Infection, Genetics and Evolution 33 (2015) 334–342

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Mechanisms of genome evolution of *Streptococcus*

Cheryl P. Andam, William P. Hanage *

Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA

ARTICLE INFO

Article history:
Received 7 August 2014
Received in revised form 5 November 2014
Accepted 7 November 2014
Available online 13 November 2014

Keywords:
Pneumococcus
Genome evolution
Recombination
Selection
Vaccination
Antibiotics

ABSTRACT

The genus *Streptococcus* contains 104 recognized species, many of which are associated with human or animal hosts. A globally prevalent human pathogen in this group is *Streptococcus pneumoniae* (the pneumococcus). While being a common resident of the upper respiratory tract, it is also a major cause of otitis media, pneumonia, bacteremia and meningitis, accounting for a high burden of morbidity and mortality worldwide. Recent findings demonstrate the importance of recombination and selection in driving the population dynamics and evolution of different pneumococcal lineages, allowing them to successfully evade the impacts of selective pressures such as vaccination and antibiotic treatment. We highlight the ability of pneumococci to respond to these pressures through processes including serotype replacement, capsular switching and horizontal gene transfer (HGT) of antibiotic resistance genes. The challenge in controlling this pathogen also lies in the exceptional genetic and phenotypic variation among different pneumococcal lineages, particularly in terms of their pathogenicity and resistance to current therapeutic strategies. The widespread use of pneumococcal conjugate vaccines, which target only a small subset of the more than 90 pneumococcal serotypes, provides us with a unique opportunity to elucidate how the processes of selection and recombination interact to generate a remarkable level of plasticity and heterogeneity in the pneumococcal genome. These processes also play an important role in the emergence and spread of multi-resistant strains, which continues to pose a challenge in disease control and/or eradication. The application of population of genomic approaches at different spatial and temporal scales will help improve strategies to control this global pathogen, and potentially other pathogenic streptococci.

1. Introduction

The genus *Streptococcus* is a diverse group of 104 recognized species (http://www.bacterio.cict.fr/s/streptococcus.html as of July 21, 2014), whose interactions with host organisms vary from commensal to pathogenic. Many of the pathogenic species cause severe, invasive infections that account for a high burden of morbidity and mortality (Gray, 1998; Mitchell, 2003), associated with high economic and health care costs. One of the most clinically relevant *Streptococcus* species colonizing humans is *Streptococcus pneumoniae* (or pneumococcus), whose control remains a challenge despite the availability of highly effective vaccines. The pneumococcus is a quiescent colonizer of the upper respiratory tract and is carried asymptomatically by a fraction of the population, ranging from 21% to 94% in high burden settings such as Sub-Saharan Africa (Usuf et al., 2014). However, it can cause diseases such as otitis media, conjunctivitis, pneumonia, bacteremia and meningitis. Pneumococcal diseases are most prevalent in children, the elderly and immuno-compromised (French et al., 2007; Gray, 1998; Rubins et al., 2008). The introduction of conjugate vaccination in children worldwide has had profound consequences for the burden of pneumococcal diseases as well as the population structure of carriage and invasive strains. However, the efficacy of these control measures has been limited by the pneumococcus' exceptional ability to circumvent them. Years after the introduction of the first pneumococcal conjugate vaccine (PCV) in 2000, we are beginning to understand the long-term impact of these vaccines.

Genomic methods have made an important contribution to our understanding of the biology and epidemiology of important bacterial pathogens, including the pneumococcus. Molecular methods historically important in pneumococcal epidemiology, and still widely used today, include multi-locus sequence typing (MLST) (Enright and Spratt, 1998), in which seven housekeeping genes are sequenced as a sample of genomic variation and are used to define the sequence type (ST). Another example of an historically important molecular method is pulsed field gel electrophoresis (PFGE) in which strains are distinguished on the basis of restriction digestion of their genomic DNA (Lefevre et al., 1993). Differences in the numbers and location of the restriction sites are then visualized by electrophoresis. By investigating genetic and phenotypic

* Corresponding author. Tel.: +1 617 432 1047.
E-mail address: whanage@hsph.harvard.edu (W.P. Hanage).

http://dx.doi.org/10.1016/j.meegid.2014.11.007
1567-1348/

variation at different spatial and temporal scales, we can also elucidate the pace and patterns of pneumococcal diversification. The control of pneumococcal disease depends on our ability to anticipate the ways in which the organism responds to interventions, and genomic methods provide the best, most extensive data for the study of pneumococcal evolution.

While genomics is a rapidly changing field, as of this writing, the pneumococcus can lay claim to be the best-studied organism from a population genomic standpoint, with multiple large samples already published (Chewapreecha et al., 2014; Croucher et al., 2014a, 2013). In this review, we discuss recent findings in pneumococcal genome evolution, with a focus on the roles of recombination and selection. We highlight the impacts of vaccine-induced selective pressures in driving the population dynamics and evolution of different pneumococcal lineages. We also discuss how the pneumococci respond to these pressures, mainly through recombination within the species as well as with other *Streptococci*. The contributions of mobile elements and phages to pneumococcal evolution will also be discussed. We conclude with future directions on how genomics can enhance our understanding of the role of vaccination and antibiotic use in pneumococcal evolution and disease control.

2. The promiscuous pneumococcus

2.1. Pneumococcal diversity, transformation and recombination

Pneumococci are competent and naturally transformable, meaning that they readily take up exogenous DNA and can incorporate it into the genome by homologous recombination. This is not rare; a study of a single lineage [designated PMEN1 by the Pneumococcal Molecular Epidemiology Network (McGee et al., 2001)] estimated that 74% of its genome had been altered by recombination in at least one isolate in a sample estimated to have diversified from a most recent common ancestor in the early 1970s (Croucher et al., 2011). This was the result of a study of a global sample, but at the opposite scale, a study of six isolates obtained within a seven-month period from the same individual found a remarkable 23 recombination replacements (Hiller et al., 2010), resulting in the replacement of 7.8% of the genome. Horizontal transfer of genetic material in general, not only by homologous recombination but also by mobile genetic elements, can substantially amplify the heterogeneity of the common gene pool of the pneumococcus, as was observed in the study of PMEN1 isolates retrieved between 1994 and 2008 mentioned above. In this work, recombinant regions varied in size from 3 bp to 72,038 bp (Croucher et al., 2011). Such high levels of recombination ensure the success of a clone over the long term, particularly in the face of environmental changes, as was reported for the multi-drug resistant PMEN2 that predominated in Iceland in the 1990s. The decline in its prevalence following the reduction in antibiotic use was largely due to the lack of sequence import through recombination that prevented the modification of core sequences associated with resistance and antigenic diversification (Croucher et al., 2014b).

Although the pneumococcus is known to be highly transformable and recombining, there is marked variation in these traits among isolates. An experimental study showed variation over four orders of magnitude (10^{-2} to 10^{-6}) in transformation frequencies in isolates from asymptomatic carriage, and this variation does not correlate with genetic relatedness of isolates (Evans and Rozen, 2013). The causes of this variation, whether essential (resulting from differences in the mechanism of DNA uptake) or ecological (recombination is limited by the isolates rarely encountering one another in nature), remain poorly understood. One possible explanation is that these observed differences might be attributed to the environmental stresses faced by different pneumococcal lineages. Chemostat evolution experiments of the costs and benefits of transformation in pneumococcus grown for 1000 generations demonstrate that transformation can benefit cells living in stressful environments and can be costly in benign environments (Engelmoer et al., 2013). Rather than being selected by stress, recombination may be directly linked to it through the overlap of the stress response with the competence pathway responsible for taking up DNA (Dagkessamanskaia et al., 2004). It is easy to understand how interactions with the immune system, as well as interventions such as antimicrobials and vaccines, could create a stressful environment for the pathogen.

Different pneumococcal lineages also demonstrate significant variation in recombination rates, measured as the rate at which nucleotide polymorphisms accumulate through recombination relative to mutation (r/m). For example, whole genome analyses of carried pneumococci from Massachusetts showed a wide range of recombination rates, with the highest value associated with a clade containing the highly resistant ST 320 and which has proliferated following vaccination despite originally having a vaccine serotype. The loci determining serotype have been replaced with those of a non-vaccine serotype (NVT) by recombination. This process, which can generate novel combinations of serotype and genomic backbone, is known as capsular switching (Croucher et al., 2013). Recombination, however, is not uniform throughout the pneumococcal population. It has been recently described to consist of two modes of action: micro-recombinations involve the frequent replacement of single, short DNA fragments, and macro-recombinations, though rarer, involve the acquisition of multiple, long fragments and is a associated with major phenotypic changes (Mostowy et al., 2014).

In addition to variation in the uptake of DNA, some strains exhibit propensity to either donate or receive exogenous DNA more often than others. Chewapreecha et al. (2014) found that a lineage of pneumococci in their sample that lacked a capsule was more likely to both gain and donate DNA to other lineages. Again, a study of isolates from a single chronic childhood infection showed similar variation at the level of evolution within a single host (Hiller et al., 2010). During this time period, one particular strain, ST 2011v4, donated DNA extensively to the other infecting strains, all of which are closely related ST 13 strains. Unambiguously identifying donors is not easy, because in a recombinogenic organism like the pneumococcus, the sequence in question may be found in multiple lineages and the ultimate origin is hard to define. It is reasonable to think that some lineages may donate more often than others, if for no other reason than that some lineages are more commonly found in the population than others, and so will have more opportunity [(for instance, it is not clear the unencapsulated lineage discussed above (Chewapreecha et al., 2014) is a more common donor or recipient than one would expect given its frequency in the population]. Moreover, lineages that are over represented in a sample may be more readily identified as donors or recipients, and it is important to recognize such possible bias.

The consequences of variation in recombination can be important; pneumococci with evidence of extensive recombination at housekeeping loci are significantly associated with antibiotic resistance (Hanage et al., 2009). Genomic sequencing in diverse environmental backgrounds can therefore help us define the causes and consequences of the variation in transformation, recombination and other phenotypic traits in pneumococcus.

2.2. Gene gain across species boundaries

The ability of the pneumococcus to donate and receive genetic material is not restricted to its own species. Novel genes can be acquired from other taxa, many of which also inhabit the same