Cell Host & Microbe
Review

Bacterial Genome Variability and Its Impact on Vaccine Design

John L. Telford[1,*]
[1]Novartis Vaccines and Diagnostics Srl, Via Fiorentina 1, 53100 Siena, Italy
*Correspondence: john.telford@novartis.com
DOI 10.1016/j.chom.2008.05.004

The majority of currently available successful vaccines induce host responses against antigens that are highly conserved in the targeted pathogens. The diphtheria, tetanus, and pertussis vaccines confer protection by inducing neutralizing antibodies to the conserved bacterial toxins that are the major virulence factors. The *Hemophilus influenzae* B vaccine induces responses to conserved epitopes in the sugar structure of the bacterial capsular polysaccharide. However, the efficacy of more recently developed vaccines is limited by antigen variation, which also presents a challenge for future vaccine development. This review will explore bacterial genome variability and its impact on vaccine development.

Introduction

It is no coincidence that the first vaccines developed were against pathogens with little variability. In fact, the concept of immunological memory on which vaccination is based was derived from the observation that survival of infection resulted in lifelong immunity to further episodes of the same disease. Hence vaccines against stable diseases such as smallpox, measles, and rubella were very effective in conferring protection. Vaccines against bacterial pathogens whose virulence depends on a single toxin have also been very effective. While the genomes of the bacteria causing diphtheria, tetanus, and pertussis (Preston et al., 2004) may in fact vary considerably, in each of these diseases the major virulence factor is a single toxin. Neutralization of these toxins by specific antibodies essentially prevents disease. *Corynebacterium diphtheriae* is a good example of this. The diphtheria toxin is carried on a prophage (Pappenheimer, 1977). In the absence of the prophage in the genome, the bacteria are totally harmless. Hence immunization with a detoxified form of the toxin that is capable of inducing neutralizing antibodies to the active toxin confers effective protection against disease. The common theme among these often-lethal pathogens is that while host survival results in immunity, they are highly infectious and can be maintained in the population because of the renewal of nonimmune cohorts of new births.

Later, vaccines became more difficult to develop. The polio vaccine required three strains of virus to give universal coverage, but at least these three strains are relatively stable (De Jesus, 2007). Influenza virus is highly variable between one year and the next, but its infection period is so short that yearly vaccination with the appropriate strains is effective (Carrat and Flahault, 2007). However, its capacity to reassort its genome with other strains during coinfection in animal hosts results in the generation of virulent pandemic strains to which we have no prior immune priming (Hilleman, 2002). Finally, viruses such as HIV vary even within a single host during a single infection and have thus resisted all attempts to date to develop a vaccine (Korber et al., 2001).

Variation in viruses with RNA genomes generally depends on inaccurate replication due to the lack of proof reading during replication or in addition, as in the case of influenza, to genome reassortment and is usually driven by immune pressure and selection. Variation in bacteria on the other hand reflects a variety of mechanisms and selection pressures for developing diversity. This review will concentrate on bacterial antigenic variation and the resulting difficulties of developing vaccines capable of conferring broad protection.

Capsular Polysaccharide Antigens

Many pathogens are encapsulated bacteria. The capsule is formed from complex polysaccharides that vary between species and even between strains of the same pathogen. While capsules often help to evade innate immunity, in several cases, antibodies that recognize the capsular polysaccharide (CPS) can activate complement deposition and lead to cell killing either by direct bacterial lysis or by complement dependent opsonophagocytic killing of the pathogen (Sood and Fattom, 1998). Hence, the CPS can be an effective antigen to induce protective immunity particularly for pathogens that invade the host to cause septicemia and meningitis. Since polysaccharides are generally poor antigens that do not induce a T cell response, they are poorly immunogenic in infants that are the major targets for the disease. To overcome this, CPS antigens are usually conjugated to a protein carrier capable of inducing a robust T cell response resulting in immunogenicity in infants and induction of long-term memory (Mond and Kokai-Kun, 2008).

Examples of successful vaccines using this strategy are those against *Streptococcus pneumoniae* (pneumococcus), *Haemophilus influenzae*, and some serogroups of *Neisseria meningitidis* (meningococcus). Unfortunately, these pathogens have highly variable capsular serotypes, and there is rarely crossprotection between serotypes. Pneumococcus, for example, has over ninety serotypes with little crossprotection (Lipsitch and O'Hagan, 2007). On the other hand, the population structure of pneumococcal strains is fairly stable, and the majority of invasive disease is caused by a few serotypes. This has led to the development of a 23-valent pneumococcal CPS vaccine that confers limited protection in adults and to a 7-valent conjugate vaccine that is very effective in both adults and infants (Galiza and Heath, 2007). The seven serotypes included in this latter vaccine account for about 80% of invasive disease in the US and Europe

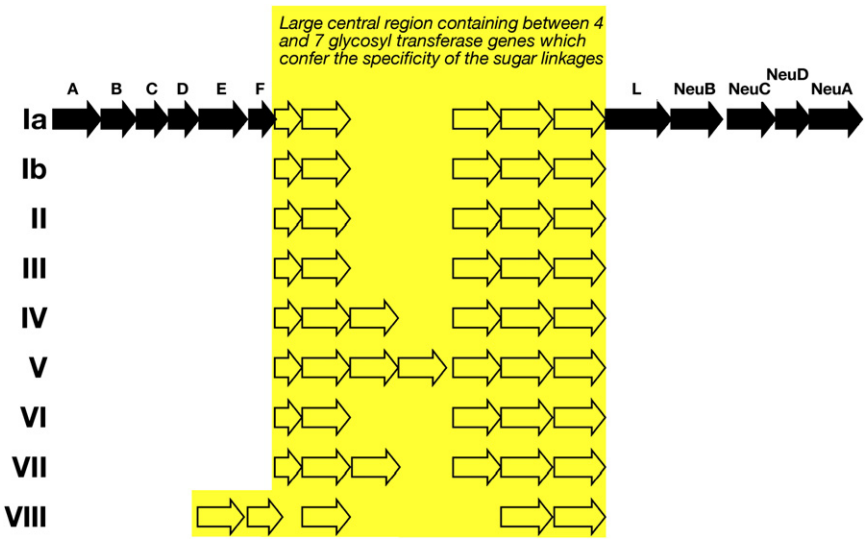

Figure 1. Schematic Representation of the Capsule Locus of Group B Streptococcus
Modified from Cieslewicz et al. (2005). Variable glycosyltransferase genes that code for enzymes which catalyze the specific sugar linkages are flanked by genes coding for enzymes involved in sialic acid synthesis and activation and genes coding for proteins involved in polysaccharide export.

(Toltzis and Jacobs, 2005). More complex pneumococcal CPS conjugates are currently in development in order to increase strain coverage. A partly expected result of the widespread use of the 7-valent pneumococcus vaccine has, however, been that serotypes not included in the vaccine are beginning to replace those that are included and disease from these non-included strains is on the increase (Toltzis and Jacobs, 2005). Having said that, the value to public health of conjugate vaccines, such as the *H. influenza* B vaccine, cannot be understated.

A similar situation exists for *N. meningitidis*. There are five serotypes (A, B, C, Y, and W135) causing the majority of disease in the world (Stephens et al., 2007). For four of these, effective vaccines have been developed or are in development. Serotype B, which is responsible for up to 50% of disease in Europe and the USA, cannot, however, be defeated by the same strategy, since its CPS is composed of polysialic acid, a polysaccharide found also in human tissues. Hence, the B group CPS is poorly immunogenic and, in principle, could induce autoantibodies (Stein et al., 2006).

CPS conjugate vaccines designed to protect neonates from intra- or postpartum bacterial infection through maternal immunization are also in development. group B Streptococcus (*S. agalactiae*, GBS) colonizes asymptomatically the ano-vaginal tracts of up to 30% of women (Schuchat, 1998). Although GBS is not pathogenic in healthy adults, newborn babies may be colonized during birth, which can lead to severe septicemia, pneumonia, or meningitis. There are nine known capsular serotypes of GBS, and up to 10% of strains are nontypeable. It has been demonstrated that risk of disease in the newborn correlates with low levels of type-specific maternal IgG antibodies (Baker and Kasper, 1976; Lin et al., 2001, 2004). It is believed that high-titer maternal IgG can cross the placenta and confer protection to the newborn. Elderly people, particularly those with underlying disease, are also at risk for serious infections by GBS (Schuchat, 1999).

A common problem with CPS-based vaccines, in addition to the serotype replacement by other known circulating serotypes observed after the introduction of the 7-valent pneumococcal vaccine, is that strains can shift in serotype and novel serotypes can emerge. In GBS, for example, the major circulating serotypes in the USA and Europe are serotypes Ia, Ib, II, III, and V, which are responsible for over 85% of disease. However, in Japan over 40% of isolates are of serotypes VI and VIII (Lachenauer et al., 1999). Furthermore, serotype V was unknown in the early 1980s and could therefore not account for more disease than the nontypables generally (Harrison et al., 1998; Zaleznik et al., 2000). Today, serotype V accounts for up to 20% of disease in neonates and up to 50% of disease in the elderly.

Serotype replacement and serotype shift are two concepts that are not necessarily related. The frequency with which different circulating serotypes occur often varies over time. This well-known observation may be due to herd immunity and may simply reflect immune selection. The possible emergence of new strains, with similar multilocus sequence typing but different serotype, i.e., seroptype shift, is a different phenomenon that may reflect the emergence of novel strains with increased fitness.

One mechanism of serotype shift is clear from the organization of the genes which code for the synthesis and assembly of the GBS capsule (Cieslewicz et al., 2005). Figure 1 shows schematically the organization of the capsule locus for GBS. The capsule locus of all nine serotypes contains a number of conserved genes likely to be involved in common functions such as sialic acid synthesis and activation and polysaccharide transport. These conserved genes flank glycosyltransferase genes that vary both in number and sequence and are responsible for catalyzing the specific sugar linkages in the polysaccharides of each serotype, thus defining the capsule composition and structure. This appears to be a common theme in the structure of capsule loci from many encapsulated bacteria (Yother, 1999). It is likely that the variation in the glycosyltransferase region is due to horizontal transfer of genes and recombination events. It has been hypothesized that the much higher diversity found in pneumococcal CPS (over 90 serotypes) may reflect a higher frequency of DNA acquisition through natural competence in pneumococcus (Cieslewicz et al., 2005).

The consequence of this mechanism of capsule diversity is that capsule switching can occur and new serotypes can emerge. It is clear that some hypervirulent lineages of *N. meningitidis*, as defined by multilocus sequence typing, contain strains with different serotypes (Brehony et al., 2007) indicating serotype switching within the lineage. This conclusion is supported by analysis of eight GBS genomes (Tettelin et al., 2002, 2005) that clearly show that the genomes cluster to some extent

independently of the capsule type. Hence, although capsular polysaccharide conjugates can be very successful vaccines, capsule diversity needs to be monitored continuously and new serotypes added to the vaccine when necessary.

Protein Antigens

While CPS conjugate-based vaccines have been very successful on the short to medium term, problems of serotype replacement and serotype switching have encouraged an approach to vaccination based on protein antigens. The paradigm, based on the obvious efficacy of early toxoid-based vaccines, was initially to try to understand the mechanisms of disease, identify the major virulence factors, and use inactivated forms to induce protective immunity. This approach has led to limited success. A very successful acellular vaccine against *Bordetella pertussis* based on the identification of major virulence factors, the pertussis toxin, the filamentous hemagglutinin, and pertactin is currently available (Rappuoli, 1996). More recently, a vaccine against the Gram-negative human gastric pathogen, *Helicobacter pylori*, containing three important virulence factors which together confer protection against intragastric *H. pylori* challenge in mouse (Marchetti et al., 1995) and beagle dog (Rossi et al., 2004) models is currently in clinical trials in humans (Ruggiero et al., 2003; more on *H. pylori* later). The identification of these three *H. pylori* virulence factors was a long and complex process. One factor, a high-molecular weight, oligomeric protein exotoxin, VacA (Reyrat et al., 1999), has been shown in mouse models to be involved in the erosion of the gastric epithelium leading to ulceration (Telford et al., 1994). A second factor, CagA, encoded within a pathogenicity island, is directly transferred to gastric epithelial cells through a type IV secretion system and contributes to morphological transformation of the host cell (Bagnoli et al., 2005; Covacci et al., 1999a). The third factor is a protein that contributes to neutrophil activation and hence epithelial inflammation (Montecucco and de Bernard, 2003). Two of these antigens in fact show some variation (Atherton, 2006), and only a combination of all three are predicted to confer protection against the most virulent strains.

On the other hand, attempts to develop a vaccine against group A Streptococcus (*S. pyogenes* or GAS) based on the M protein have encountered major hurdles due to the variability of this protein between strains (Dale, 2008). The M protein is the antigen recognized by sera used for M serotyping discovered by Lancefield and colleagues over 50 years ago (Lancefield and Perlmann, 1952). Early data, in mice and in studies in US military recruits, demonstrated unequivocally that pharyngitis caused by any single M serotype resulted in protection against infection by any other strain of the same serotype but had no effect on infection by other serotypes (Lancefield, 1959). M protein is a major virulence factor in GAS that is involved in evasion of phagocytic attack and adhesion to host tissues (Bisno et al., 2003). However, its intrinsic immunogenicity results in enormous immune pressure to vary. In fact, there are over 130 known M variants and crossprotection between M types is very limited or nonexistent.

More recently, the basis of the second serotyping system of Lancefield, the T-typing system, has, at least in part, been elucidated (Mora et al., 2005). Lancefield described approximately 21 T serotypes, and it has been shown that at least four of these specificities are due to variants in the backbone protein of a previously unidentified pilus-like structure. Immunization with components of these pili has been shown to confer protection against strains carrying the same variant. From the known T serotypes, it is likely that multiple T antigens would be required for a vaccine capable of broad coverage. Nevertheless, there are an estimated 21 T serotypes compared to over 130 M serotypes.

Pharyngitis may seem to be a problem of limited magnitude. Unfortunately GAS is also responsible for several other more serious invasive diseases such as septicemia and necrotizing fasciitis (flesh-eating disease). Furthermore, nonsupurative sequelae of GAS infection include autoimmune rheumatic fever, resulting in cardiac damage and glomerulonephritis, causing severe kidney damage (Cunningham, 2008). The epidemiology of GAS infection suggests a ray of hope in tackling this disease. GAS pharyngitis peaks between the ages of 5 and 8 years, corresponding to the beginning of wider social contact among children, and adults are rarely infected. It is unlikely that the protection in most adults can be explained by having been infected by a sufficient number of M or T serotypes leading to the possibility that there are more common but less immunodominant antigens expressed by GAS that could be potential vaccine candidates.

Other attempts at vaccine development using the components of adhesive pili of Gram-negative bacteria have failed to result in successful vaccines. Both for pathogenic *E. coli* (Morgan et al., 1978) and *Neisseria gonorrhoeae* (Johnson et al., 1991), pilin subunits have been shown to induce protective immune responses in animal models. Nevertheless, the fact that these structures are essential for initial adhesion of the pathogens has resulted in strong immune system-driven variation in the amino acid sequences of the pilin subunits. The net result is that there is so much diversity in the pili that it has not been possible to develop vaccines that could confer broad protection.

Hence, immunodominant antigens that are essential for virulence in most pathogens are under tremendous pressure to vary in order to evade the immune system of the host. This probably accounts for the majority of variation in amino acid sequence of potentially protective antigens. Novel approaches to the design of vaccines will need to be based on the search for antigens that are not exposed to immune pressure during infection and thus do not induce high levels of immunity in the course of disease but which if used together with appropriate additional antigens could lead to protective immunity.

Reverse Vaccinology: A Novel Genomic Approach to Antigen Identification

The genome revolution of the mid-1990s has led to a quite different approach to the identification of vaccine antigens. Rather than trying to identify virulence mechanisms, then identifying the factors involved, then cloning and expressing the genes coding for these antigens, reverse vaccinology makes use of the availability of the complete genome sequence of a pathogen (Rappuoli, 2001). Since the genome sequence gives the blueprint for every potential protein antigen in the bacterium, a rigorous search through the whole proteome must, by definition, identify protective antigens, if they exist. This approach was first used to identify protective antigens for serogroup B *N. meningitidis* (Pizza et al., 2000) which, as mentioned above, has a capsular polysaccharide shared with human tissues. The complete genome sequence of

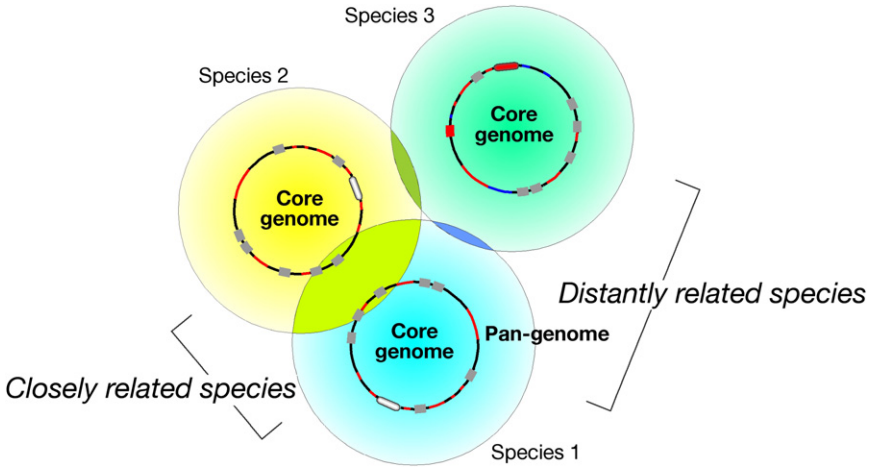

Figure 2. The Pangenome Concept
Individual strains of a species with an open pangenome share a conserved core genome. However, up to 20% of each individual genome consists of genes that are lacking in at least one genome. The number of these "dispensable" genes may be quite large, and hence the total gene pool available to the species would be correspondingly large. Furthermore, dispensable genes may be shared with other species leading to an overlap in pangenomes. This is represented in the illustration as greater or less overlap between the pangenomes of related species.

a strain of meningococcus B revealed about 2158 open reading frames (Tettelin et al., 2000). Of these, 570 were predicted to code for proteins that were either surface associated or secreted. Each of these genes was cloned in *E. coli* expression vectors and recombinant proteins from 350 of the clones were successfully purified. These recombinant antigens were used to immunize mice, and the resulting immune serum was tested for its capacity to promote bacterial killing in a complement dependent bactericidal assay. Serum bactericidal activity (SBA) in this assay is known to correlate with protection in humans immunized with CPS-conjugate vaccines against serogroups A, C, W, and Y.

From this analysis, 28 proteins were identified that could induce SBA activity in mice. However, many of these antigens were found to be highly variable as might be expected due to immune pressure. In fact, no single antigen was capable of inducing SBA activity against all strains in a panel of isolates designed to reflect current population variability. Nevertheless, antigens in combination were found that could confer protection against a broad selection of strains (Giuliani et al., 2006).

Pangenome Reverse Vaccinology

The results obtained with meningococcus B led to a broader approach to identify protein vaccine candidates for GBS: the genomes of eight different strains of GBS representing the major disease causing serotypes were obtained and analyzed (Tettelin et al., 2005). This analysis led to a rather surprising finding. Gene variation between strains was found to be much higher than expected. Only about 80% of the approximately 2100 genes found in any genome were shared by all genomes. About 20% of the genes found in the genomes were lacking in at least one strain. The total number of genes found in the eight strains was almost 50% greater than the size of the average genome. A mathematical analysis of these data indicated that further analysis of more genomes would reveal even more genes. In fact, the extrapolation of the data indicated that every time a new genome was sequenced, one would expect to find between 15 and 30 new genes not present in any of the previous genomes. How far this model will hold up is yet unclear, but it is clear that the number of total genes in the global population of GBS may be very high indeed (Medini et al., 2005). We have termed this concept the pangenome (Tettelin et al., 2005). A schematic representation of the concept of the pangenome is shown in Figure 2.

The probable mechanism of this genomic variability between strains is horizontal transfer of DNA between strains and from outside sources. But the data force the question of why the pangenome is so large. One possible answer may revolve around the requirement for genome efficiency in bacteria. The pressure to be energy efficient drives genomes toward the smallest possible size and the least number of biochemical processes required for the normal life cycle of the bacteria. Hence exclusively human pathogens, e.g., the streptococci and *H. pylori*, have relatively small genomes varying between 1.5 and 2.5 Mbp, whereas *Pseudomonas*, which colonizes many niches and must survive in a variety of environmental conditions has a genome of over 6 Mbp. Perhaps, the very large pangenome allows the global population of bacteria with small genomes to maintain sufficient variability to confront novel conditions that may arise by, for example, evolution of the host and/or increased immunity in the host.

From the eight GBS genomes, a total of 589 predicted surface exposed or secreted proteins were identified, and 312 of these were successfully expressed as soluble proteins and tested in a mouse model of infection and disease. In order to reflect the fact that most GBS neonatal disease occurs in the first week of life and that high titers of maternal IgG can cross the placenta and correlate with protection of the newborn, the model that was used involved immunization of female mice followed by mating and lethal challenge of the pups within the first 48 hr after birth. As in the case of meningococcus B, no single antigen was found that could confer protection against a broad range of strains. Nevertheless, a combination of four antigens did confer broad protection (Maione et al., 2005). Surprisingly, only one of the genes among those encoding the four antigens was shared among all eight strains sequenced, whereas the other three were only present each in a subset of strains. These latter dispensable proteins, surprisingly, were found to be components of novel pilus-like structures (Lauer et al., 2005) with limited variation (Rosini et al., 2006). Hence only one antigen belonged to the core genome, and even this antigen was not universally expressed in all strains. There is a certain sense in this pattern. One might imagine that the core genome of this pathogen, which is capable of chronic ano-genital carriage, must not, by necessity, be particularly immunogenic or confer protection, in contrast to more virulent acute pathogens that rely on new cohorts of nonimmune hosts for their maintenance.

Many more pathogens have now been found to have very large pangenomes. Examples include *E. coli* (Marrs et al., 2005), GAS, and pneumococcus (Medini et al., 2005). However, not all pathogens have open-ended genomes. After the analysis of only four genomes for *Bacillus anthracis*, for example, no new genes were found (Tettelin et al., 2005). The probable reason for this is that *B. anthracis* is believed to be a relatively recent subclone of the very much more variable *Bacillus cereus*.

The Strange Case of *Helicobacter pylori*

As mentioned above, a vaccine against *H. pylori* containing three components is currently in clinical evaluation in humans. While this vaccine is predicted to be widely protective, two of the antigens identified as virulence factors, the vacuolating cytotoxin VacA and CagA, an effector molecule transferred to host cells, are in fact variable (Atherton, 2006). There are two major alleles of VacA and expression of functional VacA varies due to variation in the signal peptide and N terminus of the molecule (Atherton et al., 1995). CagA, on the other hand, is a component of a pathogenicity island coding for a type IV secretion system which is present in only a subset of strains (Censini et al., 1996). Nevertheless, strains containing the pathogenicity island and which express VacA are associated with more serious H. pylori disease including peptic ulcer and gastric cancer (Atherton, 1997).

Genome analysis of *H. pylori* has, however, revealed a high level of diversity. Individual isolates of *H. pylori* differ in gene content, gene similarity, and even genome structure (Tomb et al., 1997; Alm et al., 1999). The major mechanism driving variability in *H. pylori* is believed to be horizontal transfer of DNA between strains and between *H. pylori* and other pathogens due to its capacity to actively acquire foreign DNA (Karnholz et al., 2006). Based on multilocus sequence typing (Karnholz et al., 2006) and microarray analysis (Salama et al., 2000), the core genome of *H. pylori* has been estimated as between 1100 to 1300 of the approximately 1600 genes in the genome of any single strain. Remarkably, individual isolates from the same patient (Israel et al., 2001) often differ in gene content, and isolates from different hosts are generally distinguishable by genome analysis (Salama et al., 2000).

To understand this diversity, one must consider the niche and lifestyle of *H. pylori*. This bacterium is an exclusively human pathogen with an exclusive niche in the stomach. Hence, *H. pylori* requires a very small genome to meet its needs. Furthermore, *H. pylori* colonizes between 30% and 90% of the human population. Colonization usually occurs in childhood, mainly from within the family or close contacts and persists for the life of the individual. Infection causes peptic ulcer disease in about 20% of those infected and is strongly associated with gastric cancer in a smaller number of cases. *H. pylori* has been a constant companion of man for at least thousands of years (Covacci et al., 1999b). Early studies of *H. pylori* indicated that variation in the genes coding for virulence factors VacA and CagA were associated with specific geographic variations. From this observation, it was suggested that *H. pylori* had colonized humans at a very early stage of our evolution and that it had evolved together with humans as they spread around the globe. More recently, this concept has been powerfully substantiated by seminal studies by Suerbaum, Achtman, and colleagues (reviewed in Suerbaum and Achtman, 2004) who have shown a direct association between the *H. pylori* population structure and that of man. They found that the worldwide *H. pylori* population could be subdivided in seven major groups originating in Africa, Central Asia, and East Asia (Falush et al., 2003) and showed that these subtypes associated with human subpopulations or even ethnic groups (Wirth et al., 2004). They conclude that the ancestral *H. pylori* arose in East Africa, where it is believed modern man arose and began his migrations, and that *H. pylori* has coevolved with its human host.

Since *H. pylori* colonizes for the life of the host, despite a vigorous humoral immune response, it is unlikely that the genomic and proteomic variation is due to immune pressure. An explanation of coevolution with its host is much more plausible. It is likely that *H. pylori* has adapted to human diversity, perhaps in major histocompatibility complex, blood group antigens (which are receptors for *H. pylori* adhesins [Aspholm et al., 2006]), and other host-specific diversity in its gastric niche. Hence, while immune pressure is certainly a driving force in bacterial variation, other mechanisms may play important roles in this phenomenon. Independently of the mechanisms driving variation, the result for vaccine development is still the same: novel vaccines most probably will need multiple antigens in order to achieve broad coverage.

Antigen Variability Due to Variations in Gene Expression

Bacteria are under constant evolutionary pressure for efficiency. Energy requirements for DNA replication, transcription, protein synthesis, and macromolecule synthesis account for the major part of energy turnover in the cell. Hence evolutionary pressure selects for small efficient genomes carrying only the genes required to code for the minimally required metabolic pathways necessary for life in the particular niche. The cost of this genome reduction is a reduced flexibility to variations in the environment such as nutrient sources, niche variation, and immune responses. This latter is particularly important for pathogenic species restricted to a particular host. As discussed above, one strategy to deal with environmental variation is to have a large and varied pangenome. In this way, individual bacteria of species such as streptococci may have a relatively small genome, but the total number of dispensable genes in the population permits sufficient flexibility for the species to adapt to environmental challenge. On the other hand, the long association of *H. pylori* with its human host has permitted a continual coevolution as the host environment has altered.

For more rapid responses to environmental changes, bacteria have evolved systems for regulated gene expression. This means that the genome must carry the necessary genes, but considerable energy can be saved by only expressing those genes required for the particular stage of its life cycle. For very rapid responses, gene regulation relies on sensing the environment and regulating the expression of appropriate genes at the transcriptional level through transcriptional regulators. These are frequently two-component systems consisting of a sensor histidine kinase that when activated phosphorylates and thus activates a transcriptional regulator (Mascher et al., 2006). These types of responses occur at the level of individual bacteria.

A second mechanism of gene regulation which acts at the bacterial population level is based on genetic switching on or off of genes by genomic changes. Genes regulated in this fashion are usually referred to as contingency genes. There are two

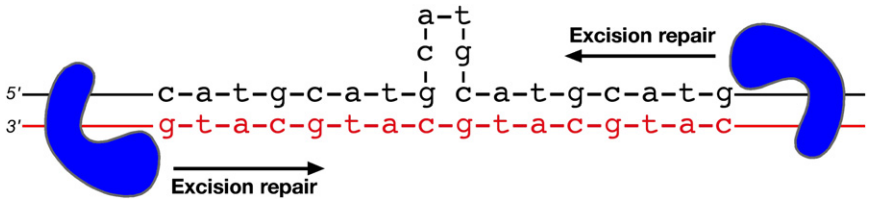

Figure 3. A Schematic Representation of Slipped-Strand Mispairing during DNA Replication

In this example of a four base repeated sequence, the polymerase skips one repeat in the template strand (black), and the nascent strand lacks one repeat. Excision repair can then either repair the nascent strand, such that the number of repeats is maintained, or the template strand, resulting in deletion of a repeat. In bacteria that use the Dam system, methylation of the template strand directs repair predominantly to the nascent strand. In the absence of Dam, the template and nascent strands can be repaired with equal likelihood, thus increasing the frequency of repeat number variation.

major mechanisms of controlling the expression of contingency genes. The first relies on homologous recombination between identical sequences flanking a cassette of gene sequences (Zieg et al., 1977). Recombination results in reversing the direction of the entire cassette such that in one orientation the gene can be transcribed whereas in the other the gene is off (Silverman et al., 1979). This mechanism is dependent on a functional RecA system.

Another mechanism of genetic gene regulation is through short sequence repeats of up to four base pairs (thoroughly reviewed [Moxon et al., 2006]). These regions are highly susceptible to mutation through slipped-strand mispairing during replication that results in either shortening or lengthening the repeat region (Figure 3). For example, extending a short homopolymeric run of a single nucleotide in a coding sequence by one or two nucleotides would result in a shift of reading frame eliminating correct translation of the mRNA. Furthermore, variation in the length of repeated sequences upstream of a coding sequence between the promoter and the transcriptional start site can have profound effects on the efficiency of transcriptional initiation, thus effectively modulating the level of expression of the gene (Dawid et al., 1999; van Ham et al., 1993). Switching genes on or off by this mechanism is independent of RecA but may be to some extent depend on the integrity of DNA repair systems (Bayliss et al., 2002; Bucci et al., 1999; Martin et al., 2004; Richardson and Stojiljkovic, 2001).

This is not a rare mechanism for the control of gene expression. In the genome of *N. meningitides*, there are about 80 different repeated segments predicted to be involved in phase variation (Snyder et al., 2001). Several of these control the on-off switching of expression of surface molecules including capsule synthesis (Hammerschmidt et al., 1996). In serogroup B meningococci, phase variation resulting in loss of capsule synthesis favors cell invasion and may be necessary for invasive disease (Stephens et al., 1993). On the other hand, capsule is an important virulence factor helping to protect the bacteria from both innate and adaptive immune responses in the blood phase of infection (Taylor and Roberts, 2005; Vogel et al., 1996). Phase variation of capsule synthesis in serogroup B meningococci involves a homopolymeric run of dC in the polysialyltransferase gene (*siaD*) (Hammerschmidt et al., 1996). Loss or gain of a single residue results in either an intact open reading frame or a frame shift leading to a truncated protein. Since phase variation at this locus is RecA independent, it is assumed that the variation is due to slipped-strand mispairing (Murphy et al., 1989).

More recently, it has been reported that capsule phase variation in serogroup B meningococcus is at least 100-fold more frequent in strains lacking a DNA adenine methyltransferase (Dam) system than in strains with functional Dam system (Bucci et al., 1999). Since the Dam methylase functions to mark the template strand such that repair of polymerase errors is focused to the newly synthesized strand, in the absence of this activity, the repair can occur equally frequently on both strands thus perpetuating the errors. Interestingly, in this study it was observed that all of 24 invasive strains tested lacked the Dam activity, whereas only 50% of 24 carrier strains lacked this enzyme, suggesting that, at least in part, increased frequency of phase variation may contribute to virulence through favoring rapid changes from cell invasive capsule negative forms to immune system resistant forms in the blood phase of infection (Bucci et al., 1999). This latter hypothesis is very attractive. However, more recent data have questioned the importance of the Dam system in frequency of capsule phase variation (Martin et al., 2004; Richardson and Stojiljkovic, 2001).

Phase variation may not necessarily exclude a protein from being a candidate to confer immune protection. By targeting antigens that are necessary for specific stages during the infectious cycle, for example, factors necessary for survival in the blood of pathogens that cause septicemia, vaccines against even phase variable genes could be successful.

Conclusion

Vaccination is one of the great medical success stories. We are, in a way, fortunate that many of the life-threatening diseases of childhood relied on new nonimmune cohorts of young children for their continued existence and thus did not evolve much variation in virulence factors. Furthermore, zoonotic diseases such as rabies only infrequently infected humans and thus were not under selective pressure. Through the study of natural immunity to these diseases, the concept of immunology was born, which led to vaccines that have essentially eradicated the diseases (at least in the industrialized world). But, the easy ones have all been done. The low-hanging fruits have been picked. Antigenic variation is the bane of vaccinologists. New strategies using multivalent vaccines are beginning to be successful against these more intractable pathogens, but the war on infectious disease has only just begun, and the challenge of viral antigen variation has still to be addressed.

ACKNOWLEDGMENTS

I would like to thank G. Corsi for the artwork. J.L.T. is an employee and shareholder of Novartis Vaccines and Diagnostics Srl.

REFERENCES

Alm, R.A., Ling, L.S., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C., Smith, D.R., Noonan, B., Guild, B.C., deJonge, B.L., et al. (1999). Genomic-sequence

comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. Nature *397*, 176–180.

Aspholm, M., Kalia, A., Ruhl, S., Schedin, S., Arnqvist, A., Linden, S., Sjostrom, R., Gerhard, M., Semino-Mora, C., Dubois, A., et al. (2006). Helicobacter pylori adhesion to carbohydrates. Methods Enzymol. *417*, 293–339.

Atherton, J.C. (1997). The clinical relevance of strain types of Helicobacter pylori. Gut *40*, 701–703.

Atherton, J.C. (2006). The pathogenesis of Helicobacter pylori-induced gastroduodenal diseases. Annu. Rev. Pathol. *1*, 63–96.

Atherton, J.C., Cao, P., Peek, R.M., Jr., Tummuru, M.K., Blaser, M.J., and Cover, T.L. (1995). Mosaicism in vacuolating cytotoxin alleles of Helicobacter pylori. Association of specific vacA types with cytotoxin production and peptic ulceration. J. Biol. Chem. *270*, 17771–17777.

Bagnoli, F., Buti, L., Tompkins, L., Covacci, A., and Amieva, M.R. (2005). Helicobacter pylori CagA induces a transition from polarized to invasive phenotypes in MDCK cells. Proc. Natl. Acad. Sci. USA *102*, 16339–16344.

Baker, C.J., and Kasper, D.L. (1976). Correlation of maternal antibody deficiency with susceptibility to neonatal group B streptococcal infection. N. Engl. J. Med. *294*, 753–756.

Bayliss, C.D., van de Ven, T., and Moxon, E.R. (2002). Mutations in poll but not mutSLH destabilize Haemophilus influenzae tetranucleotide repeats. EMBO J. *21*, 1465–1476.

Bisno, A.L., Brito, M.O., and Collins, C.M. (2003). Molecular basis of group A streptococcal virulence. Lancet Infect. Dis. *3*, 191–200.

Brehony, C., Jolley, K.A., and Maiden, M.C. (2007). Multilocus sequence typing for global surveillance of meningococcal disease. FEMS Microbiol. Rev. *31*, 15–26.

Bucci, C., Lavitola, A., Salvatore, P., Del Giudice, L., Massardo, D.R., Bruni, C.B., and Alifano, P. (1999). Hypermutation in pathogenic bacteria: frequent phase variation in meningococci is a phenotypic trait of a specialized mutator biotype. Mol. Cell *3*, 435–445.

Carrat, F., and Flahault, A. (2007). Influenza vaccine: the challenge of antigenic drift. Vaccine *25*, 6852–6862.

Censini, S., Lange, C., Xiang, Z., Crabtree, J.E., Ghiara, P., Borodovsky, M., Rappuoli, R., and Covacci, A. (1996). cag, a pathogenicity island of Helicobacter pylori, encodes type I-specific and disease-associated virulence factors. Proc. Natl. Acad. Sci. USA *93*, 14648–14653.

Cieslewicz, M.J., Chaffin, D., Glusman, G., Kasper, D., Madan, A., Rodrigues, S., Fahey, J., Wessels, M.R., and Rubens, C.E. (2005). Structural and genetic diversity of group B streptococcus capsular polysaccharides. Infect. Immun. *73*, 3096–3103.

Covacci, A., Telford, J.L., Del Giudice, G., Parsonnet, J., and Rappuoli, R. (1999a). Helicobacter pylori virulence and genetic geography. Science *284*, 1328–1333.

Covacci, A., Telford, J.L., Del Giudice, G., Parsonnet, J., and Rappuoli, R. (1999b). Helicobacter pylori virulence and genetic geography. Science *284*, 1328–1333.

Cunningham, M.W. (2008). Pathogenesis of group A streptococcal infections and their sequelae. Adv. Exp. Med. Biol. *609*, 29–42.

Dale, J.B. (2008). Current status of group A streptococcal vaccine development. Adv. Exp. Med. Biol. *609*, 53–63.

Dawid, S., Barenkamp, S.J., and St Geme, J.W., 3rd. (1999). Variation in expression of the Haemophilus influenzae HMW adhesins: a prokaryotic system reminiscent of eukaryotes. Proc. Natl. Acad. Sci. USA *96*, 1077–1082.

De Jesus, N.H. (2007). Epidemics to eradication: the modern history of poliomyelitis. Virol. J. *4*, 70.

Falush, D., Wirth, T., Linz, B., Pritchard, J.K., Stephens, M., Kidd, M., Blaser, M.J., Graham, D.Y., Vacher, S., Perez-Perez, G.I., et al. (2003). Traces of human migrations in Helicobacter pylori populations. Science *299*, 1582–1585.

Galiza, E.P., and Heath, P.T. (2007). Pneumococcal conjugate vaccines. A review. Minerva Med. *98*, 131–143.

Giuliani, M.M., Adu-Bobie, J., Comanducci, M., Arico, B., Savino, S., Santini, L., Brunelli, B., Bambini, S., Biolchi, A., Capecchi, B., et al. (2006). A universal vaccine for serogroup B meningococcus. Proc. Natl. Acad. Sci. USA *103*, 10834–10839.

Hammerschmidt, S., Muller, A., Sillmann, H., Muhlenhoff, M., Borrow, R., Fox, A., van Putten, J., Zollinger, W.D., Gerardy-Schahn, R., and Frosch, M. (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.

Harrison, L.H., Elliott, J.A., Dwyer, D.M., Libonati, J.P., Ferrieri, P., Billmann, L., and Schuchat, A. (1998). Serotype distribution of invasive group B streptococcal isolates in Maryland: implications for vaccine formulation. Maryland Emerging Infections Program. J. Infect. Dis. *177*, 998–1002.

Hilleman, M.R. (2002). Realities and enigmas of human viral influenza: pathogenesis, epidemiology and control. Vaccine *20*, 3068–3087.

Israel, D.A., Salama, N., Krishna, U., Rieger, U.M., Atherton, J.C., Falkow, S., and Peek, R.M., Jr. (2001). Helicobacter pylori genetic diversity within the gastric niche of a single human host. Proc. Natl. Acad. Sci. USA *98*, 14625–14630.

Johnson, S.C., Chung, R.C., Deal, C.D., Boslego, J.W., Sadoff, J.C., Wood, S.W., Brinton, C.C., Jr., and Tramont, E.C. (1991). Human immunization with Pgh 3–2 gonococcal pilus results in cross-reactive antibody to the cyanogen bromide fragment-2 of pilin. J. Infect. Dis. *163*, 128–134.

Karnholz, A., Hoefler, C., Odenbreit, S., Fischer, W., Hofreuter, D., and Haas, R. (2006). Functional and topological characterization of novel components of the comB DNA transformation competence system in Helicobacter pylori. J. Bacteriol. *188*, 882–893.

Korber, B., Gaschen, B., Yusim, K., Thakallapally, R., Kesmir, C., and Detours, V. (2001). Evolutionary and immunological implications of contemporary HIV-1 variation. Br. Med. Bull. *58*, 19–42.

Lachenauer, C.S., Kasper, D.L., Shimada, J., Ichiman, Y., Ohtsuka, H., Kaku, M., Paoletti, L.C., Ferrieri, P., and Madoff, L.C. (1999). Serotypes VI and VIII predominate among group B streptococci isolated from pregnant Japanese women. J. Infect. Dis. *179*, 1030–1033.

Lancefield, R.C. (1959). Persistence of type-specific antibodies in man following infection with group A streptococci. J. Exp. Med. *110*, 271–292.

Lancefield, R.C., and Perlmann, G.E. (1952). Preparation and properties of type-specific M antigen isolated from a group A, type 1 hemolytic streptococcus. J. Exp. Med. *96*, 71–82.

Lauer, P., Rinaudo, C.D., Soriani, M., Margarit, I., Maione, D., Rosini, R., Taddei, A.R., Mora, M., Rappuoli, R., Grandi, G., and Telford, J.L. (2005). Genome analysis reveals pili in Group B Streptococcus. Science *309*, 105.

Lin, F.Y., Philips, J.B., 3rd, Azimi, P.H., Weisman, L.E., Clark, P., Rhoads, G.G., Regan, J., Concepcion, N.F., Frasch, C.E., Troendle, J., et al. (2001). Level of maternal antibody required to protect neonates against early-onset disease caused by group B Streptococcus type Ia: a multicenter, seroepidemiology study. J. Infect. Dis. *184*, 1022–1028.

Lin, F.Y., Weisman, L.E., Azimi, P.H., Philips, J.B., 3rd, Clark, P., Regan, J., Rhoads, G.G., Frasch, C.E., Gray, B.M., Troendle, J., et al. (2004). Level of maternal IgG anti-group B streptococcus type III antibody correlated with protection of neonates against early-onset disease caused by this pathogen. J. Infect. Dis. *190*, 928–934.

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.

Maione, D., Margarit, I., Rinaudo, C.D., Masignani, V., Mora, M., Scarselli, M., Tettelin, H., Brettoni, C., Iacobini, E.T., Rosini, R., et al. (2005). Identification of a universal Group B streptococcus vaccine by multiple genome screen. Science *309*, 148–150.

Marchetti, M., Arico, B., Burroni, D., Figura, N., Rappuoli, R., and Ghiara, P. (1995). Development of a mouse model of Helicobacter pylori infection that mimics human disease. Science *267*, 1655–1658.

Marrs, C.F., Zhang, L., and Foxman, B. (2005). Escherichia coli mediated urinary tract infections: are there distinct uropathogenic E. coli (UPEC) pathotypes? FEMS Microbiol. Lett. *252*, 183–190.

Martin, P., Sun, L., Hood, D.W., and Moxon, E.R. (2004). Involvement of genes of genome maintenance in the regulation of phase variation frequencies in Neisseria meningitidis. Microbiology *150*, 3001–3012.

Mascher, T., Helmann, J.D., and Unden, G. (2006). Stimulus perception in bacterial signal-transducing histidine kinases. Microbiol. Mol. Biol. Rev. *70*, 910–938.

Medini, D., Donati, C., Tettelin, H., Masignani, V., and Rappuoli, R. (2005). The microbial pangenome. Curr. Opin. Genet. Dev. *15*, 589–594.

Mond, J.J., and Kokai-Kun, J.F. (2008). The multifunctional role of antibodies in the protective response to bacterial T cell-independent antigens. Curr. Top. Microbiol. Immunol. *319*, 17–40.

Montecucco, C., and de Bernard, M. (2003). Molecular and cellular mechanisms of action of the vacuolating cytotoxin (VacA) and neutrophil-activating protein (HP-NAP) virulence factors of Helicobacter pylori. Microbes Infect. *5*, 715–721.

Mora, M., Bensi, G., Capo, S., Falugi, F., Zingaretti, C., Manetti, A.G., Maggi, T., Taddei, A.R., Grandi, G., and Telford, J.L. (2005). Group A Streptococcus produce pilus-like structures containing protective antigens and Lancefield T antigens. Proc. Natl. Acad. Sci. USA *102*, 15641–15646.

Morgan, R.L., Isaacson, R.E., Moon, H.W., Brinton, C.C., and To, C.C. (1978). Immunization of suckling pigs against enterotoxigenic Escherichia coli-induced diarrheal disease by vaccinating dams with purified 987 or K99 pili: protection correlates with pilus homology of vaccine and challenge. Infect. Immun. *22*, 771–777.

Moxon, R., Bayliss, C., and Hood, D. (2006). Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. Annu. Rev. Genet. *40*, 307–333.

Murphy, G.L., Connell, T.D., Barritt, D.S., Koomey, M., and Cannon, J.G. (1989). Phase variation of gonococcal protein II: Regulation of gene expression by slipped-strand mispairing of a repetitive DNA sequence. Cell *56*, 539–547.

Pappenheimer, A.M., Jr. (1977). Diphtheria toxin. Annu. Rev. Biochem. *46*, 69–94.

Pizza, M., Scarlato, V., Masignani, V., Giuliani, M.M., Arico, B., Comanducci, M., Jennings, G.T., Baldi, L., Bartolini, E., Capecchi, B., et al. (2000). Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science *287*, 1816–1820.

Preston, A., Parkhill, J., and Maskell, D.J. (2004). The bordetellae: lessons from genomics. Nat. Rev. Microbiol. *2*, 379–390.

Rappuoli, R. (1996). Acellular pertussis vaccines: a turning point in infant and adolescent vaccination. Infect. Agents Dis. *5*, 21–28.

Rappuoli, R. (2001). Reverse vaccinology, a genome-based approach to vaccine development. Vaccine *19*, 2688–2691.

Reyrat, J.M., Pelicic, V., Papini, E., Montecucco, C., Rappuoli, R., and Telford, J.L. (1999). Towards deciphering the Helicobacter pylori cytotoxin. Mol. Microbiol. *34*, 197–204.

Richardson, A.R., and Stojiljkovic, I. (2001). Mismatch repair and the regulation of phase variation in Neisseria meningitidis. Mol. Microbiol. *40*, 645–655.

Rosini, R., Rinaudo, C.D., Soriani, M., Lauer, P., Mora, M., Maione, D., Taddei, A., Santi, I., Ghezzo, C., Brettoni, C., et al. (2006). Identification of novel genomic islands coding for antigenic pilus-like structures in Streptococcus agalactiae. Mol. Microbiol. *61*, 126–141.

Rossi, G., Ruggiero, P., Peppoloni, S., Pancotto, L., Fortuna, D., Lauretti, L., Volpini, G., Mancianti, S., Corazza, M., Taccini, E., et al. (2004). Therapeutic vaccination against Helicobacter pylori in the beagle dog experimental model: safety, immunogenicity, and efficacy. Infect. Immun. *72*, 3252–3259.

Ruggiero, P., Peppoloni, S., Rappuoli, R., and Del Giudice, G. (2003). The quest for a vaccine against Helicobacter pylori: how to move from mouse to man? Microbes Infect. *5*, 749–756.

Salama, N., Guillemin, K., McDaniel, T.K., Sherlock, G., Tompkins, L., and Falkow, S. (2000). A whole-genome microarray reveals genetic diversity among Helicobacter pylori strains. Proc. Natl. Acad. Sci. USA *97*, 14668–14673.

Schuchat, A. (1998). Epidemiology of group B streptococcal disease in the United States: shifting paradigms. Clin. Microbiol. Rev. *11*, 497–513.

Schuchat, A. (1999). Group B streptococcus. Lancet *353*, 51–56.

Silverman, M., Zieg, J., Hilmen, M., and Simon, M. (1979). Phase variation in Salmonella: genetic analysis of a recombinational switch. Proc. Natl. Acad. Sci. USA *76*, 391–395.

Snyder, L.A., Butcher, S.A., and Saunders, N.J. (2001). Comparative whole-genome analyses reveal over 100 putative phase-variable genes in the pathogenic Neisseria spp. Microbiology *147*, 2321–2332.

Sood, R.K., and Fattom, A. (1998). Capsular polysaccharide-protein conjugate vaccines and intravenous immunoglobulins. Expert Opin. Investig. Drugs *7*, 333–347.

Stein, D.M., Robbins, J., Miller, M.A., Lin, F.Y., and Schneerson, R. (2006). Are antibodies to the capsular polysaccharide of Neisseria meningitidis group B and Escherichia coli K1 associated with immunopathology? Vaccine *24*, 221–228.

Stephens, D.S., Spellman, P.A., and Swartley, J.S. (1993). Effect of the (alpha 2→8)-linked polysialic acid capsule on adherence of Neisseria meningitidis to human mucosal cells. J. Infect. Dis. *167*, 475–479.

Stephens, D.S., Greenwood, B., and Brandtzaeg, P. (2007). Epidemic meningitis, meningococcaemia, and Neisseria meningitidis. Lancet *369*, 2196–2210.

Suerbaum, S., and Achtman, M. (2004). Helicobacter pylori: recombination, population structure and human migrations. Int. J. Med. Microbiol. *294*, 133–139.

Taylor, C.M., and Roberts, I.S. (2005). Capsular polysaccharides and their role in virulence. Contrib. Microbiol. *12*, 55–66.

Telford, J.L., Ghiara, P., Dell'Orco, M., Comanducci, M., Burroni, D., Bugnoli, M., Tecce, M.F., Censini, S., Covacci, A., Xiang, Z., et al. (1994). Gene structure of the Helicobacter pylori cytotoxin and evidence of its key role in gastric disease. J. Exp. Med. *179*, 1653–1658.

Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., et al. (2000). Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. Science *287*, 1809–1815.

Tettelin, H., Masignani, V., Cieslewicz, M.J., Eisen, J.A., Peterson, S., Wessels, M.R., Paulsen, I.T., Nelson, K.E., Margarit, I., Read, T.D., et al. (2002). Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V Streptococcus agalactiae. Proc. Natl. Acad. Sci. USA *99*, 12391–12396.

Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S.V., Crabtree, J., Jones, A.L., Durkin, A.S., et al. (2005). Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pangenome". Proc. Natl. Acad. Sci. USA *102*, 13950–13955.

Toltzis, P., and Jacobs, M.R. (2005). The epidemiology of childhood pneumococcal disease in the United States in the era of conjugate vaccine use. Infect. Dis. Clin. North Am. *19*, 629–645.

Tomb, J.F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., et al. (1997). The complete genome sequence of the gastric pathogen Helicobacter pylori. Nature *388*, 539–547.

van Ham, S.M., van Alphen, L., Mooi, F.R., and van Putten, J.P. (1993). Phase variation of *H. influenzae* fimbriae: Transcriptional control of two divergent genes through a variable combined promoter region. Cell *73*, 1187–1196.

Vogel, U., Hammerschmidt, S., and Frosch, M. (1996). Sialic acids of both the capsule and the sialylated lipooligosaccharide of Neisseria meningitis serogroup B are prerequisites for virulence of meningococci in the infant rat. Med. Microbiol. Immunol. (Berl.) *185*, 81–87.

Wirth, T., Wang, X., Linz, B., Novick, R.P., Lum, J.K., Blaser, M., Morelli, G., Falush, D., and Achtman, M. (2004). Distinguishing human ethnic groups by means of sequences from Helicobacter pylori: lessons from Ladakh. Proc. Natl. Acad. Sci. USA *101*, 4746–4751.

Yother, J. (1999). Common themes in the genetics of streptococcal capsular polysaccharides. In Genetics of Bacterial Polysaccharides, J. Goldberg, ed. (Boca Raton, FL: CRC Press), pp. 161–184.

Zaleznik, D.F., Rench, M.A., Hillier, S., Krohn, M.A., Platt, R., Lee, M.L., Flores, A.E., Ferrieri, P., and Baker, C.J. (2000). Invasive disease due to group B Streptococcus in pregnant women and neonates from diverse population groups. Clin. Infect. Dis. *30*, 276–281.

Zieg, J., Silverman, M., Hilmen, M., and Simon, M. (1977). Recombinational switch for gene expression. Science *196*, 170–172.