Supporting Text 1: Balancing Selection

Strength of balancing selection

The prevalence of antibiotic resistance at a given antibiotic consumption rate τ is determined by the prevalence of D-types with durations of carriage above and below $\mu = \frac{\tau}{c_\mu c_\beta - 1}$. The strength of balancing selection therefore affects the relationship between antibiotic consumption and resistance in two ways (see SI Figure 1). Firstly, it determines the antibiotic consumption range at which coexistence is observed: if the strength of balancing selection is decreased, D-types with short durations of carriage will no longer be maintained and full resistance will occur at lower antibiotic consumption. Secondly, the strength of balancing selection determines the shape of the relationship between antibiotic consumption and resistance: the stronger the balancing selection, the more even the D-type prevalences and therefore the more linear the relationship between antibiotic consumption and resistance.

For balancing selection to maintain a D-type in the presence of a competing D-type with a longer duration of carriage, balancing selection needs to be strong enough to overcome the fitness disadvantage introduced by the shorter duration of carriage. In the case of two competing strains x and y, with $\mu_x > \mu_y$ and where $v(I_x)|_{I_x=0}$ signifies $v(I_x)$ when $I_x = 0$ at equilibrium:

$$\frac{v(I_x)|_{I_x=0}}{v(I_y)|_{I_x=0}} > \frac{\mu_x}{\mu_y}$$

Taking into account antibiotic consumption rate τ and the resistance locus, this becomes:

$$\frac{v(I_x)}{v(I_y)} > \left\{ \begin{array}{ccc} \frac{\mu_x+\tau}{\mu_y+\tau} & \text{when} & \tau < u_y(c_\beta c_\mu - 1) \\ \frac{\mu_x+\tau}{c_\beta c_\mu \mu_y} & \text{when} & u_y(c_\beta c_\mu - 1) < \tau < u_x(c_\beta c_\mu - 1) \\ \frac{\mu_x}{\mu_y} & \text{when} & \tau > u_x(c_\beta c_\mu - 1) \end{array} \right\}$$

The required strength of balancing selection therefore depends on the relative clearance rates of the D-types, the cost of resistance and the antibiotic consumption rate until both strains are resistant (SI Figure 1 C). Unless the cost of resistance is very high (say, $c_\beta c_\mu > 2$), the antibiotic consumption rates which affect the strength of balancing selection will be small compared to the clearance rate of strains ($u > \tau$). Thus in practice the strength of balancing selection necessary to maintain the D-types is mainly determined by the relative clearance rate of the D-types.

Note that if balancing selection has the effect of reducing the R_0 of both strains (for example, if balancing selection arises from acquired immunity), an additional requirement for the existence of the weaker strain is that balancing selection does not reduce the reproductive number of the strain below 1: $\beta v(I_x) \geq \mu$.

Type of balancing selection

The balancing selection presented in the main text does not have a direct mechanistic interpretation. Here, we present results for balancing selection that approximates acquired immunity for D-type. If we model this immunity as waning at rate $1/k$, the dynamics of individuals infected (I_x) with D-type x and uninfected individuals immune to x (U_x) can be approximated as:

$$\frac{dI_x}{dt} = \beta I_x(U - U_x) - \mu_x I_x = \beta I_x v(I_x) U - \mu_x I_x$$

$$\frac{dU_x}{dt} = \mu_x I_x - \frac{1}{k} U_x$$

At equilibrium (assuming both the proportion of infected and immune hosts is constant), $\beta I_x v(I_x) U = \beta I_x(U - U_x) = \frac{1}{k} U_x$, which gives $v(I_x) = \frac{1}{k\beta I_x + 1}$.

Although the shape of the relationship between antibiotic consumption and resistance changes with how balancing selection is modelled, the conclusions presented in the main text (D-type coexistence allows co-existence of sensitivity and resistance and a gradual increase in the prevalence of resistance with increasing antibiotic consumption) remain unchanged (SI Figure 1 B). The exception is if balancing selection acts to increase clearance rate with increasing strain prevalence: this abolishes differences in clearance rate between

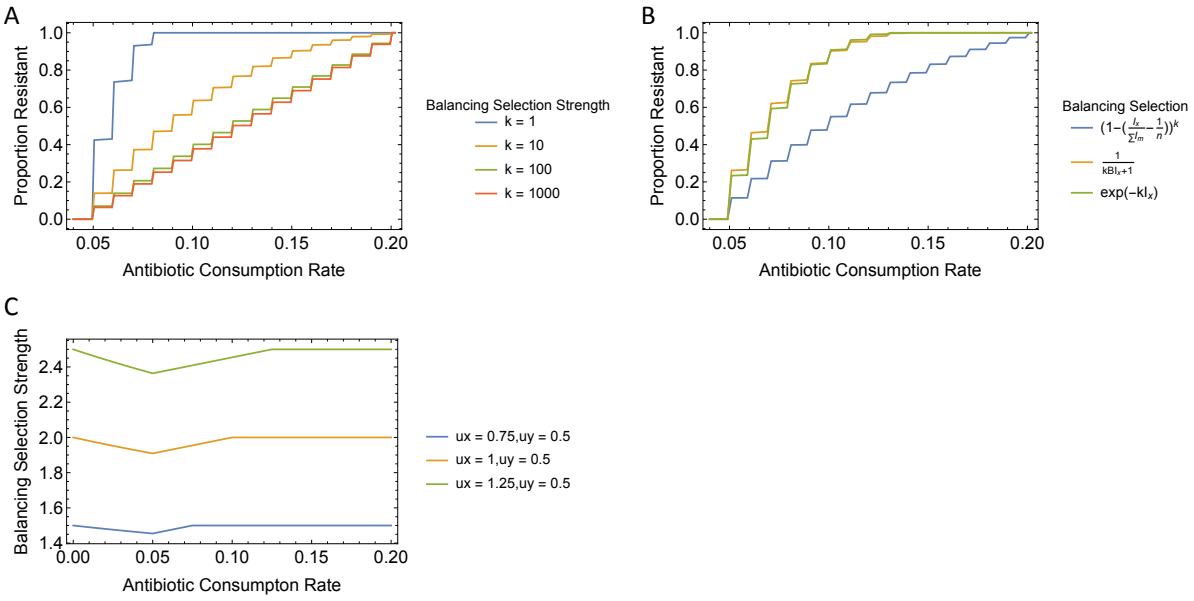

Figure 1: A: Relationship between antibiotic consumption rate and the prevalence of resistance at different strengths of balancing selection. The prevalence of resistance is determined by the prevalence of the resistant D-types. As the strength of the balancing selection increases, differences in the prevalence of D-types become smaller and the relationship between antibiotic consumption and resistance becomes increasingly linear. Balancing selection is the same as presented in the main text: $v(I_x) = (1 - [\frac{I_{xs}+I_{xr}}{\sum_y (I_{ys}+I_{yr})} - \frac{1}{n}])^k$. *The model had 16 D-types, with clearance rates evenly spaced in the range* $[0.5, 2]$. *Other parameter values were* $\beta = 2, c_\beta = 1, c_\mu = 1.1$. *B: Relationship between antibiotic consumption rate and the prevalence of resistance for different types of balancing selection. The coexistence of D-types allows coexistence of antibiotic sensitivity and resistance. The blue line corresponds to the results presented in the main text. The range of antibiotic consumption rates at which coexistence is observed is smaller when balancing selection has the effect of always decreasing transmission (orange and green lines) because the same value of* β *is not enough to sustain the D-types with higher clearance rate. The model had 16 D-types, with clearance rates evenly spaced in the range* $[0.5, 2]$. *Other parameter values were* $k = 15, \beta = 2, c_\beta = 1, c_\mu = 1.1$. *C: The relationship between antibiotic consumption, the relative fitness of D-types and the minimum strength of balancing selection required to maintain the weaker D-type. The strength of balancing selection is* $\frac{v(I_x)|_{I_x=0}}{v(I_y)|_{I_x=0}}$.

D-types and thus removes the effect D-type had on the fitness advantage of resistance. Balancing selection acting on clearance rate alone, however, does not seem like a realistic mechanism, given that strains with different durations of carriage are known to coexist.

Supporting Text 2: Effect of D-type allele change

Inclusion of the possibility of D-type allele change (through recombination or mutation) decreases, and eventually abolishes, coexistence. However recombination and mutation rates have to be unrealistically high to have a substantial effect on the model.

We modelled recombination as occurring between strains during co-infection. A strain xs can arise from recombination between strains $I_{\bullet s}$ and $I_{x\bullet}$ through transfer of either the D-type locus or the resistance locus. For strain xs, transmission occurring due to recombination is therefore given by $\theta(I_{xs} + I_{xr}) \sum_{i=1}^{n} I_{is}$, where n is the number of D-types and θ covers the probability of co-infection, the probability of recombination occurring and changing either the resistance allele or the D-type allele (we assume the two are equally likely) and the probability of transmitting the recombinant strain. For strain xs, the loss of transmission due to recombination is given by $\theta I_{xs} \sum_{i=1}^{n} (I_{is} + I_{ir})$. Terms involving xs as the recombinant cancel out; the net effect of recombination is therefore $-\theta I_{xs} \sum_{i \neq x} I_{ir} + \theta I_{xr} \sum_{i \neq x} I_{is}$. The strain dynamics for D-type x with recombination are therefore given by:

$$\frac{dI_{xs}}{dt} = \beta v(I_x) U \left[I_{xs} - \theta I_{xs} \sum_{i \neq x} I_{ir} + \theta I_{xr} \sum_{i \neq x} I_{is} \right] - (\tau + \mu_x) I_{xs}$$

$$\frac{dI_{xr}}{dt} = \frac{\beta}{c_\beta} v(I_x) U \left[I_{xr} - \theta I_{xr} \sum_{i \neq x} I_{is} + \theta I_{xs} \sum_{i \neq x} I_{ir} \right] - c_\mu \mu_x I_{xr}$$

For recombination to have a substantial effect on resistance requires unrealistically high recombination rates (SI Figure 2, $\theta > 0.01$): Mostowy et al. estimate that the PMEN1 pneumococcal lineage undergoes recombination events at a rate of approximately 0.01 events per month [SI ref: 1]. As not all of these recombination events will involve the resistance or D-type allele, θ is likely to be well below 0.01. (The model is parametrised such that rates correspond to reasonable estimates of monthly rates for *S. pneumoniae*, see Methods).

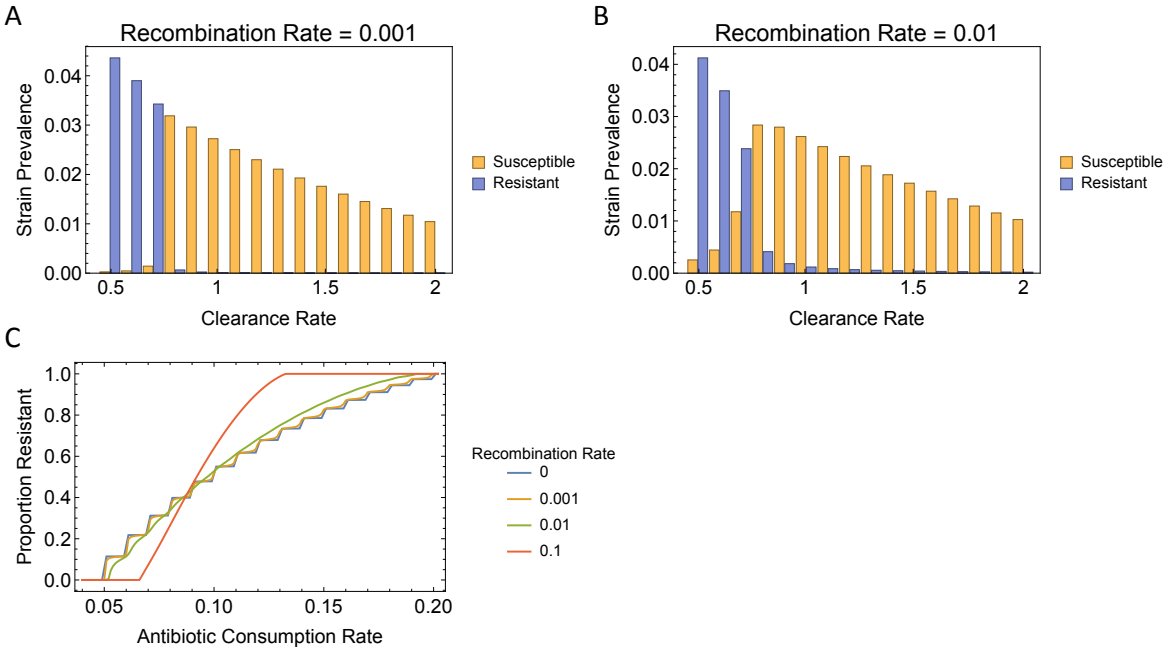

Figure 2: Inclusion of recombination allows within D-type coexistence of sensitivity and resistance (panels A and B) and narrows the range of antibiotic consumption rates at which coexistence is observed (panel C). Parameters for these plots were $\beta = 2, c_\beta = 1, c_\mu = 1.1, k = 15$, μ_x *evenly spaced in the range* $[0.5, 2]$ *and, for panels A and B* $\tau = 0.075$.

To model mutation between D-types, we introduce a mutation matrix **M**, where M_{xy} represents the probability of D-type y mutating into D-type x, and M_{xx} represents to probability of D-type x remaining x. The proportion of the population transmitting strain xs is therefore given by $\sum_{y=1}^{n} [M_{xy} I_{ys}]$. The dynamics of this model are described by:

$$\frac{dI_{xs}}{dt} = \beta \sum_{y=1}^{n} [M_{xy} I_{ys}] v(I_x) U - (\tau + \mu_x) I_{xs}$$

$$\frac{dI_{xr}}{dt} = \frac{\beta}{c_\beta} \sum_{y=1}^{n} [M_{xy} I_{yr}] v(I_x) U - c_\mu \mu_x I_{xr}$$

If D-type is modelled as determined by a single locus (SI Figure 3 B), with transition between each D-type equally likely, the mutation matrix is given by $M_{xy} = \frac{m}{n-1}$ if $x \neq y$ and $M_{xx} = 1 - m$, where m sets the mutation probability. If D-type is modelled as determined by $n-1$ two allele loci (each with a 'long' and 'short' carriage allele, the total number of long alleles defining D-type), transitioning from y 'long' alleles to x 'long' alleles requires a total gain of $x - y$ long alleles. If i out of the y 'long' alleles transition into the short allele, $i+x-y$ out of the $n-1-x$ 'short' alleles need to transition into 'long' alleles. The transition probability from state y to state x (M_{xy}) is therefore given as the product of the probability of i long to short transitions and $i+x-y$ short to long transitions, summed over all possible i: $M_{xy} = \sum_{i=0}^{y} b(i; y, m) b(i+x-y); n-1-y, m)$, where $b(i; y, m)$ is the binomial distribution with y trials and probability m (SI Figure 3 C).

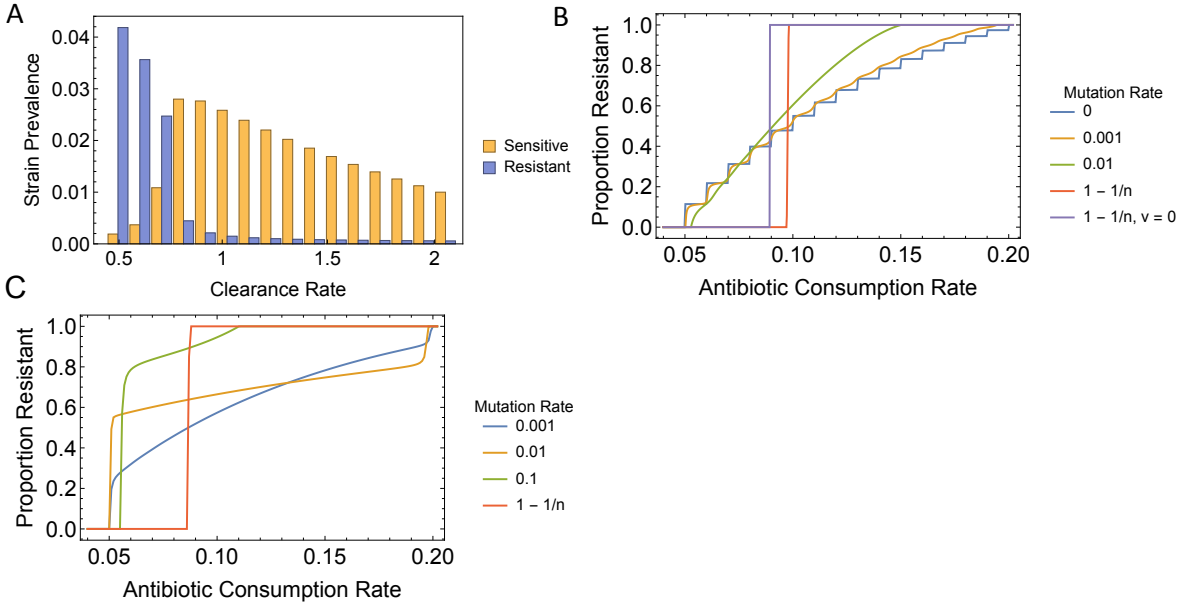

Figure 3: The effect of allowing D-type mutation. For all plots, clearance rates for the 16 D-types were evenly spaced in the range $[0.5, 2]$ and other parameters were $\beta = 2, c_\beta = 1, c_\mu = 1.1, k = 15$. A: Prevalence of sensitive and resistant D-types, in a single locus model with 16 D-types, with mutation rate 0.005 and antibiotic consumption rate $\tau = 0.075$. B: The relationship between antibiotic consumption and resistance at different mutation rates for a single-locus model with 16 D-types. Mutation rate $1 - \frac{1}{n}$ is the rate at which the D-type is not inherited (i.e. the D-type of a new infection is independent of the D-type of the transmitting strain). $v = 0$ represents the absence of balancing selection (D-type heterogeneity maintained through mutations only). For the two scenarios with non-inherited D-type, the switch to resistance occurs at a lower antibiotic consumption rate in the absence of balancing selection. This is because balancing selection flattens the D-type distribution. As a consequence, the average duration of carriage is longer in the model without balancing selection, leading to the switch to resistance occurring at a lower antibiotic consumption rate. C: The relationship between antibiotic consumption rate and the prevalence of resistance for different mutation rates using a multi-locus model of D-type. The shape of the relationship is different from the single-locus model, but the result that introducing D-type mutation decreases coexistence holds for this multi-locus model of D-type as well.

Supporting Text 3: Approximate fit to European resistance and consumption data

In Figure 4B, we show that a model with 40 D-types can approximately replicate the relationship between pneumococcal penicillin non-sensitivity and primary care beta-lactam prescriptions in Europe. Our aim was to demonstrate that the model reproduces qualitative trends observed in data (i.e a gradual, approximately linear increase in antibiotic resistance with increasing antibiotic consumption), rather than making stronger, quantitative, claims about the fit. We therefore did not choose parameters through a formal model fitting process. The range in clearance rates (0.2-4 per month) was chosen to produce coexistence over a 20-fold antibiotic consumption range (extrapolated from the best fit line) and such that clearance rates were of the same order of magnitude as the clearance rate estimates from the data, and the cost of resistance ($c_\mu = 1.075$) was chosen such that emergence of resistance occurs near the x-intercept of the best-fit line.

The range of clearance rates necessary to reproduce coexistence similar to that observed in the data is larger than the observed range of serotype clearance rates (20-fold vs 5-fold for the Maela-based estimates, for example). If, however, there is significant within-serotype diversity in clearance rate (see Discussion), averaging at the serotype level would under-estimate the true range of clearance rates.

Supporting Text 4: Choice of correlation measure

We measured the correlation between serotype resistance and duration of carriage using Kendall's rank correlation coefficient tau. We chose Kendall's tau rather than Pearson's r because we had no reason to expect the relationship between duration of carriage and resistance to be linear. We chose Kendall's tau over Spearman's rho because of the nature of uncertainty in our data: for the proportion of resistance within serotype, extreme values (i.e. 0% resistant or 100% resistant) are more likely for serotypes with few samples and thus subject to greater uncertainty. Kendall's rank correlation is calculated based on whether the relative rankings of pairs of serotypes in terms of resistance are consistent with their relative rankings in terms of duration of carriage. Spearman's rank correlation, on the other hand, takes into account the magnitude of the difference in rank. Spearman's correlation will therefore be more sensitive to error at the extremes of the distribution.

Supporting Text 5: Mean carriage duration estimates

Estimates for the Maela dataset

Estimates of carriage duration for the Maela dataset were based on longitudinal sampling in Maela and calculated as the area under the Kaplan–Meier survival function (see Turner et al [main text ref: 14] for details of the study, SI Data for estimates). Serotypes with a single instance of carriage were omitted from the analysis because of large uncertainty in the duration of carriage estimate.

Clearance rates were estimated separately for infants (based on a total of 1401 carriage episodes) and their mothers (total of 536 carriage episodes). We use the infant estimates in the results presented in the main text as pneumococcal carriage is more prevalent in children than adults: these rates are therefore more likely to reflect clearance rates in the overall infected population. The correlation between serotype resistance and duration of carriage was positive, but non-significant, when using the mothers' clearance rates (SI Table 2).

Estimates for Malawi and Massachusetts datasets

Because serotype duration of carriage may be setting-specific, ideally we would have tested the correlation between duration of carriage and resistance using duration of carriage estimated directly from each of the resistance datasets. However, such estimates were only possible for the Maela dataset (there was no longitudinal sampling in Malawi and Massachusetts datasets). We therefore 're-estimated' duration of carriage for the other two datasets from rates of immune clearance and susceptibility to replacement ('knock-out') by other serotypes calculated by Lipsitch et al. [main text ref: 17] based on data collected in Kilifi, Kenya.

Rates of immune clearance and susceptibility to knock-out may differ between setting (see main discussion), which we cannot account for. To confirm that our estimation method is reasonable, we derived duration of carriage for the Maela dataset in the way described above and compared these estimates to those derived

from the longitudinal sampling in Maela. Because Turner et al. reported carriage rates separately for mothers (16.8–29.6%) and infants (67.6–83.6%), we used an average rate of 50% to compute the I_j^R values. There is a reasonable correlation between the two estimates (Pearson's $r = 0.56$ 95%CI 0.18-0.79), Supporting Figure 4), although the Kilifi-based estimates are considerably smaller. However, as we measure the rank-correlation between duration of carriage and resistance, this discrepancy will not impact our results.

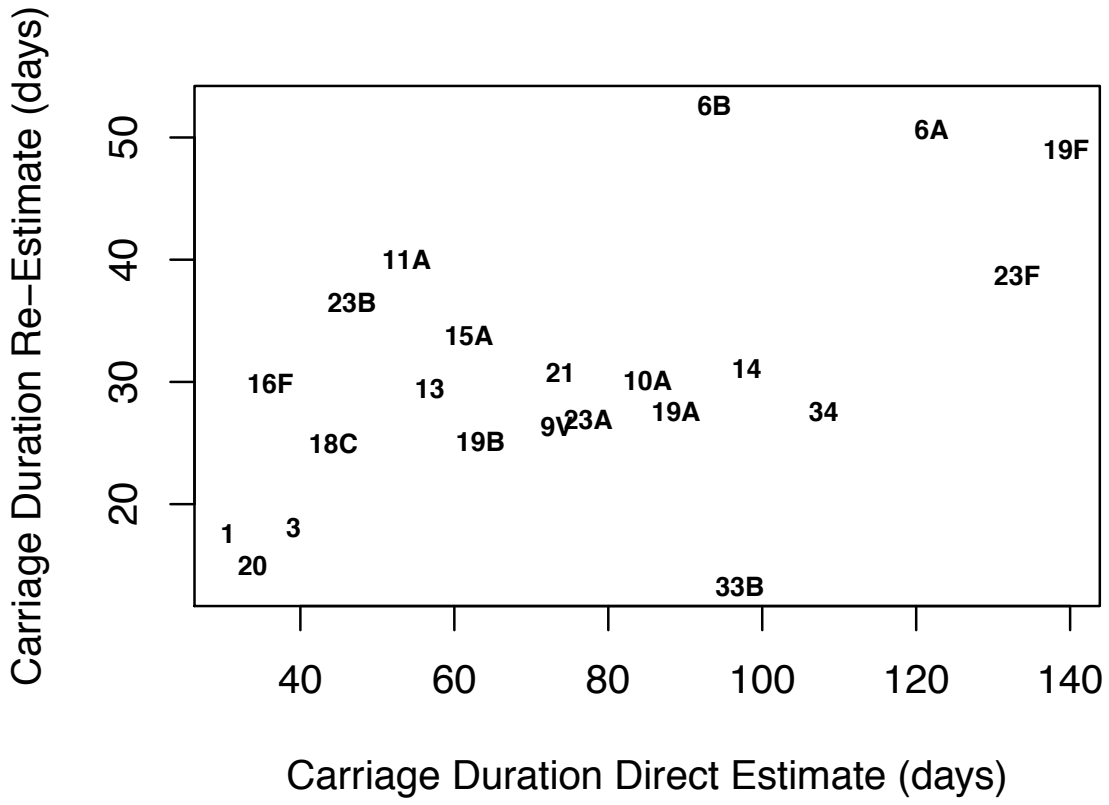

Figure 4: Serotype duration of carriage in Maela as estimated direcly from longitudinal data collected in Maela and as re-estimated using Kilifi-based rates of immune clearance and serotype knock-out.

Supporting Text 6: Strain 'knock-out' and immune clearance

We modify the model presented in the main text to differentiate between immune clearance ω and strain 'knock-out' (i.e. end of carriage through replacement by a competing strain) γ, which is a function of the prevalence of competing strains (but not I_{xs} or I_{xr}, as we assume strain x cannot replace itself) with associated costs of resistance c_ω and c_γ.

$$\frac{dI_{xs}}{dt} = \beta I_{xs} U - (\tau + \omega + \gamma) I_{xs}$$
$$\frac{dI_{xr}}{dt} = \frac{\beta}{c_\beta} I_{xr} U - c_\omega \omega I_{xr} - c_\gamma \gamma I_{xr}$$

The antibiotic consumption threshold at which strains switch therefore becomes:

$$\tau = \omega(c_\beta c_\omega - 1) + \gamma(c_\beta c_\gamma - 1)$$

Thus, if c_ω and c_γ are different, the two clearance rates cannot simply be combined when calculating the treatment threshold.

Supporting Text 7: Age-structure as a source of heterogeneity in carriage duration

A single-locus age-stratified model where age groups have different clearance rates is structurally near equivalent to the D-type model presented in the main text:

$$\frac{dI_{xs}}{dt} = \sum_{y=1}^{n}[B_{xy}I_{ys}]U_x - (\tau + \mu_x)I_{xs}$$

$$\frac{dI_{xr}}{dt} = \frac{1}{c_\beta}\sum_{y=1}^{n}[B_{xy}I_{yr}]U_x - c_\mu\mu_x I_{xr}$$

Here, I_{xs} and I_{xr} represent the proportion of hosts in age group x infected with the sensitive and resistant strain, respectively, U_x is the proportion of uninfected hosts in age group x, $U_x + I_{xs} + I_{rx} = 1$ and μ_x is the clearance rate for age group x. B_{xy} gives the rate at which group y transmits to group x. $\mathbf{B}$ is equivalent to $\beta\mathbf{M}$ in the D-type model (where $\mathbf{M}$ represents mutation between D-types). The age-stratified model differs from the D-type model presented in the main text only in the absence of the balancing selection term and in the stratification of the uninfected population.

By analogy to the D-type model, this model promotes coexistence when there is limited between-group transmission (i.e. host age and therefore clearance rate can be considered to be an inherited characteristic). This requires transmission to be sustained within age groups (i.e. $B_{xx} > \mu_x + \tau$ and/or $B_{xx} > c_\beta c_\mu \mu_x$, analogous to D-type diversity being maintained by balancing selection rather than mutation). Given that vaccination of children against a subset of serotypes has led to near complete elimination of vaccine types in both children and adults [SI ref: 2], it is unlikely transmission rates between adults are high enough to meet these criteria: if rates between adults were high enough to sustain transmission in the absence of transmission from children, vaccine-type serotypes would have continued to circulate in the adult population. It is possible, however, that differences in carriage duration combined with assortative mixing between different ages of children (as opposed to children and adults) could contribute to maintaining coexistence.

Supporting Text 8: Genome wide association studies on resistance and duration of carriage

Because of the association between duration of carriage and resistance, we expect to see substantial overlap between variants associated with a resistant phenotype and variants associated with a long duration of carriage. However, this prediction is dependent on the GWAS (genome-wide association study) methodology: duration of carriage is *not* associated with resistance conditional on the resistance allele (i.e. if a strain has the resistance allele, its duration of carriage does not affect whether it will be resistant or not). Therefore, a GWAS for resistance that accounted for genetic background and returned only variants associated with resistance *conditional on the rest of the genome*, would not detect variants associated with duration of carriage. In practice, this could occur as a result of conditioning on the clonal structure of the population if resistance alleles were strongly population stratified. Furthermore, the association between resistance and long duration of carriage only exists within the range of antibiotic consumption rates in which coexistence is observed. Whether a GWAS for duration of carriage would detect resistance variants would therefore depend on the antibiotic consumption rate at which the samples used in the GWAS were collected.

Supporting Tables

Table 1: Kendall's tau and p-value for the association between serotype frequency and resistance in Massachusetts, Malawi and Maela. Although duration of carriage is not the only factor determining serotype frequency, it is likely to be a reasonable proxy: in the Kilifi dataset, serotype frequency and carriage duration are correlated (Kendall's tau = 0.60, p = 2.1×10^{-6}). The number of serotypes was 28, 18 and 65 for the Massachusetts, Malawi and Maela datasets respectively. P values are one-tailed and bold P values are below 0.05.

Dataset	Antibiotic	tau (95% CI)	p
Massachusetts	Penicillin	0.53 (0.26, 0.66)	**0.0003**
Malawi	Penicillin	0.38 (0.03, 0.62)	**0.02**
Maela	Penicillin	0.49 (0.32, 0.58)	**1.4×10^{-7}**
Malawi	Co-trimoxazole	-0.11 (-0.43, 0.23)	0.74
Maela	Co-trimoxazole	0.20 (0.03, 0.35)	**0.01**
Massachusetts	Trimethoprim	0.40 (0.14, 0.58)	**0.004**
Massachusetts	Erythromycin	0.37 (0.11, 0.56)	**0.006**
Massachusetts	Ceftriaxone	0.27 (0.01, 0.49)	**0.04**

Table 2: Kendall's tau and P value for the association between serotype duration of carriage in Maela using Turner et al.'s duration of carriage estimates from mothers rather than infants and combined P values for penicillin and co-trimoxazole resistance (using Fisher's method, see Table 1 in main text). Bold P values are below 0.05.

Dataset	Antibiotic	tau (95% CI)	p
Maela	Penicillin	0.15 (-0.04, 0.32)	0.086
Maela	Co-trimoxazole	0.05 (-0.14, 0.23)	0.31
Combined	Penicillin	-	**0.0067**
Combined	Co-trimoxazole	-	0.54

References

[1] Mostowy R, et al. (2014) Heterogeneity in the frequency and characteristics of homologous recombination in pneumococcal evolution. *PloS Genetics* 10(5):e1004300.

[2] Davis SM, Deloria-Knoll M, Kassa HT, O'Brien KL (2013) Impact of pneumococcal conjugate vaccines on nasopharyngeal carriage and invasive disease among unvaccinated people: review of evidence on indirect effects. *Vaccine* 32(1):133–145.