion and outlook

Transformation, together with other modes of HGT, can undoubtedly have an enormous influence on bacterial genomes and ecology. It can facilitate invasion into novel habitats and permit escape from drug or immune pressure [3,46]. However, these benefits are probably more the exception than the rule. The vast majority of HGT events between different strains or species of bacteria are expected to be deleterious, and this has given rise to a complex suite of processes that limit transformation to close relatives with highly similar genomes and ecologies. Moreover, even when acting within largely clonal populations, there is currently no consistent evidence that natural transformation enables bacteria to become more fit or adapt more rapidly to their environment. Together, this leads us to suggest that transformation will tend to be a conservative mechanism, acting much like any other repair process in the cell. This expanded idea of repair can extend to specific forms of DNA damage, deleterious mutations, to more generic stress or to the removal of genomic parasites. However, even if this is the predominant role for transformation, it clearly does not preclude that transformed DNA will also occasionally provide important adaptive benefits at the individual and population levels. But this leads to important unresolved questions: (i) if transformation is beneficial because of its capacity to repair damage (defined most broadly), why is natural competence so sparsely distributed across bacteria; (ii) related to this, why are transformation rates so highly variable within naturally competent species; and (iii) what factors, if any, unite species that retain natural competence? The challenge for future work is to address these questions by integrating theoretical and bioinformatics approaches, as exemplified by the recent paper from Croucher et al. [71]. There is also further need to consider how the three modes of HGT differentially influence the dynamics of the core and accessory genomes of competent species. Finally, we encourage the development of additional long-term experimental studies that address the benefits of competence induction and transformation for stress resistance and genome maintenance at short and evolutionary timescales.

Authors’ contributions. All authors contributed equally to the structure and writing of this review and approved of the final submission. E.L.M. conducted bioinformatic analyses of pneumococcal phage.

Competing interests. We declare we have no competing interests.

Funding. Support for this work was provided by the NWO (Dutch Science Foundation) and the BBSRC (UK) to D.E.R. and E.L.M.; by the Australian Research Council through a Future Fellowship (FT140100907) to J.E. and by grants from the Norwegian Research Council (project number 204263) and the Northern Norway Regional Health Authority to P.J.


