Chapter 5: Recombination allows rapid adaptation in response to local selective pressure

---

Declaration of work contributions:

Clinical records on antibiotic usage described in 5.2.1.1 were kindly provided by Dr Paul and Claudia Turner.

5. Recombination allows rapid adaptation in response to local selective pressure

5.1 Introduction and aims

Homologous recombination is a major driving force in the evolution of *Streptococcus pneumoniae* genomes and a key contributor to genetic diversity and population dynamics. Its importance in the failure of clinical intervention such as antibiotics has been increasingly recognised (Dowson, Hutchison *et al.* 1989, Laible, Spratt *et al.* 1991, Hanage, Fraser *et al.* 2009, Croucher, Harris *et al.* 2011). Following the identification of population-wide sequence exchange discussed in chapter 4, this chapter explores the contents of the exchanged genetic materials, many of which are likely to be a reflection of host immunity and clinical practices. The analysis shows that highly recombined regions in the Maela pneumococcal population are concentrated on genes encoding cell surface antigens and genes associated with resistance to antibiotics. The analyses will next focus on the directionality of recombination in response to changes in selection pressure: monitoring bacterial response to an increase as well as a decrease in consumption of two classes of antibiotics, beta-lactams and co-trimoxazole respectively. An increase in beta-lactam consumption coincided with recombination events that lead to the population becoming beta-lactam non-susceptible, providing genomic evidence of adaptation in response to antibiotic usage. On the other hand, a reduction in co-trimoxazole consumption has been found to coincide with a decrease in the number of co-trimoxazole resistant isolates compared to sensitive isolates in strains that have undergone recent recombination. Together, these observations further support the role of recombination in pneumococcal adaptations, particularly in the presence and absence of clinical interventions.

This chapter aimed at:

i) Defining the regions where recombination was observed at heightened frequencies and their biological relevance.

ii) Associating the recombination trends observed in the population with known changes in selection pressure.

5.2 Methods

5.2.1 Preparation of nucleotide sequences for penicillin-binding proteins (*pbps*), dihydrofolate reductase (*dhfR*) and dihydropterpate synthase (*folP*)

The DNA sequences of the above genes, whose allelic forms are known to confer resistant to β-lactam (*pbp1a, pbp2b, pbp2x*) and co-trimoxazole (*dhfR, folP*), were extracted from the draft genome assembly using Glimmer3 (Delcher, Bratke *et al.* 2007) for gene prediction. Predicted genes were searched for similarity against all *pbp1a, pbp2b, pbp2x, dhfR, folP* genes available in the public dataset using blastall v 2.2.15 (Altschul, Gish *et al.* 1990) with default settings, thereby allowing large diversity of these genes especially for *pbp1a*, *pbp2b* and *pbp2x* to be captured.

5.2.2 Phylogenetic analyses

Phylogenies of individual gene trees were estimated with RAxML v7.0.4 (Stamatakis 2006) using a GTR model with a gamma correction for site rate variation using 100 bootstraps.

5.2.3 Statistical tests

5.2.3.1 Associations between recombination and resistant phenotypes

The trend of recombination was estimated through the detected phenotypes observed in the presence and absence of recombination in the sub-population including 7 most prevalent clusters. Please note that 5 isolates with missing phenotypes (**Appendix A**) were not included in this analysis. Based on the prediction of recombination from 7 dominant clusters, strains undergoing recombination at *pbp1a, pbp2b, pbp2x, dhfR* or *folP* and their phenotypic resistance to β-lactam and co-trimoxazole were compared against the strains with no recombination events observed at these sites. The statistical difference between the recombining group and the non-recombining group was estimated with two-tailed Fisher's exact test.

Please note that beta-lactam resistance phenotype were considered as the binary outcomes with the breakpoint defining the categories of susceptible and non-susceptible to beta-lactams as ≤ 0.06, and > 0.06 μg/ml respectively. The breakpoint defining categories sensitive, intermediate and resistant for co-trimoxazole were given at a 1/20 ratio of trimethoprim/sulfamethoxazole as ≤ 0.5/9.5, 1/19-2/38, and ≥ 4/76 respectively.

5.2.3.2 Associations between past and recent recombination events and resistant phenotypes

Temporal trends were determined by comparing phenotype differences in isolates showing evidence of recent recombination (events predicted at external branch using algorithm described previously (Croucher, Harris *et al.* 2011)) to isolates whose ancestors had undergone recombination (events predicted at internal nodes). Two-tailed Fisher's exact test was used to test the significant difference in trend between the two temporal groups.

Note that alternative *murM* and *murN* genes associated with high beta-lactam resistance (Smith and Klugman 2001) were also considered. However, only two candidates with partial matches were observed and are thus less likely to explain trends in beta-lactam resistance.

5.3 Results

5.3.1 Biological relevance of sequences that have undergone recombination

To investigate the impact of recombination introducing selectively advantageous sequences, the frequency of recombination events across all genomic sites for the seven dominant clusters (BC1-BC7, see 4.1.1) was analysed and correlated to potential selection pressures applied to the community.

Here, recombination hotspots are defined as the genome location where recombination events were observed with higher frequency as a result of selection. Based on the recombination events identified in 4.1.1, the recombination frequency for each site was counted from a range of genome coordinates where homologous recombination events were predicted (**Figure 5.1**). Recombination occurring at sites of mobile genetic elements, although having the potential to have been generated through both site-specific recombination and homologous recombination, was not considered here. Hotspots were defined as a recombination frequency above the 95th percentile of site frequencies detected for the cluster as a whole, thus accounting for both recombination frequency and population size of each cluster.

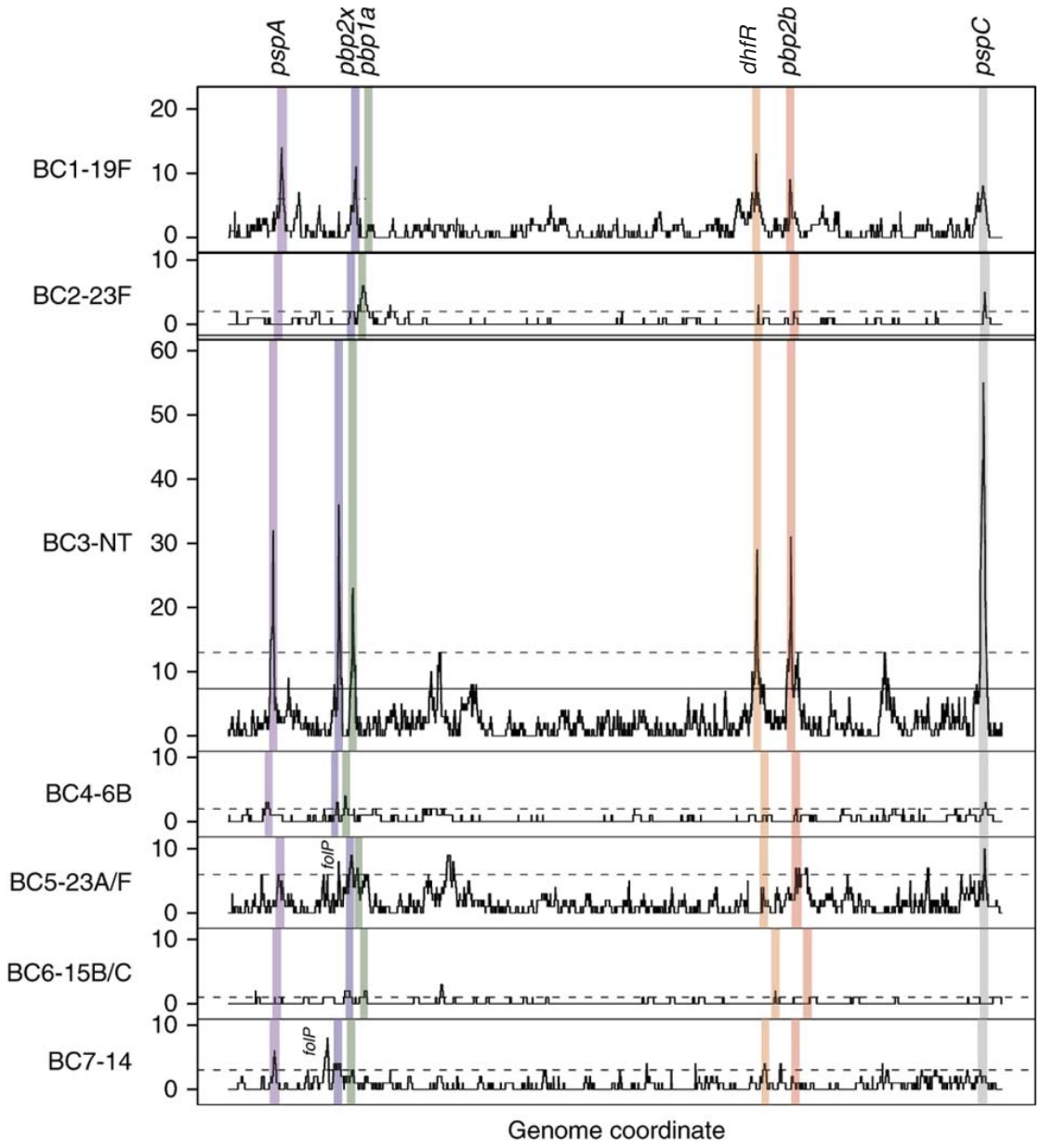
Consistent with the observed highly recombinogenic nature of the nontypable strains discussed in 4.1.3, the subpopulation cluster BC3-NT contained the highest frequency hotspots. This is despite its being third in term of population size (n=202 isolates), after two dominant encapsulated clusters, BC1-19F (n=365 isolates) and BC2-23F (n=213 isolates).

For each dominant cluster, a non-random distribution of detected recombination events across genomic sites was observed (**Figure 5.1**). Although recombination frequencies vary between clusters, the positions of recombination hotspots in most clusters were remarkably similar. This consistency within the Maela population potentially highlights two important points. Firstly, despite high levels of species-wide genotype diversity, major selective pressures appear to have equal impact across investigated clusters, resulting in similar patterns of selection. Secondly, the force of selection for these recombination sites must have acted rapidly enough for clusters to have similar hotspot patterns regardless of their time since transmission to the Maela community. Host immunity and clinical practices may account for these hotspot patterns. Indeed, the six leading hotspots in the Maela pneumococcal population are among genes encoding cell surface antigens (*pspA, pspC*) and genes associated with resistance to antibiotics (*pbp1a, pbp2b, pbp2x* and *dhfR*).

Figure 5.1 Recombination hotspots

Panels (top to bottom) are ordered by decreasing cluster population size. For each cluster, recombination hotspots were identified as sites with recombination frequency above the 95th percentile of the homologous recombination frequency detected in that cluster. The 95th percentile levels are represented as horizontal dashed lines. Recombination hotspots detected in at least four of the seven studied clusters are shaded in different colours. These common hotspots (in order by genomic coordinates) encode pneumococcal surface protein A (*pspA*, purple), penicillin-binding protein 2x (*pbp2x*, blue), penicillin-binding protein 1a (*pbp1a*, green), dihydrofolate reductase (*dhfR*, orange), penicillin-binding protein 2b (*pbp2b*, red) and pneumococcal surface protein C (*pspC*, grey). The figure includes 2,077 recombination events; the 132 events associated with mobile genetic elements are not displayed.

(**Figure is shown on the next page**)


pspA
pbp2x
pbp1a
dhfR
pbp2b
pspC
BC1-19F
BC2-23F
BC3-NT
BC4-6B
BC5-23A/F
BC6-15B/C
BC7-14
folP
folP
Genome coordinate

5.3.1.1 Hotspots associated with surface antigens

To allow asymptomatic carriage, it is essential for pneumococci to be transmitted without being recognised by the host immune system. Two evolutionary models have been proposed to explain the co-evolutionary processes taking place between bacteria and their hosts. The balanced polymorphism model, also called trench warfare model (Bergelson, Dwyer *et al.* 2001) predicts constant diversity in the population due to selective pressure to maintain polymorphisms especially to evade the host immune system. The model predicts that the fitness of the phenotypes increases as it becomes rarer. Another model, the evolutionary arms race or the red queen hypothesis suggests that there is co-evolution between the bacterial protein and the targeted host molecule (Dawkins and Krebs 1979). Therefore, bacterial surface proteins, which are directly involved in host-pathogen interactions, are likely to show greater diversity in sequence patterns than sequences that evolve in neutral manner according to these two models (Toft and Andersson 2010). This analysis suggests that recombination seems to be the tool to generate and spread the diversity in pneumococcal surface proteins.

Indeed, pneumococcal surface protein A (*pspA*) and pneumococcal surface protein C (*pspC*) have undergone recombination at a higher frequency, manifested as recombination hotspots in four and six of seven main Maela clusters respectively (**Figure 5.1**). This observation is consistent with hotspots reported in PMEN-1 (Croucher, Harris *et al.* 2011) and PMEN-14 (Croucher, Chewapreecha *et al.* 2014). PspA has been shown to be an important virulence factor for both invasive infections and nasopharyngeal colonisation. Pre-existing antibodies against PspA were shown to prevent carriage completely, suggesting that the protein is essential for colonisation in the human host (McCool, Cate *et al.* 2002). The protein acts to prevent the binding of host complement component C3 to the bacterial surface (Shaper, Hollingshead *et al.* 2004), the process by which the bacterium is marked for ingestion by host phagocyte known as opsonophagocytosis. This allows the bacterium to evade detection and destruction by the host immune system. Given its role in host-pathogen interaction, it

is not surprising to observe high recombination frequency at *pspA*, in consistent with its high sequence diversity reported previously (Hollingshead, Becker *et al.* 2000).

A paralogue of PspA, PspC also helps the bacterium to avoid host immune detection. It prevents the binding of host factor H to pneumococci, which again leads to the inhibition of complement-mediated opsonophagocytosis (Dave, Pangburn *et al.* 2004, Quin, Onwubiko *et al.* 2007). Moreover, a binding of a specific PspC subclass to a polymeric immunoglobulin receptor, a host protein highly expressed in the nasopharyngeal epithelium, results in an internalisation of receptor-bacterium complex into the host cell (Brock, McGraw *et al.* 2002). This process facilitates the invasion of pneumococci across mucosa epithelia.

Recombination hotspots were also observed in other surface proteins, but were limited to one or two clusters. These extra hotspots include Immunoglobulin A1 metalloprotease (*zmpA*) and pneumococcal serine-rich repeat protein (*psrP*). ZmpA is a protease that cleaves host immunoglobulin molecules attached to the surface of the bacterium and thereby prevents triggering of the host inflammatory response (Bek-Thomsen, Poulsen *et al.* 2012). Another integral membrane protein, PsrP is known to mediate pneumococcal aggregation in biofilms (Shivshankar, Sanchez *et al.* 2009). Although recombination frequency observed in *zmpA* and *psrP* did not reach the 95th percentile cut-off population-wide, their biological functions imply that diversifying selection, to a certain level, may well have acted on these genes to help the pneumococci thrive in their host.

5.3.1.2 Hotspots associated with antibiotic resistance determinants

Genes where allelic forms are known to confer resistance to beta-lactams and co-trimoxazole were also identified as recombination hotspots. These include penicillin binding proteins encoded by *pbp1a, pbp2b* and *pbp2x* which are known targets for a group of beta-lactam antibiotics (for an excellent review, see Hakenbeck, Bruckner *et al.* 2012). Beta-lactam functions as an inhibitor of penicillin-binding proteins (PBPs),

thereby preventing the complete formation of peptidoglycan in the cell wall during cell growth. The resistance to beta-lactams occurs with variant penicillin-binding proteins that have lower affinity for beta-lactams. This abolishes the inhibition and allows the formation of a complete cell wall. It has been long recognised that penicillin-binding proteins are highly altered in clinical isolates with mosaic gene structure; thereby indicating interspecies gene transfer from closely related commensal species mediated by recombination events (Dowson and Hutchison *et al.* 1989, Laible and Spratt *et al.* 1991).

Other antibiotic-resistance associated hotspots were detected in dihydrofolate reductase (*dhfR*) and, to a lesser extent dihydropteroate synthase (*folP*) (**Figure 5.1**). The allelic forms of *dhfR* and *folP* are known to confer resistance to trimethoprim and sulfamethoxazole, drugs that are commonly prescribed in combination as co-trimoxazole due to their synergistic effects (Adrian and Klugman 1997, Padayachee and Klugman 1999, Silver 2007). Both trimethoprim and sulfamethoxazole act by interfering with the synthesis of precursors in the thymidine synthesis pathway, thereby inhibiting bacterial DNA synthesis. Similar to beta-lactams, co-trimoxazole resistance develops with variant *dhfR* and *folP*, which have lower affinity to the drugs (Maskell, Sefton *et al.* 2001). This removes the inhibition and allows DNA synthesis to continue.

High recombination frequencies observed in *pbp1a*, *pbp2b*, *pbp2x*, *dhfR* and *folP* suggests that there have been extensive exchanges of different alleles at these gene loci. While replacements of sensitive by resistant alleles potentially lead to more isolates becoming non-susceptible, replacements of resistant by sensitive alleles may lead to a resistant being superseded by a sensitive population. The directionality of these exchanges cannot be determined by the data presented in **Figure 5.1** alone. This topic, as well as potential selection pressures driving the directionality will be discussed in 5.2.

5.3.1.3 Hotspot patterns vary over time

Not surprisingly, genetic reshuffling patterns resulting from recent recombination (events predicted at the external node, previously described in 4.2.1.1) appear to be more random compared to older recombination events (predicted at internal nodes) where more pronounced hotspot patterns can be observed (**Table 5.1**, Chi-square test comparing the ratio of hotspot patterns in recent and older recombination events = 13.979, two-tailed p value = 2.00×10^{-4}). This is expected given that the recombination process occurs at random. Under neutral evolution, recombination results in the reshuffling of alleles all over the genome. However, selection for advantageous alleles results in the pattern of hotspots where only a limited number of genes show a heightened recombination frequency relative to the rest of the genome. The selection process itself takes time. Therefore, the random pattern can still be observed in recent recombination events, in contrast to older recombination events where hotspot patterns are more pronounced, as there has been enough time for selection to act. This is consistent with the definition of recombination hotspots as a result of the combined action of recombination events in the past and selection pressure.

Table 5.1 Recombination signals have been refined through time

	Number of recombination events detected in seven dominant clusters (BC1-BC7) of method described in Croucher *et al.*		
	Total events	Events associated with recombination hotspots discussed in 5.1.1 and 5.1.2	Events occurring outside recombination hotspots
Recombination events predicted at internal nodes, which are likely to represent older recombination	1,589	356 (22.4%)	1,233 (77.6%)
Recombination events predicted at external nodes, which are likely to represent recent recombination	620	94 (15.2%)	526 (84.8%)

5.3.2 Changes in recombination trends reflect changes in selection pressure

Next, hotspots associated with antibiotic resistance described in 5.1.2 were used to demonstrate the associations between trends in recombination and temporal changes in selection pressures. The analysis focused on bacterial genetic and phenotypic responses to an increase in beta-lactam and a reduction in co-trimoxazole consumption.

5.3.2.1 Trend in antibiotic consumption in Maela

5.3.2.1.1 Measurable clinical prescriptions

With the availability of clinical records in Maela, the selection pressure applied to pneumococcal population can be estimated based on the trends in antibiotic prescriptions. **Table 5.2** summarises the trends in consumption of two antibiotics frequently used in the SMRU clinic, beta-lactams and co-trimoxazole. This local record (1994 – 2010) predated the pneumococcal collection used in the study (2007-2010), allowing genomic signals from evolutionary events years prior to the sampling time to be interpreted. Co-trimoxazole was recommended as a primary treatment for non-severe pneumonia in Maela from 1994 until 2002. However, due to increasing resistance across the region (Hoge, Gambel *et al.* 1998) and several side-effects (Medscape 2014), its use has been in decline. This is contrast to the rise in beta-lactam consumption. Based on this clinical information, it is likely that the pneumococci have been subjected to reducing pressure to become co-trimoxazole resistant. However, the population appeared to be under continuous pressure to be beta-lactam non-susceptible.

Table 5.2 Trend in antibiotic consumption based on the Burmese border guidelines (1994-2010)

Year	Co-trimoxazole consumption	Beta-lactam consumption
1994	Co-trimoxazole was the primary treatment for non-severe pneumonia, otitis media, urinary tract infection, and dysentery.	Ampicillin was used for severe pneumonia, meningitis and for infections in pregnant women where co-trimoxazole was contra-indicated.
1999	As in 1994.	As in 1994, but with amoxicillin being used as second line treatment for non-severe pneumonia (if no improvement with co-trimoxazole).
2002/2003	As in 1999 but ciprofloxacin replaced co-trimoxazole for dysentery.	As in 1999 with ceftriaxone appearing as an alternative drug for meningitis and typhoid.
2007	Amoxicillin now replaced co-trimoxazole for non-severe pneumonia.	Amoxicillin was recommended as primary treatment for non-severe pneumonia while ceftriaxone was first line for meningitis.

5.3.2.1.2 Non measurable self-prescriptions

While 5.2.1.1 describes predicted selection pressure from known clinical prescriptions, it is important to note that there might have been unknown pressure from self-prescriptions. Self-medication has been common in rural Thailand (Osaka and Nanakorn 1996). Based on a survey conducted in 1996 from Thai rural communities: only 42.2% of patients had proper clinical prescriptions from physicians; 37.5% of patients were self-prescribed with antibiotics purchased from local markets; and 9% were reported to “wait and see”. Although choices of drugs

used for self-medication have not been recorded, the drugs available in the markets largely matched those used in the clinic (personal communication from Dr Paul and Claudia Turner). Therefore it is highly likely that this unknown selective pressure is similar to that described from clinical prescriptions in 5.2.1.1.

5.3.2.2 Maela pneumococci have become more resistant to beta-lactams

Ampicillin (which has strong binding affinity to *pbp1a*, *pbp2x* and *pbp2b*), amoxycillin (strong binding affinity to *pbp2b*) and ceftriaxone (strong binding affinity to *pbp2x*) were first prescribed in Maela in 1994, 1999 and 2002, respectively (**Table 5.2)**. The usage of this group of beta-lactams has since become common, especially for amoxicillin. This has exerted a selective pressure with observable genetic signals from the phylogenies of *pbp* genes.

Figure 5.2 represents a phylogeny of each *pbp1a* (top), *pbp2b* (middle) and *pbp2x* (bottom) gene from all genomes in the study (n=3,085). Each phylogenetic tree reveals similar pattern, allowing two observations to be drawn. First, the sparseness of inner ring structure suggests that allelic groups are not exclusively linked to genomic population clusters. This indicates that the alleles have been distributed throughout the population, possibly through recombination. Notably, alleles found in the BC3-NT cluster (highlighted in red in **Figure 5.2**) are the most widely distributed on the tree. Second, the tree contains a mixture of short and long branches with beta-lactam non-susceptibility associated with longer branches. Long branches generally indicate recombination has occurred, thereby suggesting recombination as the main mechanism for acquisition of beta-lactam resistance.

To further test this hypothesis, the association between beta-lactam susceptibility and recombination was measured for the seven dominant clusters (1,126 genomes) and showed that strains that have undergone recombination at either *pbp1a*, *pbp2b* and *pbp2x* appear to be more phenotypically non-susceptible compared to strains without recombination (**Table 5.3**, Fisher's exact test p-value $< 2.20 \times 10^{-16}$).

Long-term consumption of beta-lactam antibiotics implies an on-going selection pressure for resistance. The data show that both recent recombination events (positioned on the external branches of the tree phylogeny) and older recombination events (on the internal nodes) are equally associated with resistance (no significant difference in recombination trend, Fisher's exact test, p-value = 0.6178), suggesting that selection pressure for recombination at these beta-lactam resistance loci has been continuous in the Maela community. Given the high beta-lactam consumption, these results are consistent with what was predicted earlier.

Figure 5.2 Association between recombining *pbp* genes and resistance phenotypes

(a) *pbp1a* gene tree. (b) *pbp2b* gene tree (c) *pbp2x* gene tree. The centre of each diagram shows a SNP-based phylogeny from 3,085 strains rooted on *Streptococcus mitis*. The inner ring is coloured according to dominant population clusters (BC1-7), with the rest of the population appearing in white. The outer ring is coloured according to resistance to penicillin with black and white showing non-susceptibility and susceptibility respectively.

(**Figure is shown on the next page**)

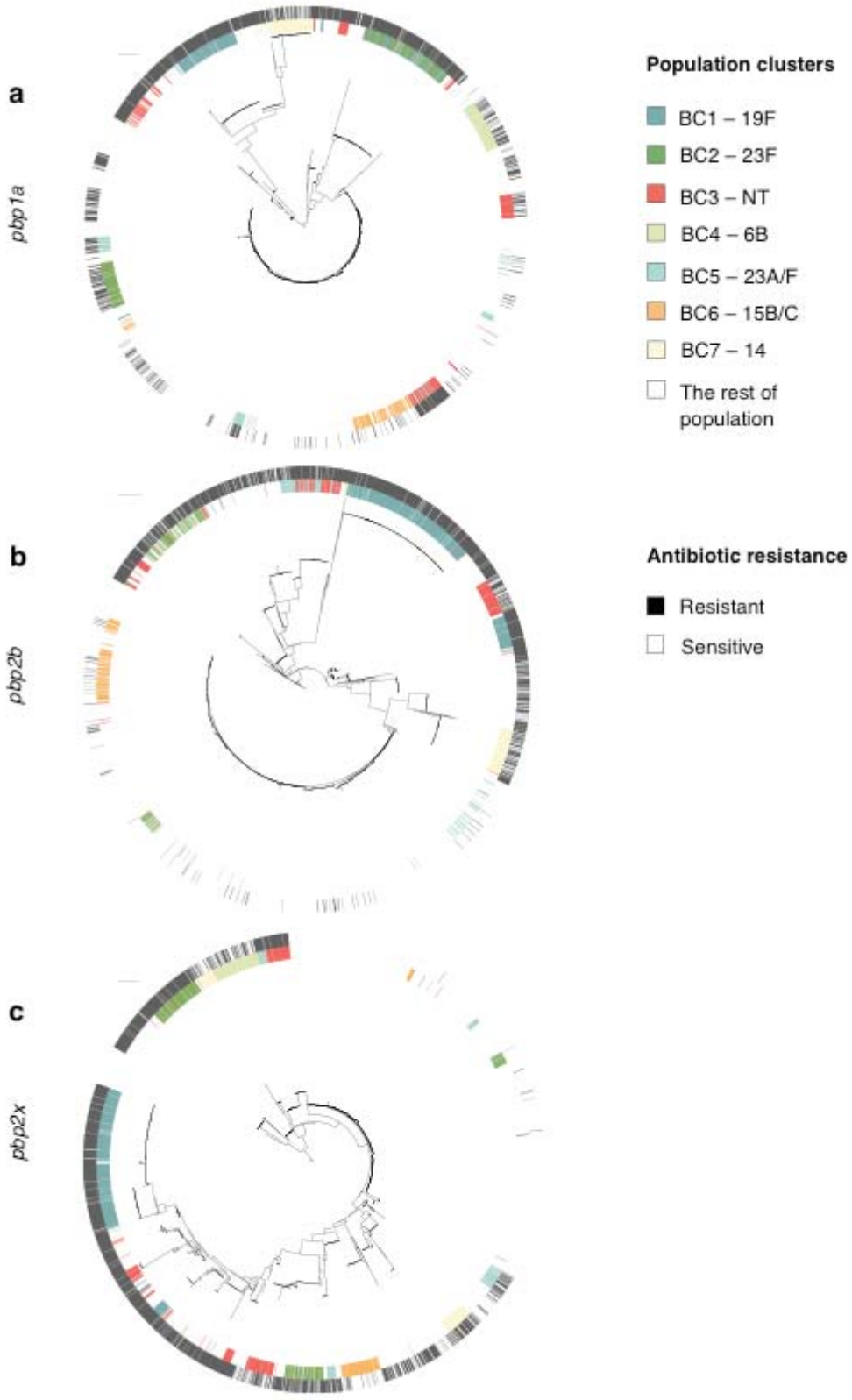

a
pbp1a
b
pbp2b
c
pbp2x
Population clusters
BC1 – 19F
BC2 – 23F
BC3 – NT
BC4 – 6B
BC5 – 23A/F
BC6 – 15B/C
BC7 – 14
The rest of population
Antibiotic resistance
Resistant
Sensitive

Table 5.3 Association between recombination, resistance phenotypes and temporal changes in recombination from seven dominant clusters

Observed phenotypes	No recombination at loci of interest	Recombination at loci of interest	Recent recombination (external node) at loci of interest	Older recombination (internal node) at loci of interest
Beta-lactam resistance: Resistant/sensitive (ratio)	120/146[a] (0.82)	795/150[a] (5.30)	25/6 (4.17)	770/144 (5.35)
Co-trimoxazole resistance: Resistant/ sensitive+intermediate (ratio)	210/28 (7.50)	873/100 (8.73)	10/9[b] (1.11)	863/91[b] (9.84)

[a] Significant difference between beta-lactam resistance phenotypes observed in strains with recombination at *pbp* genes and those without recombination (p-value $< 2.2 \times 10^{-6}$)

[b] Significant difference in cotrimoxazole resistance phenotypes between recent recombination and older recombination at *dhfR* and *folP* genes (p value = 3.49×10^{-5}). Note that the difference is still significant when ratios are grouped by resistant + intermediate/sensitive phenotype (p value = 0.00931)

5.3.2.3 Maela pneumococci have become less resistant to co-trimoxazole

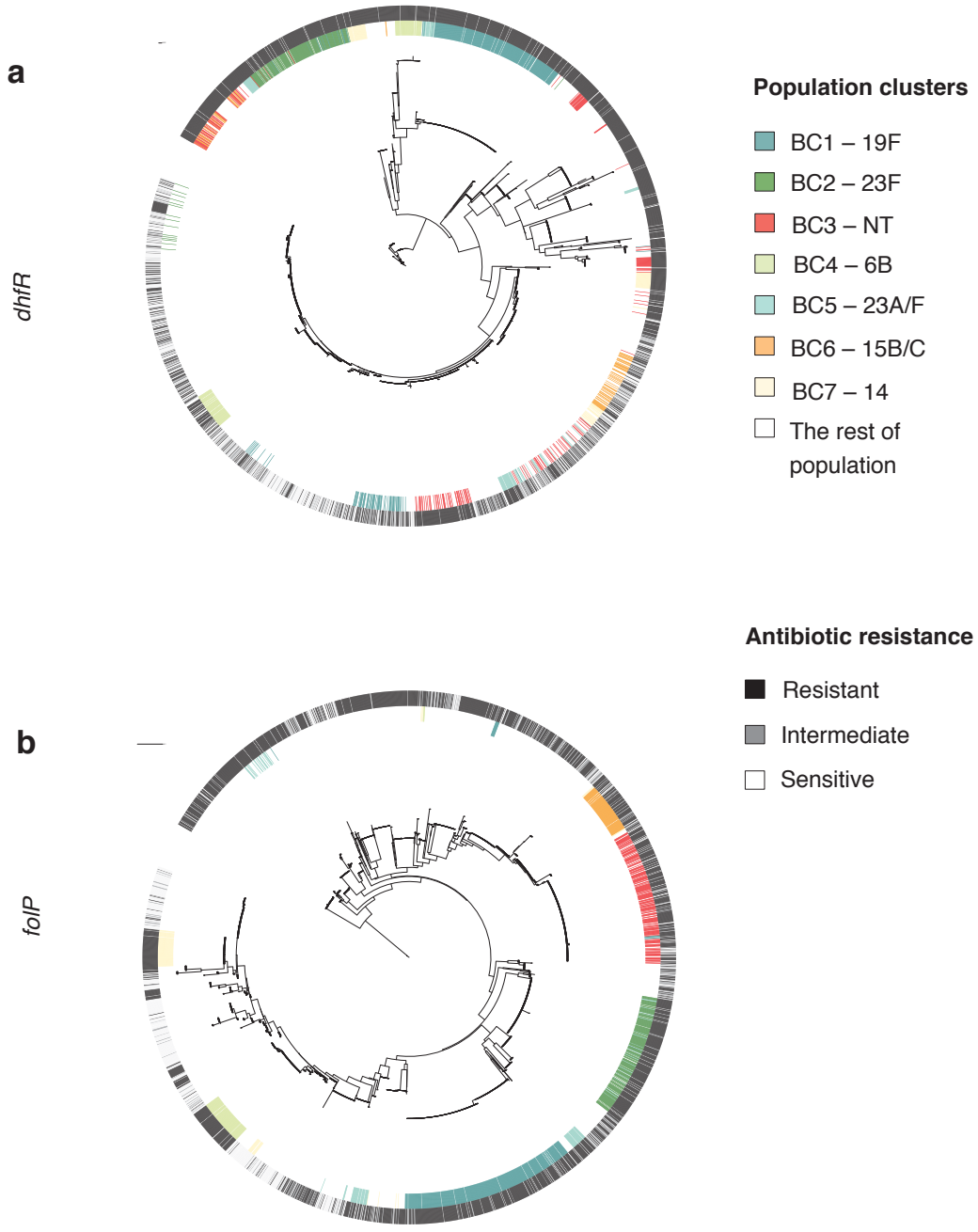
Based on clinical records described earlier, a reduced selection pressure for co-trimoxazole resistant was predicted. The results discussed next potentially provide genetic evidence in favour of this model.

Similar to the *pbp* genes, the phylogeny of *dhfR* (**Figure 5.3**) shows that alleles have been distributed through the population, suggesting a feasible role of recombination in

driving the evolution of *dhfR*. In contrast, the *folP* phylogeny revealed that the alleles were clustered according to their population clusters, suggesting a clonal expansion as the main mechanism for the spread of *folP*. Longer branches in both gene trees appeared to have a weaker association with phenotypic resistance than was seen for the *pbp* genes. Accordingly, there is no association between recombination at *dhfR* and *folP* genes and resistance (**Table 5.3**, Fisher exact test, p-value = 0.2229). It is possible that the lack of association might be due to the acquisition of resistance through nucleotide substitution followed by a clonal expansion. Another possible explanation could be that the signal is distorted due to changes in selective pressure over time. More specifically, the signal for positive selection for resistance before 2002 could be masked by a lack of strong pressure for resistance alleles thereafter (**Table 5.2**). A lack of selection pressure for resistance genes means that all alleles can be neutrally incorporated into the genome via recombination. This would be reflected by a significant difference in trend between recent and past recombination events, which can be seen (**Table 5.3,** Fisher exact test, p-value = 3.49×10^{-5}). While recent recombination events (on external branches) were associated with sensitive/intermediate susceptibility phenotypes, older recombination events (internal nodes) were associated with fully resistant phenotypes. This provides evidence that reduction in the use of an antibiotic may lead to a lower requirement for the isolates to be resistant. However, the magnitude by which the resistant phenotypes were observed to decrease after a decrease in co-trimoxazole usage is likely due to the cost of being resistant.

Figure 5.4 Association between recombining *fol* genes and resistant phenotypes.

(a) *dhfR* gene tree, (b) *folP* gene tree. The centre of each diagram represents a SNP-based phylogeny from 3,085 strains rooted on *Streptococcus mitis*. The inner ring is coloured according to the membership of the seven dominant population clusters (BC1-BC7). The outer ring is coloured on the basis of co-trimoxazole resistance phenotypes (white, sensitive; grey, intermediate; black, resistant). (**Figure is shown on the next page**)


a
dhfR
Population clusters
BC1 – 19F
BC2 – 23F
BC3 – NT
BC4 – 6B
BC5 – 23A/F
BC6 – 15B/C
BC7 – 14
The rest of population
b
folP
Antibiotic resistance
Resistant
Intermediate
Sensitive

5.4 Conclusion

The rise of resistant pneumococci in clinics has been observed since the late 1970s following the use of antibiotics (Kislak and Razavi *et al.* 1965, Hansman 1975). Moreover, high-level cephalosporin-resistant strains were observed in areas where the use of new-generation cephalosporins, which are a class of beta-lactams, was encouraged (McDougal and Rasheed *et al.* 1995, Smith and Botha *et al.* 2001). Together, these provided evidence that the use of antibiotics has driven the emergence of antibiotic resistance pneumococci (Hakenbeck and Bruckner *et al* 2012). Genetic culprits for the rise of resistance were soon after identified. Highly altered penicillin-binding protein genes, also known as mosaic structures, were observed among penicillin resistant pneumococci. The mosaic patterns were caused by recombination, thereby highlighting the role of recombination in spreading beta-lactam resistance (Dowson and Hutchison *et al.* 1989, Laible and Spratt *et al.* 1991). The genetic cause of co-trimoxazole resistance was also identified (Adrian and Klugman 1997, Padayachee and Klugman 1999, Silver 2007), but it was less clear which mechanism help disseminate the resistance alleles.

The depth of genomic sampling in this study allowed the identification of highly exchanged genes, which further led to the association of these genetic signals with antibiotic resistance patterns, and clinical records on antibiotic consumption. The results were consistent with the use of antibiotics and resistance patterns described previously. Genetic signals associated with antibiotic resistance could be linked to trends in antibiotic consumption. A reduction in co-trimoxazole consumption over time was reflected in the patterns of recombination, with more recent events having a weaker association with resistance phenotypes. On the other hand, a steady increase in the consumption of beta-lactams resulted in continuous acquisitions of resistance alleles via recombination.

The analyses also highlighted the remarkable similarity of sites under selection between dominant clusters, suggesting some uniformity in response to common selection pressure across the species. Aside signals related to antibiotic resistance,

recombination hotspots are also functionally linked to host immunity, reflecting the importance of this force in the evolution of *Streptococcus pneumoniae*.

Overall, this chapter highlights the role of recombination in mediating rapid adaptation of pneumococci. The nontypable clusters were described as a hub of gene flow in chapter 4. Therefore, it is plausible that the nontypable strains may be a principal driver of adaptation for a wider population, especially for antibiotic resistance. However, this role in mediating antibiotic resistance cannot be established without looking at the actual distributions of resistance alleles across the population. This will be explored in the next chapter where potential resistance markers for beta-lactams and their distributions in the wider population will be determined.