

Review

Diversity-Generating Machines: Genetics of Bacterial Sugar-Coating

Rafał J. Mostowy^{1,*} and Kathryn E. Holt^{2,3}

Bacterial pathogens and commensals are surrounded by diverse surface polysaccharides which include capsules and lipopolysaccharides. These carbohydrates play a vital role in bacterial ecology and interactions with the environment. Here, we review recent rapid advancements in this field, which have improved our understanding of the roles, structures, and genetics of bacterial polysaccharide antigens. Genetic loci encoding the biosynthesis of these antigens may have evolved as bacterial diversity-generating machines, driven by selection from a variety of forces, including host immunity, bacteriophages, and cell–cell interactions. We argue that the high adaptive potential of polysaccharide antigens should be taken into account in the design of polysaccharide-targeting medical interventions like conjugate vaccines and phage-based therapies.

Diversity-Generating Machinery

The world of bacteria is extraordinarily diverse, and even the most superficial understanding of each species alone takes years of research. Therefore, even though most bacteria produce extracellular polysaccharides, the biology of polysaccharide production has more exceptions than rules. Nevertheless, a cross-sectional look at different bacterial species reveals many similar characteristics of carbohydrate antigens, and these similarities have important implications for their evolution. Here we compare the main aspects of polysaccharide antigen biology between different antibiotic-resistant global priority pathogens as defined by the World Health Organization in 2017 (www.who.int/mediacentre/news/releases/2017/bacteria-antibiotics-needed), particularly in relation to their functionality, genetics, and phenotypic variation.

Functions of Polysaccharide Antigens and Their Role in Disease

Bacteria produce several extracellular polysaccharides which are crucial for their ability to colonise and cause disease. Gram-negative bacteria produce a lipopolysaccharide (LPS), which is an important component of the outer cell membrane and often includes a highly variable O-antigen at the end. Gram-negatives can also produce an additional capsular polysaccharide (CPS), which forms a surface layer. Gram-positive bacteria can synthesise teichoic acids, and like Gram-negatives, can also produce CPS, but they do not produce LPS. In addition to these major antigens, both Gram-negative and Gram-positive bacteria can synthesise exopolysaccharides (EPS), which are released into the environment and are not attached to the cell.

All of these extracellular polysaccharides play a variety of important roles in bacterial lifestyle and pathogenesis. Polysaccharide capsules enable bacterial evasion of the host immune system by shielding bacteria from the complement system, antibodies, or engulfment by macrophages [14]. Consequently, bacterial capsules are widely recognised as important virulence determinants [15–20]. A brief summary of major polysaccharide antigens in WHO

Highlights

Recent advances in glycobiology emphasise the importance of surface carbohydrates in the evolution of bacterial commensals and pathogens.

Epidemiological and genomic analyses highlight the remarkable diversity and adaptability of bacterial surface polysaccharides, often driven by horizontal DNA transfer.

Functional heterogeneity of bacterial polysaccharide antigens points to multiple factors which have shaped their diversity.

Antigenic diversity is a challenge for polysaccharide conjugate vaccines which strive to eliminate some bacterial diseases, like the pneumococcal disease.

¹Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, UK

²Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Victoria, Australia

³The London School of Hygiene and Tropical Medicine, London, United Kingdom

*Correspondence: r.mostowy@imperial.ac.uk (R.J. Mostowy).

Table 1. Comparison of Polysaccharide Antigens among WHO Priority Pathogens in Need of New Antibiotics^a

Species	Class	Common habitat	Major PS antigens	Diversity	Biosynthesis	Refs
CRITICAL						
<i>Acinetobacter baumannii</i>	Gram–	Diverse environments	Capsule	38 serovars, 25 genetic clusters	wzy	[1]
<i>Pseudomonas aeruginosa</i>	Gram–	Diverse environments	O-antigens (2 types)	20 serogroups	wzy, ABC	[2,3]
<i>Klebsiella pneumoniae</i>	Gram–	Diverse environments	Capsule, O-antigen	80 K-antigens, 8 O-antigens	wzy, ABC	[4]
<i>Escherichia coli</i>	Gram–	Animals	Capsule, O-antigen	180 O-antigens, 80 K-antigens	wzy, ABC	[5,6]
HIGH						
<i>Enterococcus faecium</i>	Gram+	Animals	Teichoic acid, capsule	(Understudied)	Probably wzy	[7]
<i>Staphylococcus aureus</i>	Gram+	Animals	Capsule	11 serotypes	wzy	[8]
<i>Helicobacter pylori</i>	Gram–	Human stomach	O-antigen	At least eight types	wzk	[9]
<i>Campylobacter jejuni</i>	Gram–	Animals	Capsule, LOS	47 serotypes, LOS phase variation	ABC	[10]
<i>Salmonella enterica</i> ^b	Gram–	Diverse environments	O-antigen	46 O-antigens	wzy, ABC	[11]
<i>Neisseria gonorrhoeae</i>	Gram–	Humans	LOS	LOS phase variation	(none)	[12]
MEDIUM						
<i>Streptococcus pneumoniae</i>	Gram+	Human respiratory tract	Capsule, teichoic acid	~100 serotypes	wzy, synth	[13]
<i>Haemophilus influenzae</i>	Gram–	Human respiratory tract	Capsule, LOS	Six serotypes, LOS phase variation	ABC	[12]

^aMajor polysaccharide antigens of WHO global priority pathogens in need of new antibiotics, divided into three categories according to the urgency of need for new antibiotics: critical, high and medium priority. In critical-priority Enterobacteriaceae, only *K. pneumoniae* and *E. coli* are shown. PS, polysaccharide; LOS, lipo-oligosaccharide; wzy, wzy-dependent; synth, synthase-dependent.

^bTyphi serovar also carries a capsule.

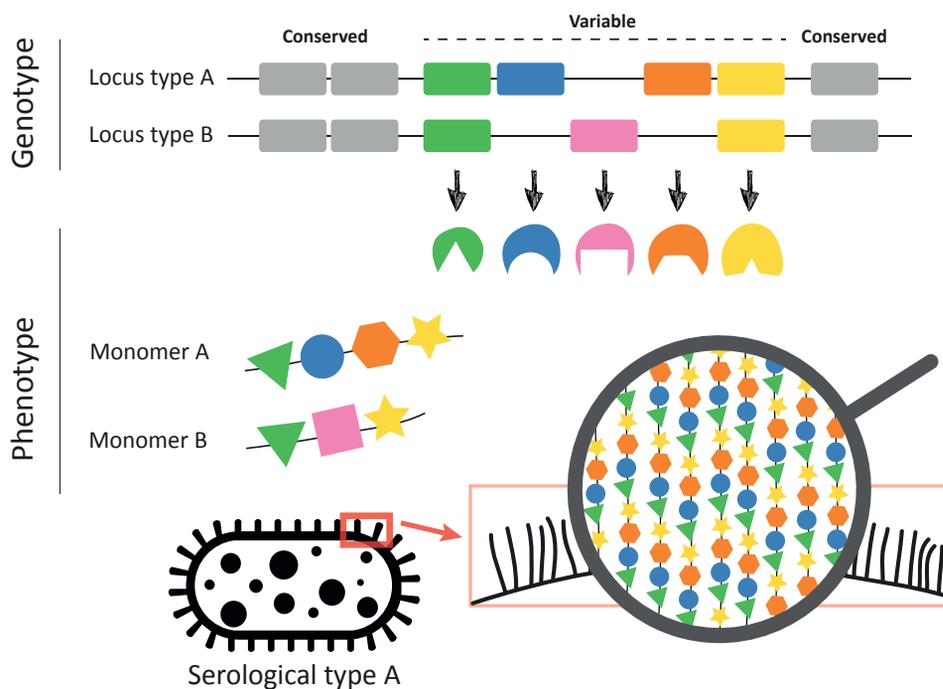
priority pathogens is presented in Table 1. (Note that *Neisseria meningitidis* is not included as it does not appear on the WHO list; however, this is also an important encapsulated pathogen for which capsule-targeting vaccines are in use [21]). *Streptococcus pneumoniae* capsules facilitate escape from mucous entrapment, thus affecting bacterial ability to colonise and transmit [22,23]. The LPS, together with its external component, O-antigen, plays an important role in bacterial colonisation by protecting the cell from hydrophobic antibiotics [24], or inducing resistance against bacteriophages [25]. LPS and LOS (lipo-oligosaccharide, which is LPS without its external O-antigen) have both been shown to be virulence determinants in *Haemophilus*, *Neisseria*, *Campylobacter*, or Enterobacteriaceae [19,21,26–29]. Furthermore, both LPS and teichoic acid can contribute to adherence to epithelial cells [30,31]. Finally, the externally secreted EPS, like colanic acid in *Escherichia coli* or the alginate in *Pseudomonas aeruginosa*, forms the basis of biofilms – a self-produced polymeric matrix which enables bacterial adhesion to surfaces and protects colonies from environmental dangers, including desiccation, bacteriophages, antibiotics, or the host immune system [32].

Genetic Architecture and Evolution

In spite of its length, a polysaccharide chain consists of a relatively small number of sugar molecules. The polymer can be biosynthesised in different ways, but in the large majority of cases it is done via one of three different mechanisms: the wzy-dependent, ABC-dependent, or synthase-dependent pathway (reviewed in [5,33]). These pathways engage sugar-specific enzymes to synthesise the polysaccharide, and the specific combination of these enzymes determines the sugar structure. There is a notable architectural similarity between genetic loci which synthesise polysaccharide chains (particularly wzy-synthesis and ABC-synthesis operons) in that the highly variable, polymer-specific region is located in the middle of the locus, and is surrounded by conserved genes that usually have roles in transport, assembly,

export, or synthesis of sugars (in *wzy*-synthesised operons this also includes *wzx* and *wzy* genes). The typical design of a polysaccharide antigen locus in bacteria is shown in Figure 1, and its consequences for epidemiological serotyping are discussed in Box 1.

The strikingly consistent architectural design of polysaccharide biosynthesis loci has important implications for the generation of antigenic diversity in bacteria. First, the synthesis of a monosaccharide via serotype-specific genes means that shuffling of these genes can alter the serotype. This can be achieved either via inactivation of one or more enzymes (gene loss), or an exchange leading to a new combination of enzymes (horizontal gene transfer) – often with the help of transposable elements – as has been widely observed in well studied *S. pneumoniae*, *E. coli*, or *Klebsiella pneumoniae* [4,6,34]. Second, the presence of conserved genes at the flanking regions of the synthesis loci, like *dexB/aliA* in *S. pneumoniae* or *galF/gnd* in *Klebsiella*, promotes exchange of the entire locus by homologous recombination, which requires homology only at the flanks. Such changes facilitate serotype/antigen alterations between distant lineages without having to ‘invent’ a new combination, which has been widely documented in epidemiological studies [35–43]. Third, genetic and epigenetic changes in regulatory genes (commonly referred to as ‘phase variation’) can affect capsule expression, and the resulting ability to colonise or infect the host [44,45]. Finally, polysaccharide antigens are



Trends in Microbiology

Figure 1. Generation of Polysaccharide Diversity in Bacteria. Polysaccharide antigens, like capsules and O-antigens, are usually synthesised by a specialist group of enzymes which are encoded by genes located in an antigen-biosynthesis locus. The genetic architecture of these loci is often similar between different, even distantly related, bacterial species. The specialised polymer-specific genes (coloured cassettes), which encode transferase enzymes (coloured shapes), are typically located in the middle of the locus. They are flanked by conserved, regulatory or transport genes (grey cassettes). The polymer-specific genes synthesise a monomer (so-called repeat unit), which is then polymerised to a polysaccharide chain and transported outside the cell. The order of these two events depends on the synthesis pathway, which, in the majority of studied cases, belongs to either the *wzy*-dependent or the ABC-dependent class. A given combination of the polymer-specific genes is a strong predictor of the polysaccharide structure, and thus bacterial serological type (serotype).

Box 1. The Future of Serotyping

Even though the sequence of genes located in the polysaccharide biosynthesis locus is highly predictive of the expressed polysaccharide, the genotype is not a perfect predictor of the phenotype for two main reasons. First, genetic mutations within those genes can alter specificity of enzymes encoded by them, thus altering a biochemical linkage and producing a new serotype. Second, additional genes located outside the synthesis locus can influence or direct the final sugar structure. This has important consequences for public health. With increasing adoption of high-throughput sequencing for strain characterisation by research and public health labs, *in silico* serotyping has now largely overtaken the standard serotyping methods, which require specialised reagents and expertise. (Table 1 summarises currently publicly available tools for *in silico* serotyping.) The flip side of this is that the gap between serologically-determined diversity and genetically-determined diversity is widening. Thus, biochemical characterisation of polysaccharide antigens remains important, and a good understanding of the complex genotype-phenotype map in polysaccharide antigens remains one of the great challenges of polysaccharide biology (see also Outstanding Questions).

Table 1. Public Tools for *In Silico* Serotyping Using Whole-Genome Data^a

Name	Species	Data input	Precision	Website	Refs
PAst	<i>Pseudomonas aeruginosa</i>	SR or A	Entire locus	github.com/Sandramses/PAst	[102]
Kaptive	<i>Klebsiella pneumoniae</i>	A	Entire locus	kaptive.holtlab.net	[4]
SerotypeFinder	<i>Escherichia coli</i>	SR or A	wzx/wzy/wzt/wzm	cge.cbs.dtu.dk/services/SerotypeFinder	[103]
SRST2 + EcoH	<i>E. coli</i>	SR	wzx/wzy/wzt/wzm	github.com/katholt/srst2	[104]
SeqSero	<i>Salmonella</i>	SR or A	Entire locus	www.denglab.info/SeqSero	[105]
SISTR	<i>Salmonella</i>	A	Core genes	lfz.corefacility.ca/sistr-app	[106]
PneumoCat	<i>Streptococcus pneumoniae</i>	A	Entire locus	github.com/phe-bioinformatics/PneumoCaT	[107]
SeroBA	<i>S. pneumoniae</i>	SR	Entire locus	github.com/sanger-pathogens/seroba	[108]

^aThe list contains publicly available tools for *in silico* serotyping of WHO priority pathogens listed in Table 1. SR, short-reads; A, assembly.

more robust than protein antigens because biosynthesis loci do not include housekeeping genes that are required for other functions, and monosaccharides – unlike proteins – do not have specific binding functions which can be disrupted by single amino acid changes. Thus, polysaccharides are much more flexible than proteins in their changing ability, and modifications to their structure are unlikely to incur a major fitness cost to the organism. This is evident from successful persistence of isolates that have lost the capsule [46,47], or loss of polysaccharide antigens in historical, *in vitro* passaged bacterial isolates [4,48].

The evidence therefore points to polysaccharide antigen loci having universally evolved as unique bacterial adaptive weapons, able to diversify at different speeds [49]. Over short timescales, diversity can be produced within populations via regulation and phase variation; over intermediate timescales available serotypes are exchanged between lineages via recombinational events spanning entire loci; over long timescales, gene loss and gene gain via horizontal gene transfer produce novel, polymer-specific combinations, giving rise to new serotypes. Polysaccharide antigen loci thus allow rapid bacterial adaptation to constantly changing selective pressures, thereby maintaining fitness relative to other competing strains or species.

Factors Shaping the Diversity of Polysaccharide Antigens

The genetic and phenotypic plasticity of polysaccharide biosynthesis loci can give rise to many serological types within a single species (Table 1), but it is not obvious why such diversity arose in the first place. Arguably, the simplest explanation is that diversity is randomly generated over time, but the drift alone is unlikely to be the driving evolutionary force. The sheer diversity of polysaccharide antigens in some species – much greater than expected based on the rest of the genome [4,6,50] – undoubtedly suggests the role of diversifying selection in generating new structures. The existence of such selection is also supported by the fact that polysaccharide synthesis loci are genetic variability hotspots between close bacterial species in the human gut [51], and they are often recombination hotspots, for which we have direct [41,42,52–57] and indirect evidence

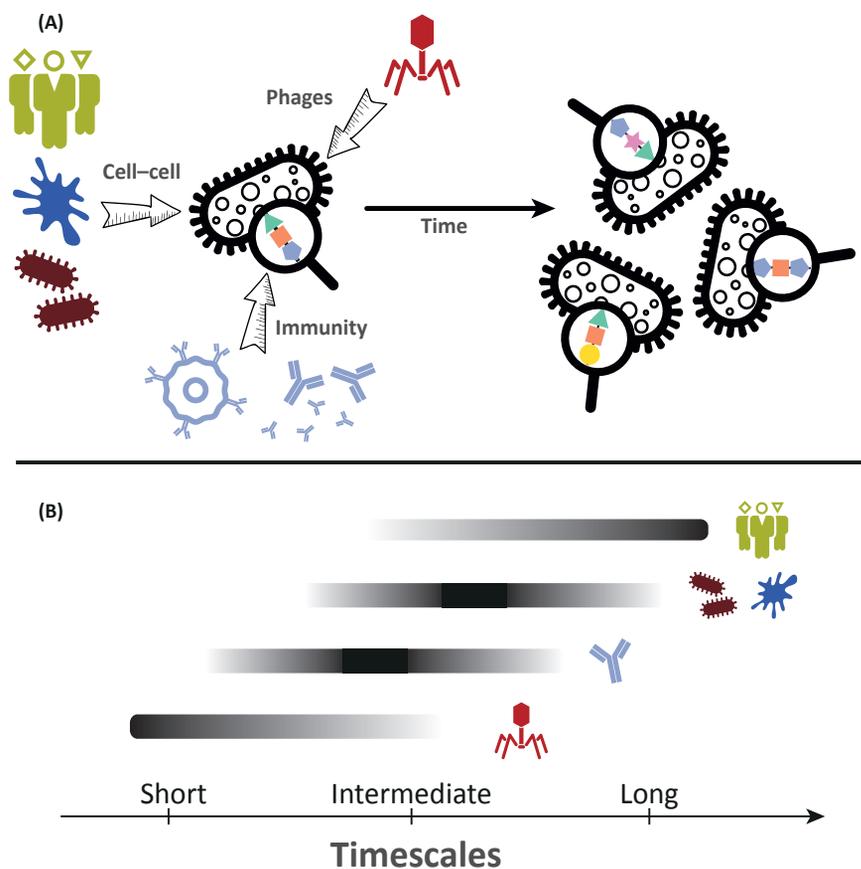
[4,35,58–61]. Finally, bacterial polysaccharides often involve synthesis of rare sugars, like L-rhamnose or L-fucose, which are not typically found in animal cells [62]. This suggests that the benefit of the possibility of generating a greater number of polysaccharide combinations overcomes the cost of being more visible to the immune system. Therefore, it is now widely accepted that polysaccharide antigens can be, and often are, under strong diversifying selection.

The strength of diversifying selection will depend on the species ecology, which can be subdivided into three different classes of ecological factors: host immunity, bacteriophages, and cell–cell interactions (Figure 2). Importantly, each of these factors contains a coevolutionary component in that each involves interactions with other evolving entities. The strength of selection will also vary with the antigen type, and the biological role and interactions of that antigen are context-specific in each organism. For example, *K. pneumoniae* expresses both K- and O-antigens, but exhibits far more capsular variation than LPS variation; whereas the closely related enterobacterial species *Salmonella enterica* (mostly unencapsulated) and *E. coli* (variably capsulated) display extensive LPS diversity. However, in general, our understanding of the relative importance of the three ecological forces for bacterial evolution and the resulting antigenic diversity is far from complete.

Host Immunity

Both capsules and O-antigens are known to interact with the immune system and can be immunogenic [14,63]. Correspondingly, host immunity has been the most popular candidate for diversifying selection of polysaccharide antigens in host-adapted bacteria. A classic model of immune-driven diversification is via negative frequency-dependent selection (Red Queen). Under this model, a new antigenic variant is under positive selection as it encounters naive hosts, but over time hosts develop immune memory against it and it is replaced by another, novel variant. This cycle continues, thus promoting antigenic diversity. However, under this scenario one would expect that novel diversity is constantly generated and dynamically changes over time. It has been argued that this model might explain the ability of some bacteria to alter their antigens on a generation timescale (like LPS variation in *N. meningitidis* and *Haemophilus influenzae*, or variable capsule expression in *Bacteroides*), which may promote colonisation and transmission by escaping the immune system [64]. But in many species the Red Queen model does not fit empirically observed temporal and spatial timescales at which novel serotypes emerge in many bacteria [65]. This conundrum has been partially resolved by mathematical modelling, which has by now quite convincingly demonstrated that immunity can play an important role in shaping the population structure of some pathogens and help maintain their antigenic diversity [66]. One class of multistrain models suggests that strain-specific immunity forces pathogens competing for hosts via cross-reactive responses to organise into non-overlapping antigenic repertoires [67]. It was argued that this framework explains the population structure of *N. meningitidis* and *S. pneumoniae* [68,69]. Furthermore, a balance between strain-specific immunity in naive infants (stabilising selection) and nonspecific immunity in adults (equalising selection) was suggested to explain the observed patterns of serotype diversity in the pneumococcus [70].

Nevertheless, immune-driven diversification of polysaccharide antigens is probably the exception rather than the rule in the bacterial kingdom. A recent study has shown that capsules are more common in environmental than in pathogenic bacteria, and also in facultative than in obligate pathogens [71]. The authors have argued that, for this reason, capsules have evolved as tools for environmental adaptation rather than as host-associated virulence factors. This is also consistent with the fact that, amongst all discussed bacterial species, the most extensive capsular variation is found in gut microbes that behave as opportunistic pathogens, such as *K. pneumoniae* and



Trends in Microbiology

Figure 2. Factors Driving and Maintaining Diversity of Polysaccharide Antigens. (A) Major diversifying forces in the world of bacterial polysaccharide antigens: host immunity, bacteriophages, and cell–cell interactions (including host glycan diversity, eukaryotic predators and other host commensals). These forces are likely to drive and maintain the polysaccharide antigen diversity we observe today. (B) These factors should not be viewed as mutually exclusive, but rather as different forces operating at different scales of time and space. Coevolution with bacteriophages could select for novel polysaccharide diversity on short timescales from just a few bacterial generations, could occur within virtually any ecological niche including host associated or non-host associated, and could have transient or long-lasting impacts on both bacterial and phage population dynamics. The impact of genetic variation in host glycan diversity is expected to take much longer to affect bacterial population structures, depending on host diversity and generation times. Host immunity, the diversity of other host commensals, and the impact of predators are likely to operate somewhere between the two. Phages may promote within-host diversity of antigens, but a serotype which provides resistance against a given phage population may not spread in the population due to its low between-host fitness. Likewise, glycan diversity in different host populations may promote diversity over space: different populations found in different locations may promote different bacterial antigens, but a single type within each population.

E. coli. Furthermore, extensive capsule diversity is also found in commensal bacteria (e.g., oral and pharyngeal streptococci [60], some gut microbiota species [51]), which are not affected by antibodies or phagocytes due to immunological tolerance to mutualistic members of our microbiome. All of this suggests that, while immunity can play a role in shaping polysaccharide antigen diversity, it cannot be the single and most important factor driving such diversity in all bacteria.

Bacteriophages

Another factor that can act as an important driver of antigen polysaccharide diversity is bacterial coevolution with their viruses, namely bacteriophages (phages). Phages often encounter

physical barriers preventing them from infecting bacterial cells: capsules, peptidoglycans, or extracellular polysaccharides in biofilms. Thus, phages have evolved strategies to overcome these barriers, and often carry receptors containing depolymerase – an enzyme which degrades polymers like extracellular polysaccharides present on bacterial surfaces. These receptors are typically encoded in tail fibres or base plates, and are highly specific to the antigen type. Consequently, diverse serotype-specific depolymerases have been identified in phages isolated from many different bacterial species [72]. Microbiological experiments have demonstrated that, over time, antagonistic coevolution between bacteria and their viruses will lead to an increased diversity of both bacteria and phages over time [73,74]. This hypothesis is in line with the dynamics observed in longitudinal analyses of the gut microbiome, which have shown dynamic shifts in bacterial and phage communities over time [75] and capsular polysaccharides as variability hotspots [51]. Temperate phages could also play a role in polysaccharide antigen diversification. They are known to coinfect bacteria [76], undergo competition [77,78], and alter the chemical composition of O-antigen upon entry to prevent other phages from coinfecting, which has been well documented in several species [79–81].

Cell–Cell Interactions

Cells are surrounded by carbohydrate structures known as glycans, and interaction between those glycans and proteins of other cells is expected to be an important ecological factor shaping their diversity [82]. Given the plethora of different cell–cell interactions in nature, we here focus on three important types of interaction which are central to the survival of bacterial pathogens and commensals inside hosts: interactions with host cells, with other colonising bacteria, and with predatory eukaryotes.

Host tissues, such as gut or respiratory epithelia (but also plant cell walls), are covered by a dense layer of glycans called the glycocalyx, which impacts many colonising bacteria. Consequently, microbes have evolved various ways of interacting with these structures to facilitate colonisation, including attachment to host glycans [82], modification of host glycans [83,84], and biofilm formation [85]. As these interactions are necessary for successful colonisation, bacteria have been under evolutionary pressure to adapt to the spectrum of glycans expressed by the cells of their hosts. One consequence of this is the phenomenon of ‘molecular mimicry’, whereby some bacterial polysaccharide antigens have evolved to resemble the glycan structures of their hosts [86]. Such resemblance allows the bacteria to take advantage of self-tolerance in order to evade the host immune system. Examples include production of polysialic acid in *N. meningitidis* serogroup B, chondroitin or heparosan in *E. coli*, hyaluronic acid in some species of *Streptococcus* (including *S. pyogenes*, *S. equi*, *S. dysgalactiae*, and *S. uberis*) [86], or blood-group-antigen-resembling LPS in *Helicobacter pylori* or *Campylobacter jejuni* [87,88]. These adaptations are part of a larger coevolutionary dynamic between hosts and bacteria, whereby hosts diversify their glycans over time to escape the selective pressure of bacteria, which in turn coadapt (for example via molecular mimicry) [89]. In vertebrates, such diversification can also rely on somatic mechanisms (like hypermutation or recombination), which permit keeping up with rapidly evolving prokaryotes [82]. All of this generates not just large between-host diversity of glycans, but also diversity within individual hosts. Such diversity can be beneficial by inhibiting the spread of pathogens within a host population, due to the pathogen’s differential recognition of distinct glycan structures expressed by members of the host population [82].

In addition to host cells, bacteria colonising new hosts may encounter other prokaryotic and eukaryotic cells sharing the same niche. First, successful colonisation requires compatibility with host-associated commensal bacteria – for example, the mutualistic streptococcal communities found on oral mucosal surfaces. These communities are shaped by interactions

between lectin-like adhesins and polysaccharides which act like receptors, thus connecting multiple cell–cell adhesions into large complex networks known as biofilms [90]. Since joining a biofilm substantially increases the chances of survival, new colonisers will be under selective pressure to attach and join the biofilm, or perish. This, in turn, will affect the population diversity of spreading bacteria [90]. Second, other single-celled species that feed on bacteria may also be present. Well known examples are the predatory eukaryotes, such as amoebae, which are sometimes found in animal intestines. As demonstrated in *S. enterica*, these predators exhibit different feeding preferences towards different O-antigens [91]. The specificity of such recognitions between a bacterivorous amoeba, *Acanthamoeba castellanii*, and *E. coli* has been attributed to the interaction between LPS and the predator's surface mannose-binding protein [92]. Given the genetically characterised diversity of mannose N-glycans between different species of amoebae [93], it is plausible that such interactions have shaped (at least partly) the diversity of O-antigens in some bacteria, as has been argued previously [64].

Selection over Time and Space

A schematic summary of the factors discussed above, and their expected timescales, is given in Figure 2. Importantly, one needs to be cautious when attempting to find a single explanation for the diversity of polysaccharide antigens in a given bacterial species, as forces maintaining it might be different from those which have driven it in the first place. For example, in *S. pneumoniae* it is widely accepted that immune selection is an important driver of antigenic population structure. However, it is not entirely clear whether occasionally appearing capsule switches are evolutionary 'mistakes', or whether they are (at least partly) driven by interactions with bacteriophages. Pneumococcal capsule-specific depolymerase enzymes have been isolated before [94,95], but almost nothing is known about capsule–bacteriophage interactions in *S. pneumoniae*. In *Klebsiella*, some K-antigens are known to interact with the adaptive immune system [96], but frequent capsule switches in clinical lineages are unlikely to be explained by host immunity alone, and interaction with phages and protists are considered to be important [4], not least because these bacteria are found in the guts of many animals. It is also possible that, for a facultative pathogen like *Klebsiella*, the ability to exchange K-antigens permits adaptation during environmental shifts. In *S. enterica*, host-specific distribution of O-antigens was previously argued to be a consequence of the the varying feeding preferences of intestinal predators present in different animal hosts [91]. However, it has also been speculated that such diversity could be (at least partly) driven by the diversity of mucins (intestinal polysaccharides) which amoebae use for attachment: O-antigens resembling these attachments would be much less likely to get predated, thus providing bacteria carrying those polysaccharides with a selective advantage [97]. All of this emphasises how little we know about the relative importance of various ecological and evolutionary forces interacting with bacterial polysaccharide antigens. Their characterisation will be important for an accurate prediction of the short-term and long-term impact of therapeutic interventions (see also Box 2).

Medical Implications

The adaptive potential of polysaccharide antigen loci to spread within bacterial populations or to evolve new types has potentially important consequences for public health. The antibiotic resistance crisis has led to an increased interest in polysaccharide-based therapeutic interventions (e.g., polysaccharide conjugate vaccines or phage-based therapies). These approaches impose strong selective pressures on a fraction of the bacterial population by targeting only a small subset of the capsular repertoire. Given the discussed potential of polysaccharide antigens to rapidly generate novel diversity, a question arises about the long-term efficacy of such interventions (see Outstanding Questions). We have seen some evidence

Box 2. Predicting the Impact of Medical Interventions with Mathematical Models

Mathematical models serve as powerful tools for predicting the effects of medical interventions. For a model to be a successful prediction tool, one needs to choose the right balance between the model simplicity and its biological accuracy.

Polysaccharide Conjugate Vaccines

Since the development and introduction of the conjugate vaccination programmes against *H. influenzae* and *S. pneumoniae*, mathematical models have been widely used to predict their effect on the control of bacterial disease [109–111]. Such models are based on compartmental modelling, in which the human population is represented by different classes (for example vaccinated and unvaccinated), and rates of changes between these compartments are defined by a set of mathematical equations. Disease dynamics can be described deterministically (via ordinary differential equations) or stochastically (via parameters drawn from random distributions). These models have varying levels of complexity, depending on the underlying assumptions about the complexity of the host structure, pathogen biology, and interactions between the two.

Studies modelling the impact of conjugate vaccines have mostly focused on their short-term effects, namely their effect on bacterial transmission and disease. Some papers have examined the impact of conjugate vaccines on the bacterial population structure [112]; however, none have investigated the role of emergence of novel bacterial diversity on long-term vaccine effectiveness. This is difficult and requires models which can merge the epidemiological framework with pathogen genetics, as has been done with some viral disease models such as Influenza or HIV-1 [113,114]. The predictive value of such models will depend not only on accurate estimates of epidemiological parameters (e.g., transmission rates, between-strain competition, strength of strain-specific and nonspecific immunity) but will also rely on good estimates of evolutionary rates driving diversification of bacterial genomes, including (but not limited to) acquisition of novel diversity from nonclinical reservoirs.

Phage Therapy

While phage therapy is fundamentally different from vaccination, previous studies have mostly relied on epidemiological models to predict its dynamics, and thus likely success [99,115]. However, these models still do not accurately capture the complexity of *in vivo* dynamics [116], nor do they address the question of coevolution between bacteria and viruses, or their potential epidemiological implications (e.g., transmission of bacterial resistance). Using mathematical models to predict the long-term success of phage therapy thus relies on quantifying parameters of *in vivo* bacteria–phage coevolution. Such quantification will be facilitated by recent advances in microbiome research and genomics [98].

of vaccine-driven adaptation in the antigenically diverse *S. pneumoniae*, which emphasises the challenge of robust vaccine design against other diverse bacteria like *Klebsiella* or *Acinetobacter* (Box 3). However, predicting the consequences of phage therapy will be arduous as current understanding of bacteria–phage coevolution in the complex environment of the human gut remains limited [98]. Also, since phage therapy is currently under active research and development, it is unclear what approaches will end up being used in practice. Using viruses as a public health control strategy has a long history of controversy, and indeed the implications of using bacteriophages for medical purposes have been actively debated for many years [99]. The rapidity with which some phage can kill their target suggests that the targeted bacteria would have a very narrow window of opportunity to evolve escape, for example by changes in the capsule locus; however, evolution experiments and observational microbiome data show that this may be possible or even likely [98,100]. One alternative might be to use phage-derived depolymerases which degrade bacterial capsules [101]; they alone would not kill bacteria and thus would impose very different selective pressures compared to lytic treatments.

Ultimately, long-term effectiveness of any polysaccharide-targeted medical intervention will depend on a good understanding of the evolutionary dynamics of the relevant bacterial system, as well as on clever formulation of mathematical models to predict the impact of such approaches (see also Box 2). However, neither can be achieved without viewing bacterial pathogens in the wider context of their ecological interactions as such pathogens in reality represent a ‘tip of the iceberg’ of the entire bacterial population. As medical interventions – like antibiotics, vaccines, or phage therapies – will impact the bacterial ecosystem as a whole, more attention in the future should be devoted to isolate sampling designs that overcome such

Box 3. Implications of Polysaccharide Diversity for Treatment Design: Lessons from Conjugate Vaccines

While the significance of natural polysaccharide variation for vaccine escape has been appreciated for decades, recent advances in genomic epidemiology have improved our understanding about its actual importance. Our knowledge comes predominantly from vaccines against three nasopharyngeal pathogens: *Haemophilus influenzae*, *Streptococcus pneumoniae* and *Neisseria meningitidis*, the impact of which has been studied for many years.

The simplest form of vaccine escape is by losing the capsule, which is an acceptable outcome since unencapsulated bacteria are unlikely to cause disease. This is most likely to happen when a vaccine spans the entire capsule repertoire (as in cases where a single capsule serotype is associated with human-adapted invasive disease, such as in *Salmonella enterica* serovar Typhi or *H. influenzae* type B), but in fact in all three species there has been a rise in the reported frequency of unencapsulated strains [47,117,118]. By contrast, when the vaccine spans only a fraction of the capsule repertoire, from an evolutionary point of view it becomes an ecological experiment where we can observe the impact of the removal of selected serotypes on the bacterial population (see also Figure 1A). Interestingly, recombination-driven vaccine escapes have only been reported in the species which, incidentally, has the highest capsule diversity: the pneumococcus [119]. There, we have seen persistence of some lineages associated with vaccine serotypes thanks to exchange of those serotypes into nonvaccine serotypes via recombination [39,54]. Genetic sequencing from densely sampled areas also revealed the existence of previously unseen capsule loci that were recombinations of other serotypes [56]. Importantly, phylogenetic and phylodynamic analyses revealed that these recombinants emerged prior to the introduction of the vaccine [39,54], implying that the polysaccharide capsule diversity evolves over time and that selection acts on this diversity. This process could be enhanced by interactions with closely related commensal bacteria, as we have seen that capsule genes are often shared between pathogens and commensals in all three bacterial species [59–61]. Furthermore, in the pneumococcus there is evidence that commensals can act as a source of horizontally acquired capsule diversity [56,120]. It thus seems that the dynamic microbial ecosystem of the nasopharynx could be an evolutionary hub where novel serotypes occasionally emerge and, in the presence of a vaccine, gain a selective advantage and rise in frequency (Figure 1B). However, as seen in the case of the three major pathogens, sharing a similar ecological niche, it is difficult to gauge the likelihood of such a scenario due to our limited understanding of the complex microbial interactions in the nasopharynx and the probability of new serotypes rising in frequency (see Outstanding Questions).

It will be even more challenging to predict the impact that polysaccharide conjugate vaccines in humans would have on populations of opportunistic pathogens with very different host ranges and ecologies, like *Klebsiella* or *Acinetobacter*. However, vaccines against these bacteria would most likely be used in a much more targeted way, such as aiming to protect at-risk patients in hospitals that are experiencing problems with outbreaks of antibiotic-resistant strains, or patients who are known to be colonised with such strains. Ultimately, estimating such risks requires a good understanding of the bacterial colonisation dynamics, which, in turn, emphasises the importance of routine bacterial carriage studies and quantifying bacterial evolution in real time.

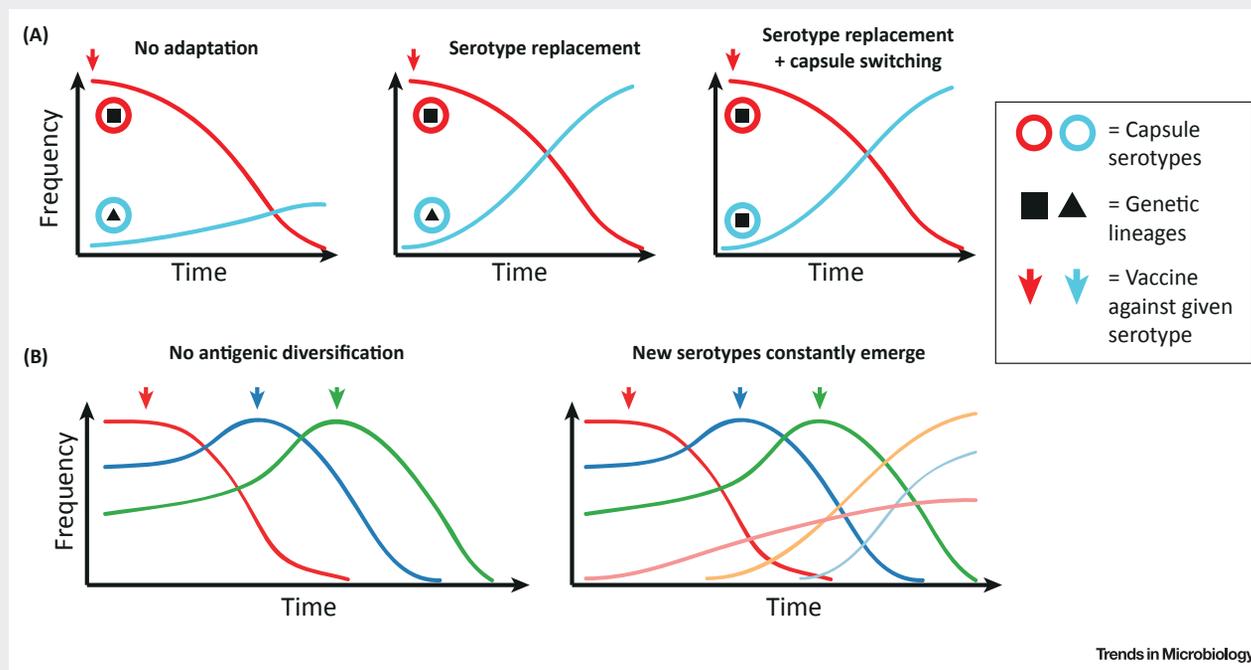


Figure 1. Mechanisms of Bacterial Adaptation against Polysaccharide Conjugate Vaccines. (A) Impact of polysaccharide conjugate vaccines on bacterial population structure. On the left, introduction of the vaccine against the red serotype is followed by the decline of this serotype but no replacement by another strain with the blue serotype, and thus overall reduction in carriage rates. This is similar to the situation in *Haemophilus influenzae* [121]. In the middle, vaccination against the red serotype is followed by the rise of another lineage (triangle) with the blue serotype with no significant reduction in carriage, known as ‘serotype replacement’. On the right, vaccination against the red serotype is followed by the rise of the same lineage (square) with another serotype (blue). This is a result of an acquisition of the blue serotype by the square lineage (known as ‘serotype switching’), which had occurred prior to the introduction of the vaccine. The latter two situations are

frequently observed in *Streptococcus pneumoniae*. (B) Potential impact of antigenic diversification on multivalent vaccine strategies. On the left, it is assumed that serotypes do not diversify over time. In this theoretical scenario, broader, multivalent vaccines could eventually lead to the eradication of the bacterial disease. On the right, new serotypes constantly emerge at low frequencies and are selected for by the vaccine due to serotype replacement. This scenario represents a Red Queen race between the vaccines and bacteria. In such a case, broader vaccines could select for novel, previously unseen serotypes to rise in frequency.

clinical bias. Fortunately, with the advancement of next-generation sequencing such studies are increasingly likely to become routine in the future.

Concluding Remarks

Polysaccharide antigens, such as capsules, O-antigens, or teichoic acids, are common in pathogenic and commensal bacteria. Our understanding of how these structures are synthesised and how they interact with hosts, viruses, and the environment have largely improved over the last two decades. Altogether, they paint a picture of genetic loci which are highly adapted to become modified and exchanged between different bacteria. Such diversity is likely a result of millions of years of coevolution with bacterial viruses, hosts, predators, and other bacteria, and is unlikely to change drastically on epidemiological time-scales. However, the introduction of vaccines, antibiotics, or phage therapies may dramatically alter the structure of bacterial populations due to the resulting strong selective pressures. It is thus conceivable that these antigens could rapidly evolve, undermining the long-term efficacy of therapeutic interventions. Therefore, the ability to predict the long-term consequences of these interventions and inform public health will be conditional on our understanding of the biology, epidemiology, and ecology of the system in question, quantification of evolutionary and epidemiological parameters, and the resulting accurate design of mathematical models.

Acknowledgments

This work has been funded by the Imperial College Research Fellowship (RM), the Bill and Melinda Gates Foundation (KEH) and a Senior Medical Research Fellowship from the Viertel Foundation of Australia (KEH). Authors thank Professor Paul Rainey for his helpful comments on the manuscript.

References

- Hu, D. *et al.* (2013) Diversity in the major polysaccharide antigen of *Acinetobacter baumannii* assessed by DNA sequencing, and development of a molecular serotyping scheme. *PLoS One* 8, e70329
- Lam, J.S. *et al.* (2011) Genetic and functional diversity of *Pseudomonas aeruginosa* lipopolysaccharide. *Front. Microbiol.* 2, 118
- Islam, S.T. and Lam, J.S. (2014) Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. *Can. J. Microbiol.* 60, 697–716
- Wyres, K.L. *et al.* (2016) Identification of *Klebsiella* capsule synthesis loci from whole genome data. *Microb. Genom.* 2, e000102
- Whitfield, C. (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu. Rev. Biochem.* 75, 39–68
- Iguchi, A. *et al.* (2015) A complete view of the genetic diversity of the *Escherichia coli* O-antigen biosynthesis gene cluster. *DNA Res.* 22, 101–107
- Thurlow, L.R. *et al.* (2009) Capsular polysaccharide production in *Enterococcus faecalis* and contribution of CpsF to capsule serospecificity. *J. Bacteriol.* 191, 6203–6210
- O’Riordan, K. and Lee, J.C. (2004) *Staphylococcus aureus* capsular polysaccharides. *Clin. Microbiol. Rev.* 17, 218–234
- Li, H. *et al.* (2017) The redefinition of *Helicobacter pylori* lipopolysaccharide O-antigen and core-oligosaccharide domains. *PLoS Pathog.* 13, e1006280
- Guerry, P. *et al.* (2012) *Campylobacter* polysaccharide capsules: virulence and vaccines. *Front. Cell. Infect. Microbiol.* 2, 7
- Liu, B. *et al.* (2014) Structural diversity in *Salmonella* O antigens and its genetic basis. *FEMS Microbiol. Rev.* 38, 56–89
- Goldberg, J.B. (1999) *Genetics of Bacterial Polysaccharides*, Taylor & Francis
- Bentley, S.D. *et al.* (2006) Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet.* 2, e31
- Roberts, I.S. (1996) The biochemistry and genetics of capsular polysaccharide production in bacteria. *Annu. Rev. Microbiol.* 50, 285–315
- Achtman, M. and Pluschke, G. (1986) Clonal analysis of descent and virulence among selected *Escherichia coli*. *Annu. Rev. Microbiol.* 40, 185–210
- Kahler, C.M. *et al.* (1998) The (alpha2->8)-linked polysialic acid capsule and lipooligosaccharide structure both contribute to the ability of serogroup B *Neisseria meningitidis* to resist the bactericidal activity of normal human serum. *Infect. Immun.* 66, 5939–5947
- Kadioglu, A. *et al.* (2008) The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat. Rev. Microbiol.* 6, 288–301
- Rajagopal, L. (2009) Understanding the regulation of group B streptococcal virulence factors. *Future Microbiol.* 4, 201–221

Outstanding Questions

What is the impact of within-host evolutionary dynamics on bacterial population structure? We know relatively little about within-host forces driving antigenic diversification, and about trade-offs between within-host and between-host bacterial fitness and transmission.

How important are polysaccharides in carriage and disease? Some capsule types are strong predictors of carriage and virulence, but their role in pathogenesis is not fully understood.

Can we efficiently predict phenotypes from genotypes? Genetic typing permits rapid identification of well known serotypes but it tells us little about the genotype–phenotype map.

How important are commensal and nonclinical strains in driving evolution of pathogens? Nonpathogenic bacteria are largely understudied, and we need to understand their role in pathogen evolution and diversification.

How fit are different serotypes in different environments? Environment-dependent measures of serotype fitness would help us predict the emergence of new serotypes in the future.

What impact do bacteriophages have on bacterial population dynamics? We need to understand how phages interact with bacteria and how they shape population structures of bacterial populations.

How common are interactions between eukaryotic predators and bacteria? We need more studies highlighting the nature of interactions between protists, bacteria, and their ecology, and need to understand how they impact evolution of non-host-associated bacteria.

What is the impact of host glycan diversity on pathogen evolution? Glycan diversity may affect the structuring of bacterial populations, both within-host and between-host.

19. Shon, A.S. *et al.* (2013) Hypervirulent (hypermucoviscous) *Klebsiella pneumoniae*: a new and dangerous breed. *Virulence* 4, 107–118
20. Sukupolvi-Petty, S. *et al.* (2006) The *Haemophilus influenzae* Type b hcsA and hcsB gene products facilitate transport of capsular polysaccharide across the outer membrane and are essential for virulence. *J. Bacteriol.* 188, 3870–3877
21. Harrison, O.B. *et al.* (2013) Description and nomenclature of *Neisseria meningitidis* capsule locus. *Emerg. Infect. Dis.* 19, 566–573
22. Nelson, A.L. *et al.* (2007) Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infect. Immun.* 75, 83–90
23. Zafar, M.A. *et al.* (2017) Capsule type and amount affect shedding and transmission of *Streptococcus pneumoniae*. *mBio* 8, e00989-17
24. Delcour, A.H. (2009) Outer membrane permeability and antibiotic resistance. *Biochim. Biophys. Acta* 1794, 808–816
25. Leon, M. and Bastias, R. (2015) Virulence reduction in bacteriophage resistant bacteria. *Front. Microbiol.* 6, 343
26. Zwahlen, A. *et al.* (1986) Contribution of lipopolysaccharide to pathogenicity of *Haemophilus influenzae*: comparative virulence of genetically-related strains in rats. *Microb. Pathog.* 1, 465–473
27. Plant, L. *et al.* (2006) Lipooligosaccharide structure contributes to multiple steps in the virulence of *Neisseria meningitidis*. *Infect. Immun.* 74, 1360–1367
28. Zarantonelli, M.L. *et al.* (2006) Differential role of lipooligosaccharide of *Neisseria meningitidis* in virulence and inflammatory response during respiratory infection in mice. *Infect. Immun.* 74, 5506–5512
29. Louwen, R. *et al.* (2008) The sialylated lipooligosaccharide outer core in *Campylobacter jejuni* is an important determinant for epithelial cell invasion. *Infect. Immun.* 76, 4431–4438
30. Aly, R. *et al.* (1980) Role of teichoic acid in the binding of *Staphylococcus aureus* to nasal epithelial cells. *J. Infect. Dis.* 141, 463–465
31. Bravo, D. *et al.* (2011) Different sugar residues of the lipopolysaccharide outer core are required for early interactions of *Salmonella enterica* serovars Typhi and Typhimurium with epithelial cells. *Microb. Pathog.* 50, 70–80
32. Okshevsky, M. and Meyer, R.L. (2015) The role of extracellular DNA in the establishment, maintenance and perpetuation of bacterial biofilms. *Crit. Rev. Microbiol.* 41, 341–352
33. Schmid, J. *et al.* (2015) Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. *Front. Microbiol.* 6, 496
34. Aanensen, D.M. *et al.* (2007) Predicted functions and linkage specificities of the products of the *Streptococcus pneumoniae* capsular biosynthetic loci. *J. Bacteriol.* 189, 7856–7876
35. Kroll, J.S. and Moxon, E.R. (1990) Capsulation in distantly related strains of *Haemophilus influenzae* type b: genetic drift and gene transfer at the capsulation locus. *J. Bacteriol.* 172, 1374–1379
36. Achtman, M. *et al.* (2012) Multilocus sequence typing as a replacement for serotyping in *Salmonella enterica*. *PLoS Pathog.* 8, e1002776
37. Alqasim, A. *et al.* (2014) Comparative genome analysis identifies few traits unique to the *Escherichia coli* ST131 H30Px clade and extensive mosaicism at the capsule locus. *BMC Genomics* 15, 830
38. Thrane, S.W. *et al.* (2015) The widespread multidrug-resistant serotype O12 *Pseudomonas aeruginosa* clone emerged through concomitant horizontal transfer of serotype antigen and antibiotic resistance gene clusters. *mBio* 6, e01396-01315
39. Croucher, N.J. *et al.* (2015) Selective and genetic constraints on pneumococcal serotype switching. *PLoS Genet.* 11, e1005095
40. Holt, K.E. *et al.* (2015) Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. *Proc. Natl. Acad. Sci. U. S. A.* 112, E3574–E3581
41. Holt, K. *et al.* (2016) Five decades of genome evolution in the globally distributed, extensively antibiotic-resistant *Acinetobacter baumannii* global clone 1. *Microb. Genom.* 2, e000052
42. Neemuchwala, A. *et al.* (2016) Capsular switching and other large-scale recombination events in invasive sequence Type 1 Group B *Streptococcus*. *Emerg. Infect. Dis.* 22, 1941–1944
43. Perrin, A. *et al.* (2017) Evolutionary dynamics and genomic features of the *Elizabethkingia anophelis* 2015 to 2016 Wisconsin outbreak strain. *Nat. Commun.* 8, 15483
44. Hammerschmidt, S. *et al.* (1996) Capsule phase variation in *Neisseria meningitidis* serogroup B by slipped-strand mispairing in the polysialyltransferase gene (siaD): correlation with bacterial invasion and the outbreak of meningococcal disease. *Mol. Microbiol.* 20, 1211–1220
45. Weiser, J.N. *et al.* (1996) Relationship between phase variation in colony morphology, intrastrain variation in cell wall physiology, and nasopharyngeal colonization by *Streptococcus pneumoniae*. *Infect. Immun.* 64, 2240–2245
46. Dolan-Livengood, J.M. *et al.* (2003) Genetic basis for nongroupable *Neisseria meningitidis*. *J. Infect. Dis.* 187, 1616–1628
47. Chewapreecha, C. *et al.* (2014) Dense genomic sampling identifies highways of pneumococcal recombination. *Nat. Genet.* 46, 305–309
48. Hobman, J.L. *et al.* (2007) Laboratory strains of *Escherichia coli*: model citizens or deceitful delinquents growing old disgracefully? *Mol. Microbiol.* 64, 881–885
49. Moxon, E.R. *et al.* (1994) Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr. Biol.* 4, 24–33
50. Mavroidi, A. *et al.* (2007) Genetic relatedness of the *Streptococcus pneumoniae* capsular biosynthetic loci. *J. Bacteriol.* 189, 7841–7855
51. Zhu, A. *et al.* (2015) Inter-individual differences in the gene content of human gut bacterial species. *Genome Biol.* 16, 82
52. Swartley, J.S. *et al.* (1997) Capsule switching of *Neisseria meningitidis*. *Proc. Natl. Acad. Sci. U. S. A.* 94, 271–276
53. Milkman, R. *et al.* (2003) Molecular evolution of the *Escherichia coli* chromosome. VI. Two regions of high effective recombination. *Genetics* 163, 475–483
54. Croucher, N.J. *et al.* (2013) Population genomics of post-vaccine changes in pneumococcal epidemiology. *Nat. Genet.* 45, 656–663
55. Wyres, K.L. *et al.* (2013) Pneumococcal capsular switching: a historical perspective. *J. Infect. Dis.* 207, 439–449
56. Mostowy, R.J. *et al.* (2017) Pneumococcal capsule synthesis locus cps as evolutionary hotspot with potential to generate novel serotypes by recombination. *Mol. Biol. Evol.* 34, 2537–2554
57. David, S. *et al.* (2017) Dynamics and impact of homologous recombination on the evolution of *Legionella pneumophila*. *PLoS Genet.* 13, e1006855
58. Brisse, S. *et al.* (2009) Virulent clones of *Klebsiella pneumoniae*: identification and evolutionary scenario based on genomic and phenotypic characterization. *PLoS One* 4, e4982
59. Nielsen, S.M. *et al.* (2015) The capsule biosynthesis locus of *Haemophilus influenzae* shows conspicuous similarity to the corresponding locus in *Haemophilus sputorum* and may have been recruited from this species by horizontal gene transfer. *Microbiology (Reading, England)* 161, 1182–1188
60. Sorensen, U.B. *et al.* (2016) Capsular polysaccharide expression in commensal *Streptococcus* species: genetic and antigenic similarities to *Streptococcus pneumoniae*. *mBio* 7, e01844-16
61. Bartley, S.N. *et al.* (2017) Acquisition of the capsule locus by horizontal gene transfer in *Neisseria meningitidis* is often accompanied by the loss of UDP-GalNAc synthesis. *Sci. Rep.* 7, 44442

62. Clark, E.L. *et al.* (2016) Development of rare bacterial monosaccharide analogs for metabolic glycan labeling in pathogenic bacteria. *ACS Chem. Biol.* 11, 3365–3373
63. Reeves, P. (1995) Role of O-antigen variation in the immune response. *Trends Microbiol.* 3, 381–386
64. Butela, K. and Lawrence, J. (2010) Population genetics of *Salmonella*: selection for antigenic diversity. In *Bacterial Population Genetics in Infectious Disease* (Robinson, D.A. and Feil, E. J., eds), pp. 287–320, Wiley-Blackwell
65. Lipsitch, M. and O'Hagan, J.J. (2007) Patterns of antigenic diversity and the mechanisms that maintain them. *J. R. Soc. Interface* 4, 787–802
66. Cobey, S. (2014) Pathogen evolution and the immunological niche. *Ann. N. Y. Acad. Sci.* 1320, 1–15
67. Gupta, S. *et al.* (1998) Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. *Science* 280, 912–915
68. Gupta, S. and Anderson, R.M. (1999) Population structure of pathogens: the role of immune selection. *Parasitol. Today (Regul Ed)* 15, 497–501
69. Watkins, E.R. *et al.* (2016) Metabolic competition as a driver of bacterial population structure. *Future Microbiol.* 11, 1339–1357
70. Cobey, S. and Lipsitch, M. (2012) Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. *Science* 335, 1376–1380
71. Rendueles, O. *et al.* (2017) Abundance and co-occurrence of extracellular capsules increases environmental breadth: Implications for the emergence of pathogens. *PLoS Pathog.* 13, e1006525
72. Latka, A. *et al.* (2017) Bacteriophage-encoded virion-associated enzymes to overcome the carbohydrate barriers during the infection process. *Appl. Microbiol. Biotechnol.* 101, 3103–3119
73. Benmayor, R. *et al.* (2008) The interactive effects of parasites, disturbance, and productivity on experimental adaptive radiations. *Evolution* 62, 467–477
74. Brockhurst, M.A. *et al.* (2004) The effect of spatial heterogeneity and parasites on the evolution of host diversity. *Proc. Biol. Sci.* 271, 107–111
75. Minot, S. *et al.* (2013) Rapid evolution of the human gut virome. *Proc. Natl. Acad. Sci. U. S. A.* 110, 12450–12455
76. Bobay, L.M. *et al.* (2013) The adaptation of temperate bacteriophages to their host genomes. *Mol. Biol. Evol.* 30, 737–751
77. Refardt, D. (2011) Within-host competition determines reproductive success of temperate bacteriophages. *ISME J.* 5, 1451–1460
78. Bobay, L.M. *et al.* (2014) Pervasive domestication of defective prophages by bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 111, 12127–12132
79. Davies, M.R. *et al.* (2013) Horizontally acquired glycosyltransferase operons drive salmonellae lipopolysaccharide diversity. *PLoS Genet.* 9, e1003568
80. Tang, S.S. *et al.* (2016) *Shigella flexneri* serotype 1c derived from serotype 1a by acquisition of gtrIC gene cluster via a bacteriophage. *BMC Microbiol.* 16, 127
81. De Smet, J. *et al.* (2017) *Pseudomonas* predators: understanding and exploiting phage–host interactions. *Nat. Rev. Microbiol.* 15, 517–530
82. Gagneux, P. and Varki, A. (1999) Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* 9, 747–755
83. King, S.J. *et al.* (2006) Deglycosylation of human glycoconjugates by the sequential activities of exoglycosidases expressed by *Streptococcus pneumoniae*. *Mol. Microbiol.* 59, 961–974
84. Mubaiwa, T.D. *et al.* (2017) The sweet side of the pathogenic *Neisseria*: the role of glycan interactions in colonisation and disease. *Pathog. Dis.* 75, 5
85. Garbe, J. *et al.* (2014) EndoE from *Enterococcus faecalis* hydrolyzes the glycans of the biofilm inhibiting protein lactoferrin and mediates growth. *PLoS One* 9, e91035
86. Cress, B.F. *et al.* (2014) Masquerading microbial pathogens: capsular polysaccharides mimic host-tissue molecules. *FEMS Microbiol. Rev.* 38, 660–697
87. Moran, A.P. (2008) Relevance of fucosylation and Lewis antigen expression in the bacterial gastroduodenal pathogen *Helicobacter pylori*. *Carbohydr. Res.* 343, 1952–1965
88. Houliston, R.S. *et al.* (2009) Complete chemoenzymatic synthesis of the Forssman antigen using novel glycosyltransferases identified in *Campylobacter jejuni* and *Pasteurella multocida*. *Glycobiology* 19, 153–159
89. Varki, A. (2006) Nothing in glycobiology makes sense, except in the light of evolution. *Cell* 126, 841–845
90. Yoshida, Y. *et al.* (2008) Molecular and antigenic characterization of a *Streptococcus oralis* coaggregation receptor polysaccharide by carbohydrate engineering in *Streptococcus gordonii*. *J. Biol. Chem.* 283, 12654–12664
91. Wildschutte, H. *et al.* (2004) Protozoan predation, diversifying selection, and the evolution of antigenic diversity in *Salmonella*. *Proc. Natl. Acad. Sci. U. S. A.* 101, 10644–10649
92. Arnold, J.W. *et al.* (2016) Determinants that govern the recognition and uptake of *Escherichia coli* O157:H7 by *Acanthamoeba castellanii*. *Cell Microbiol.* 18, 1459–1470
93. Feasley, C.L. *et al.* (2015) Evolutionary diversity of social amoebae N-glycomes may support interspecific autonomy. *Glycoconj. J.* 32, 345–359
94. Avery, O.T. and Dubos, R. (1930) The specific action of a bacterial enzyme on pneumococci of type III. *Science* 72, 151–152
95. Torriani, A. and Pappenheimer, A.M. (1962) Inducible polysaccharide depolymerases of *Bacillus palustris*. *J. Biol. Chem.* 237, 3–13
96. Lepper, P.M. *et al.* (2003) K-antigen-specific, but not O-antigen-specific natural human serum antibodies promote phagocytosis of *Klebsiella pneumoniae*. *FEMS Immunol. Med. Microbiol.* 35, 93–98
97. Wildschutte, H. and Lawrence, J.G. (2007) Differential *Salmonella* survival against communities of intestinal amoebae. *Microbiology (Reading, England)* 153, 1781–1789
98. Scanlan, P.D. (2017) Bacteria–bacteriophage coevolution in the human gut: implications for microbial diversity and functionality. *Trends Microbiol.* 25, 614–623
99. Levin, B.R. and Bull, J.J. (2004) Population and evolutionary dynamics of phage therapy. *Nat. Rev. Microbiol.* 2, 166–173
100. Schooley, R.T. *et al.* (2017) Development and use of personalized bacteriophage-based therapeutic cocktails to treat a patient with a disseminated resistant *Acinetobacter baumannii* infection. *Antimicrob. Agents Chemother.* 61, e00954-17
101. Roach, D.R. and Donovan, D.M. (2015) Antimicrobial bacteriophage-derived proteins and therapeutic applications. *Bacteriophage* 5, e1062590
102. Thrane, S.W. *et al.* (2016) Application of whole-genome sequencing data for O-specific antigen analysis and *in silico* serotyping of *Pseudomonas aeruginosa* isolates. *J. Clin. Microbiol.* 54, 1782–1788
103. Joensen, K.G. *et al.* (2015) Rapid and easy *in silico* serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. *J. Clin. Microbiol.* 53, 2410–2426
104. Ingle, D.J. *et al.* (2016) *In silico* serotyping of *E. coli* from short read data identifies limited novel O-loci but extensive diversity of O:H serotype combinations within and between pathogenic lineages. *Microb. Genom.* 2, e000064
105. Zhang, S. *et al.* (2015) *Salmonella* serotype determination utilizing high-throughput genome sequencing data. *J. Clin. Microbiol.* 53, 1685–1692
106. Yoshida, C.E. *et al.* (2016) The *Salmonella in silico* typing resource (SISTR): an open web-accessible tool for rapidly typing and subtyping draft salmonella genome assemblies. *PLoS One* 11, e0147101
107. Kapatai, G. *et al.* (2016) Whole genome sequencing of *Streptococcus pneumoniae*: development, evaluation and verification

- of targets for serogroup and serotype prediction using an automated pipeline. *PeerJ* 4, e2477
108. Epping, L. *et al.* (2017) SeroBA: rapid high-throughput serotyping of *Streptococcus pneumoniae* from whole genome sequence data. *Microb. Genom.* 2018, 4
109. Lipsitch, M. (1997) Vaccination against colonizing bacteria with multiple serotypes. *Proc. Natl. Acad. Sci. U. S. A.* 94, 6571–6576
110. Coen, P.G. *et al.* (1998) Mathematical models of *Haemophilus influenzae* type b. *Epidemiol. Infect.* 120, 281–295
111. Jackson, M.L. *et al.* (2012) Modeling insights into *Haemophilus influenzae* type b disease, transmission, and vaccine programs. *Emerg. Infect. Dis.* 18, 13–20
112. Watkins, E.R. *et al.* (2015) Vaccination drives changes in metabolic and virulence profiles of *Streptococcus pneumoniae*. *PLoS Pathog.* 11, e1005034
113. Bedford, T. *et al.* (2015) Global circulation patterns of seasonal influenza viruses vary with antigenic drift. *Nature* 523, 217–220
114. Dorratoltaj, N. *et al.* (2017) Multi-scale immunoepidemiological modeling of within-host and between-host HIV dynamics: systematic review of mathematical models. *PeerJ* 5, e3877
115. Cairns, B.J. *et al.* (2009) Quantitative models of *in vitro* bacteriophage-host dynamics and their application to phage therapy. *PLoS Pathog.* 5, e1000253
116. Bull, J.J. and Gill, J.J. (2014) The habits of highly effective phages: population dynamics as a framework for identifying therapeutic phages. *Front. Microbiol.* 5, 618
117. Johsrich, K.O. *et al.* (2012) Invasive potential of nonencapsulated disease isolates of *Neisseria meningitidis*. *Infect. Immun.* 80, 2346–2353
118. Van Eldere, J. *et al.* (2014) Non-typeable *Haemophilus influenzae*, an under-recognised pathogen. *Lancet Infect. Dis.* 14, 1281–1292
119. Maiden, M.C. (2013) The impact of protein-conjugate polysaccharide vaccines: an endgame for meningitis? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368, 20120147
120. Kilian, M. *et al.* (2014) Parallel evolution of *Streptococcus pneumoniae* and *Streptococcus mitis* to pathogenic and mutualistic lifestyles. *mBio* 5, e01490–01414
121. Ladhani, S. *et al.* (2008) No evidence for *Haemophilus influenzae* serotype replacement in Europe after introduction of the Hib conjugate vaccine. *Lancet Infect. Dis.* 8, 275–276