EXTENDED PDF FORMAT
SPONSORED BY
Travel Grants to the
15th International Congress
of Immunology
Apply Now!
R&D SYSTEMS
www.rndsystems.com

Rapid Pneumococcal Evolution in Response to Clinical Interventions

Nicholas J. Croucher *et al.*

Science **331**, 430 (2011);
DOI: 10.1126/science.1198545

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of April 22, 2013):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/331/6016/430.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2011/01/24/331.6016.430.DC1.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/331/6016/430.full.html#related

This article **cites 31 articles**, 19 of which can be accessed free:
http://www.sciencemag.org/content/331/6016/430.full.html#ref-list-1

This article has been **cited by** 61 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/331/6016/430.full.html#related-urls

This article appears in the following **subject collections:**
Microbiology
http://www.sciencemag.org/cgi/collection/microbio

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2011 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

Rapid Pneumococcal Evolution in Response to Clinical Interventions

Nicholas J. Croucher,[1] Simon R. Harris,[1] Christophe Fraser,[2] Michael A. Quail,[1] John Burton,[1] Mark van der Linden,[3] Lesley McGee,[4] Anne von Gottberg,[5] Jae Hoon Song,[6] Kwan Soo Ko,[7] Bruno Pichon,[8] Stephen Baker,[9] Christopher M. Parry,[9] Lotte M. Lambertsen,[10] Dea Shahinas,[11] Dylan R. Pillai,[11] Timothy J. Mitchell,[12] Gordon Dougan,[1] Alexander Tomasz,[13] Keith P. Klugman,[4,5,14] Julian Parkhill,[1] William P. Hanage,[2,15] Stephen D. Bentley[1]*

Epidemiological studies of the naturally transformable bacterial pathogen *Streptococcus pneumoniae* have previously been confounded by high rates of recombination. Sequencing 240 isolates of the PMEN1 (Spain^{23F}-1) multidrug-resistant lineage enabled base substitutions to be distinguished from polymorphisms arising through horizontal sequence transfer. More than 700 recombinations were detected, with genes encoding major antigens frequently affected. Among these were 10 capsule-switching events, one of which accompanied a population shift as vaccine-escape serotype 19A isolates emerged in the USA after the introduction of the conjugate polysaccharide vaccine. The evolution of resistance to fluoroquinolones, rifampicin, and macrolides was observed to occur on multiple occasions. This study details how genomic plasticity within lineages of recombinogenic bacteria can permit adaptation to clinical interventions over remarkably short time scales.

Streptococcus pneumoniae is a highly recombinogenic human nasopharyngeal commensal and respiratory pathogen estimated to be responsible for a global burden of almost 15 million cases of invasive disease in 2000 (*1*). Since the 1970s, the susceptibility of the pneumococcal population to antibiotics has decreased, largely as a consequence of the emergence and spread of a few multidrug-resistant clones (*2*). The first recognized example was Pneumococcal Molecular Epidemiology Network clone 1 (PMEN1), an *S. pneumoniae* lineage typically identified as being sequence type (ST) 81 and serotype 23F, as well as exhibiting resistance to multiple antibiotics, including penicillin. The genome sequence of the first identified member of the clone, isolated in a hospital in Barcelona in 1984, revealed that it had acquired a Tn*5252*-type integrative and conjugative element (ICE) that carries a linearized chloramphenicol resistance plasmid and a Tn*916*-type element with a *tetM* tetracycline-resistance gene (*3*).

This lineage was subsequently found to be present in Africa, Asia, and America (*4*–*8*) and, by the late 1990s, was estimated to be causing almost 40% of penicillin-resistant pneumococcal disease in the USA (*9*). Following the introduction of a heptavalent conjugate polysaccharide vaccine (PCV7) in many countries since 2000, which includes capsule type 23F as one of its seven antigens, a decrease in the frequency of serotype 23F invasive disease and carriage has been observed (*10*). However, this has been accompanied by a rise in disease caused by nonvaccine-serotype pneumococci, such as the multidrug-resistant serotype 19A strains now common in the USA (*11*). Based on evidence from multilocus sequence typing (MLST) from America (*12*) and Europe (*13*), some of these are thought to include capsule switch variants of PMEN1 lineage.

To study how this lineage has evolved as it has spread, we used Illumina sequencing of multiplexed genomic DNA libraries to characterize a global collection of 240 PMEN1 strains isolated between 1984 and 2008. Strains were identified either by using MLST or on the basis of serotype, drug-resistance profile, and targeted polymerase chain reaction (*14*). Selected isolates were distributed among Europe (seven countries, 81 strains); South Africa (37 strains); America (six countries, 54 strains); and Asia (eight countries, 68 strains) (table S1) and included a variety of drug-resistance profiles, as well as five serotypes distinct from the ancestral 23F: 19F (also included in PCV7), 19A, 6A, 15B, and 3.

Construction of the phylogeny. Sequence reads were mapped against the complete reference chromosome of *S. pneumoniae* ATCC 700669 (*3*) and, by using the criteria described in Harris *et al*. (*15*), 39,107 polymorphic sites were identified within the PMEN1 lineage. Maximum likelihood analysis produced a phylogeny with a high proportion of homoplasic sites (23%) and a weak correlation between the date of a strain's isolation and its distance from the root of the tree (Pearson correlation, $N = 222$, $R^2 = 0.05$, $p = 0.001$) (fig. S1), which suggested that variation was primarily arising through incorporation of imported DNA and not through steady accumulation of base substitutions. As these strains are closely related, sequences acquired by recombination could be identified as loci with a high density of polymorphisms. These events were reconstructed onto the phylogeny and, by using an iterative algorithm (*14*), an alignment and tree based on vertically inherited base substitutions alone was generated.

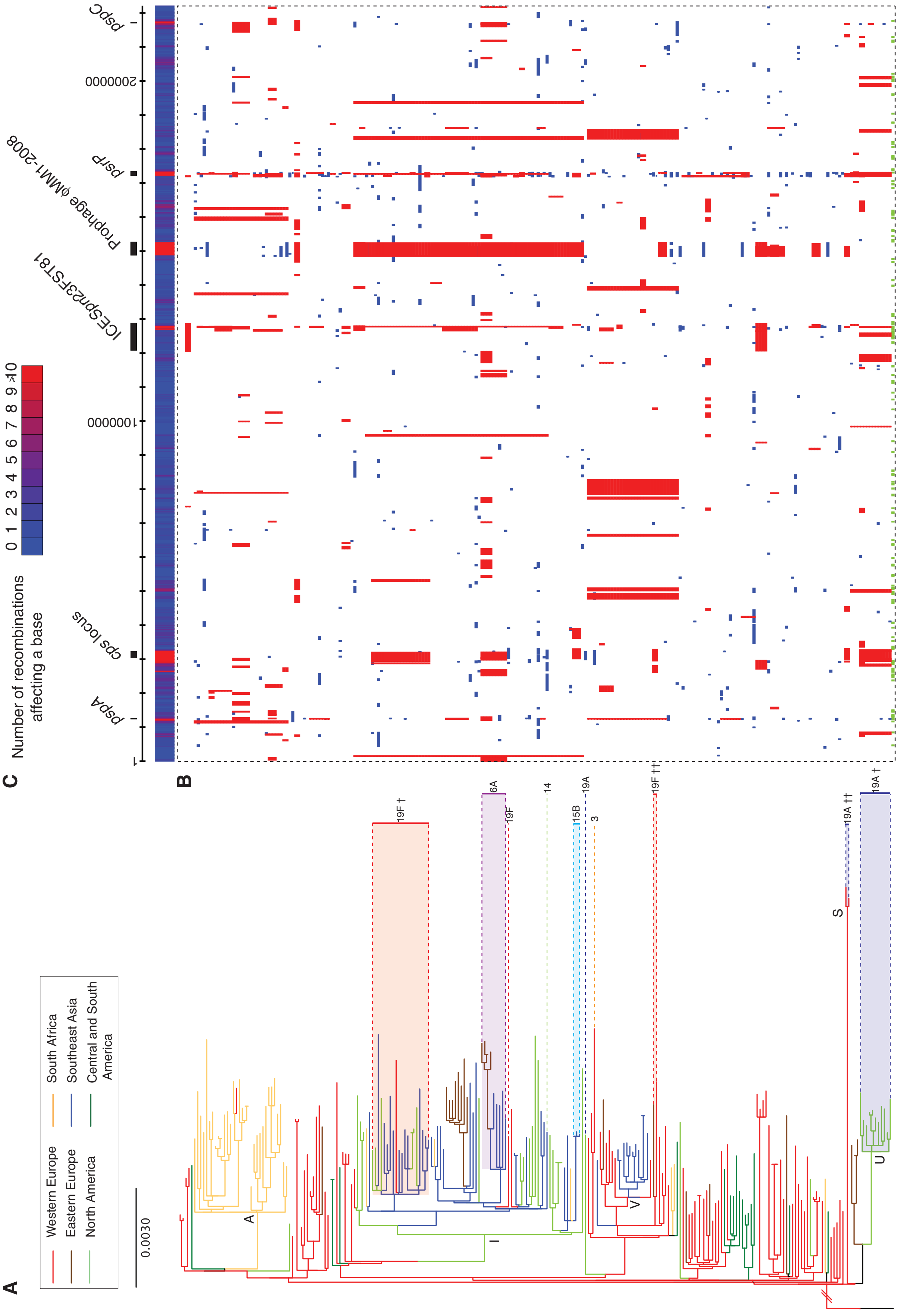
From this analysis (Fig. 1), a total of 57,736 single-nucleotide polymorphisms (SNPs) were identified, 50,720 (88%) of which were introduced by 702 recombination events. This gives a

Fig. 1. Phylogeography and sequence variation of PMEN1. (**A**) Global phylogeny of PMEN1. The maximum likelihood tree, constructed using substitutions outside of recombination events, is colored according to location, as reconstructed through the phylogeny by using parsimony. Shaded boxes and dashed lines indicate isolates that have switched capsule type from the ancestral 23F serotype. †Independent switches to the same serotype are distinguished by annotation with daggers. Specific clades referred to in the text are marked on the tree: A (South Africa), I (International), V (Vietnam), S (Spain 19A), and U (USA 19A). (**B**) Recombinations detected in PMEN1. The panel shows the chromosomal locations of the putative recombination events detected in each terminal taxon. Red blocks are recombinations predicted to have occurred on an internal branch and, therefore, are shared by multiple isolates through common descent. Blue blocks are recombinations predicted to occur on terminal branches and hence are present in only one strain. The green blocks indicate recombinations predicted to have occurred along the branch to the outgroup (*S. pneumoniae* BM4200), used to root the tree. (**C**) Biological relevance of recombination. The heat map shows the density of independent recombination events within PMEN1 in relation to the annotation of the reference genome. All regions that have undergone 10 or more recombination events are marked and annotated (Tn*916* is encompassed within ICE*Spn*23FST81).

---

[1] The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK. [2] Department of Infectious Disease Epidemiology, Imperial College, St Mary's Campus, Norfolk Place, London W2 1PG, UK. [3] Institute for Medical Microbiology, National Reference Center for Streptococci, University Hospital, RWTH Aachen, Pauwelsstrasse 30, 52074 Aachen, Germany. [4] Respiratory Diseases Branch, Centers for Disease Control and Prevention, Atlanta, GA 30333, USA. [5] Respiratory and Meningeal Pathogens Research Unit, National Institute for Communicable Diseases of the National Health Laboratory Service and University of Witwatersrand, Johannesburg 2000, South Africa. [6] Samsung Medical Centre, Sungkyunkwan University School of Medicine and Asia Pacific Foundation for Infectious Disease, Seoul, South Korea. [7] Department of Molecular Cell Biology, Sungkyunkwan University School of Medicine, Suwon 440-746, South Korea. [8] Respiratory and Systemic Infection Laboratory, Health Protection Agency Centre for Infections, London NW9 5HT, UK. [9] The Hospital for Tropical Diseases, Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam. [10] Department of Microbiological Surveillance and Research, Statens Serum Institut, 2300 Copenhagen S, Denmark. [11] Department of Laboratory Medicine and Pathobiology, University of Toronto and Ontario Agency for Health Protection and Promotion, Toronto, Ontario, M5G 1V2, Canada. [12] Institute of Infection, Immunity and Inflammation, University of Glasgow, Glasgow G12 8TA, UK. [13] Laboratory of Microbiology, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA. [14] Hubert Department of Global Health, Rollins School of Public Health. and Division of Infectious Diseases, School of Medicine, Emory University, Atlanta, GA 30322, USA. [15] Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA.

*To whom correspondence should be addressed. E-mail sdb@sanger.ac.uk.


A
Western Europe
Eastern Europe
North America
South Africa
Southeast Asia
Central and South America
0.0030
A
I
V
S
U
19F †
6A
19F
14
15B
19A
3
19F ††
19A ††
19A †
B
C
Number of recombinations affecting a base
0 1 2 3 4 5 6 7 8 9 >10
1
pspA
cps locus
1000000
ICESpn23FST81
Prophage φMM1-2008
psrP
2000000
pspC

per site r/m ratio (the relative likelihood that a polymorphism was introduced through recombination rather than point mutation) of 7.2, less than the previously calculated value of ~66 from MLST data (*16*). By removing recombination events from the phylogeny, the number of homoplasic sites is reduced by 97%, and the tree has significantly shortened branches, such that root-to-tip distance more strongly correlates with date of isolation ($R^2 = 0.46$, $p = <2.2 \times 10^{-16}$; figs. S1 and S2). The rate at which base substitutions occur outside of recombinations suggests a mutation rate of 1.57×10^{-6} substitutions per site per year (95% confidence interval 1.34 to 1.79×10^{-6}), close to the estimate of 3.3×10^{-6} substitutions per site per year from *Staphylococcus aureus* ST239 (*15*) and much higher than that found between more distantly related isolates (*17*). Furthermore, by excluding SNPs introduced through recombinations, the date of origin of the lineage implied by the tree moved from about 1930—which predates the introduction of penicillin, chloramphenicol, and tetracycline—to about 1970 (fig. S1).

Widespread recombination and antigenic variation. Even in this sample of a single lineage, 74% of the reference genome length has undergone recombination in at least one isolate, with a mean of 74,097 base pairs (bp) of sequence affected by recombination in each strain. This encompasses both site-specific integrations of prophage and conjugative elements and homologous recombinations mediated by the competence system. The 615 recombinations outside of the prophage and ICE vary in size from 3 bp to 72,038 bp, with a mean of 6.3 kb (fig. S3). Within these homologous recombinations, there is a distinct heterogeneity in the density of polymorphisms, although it is unclear whether this represents a consequence of the mechanism by which horizontally acquired DNA is incorporated or a property of the donor sequence.

Recombination hotspots are evident in the genome where horizontal sequence transfers are detected abnormally frequently (Fig. 1). One of the most noticeable is within Tn*916*, concentrated around the *tetM* gene. Excepting the prophage, the other loci—*pspA*, *pspC*, *psrP*, and the capsule biosynthesis (*cps*) locus—are all major surface structures. PspA and PspC are potential protein vaccine targets implicated in pneumococcal pathogenesis that are targeted by antibodies produced during experimental human carriage studies (*18*). PsrP is a large ~4500 amino acid serine-rich protein, present in a subset of pneumococci, that is likely to be modified by a number of glycosyltransferases that are encoded on the same genomic island. In a mouse model of infection, PsrP-targeting antibodies can block pneumococcal infection (*19*). Hence, it seems likely that these loci are under diversifying selection driven by the human immune system, and consequently, the apparent increase in the frequency of recombination in these regions is due to the selective advantage that is offered by the divergent sequence introduced by such recombination events.

In addition to base substitutions, 1032 small (<6-bp) insertion and deletion events can be reconstructed onto the phylogeny, of which 61% are concentrated in the 13% of the genome that does not encode for protein-coding sequences (CDSs), probably because of selection against the introduction of frameshift mutations. Throughout the phylogeny, 331 CDSs are predicted to be affected by either frameshift or premature stop codon mutations. Modeling these disruptive events as a Poisson distributed process occurring at a rate proportional to the length of the CDS, 11 CDSs were significantly enriched for disruptive mutations after correction for multiple testing (table S5). These included *pspA* and a glycosyltransferase posited to act on *psrP* (SPN23F17730). This again suggests there may be a selective pressure acting either to remove (*pspA*) or alter (*psrP*) two major surface antigens. Furthermore, the longest recombination in the data set spans, and deletes, the *psrP*-encoding island, which shows that such nonessential antigens can be quickly removed from the chromosome. These data imply that the pneumococcal population is likely to be able to respond very rapidly to the introduction of some of the protein antigen–based pneumococcal vaccines currently under development.

Population and serotype dynamics. The spread of PMEN1 can be tracked by using the phylogeography indicated by the tree (Fig. 1). There are several European clades with their base near the root of the tree, and a parsimony-based reconstruction of location supports a European origin for the lineage. Interspersed among the European isolates are samples from Central and South America, which may represent an early transmission from Spain, where the clone was first isolated, to Latin America, a route previously suggested to occur by data from *S. aureus* (*15*). One clade (labeled A in Fig. 1), containing South African isolates from 1989 to 2006, appears to have originated from a single highly successful intercontinental transmission event. There is also a cluster of isolates from Ho Chi Minh City (labeled V), representing a transmission to Southeast (SE) Asia. However, the predominant clade found outside of Europe (labeled I) appears to have spread quite freely throughout North America, SE Asia, and Eastern Europe, which implies that there are few barriers to intercontinental transmission of *S. pneumoniae* between these regions.

The final non-European group consists of serotype 19A U.S. isolates (labeled U). These all date from between 2005 and 2007 and are distinct from all other U.S. PMEN1 isolates, which have capsular types included in PCV7 (fig. S4). This is evidence of a shift in the PMEN1 population in the USA: Rather than a change in capsule type occurring among the resident population, it seems that it has been eliminated by the vaccine and replaced by a different subpopulation within the lineage that has expanded to fill the vacated niche. Similarly, a pair of Spanish isolates from 2001 (labeled S), the year in which PCV7 was introduced in Spain, that have independently acquired a 19A capsule are not closely associated with any

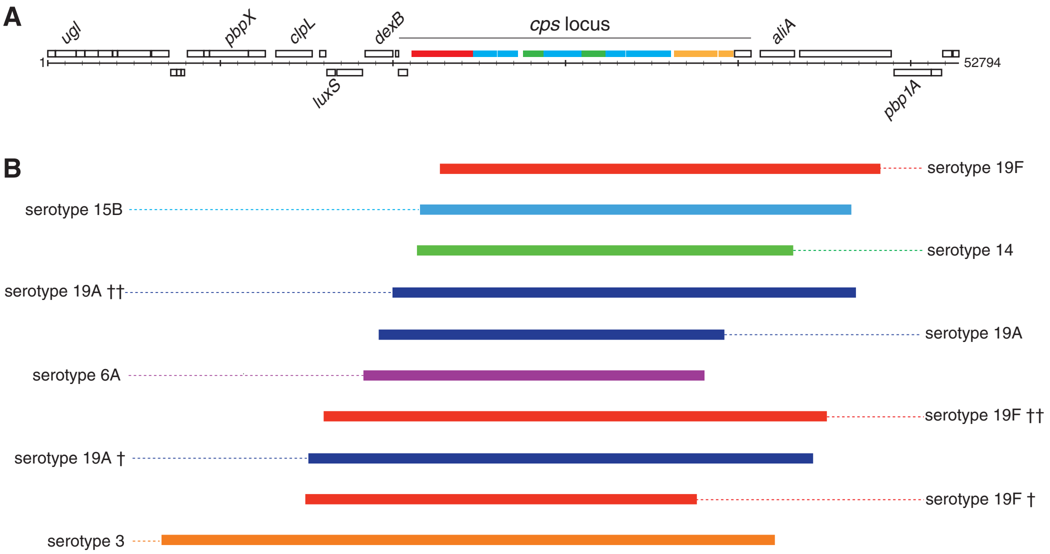


Fig. 2. Recombinations causing serotype-switching events. (**A**) The annotated *cps* locus of the reference strain. CDSs involved in capsule biosynthesis are colored according to their role. Genes in red are regulatory, those in blue synthesize and modify the oligosaccharide subunits, those in green are involved in polymerization and transport, and those in orange are required for the synthesis of rhamnose. (**B**) Below are delineated the recombinations leading to changes in serotype, colored according to the serotype of the sequence donor. The different events are labeled to correspond with Fig. 1.

other European isolates. The estimated times of origin for clades U (1996; 95% credible interval 1992–1999) and S (1998; 95% credible interval 1996–1999) both predate the introduction of PCV7, and accordingly a third 19A switch, from Canada, was isolated in 1994. Hence, it appears that these changes in serotype after vaccine introduction result from an expansion of preexisting capsular variants, which were relatively uncommon and not part of the predominant population, and would have therefore been difficult to detect before the existence of the selection pressure exerted by the vaccine.

Seven further serotype-switching events can be detected in the data (Fig. 2), including three switches to serotype 19F. The polyphyletic nature of these 19F isolates is supported by the variation observed between the acquired *cps* loci, as is also the case for the 19A isolates (fig. S5). The previously known switches to serotypes 3, 6A, and 15B are only found to occur once each in the phylogeny, and in addition, a single Korean sample that had not been typed was identified as a serotype 14 variant by mapping reads to known *cps* loci (*20*). The recombination events leading to these switches ranged from 21,780 bp to 39,182 bp in size, with a mean of 28.2 kb. Only 35 homologous recombinations of an equivalent size or larger occur elsewhere in the genome; most such events are much smaller (fig. S3), which makes it surprising that serotype switching occurs with such frequency and which indicates a role for balancing selection at this locus. Additionally, the span of these events appears to be limited by the flanking penicillin-binding protein genes, the sequences of which are crucial in determining β-lactam resistance in pneumococci (*21*). Only the recombination causing the switch to serotype 3 affects one of these, and it introduces just a single SNP into the *pbpX* CDS, which does not appear to compromise the strain's penicillin resistance (table S1). Hence, the positioning of these two genes may hinder the transfer of capsule biosynthesis operons from penicillin-sensitive to penicillin-resistant pneumococci via larger recombinations, although size constraints alone could also cause such a distribution.

Resistance to non–β-lactam antibiotics. The strong selection pressures exerted by antibiotics on the PMEN1 lineage are manifest as multiple examples of geographically disparate isolates converging on common resistance mechanisms. Single base substitutions causing reduced susceptibility to some classes of antibiotics have occurred multiple times throughout the phylogeny, as observed in *S. aureus* (*15*) and *Salmonella* Typhi (*22*) populations, including mutations in *parC*, *parE*, and *gyrA*, which cause increased resistance to fluoroquinolone antibiotics (*23*), and changes in *rpoB* causing resistance to rifampicin (*24*). The S79F, S79Y, and D83N mutations (*25*) in *parC* are estimated to occur nine, three, and five times, respectively, in PMEN1; additionally, D435N in the adjacent *parE* gene is found to happen three times. The S81F and S81Y substitutions, in the same position of *gyrA*, are found four and two times, respectively. None of these mutations are predicted to have been introduced by recombination, whereas changes at position H499 of *rpoB* causing rifampicin resistance are introduced twice by horizontal transfer and three times by means of base substitution.

Resistance to macrolide antibiotics tends not to derive from SNPs, but from acquisition of CDSs facilitating one of the two common resistance mechanisms: methylation of the target ribosomal RNA by *erm* genes and removal of the drug from the cell by the macrolide efflux (*mef*)-type efflux pumps. Both can be found in the PMEN1 population, and in all cases, the genes appear to be integrated into the Tn*916* transposon (Fig. 3). They are carried by three different elements. Tn*917*, consisting of an *ermB* gene with an associated transposon and resolvase, inserts into open reading frame *orf9* of Tn*916* (*26*). A second has been characterized as the macrolide efflux genetic assembly (mega) element (*27*), which carries a *mef*/*mel* efflux pump system and, in PMEN1, inserts upstream of *orf9*. A third element (henceforth referred to as an omega element, for omega and multidrug-resistance

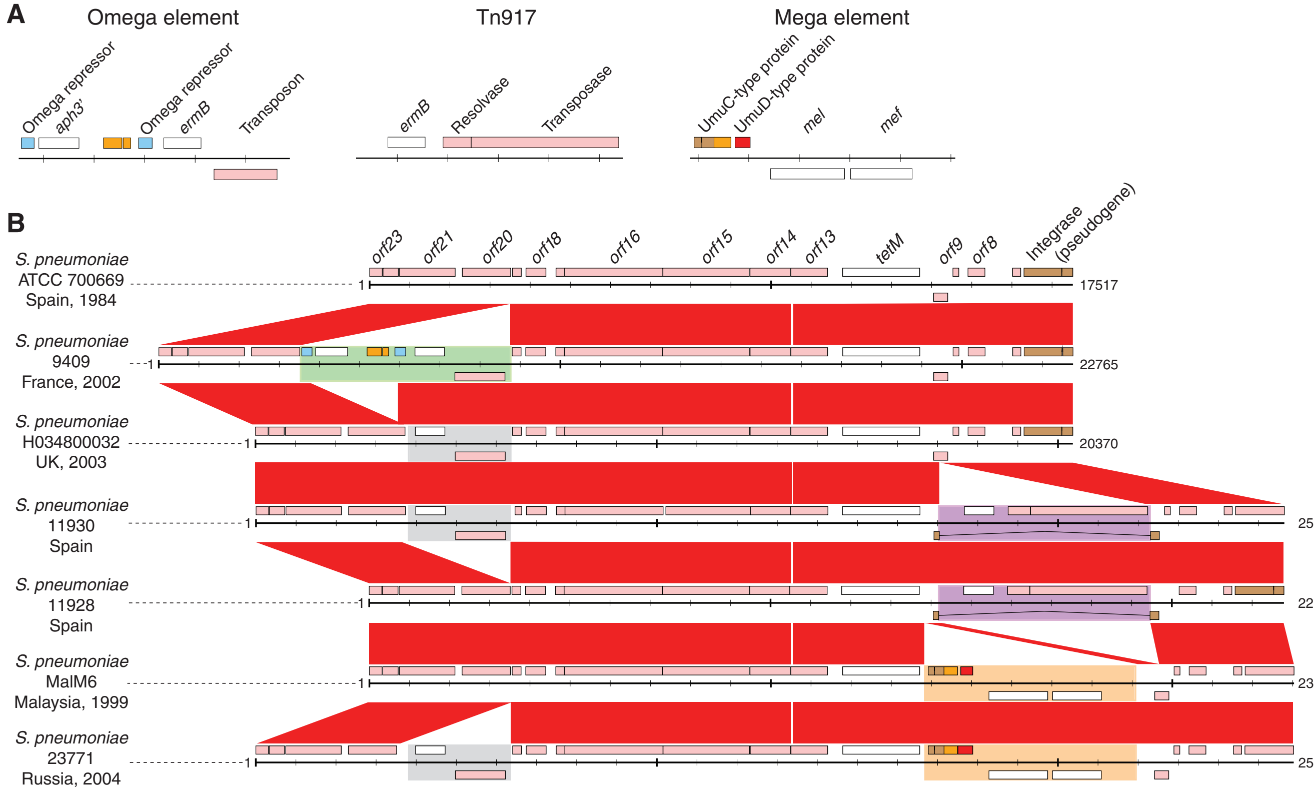


Fig. 3. Acquisition of macrolide-resistance cassettes. The three full-length resistance cassettes are shown in (**A**): the omega element, which carries an *aph3'* aminoglycoside-resistance gene and an *ermB* macrolide-resistance gene; Tn*917*, which carries just the *ermB* methylase; and the mega element, which carries the *mel*/*mef* macrolide efflux system. (**B**) A comparison of the different Tn*916* variants in the PMEN1 lineage. Red bands between the sequences indicate BLASTN matches. The omega element is shaded green when present at full length and shaded gray where present as a remnant resulting from a recombination between the omega repressor–encoding genes, which concomitantly leads to the fusion of an omega transcriptional repressor domain to the 3′ of the *orf20* CDS. Tn*917* is boxed in purple, with the two parts of *orf9*, into which it inserts, indicated on either side. The mega element is boxed in orange.

encoding genetic assembly) carries both an *ermB* gene and an aminoglycoside phosphotransferase, with the latter flanked by direct repeats of omega transcriptional repressor genes, and is found just downstream of *orf20*.

Rather than a single acquisition of these elements occurring, and the resulting clones spreading and replacing macrolide-sensitive isolates, all three elements appear to have been acquired multiple times across the phylogeny (fig. S10). The mega element is predominantly shared by isolates in clade I, although the *ermB*-encoding omega element appears to have been subsequently acquired on two occasions, and Tn*917* has entirely superseded the mega element in one isolate. This is congruent with the known advantages of target methylation over drug efflux as a broader-spectrum resistance mechanism (*28*). In most instances of the omega element, only the *ermB*-encoding part remains; the aminoglycoside phosphotransferase appears to have been deleted through a recombination between the omega-encoding genes, which leaves only an omega domain–encoding open reading frame fused to *orf20* as a scar. This implies that the benefit of the aminoglycoside-resistance element may have not been sufficient to maintain it on the ICE.

Components of the accessory genome. Other than the insertion of these cassettes, the ICE itself is otherwise relatively unchanged throughout the population. In two cases, the 5′ region of the element up to, and including, the lantibiotic synthesis machinery is deleted, whereas the self-immunity genes are retained (fig. S6). This deletion, which also removes the integrated chloramphenicol-resistance plasmid, is analogous to that observed in the pneumococcal pathogenicity island–1 of the PMEN1 lineage, in which all that remains are the immunity genes from a once-intact lantibiotic synthesis machinery (*3*). In two other cases, the ICE has been supplanted by alternative transposons, both of which are similar composites of Tn*5252*- and Tn*916*-type elements: In *S. pneumoniae* 11876, a wholesale replacement at the same locus entails the gain of an omega element at the expense of losing resistance to chloramphenicol (fig. S7), whereas, in isolate 11930, the new ICE inserts elsewhere in the chromosome and carries two *ermB* genes, as well as a chloramphenicol acetyltransferase (fig. S8). The only other identified conjugative element was an ICE*St*1-type transposon shared by isolates 8140 and 8143 (fig. S9), and the only extrachromosomal element present in the data set was the plasmid p*Spn*1 (*29*), found in isolate SA8.

The accessory genome is primarily composed of prophage sequence (fig. S11), with little evidence of much variation in the complement of metabolic genes. Viral sequences appear to be a transient feature of the pneumococcal chromosome (fig. S12), with few persisting long enough to be detected in related isolates. Four of the new prophage that could be assembled were found to insert into the competence pilus structural gene *comYC*, which lies within an operon shown to be essential for competence in *S. pneumoniae* (*30*). In two cases where such phage appear to be shared through common descent by pairs of isolates, no recombination events can be detected that are unique to either member of the pair, consistent with a nonfunctional competence system in these isolates. Furthermore, assaying the competence of available lysogenic strains in vitro also suggested that these phage insertions abrogate the ability of their host to take up exogenous DNA (fig. S13).

Discussion. The ability to distinguish vertically acquired substitutions from horizontally acquired sequences is crucial to successfully reconstructing phylogenies for recombinogenic organisms such as *S. pneumoniae*. Phylogenies are in turn essential for detailed studies of events such as intercontinental transmission, capsule type switching, and antibiotic-resistance acquisition. Although current epidemiological typing methods have indicated that recombination is frequent among the pneumococcal population, they cannot sufficiently account for its impact on relations between strains at such high resolution. Only the availability of such a sample of whole-genome sequences makes it possible to adequately reconstruct the natural history of a lineage. The base substitutions used to construct the phylogeny have accumulated over about 40 years and occur, on average, once every 15 weeks. Recombinations happen at a rate about 1/10th as fast but introduce a mean of 72 SNPs each. The responses to the different anthropogenic selection pressures acting on this variation are distinct. The apparently weak selection by aminoglycosides and chloramphenicol has led to the occasional deletion of loci encoding resistance to these antibiotics. By contrast, resistance to macrolide antibiotics has been acquired frequently throughout the phylogeny, with selection strong enough to drive supplementation or replacement of the resistance afforded by the *mef* efflux pump with the broader-range resistance provided by *ermB*-mediated target modification. The response to vaccine selection is different and involves the depletion of the resident population before it can respond to the selection pressure and thereby opens the niche to isolates that already expressed nonvaccine serotypes. This is likely to reflect the high host population coverage of PCV7 in the USA, as opposed to macrolides or other antibiotics, and the relative likelihood of the recombination events that underlie these responses.

Over a few decades, this single pneumococcal lineage has acquired drug resistance and the ability to evade vaccine pressure multiple times, demonstrating the remarkable adaptability of recombinogenic bacteria such as the pneumococcus. PMEN1 is, nevertheless, only one lineage of this pathogen. Our relative ignorance of the forces that affect bacterial evolution over the long term is illustrated by BM4200 (*31*), a multidrug-resistant serotype 23F isolate of ST1010 sequenced as the outgroup for this analysis (Fig. 1). This isolate dates to 1978 but, despite its apparent similarity to PMEN1 strains, has been found very rarely since then. Hence, this phenotype is not sufficient to guarantee success, an observation supported by the continued presence of successful, but susceptible, pneumococci in the population (*32*, *33*). Improved understanding of the interplay between ecology and adaptation in other lineages through further focused sequencing programs may prove crucial to the future control of this, and other, diverse bacterial pathogens.

References and Notes

1. K. L. O'Brien *et al.*; Hib and Pneumococcal Global Burden of Disease Study Team, *Lancet* **374**, 893 (2009).
2. L. McGee *et al.*, *J. Clin. Microbiol.* **39**, 2565 (2001).
3. N. J. Croucher *et al.*, *J. Bacteriol.* **191**, 1480 (2009).
4. R. Muñoz *et al.*, *J. Infect. Dis.* **164**, 302 (1991).
5. C. M. Parry *et al.*, *Antimicrob. Agents Chemother.* **46**, 3512 (2002).
6. K. P. Klugman *et al.*, *Eur. J. Clin. Microbiol. Infect. Dis.* **13**, 171 (1994).
7. L. McGee, K. P. Klugman, D. Friedland, H. J. Lee, *Microb. Drug Resist.* **3**, 253 (1997).
8. A. Tarasi, Y. Chong, K. Lee, A. Tomasz, *Microb. Drug Resist.* **3**, 105 (1997).
9. A. Corso, E. P. Severina, V. F. Petruk, Y. R. Mauriz, A. Tomasz, *Microb. Drug Resist.* **4**, 325 (1998).
10. R. Dagan, K. P. Klugman, *Lancet Infect. Dis.* **8**, 785 (2008).
11. M. R. Moore *et al.*, *J. Infect. Dis.* **197**, 1016 (2008).
12. C. Muñoz-Almagro *et al.*, *Clin. Infect. Dis.* **46**, 174 (2008).
13. C. Ardanuy *et al.*, *J. Antimicrob. Chemother.* **64**, 507 (2009).
14. Materials and methods are available as supporting material on *Science* Online.
15. S. R. Harris *et al.*, *Science* **327**, 469 (2010).
16. E. J. Feil, J. M. Smith, M. C. Enright, B. G. Spratt, *Genetics* **154**, 1439 (2000).
17. H. Ochman, S. Elwyn, N. A. Moran, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12638 (1999).
18. T. L. McCool *et al.*, *Infect. Immun.* **71**, 5724 (2003).
19. L. Rose *et al.*, *J. Infect. Dis.* **198**, 375 (2008).
20. S. D. Bentley *et al.*, *PLoS Genet.* **2**, e31 (2006).
21. K. Trzciński, C. M. Thompson, M. Lipsitch, *J. Bacteriol.* **186**, 3447 (2004).
22. K. E. Holt *et al.*, *Nat. Genet.* **40**, 987 (2008).
23. M. W. Pletz *et al.*, *Emerg. Infect. Dis.* **12**, 1462 (2006).
24. M. J. Ferrándiz *et al.*; Spanish Pneumococcal Infection Study Network, *Antimicrob. Agents Chemother.* **49**, 2237 (2005).
25. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
26. J. H. Shaw, D. B. Clewell, *J. Bacteriol.* **164**, 782 (1985).
27. M. Del Grosso, R. Camilli, F. Iannelli, G. Pozzi, A. Pantosti, *Antimicrob. Agents Chemother.* **50**, 3361 (2006).
28. M. Del Grosso, J. G. Northwood, D. J. Farrell, A. Pantosti, *Antimicrob. Agents Chemother.* **51**, 4184 (2007).
29. P. Romero *et al.*, *Plasmid* **58**, 51 (2007).
30. E. V. Pestova, D. A. Morrison, *J. Bacteriol.* **180**, 2701 (1998).
31. A. Buu-Hoï, T. Horodniceanu, *J. Bacteriol.* **143**, 313 (1980).
32. W. P. Hanage, C. Fraser, J. Tang, T. R. Connor, J. Corander, *Science* **324**, 1454 (2009).
33. C. Colijn *et al.*, *J. R. Soc. Interface* **7**, 905 (2010).

34. We thank the participating surveillance networks, listed in table S1, and the core informatics, library-making, and sequencing teams at the Wellcome Trust Sanger Institute. Attending authors were grateful for the opportunity to discuss this project at the Permafrost conference. Sequence accession codes are given in tables S1 and S2. This work was funded by the Wellcome Trust.

Supporting Online Material

www.sciencemag.org/cgi/content/full/331/6016/430/DC1
Materials and Methods
Figs. S1 to S13
Tables S1 to S5
References

1 October 2010; accepted 8 December 2010
10.1126/science.1198545