Crosstalk and eavesdropping among quorum sensing peptide signals that regulate bacteriocin production in *Streptococcus pneumoniae*

Eric L. Miller[a,b,§], Morten Kjos[c,d,§], Monica Abrudan[e], Ian S. Roberts[a,1], Jan-Willem Veening[c,f,1] and Daniel E. Rozen[a,b,1]

[a] Faculty of Biology, Medicine, and Health, University of Manchester, Manchester, M13 9PL, UK.

[b] Institute of Biology Leiden, Leiden University, Leiden, 2333 BE, The Netherlands.

[c] Molecular Genetics Group, Groningen Biomolecular Sciences and Biotechnology Institute, Centre for Synthetic Biology, University of Groningen, Groningen, 9700 AE, The Netherlands.

[d] Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, N-1432 Ås, Norway

[e] Wellcome Trust Sanger Institute, Genome Campus, Cambridge, CB10 1SA, UK

[f] Department of Fundamental Microbiology, Faculty of Biology and Medicine, University of Lausanne, Biophore Building, CH-1015 Lausanne, Switzerland

[§] Both authors contributed equally to this manuscript.

[1] Correspondence to: d.e.rozen@biology.leidenuniv.nl

Jan-Willem.Veening@unil.ch

i.s.roberts@manchester.ac.uk

Classification: Biological Sciences; Genetics / Microbiology

Short title: Intraspecific crosstalk in *blp* quorum sensing

Key words: Streptococcus, quorum sensing, intra-species competition, signaling, eavesdropping, crosstalk

Abstract

During colonization of the human nasopharynx, multiple strains of the Gram-positive pathogen *Streptococcus pneumoniae* coexist and compete with each other using secreted antimicrobial peptides called bacteriocins. The major class of pneumococcal bacteriocins is encoded by the *blp* operon, whose transcription is controlled by the secretion and detection of a polymorphic family of quorum sensing (QS) peptides. We examined interactions between the *blp* QS signal (BlpC) and receptor (BlpH) across 4,096 pneumococcal genomes. Imperfect genomic concordance between nine QS signal peptides and five phylogenetically-related QS receptor groups suggested extensive signal crosstalk (where cells produce signals that non-clonal cells detect) and eavesdropping (where cells respond to signals that they do not produce). To test this, we quantified the response of reporter strains containing one of six different *blp* QS receptors to cognate and non-cognate synthetic peptide signals. As predicted, we found evidence for eavesdropping in four of these receptors and for crosstalk in five of six tested signals. These *in vitro* results were confirmed during interactions between adjacent pneumococcal colonies, providing evidence that crosstalk and eavesdropping occur at endogenous, ecologically-relevant, levels of signal secretion. Using a spatially explicit stochastic model, we show that eavesdropping genotypes gain evolutionary advantages during inter-strain competition, even when their affinity to non-cognate signals is as low as 10% of the affinity to their cognate signal. Our results highlight the importance of social interactions in mediating intraspecific competition among bacteria and clarify that diverse competitive interactions can be mediated by polymorphism in QS systems.

Significance Statement

Quorum sensing (QS), where bacteria secrete and respond to chemical signals to coordinate population-wide behaviors, has revealed that bacteria are highly social. Here, we use bioinformatics, experiments, and simulations to investigate how diversity in QS signals and receptors can modify social interactions controlled by the QS system regulating antimicrobial peptide secretion in *Streptococcus pneumoniae*. We experimentally confirmed that single receptors can detect multiple signals (eavesdropping) and single signals activate multiple receptors (cross-talk), while simulations revealed that eavesdropping can be evolutionarily beneficial even when the affinity for non-cognate signals is very weak. Our results highlight the importance of eavesdropping and crosstalk as drivers of the outcome of competitive interactions mediated by bacterial quorum sensing.

Introduction

Quorum sensing (QS) is a mechanism of intercellular communication that allows bacterial populations to coordinately regulate gene expression in response to changes in population density. QS is controlled by the secretion and detection of diffusible signaling molecules that, at threshold concentrations, lead to increased signal secretion as well as the induction of coupled downstream pathways (1, 2). By this process, QS ensures that pathways for metabolically costly products are only induced when bacterial populations would benefit from their production, i.e. when they are at high concentrations (2, 3). Signaling in QS systems is coordinated by the fact that cells are simultaneously capable of sending and detecting a specific signal (2–4), a characteristic that serves to increase the likelihood that QS functions as a private message among clonemates, thereby coordinating gene regulation between cells sharing evolutionary interests (5, 6). However, outside the laboratory, bacteria often reside in multispecies communities where secreted QS signals can be detected by any cell, not just their clonemates (3). Thus, although QS works as an effective means of gene regulation in the laboratory in single strain cultures, QS in nature may be less reliable because it is susceptible to signal eavesdropping (i.e. where a promiscuous QS receptor can detect a QS signal not produced by that genotype) and signal crosstalk (i.e. where a non-specific QS signal can activate QS receptors in genotypes that produce other QS signals) (Fig. 1A; 3, 7). Signal-blind bacteria that produce, but are incapable of responding to, QS signals can engage in signal crosstalk in order to manipulate the behavior of other cells, e.g. by inducing them to produce expensive public goods (8). However, crosstalk and eavesdropping can occur even if all cells within a population are otherwise phenotypically wild-type if (i) QS signals and receptors are polymorphic and (ii) signals can bind and activate more than one receptor variant. At present, the incidence and consequences of these deviations from canonical QS systems remain poorly understood. Here we examine these issues using the polymorphic QS system regulating bacteriocin production in the Gram-positive opportunistic pathogen *Streptococcus pneumoniae*, where QS is integral for mediating intraspecific competition.

To initiate infection, *S. pneumoniae* must first successfully colonize the nasopharynx and then persist during subsequent colonization attempts from other strains. Commensal carriage of *S. pneumoniae* is ubiquitous, affecting up to 88% of children worldwide (9, 10), while between 5-52% of individuals are co-colonized with multiple strains (10–13). The interactions between different strains during colonization are thus widespread and dynamic, leading to rates of clonal turnover — where one strain displaces another — that occur on a timescale of days to months (14, 15). Among the key factors thought to mediate intraspecific competition among pneumococcal strains are small anti-microbial peptides with narrow target ranges called bacteriocins, many of which are regulated by QS. While some studies in this species have examined the dynamics of bacteriocin-mediated competition *in vitro* (16, 17) and *in vivo* (16) between strains either lacking or carrying a single QS type, there is limited understanding of the evolution and ecology of polymorphic QS systems that regulate bacteriocin production and immunity between competing strains. *S. pneumoniae*

91 presents an ideal opportunity to study the evolution of QS systems in an easily manipulated, highly relevant study
92 system.

93 The genome of *S. pneumoniae* encodes several bacteriocin families, the most diverse of which is the bacteriocin
94 suite encoded by the *blp* (*b*acteriocin-*l*ike *p*eptides) operon (16, 17). Our recent work revealed that the number of
95 distinct combinations of bacteriocins and immunity genes can extend into the trillions, although phylogenetic and
96 functional constraints reduce this number to several hundred realized combinations (18). As with other Gram-positive
97 peptide signals, the QS signal peptide (BlpC) regulating the *blp* operon is constitutively produced at low levels, but is
98 auto-induced at high levels once a threshold concentration has been reached (17). Secreted BlpC binds to the
99 extracellular domain of the membrane-bound histidine kinase BlpH, and upon binding the kinase phosphorylates the
100 response regulator BlpR (Fig. 1B; 14, 15), which initiates production of the *blp* bacteriocin genes and increases
101 production of the BlpC signal. Additionally, *blpC* expression is enhanced by the induction of competence for genetic
102 transformation, which is regulated by the paralogous *com* QS signaling system (21). Both ABC transporters BlpAB (22)
103 and ComAB (21, 23) cleave the N-terminal, double-glycine leader sequence of BlpC before export of the mature
104 peptide signal by the same transporters (Fig. 1B). Using QS to regulate their secretion ensures that Blp peptides are only
105 produced when there is a sufficiently high cell number to allow the bacteriocins to reach effective concentrations.

106 Importantly, both the BlpC signal and its dedicated receptor, BlpH, are highly polymorphic. Our survey of 4,096
107 genomes identified 29 amino acid variants of the BlpC gene and 156 BlpH amino acid variants (18). What are the
108 effects of this variation, and how does this diversity influence the competitive interactions between strains that are
109 affected by *blp* bacteriocins? One extreme option is that each BlpC signal corresponds to a non-overlapping set of BlpH
110 receptors to which it specifically and exclusively binds. By this explanation, strains respond only to their own signal in
111 determining the threshold at which they induce the *blp* operon. Such exclusivity is found in the competence signaling
112 system where the two dominant peptide signals, CSP1 and CSP2, only induce cells expressing the cognate receptor (24).
113 Similarly, there is near absolute concordance between the signal and receptors carried by any single genome, suggesting
114 that tight coupling of these loci is crucial for the activation of competence (25). An alternative possibility is that BlpC
115 peptides cross-react via crosstalk or eavesdropping with different BlpH receptors with varying affinity levels, thereby
116 leading to a scenario whereby strains interact and possibly cause competing strains to induce the production of either
117 immunity or bacteriocins at densities that would be insufficient for activation by auto-induction. Previous experimental
118 research supported restricted crosstalk in the BlpH receptor, with a trade-off between signal sensitivity and signal
119 specificity (26). By this explanation, strains may benefit if they are forewarned of the threats from others, allowing
120 them to induce their own bacteriocins or immunity. Alternatively, eavesdropping may be costly if strains with
121 promiscuous receptors are induced to secrete bacteriocins at densities that are too low to provide sufficient benefits to
122 offset the costs of their production.

To understand the incidence and consequences of crosstalk and eavesdropping in a QS system that contains functional diversity, we investigated the fidelity between BlpC signals and BlpH histidine kinases across thousands of genomes using a bioinformatics approach. These results then informed experiments that quantified the response of bacterial strains to cognate and non-cognate peptide signals across the major signaling classes. Finally, these results were then examined in light of a stochastic model that investigated the consequences of QS eavesdropping for bacteriocin regulation. Our results reveal the importance of QS signaling polymorphism on *blp* operon regulation and clarify its ecological effects on *S. pneumoniae* intraspecific interactions. More broadly, we explore crosstalk and eavesdropping in a QS system that contains functional diversity.

Results

Molecular diversity of *blpH* and *blpC*

We examined 4,096 *S. pneumoniae* genomes across six unbiased data sets (Maela, Massachusetts Asymptomatic, GenBank, Hermans, Georgia GenBank, and PMEN) alongside two additional data sets that represent intentionally biased samples of clonal sub-groups (Complex 3 and PMEN-1, 322 genomes combined). We identified *blpC* in 99.0%, *blpH* in 99.0%, and both *blpC* and *blpH* in 98.2% of the 4,418 genomes using a DNA reciprocal BLAST algorithm (18). We note that the few genomes apparently lacking a *blp* operon gene may still contain these genes, as the data sets contain incomplete draft genomes.

We observed extensive allelic variation within *blpC*, which contains 37 alleles at the nucleotide level, 29 protein variants, and 20 different mature BlpC signal peptides, including signal peptides lacking a double-glycine cleavage site. Nine of these peptide signal sequences were found in more than 0.5% of genomes (i.e., over 20 genomes; Table 1), and together these nine comprise ~98% of all signal variants. All signals under this 0.5% threshold were each confined to a single clade in the whole-genome phylogeny (Fig. S1A). Each unique BlpC signal was designated with a letter from the NATO phonetic alphabet. As expected for the genomes from intentionally biased samples, the PMEN-1 data set almost exclusively carried the Golf signal (93.8%; Table 1), while the Clonal Complex 3 data set almost exclusively carried the Kilo signal (97.6%; Table 1). The Bravo and Hotel signal peptides were exclusively found in strains collected as part of the Maela data set. Even though the Maela genomes composed the majority of genomes in our data, the relatively high proportion of Maela genomes with the Bravo and Hotel signals (3.8% and 8.2%, respectively) suggests that either natural selection or limited admixture prevented these signals from appearing in the other data sets. There was more variation in *blpH* (194 alleles at the nucleotide level across 156 amino acid variants) than in *blpC*, although this is a likely consequence of the fact that *blpH* is longer than *blpC*. Rarefaction curves of non-singleton protein variants of BlpC and BlpH show that the diversity of protein variants reached saturation after ~2,000 and ~3,000 sampled genomes, respectively (Fig. S1B).

Indicators of molecular evolution showed that *blpH* and *blpC* are evolving rapidly (Table 2). *blpH* had the highest peaks of nucleotide diversity across the *blp* regulatory genes (Fig. S2A). Additionally, higher d_N / d_S ratios (i.e. a higher non-synonymous to synonymous mutation ratio) in the receptor versus the kinase region of this protein (Table 2) suggests stronger diversifying selection acting on peptide:receptor binding than on the intracellular kinase domain responsible for downstream signal transfer by phosphorylation. Inferred rates of recombination peaked within the second transmembrane domain of BlpH; however, specific peaks of either nucleotide diversity or recombination were not distinguished by predicted transmembrane regions (Fig. S2B). Recombination also locally peaked within *blpC* (Fig. S2B), which also had an increased d_N / d_S ratio when compared to estimates for housekeeping genes (Table 2).

blpC/blpH intragenomic pairing is highly biased but not exclusive

Phylogenetic analysis of *blpC* revealed four well-supported clades (Fig. 2) containing the following signals: 1) Alpha, Bravo, and Kilo; 2) Golf and Hotel; 3) Charlie; and 4) Delta, Echo, and Foxtrot. The relationships between signaling groups within these major clades are uncertain, although there is evidence (0.75 < posterior probability < 0.95) that the Hotel, Bravo, and Delta signals are each monophyletic within their respective larger clades.

After accounting for recombination, phylogenetic analysis of *blpH* identified five paraphyletic clades that are broadly concordant with the divisions observed for BlpC signals (Fig. 3), although there are clear exceptions to this correspondence. Across the five clades, the classification of *blpH* alleles correlated with the mature BlpC signal in at least 75% of cases: (Alpha / Bravo / Kilo Clade: 88.5%; Echo / Foxtrot Clade: 94.6%; Delta Clade: 100%: Charlie Clade: 86.3%; Golf / Hotel Clade: 75.5%). Examples of a mismatched BlpC signal and *blpH* receptor occur at 27 distinct clades on the *blpH* tree. Evidence of extensive recombination affecting the *blpH* kinase, intergenic region, and *blpC* signal (Fig. S2) suggests that recombination has caused some of these mismatches; however, multiple recombination events in the same region obscure reconstructing these evolutionary events with confidence. Overall, from the 4,096 genomes in this analysis, 596 genomes (14.6%) show a lack of correspondence between signal and peptide, suggesting either that these strains are deficient in *blp* signaling or that these BlpH histidine kinase receptors can be cross-induced by non-cognate BlpC signals. Overall frequencies by signal and receptor class are summarized in Fig. 4A.

Crosstalk and eavesdropping between BlpC signals and BlpH receptors

To examine the incidence of crosstalk and eavesdropping between signals and receptors experimentally, we measured the responsiveness of each of the major BlpH clades to synthetic peptides from each BlpC class. We transformed a *S. pneumoniae* D39 strain lacking the native *blp* regulatory genes (*blpSRHC*) with constructs expressing one of six different BlpH histidine kinases alleles: $blpSRH^{D39}$ from the Alpha/Bravo/Kilo clade, $blpSRH^{PMEN\text{-}2}$ from the

Echo/Foxtrot clade, *blpSRH*$^{\text{Hermans-33}}$ from the Delta clade, *blpSRH*$^{\text{Hermans-1012}}$ and *blpSRH*$^{\text{PMEN-14}}$ from the Charlie clade, and *blpSRH*$^{\text{PMEN-18}}$ from the Golf/Hotel clade. The respective strains also contained a reporter cassette, in which a typical *blp*-promoter (P_{blpK} or P_{blpT}) was controlling expression of firefly luciferase (*luc*), GFP ((*sf*)*gfp*), and β-galactosidase (*lacZ*) (21). Deletion of the *blpC* signal gene and the native *blpSRH* genes from the D39 ancestor ensured that the reporter strains would only be induced in response to exogenously added signal via the introduced *blpSRH* systems. By exposing cells to a concentration gradient of exogenous peptide, we could estimate the peptide concentration that induced the maximum response as well the minimum concentration required to elicit a response. While the maximum response indicates the overall influence of a given peptide on each receptor, the minimal concentration required to induce a response provides an indication of the sensitivity of each receptor to every potential peptide partner. Figures 4A-B shows that five of six P_{blpK} reporter strains were maximally induced by the BlpC signal carried by a significant majority of their wild type counterparts. However, we also see extensive evidence for crosstalk and eavesdropping between mismatched peptide:receptor pairs, demonstrating that some BlpH receptors are highly promiscuous while equally, several BlpC peptides can induce the *blp* operon in strains carrying non-complementary BlpH receptors. For example, *blpSRH*$^{\text{PMEN-2}}$ (Echo / Foxtrot BlpH clade) could be induced by 4 out of 6 synthetic peptides, and the strain with *blpSRH*$^{\text{Hermans-1012}}$ (Charlie BlpH clade) was strongly induced by the Echo and Foxtrot signals at 65% and 71% expression of its cognate signal. While there is clear evidence for cross-induction, these responses tended to be less sensitive to non-cognate peptides, with a minimum concentration required for induction of between 2-500-fold greater than with the cognate signal (Fig. 4C). By contrast, the strain with *blpSRH*$^{\text{Hermans-1012}}$ (Charlie BlpH clade) was more sensitive to the non-cognate Echo and Foxtrot signals (1 ng/ml and 3.9 ng/ml) than to its complementary Charlie signal (7.8 ng/ml; Fig. 4C). The reporter strain carrying *blpSRH*$^{\text{Hermans-33}}$ did not respond to any of the BlpC peptides, not even its cognate Delta BlpC (Fig. 4B-C). Interestingly, *blpSRH*$^{\text{Hermans-33}}$, as well as all other strains with *blpH* alleles in the Delta clade, contains a frameshift in the *blpR* gene, encoding the response regulator, thus preventing expression of the full-length *blpR*. This probably renders the QS systems non-functional and therefore not responsive to added peptide. All results were mirrored with a different set of reporter strains that used the *blpT* promoter for the reporter cassette (Fig. S3A). We conclude from these results that crosstalk among quorum-dependent peptide BlpC signals is widespread and concentration dependent, with strains able to eavesdrop onto multiple signals using cross-responsive receptors. Furthermore, these results are highly concordant with the patterns of co-association observed in our bioinformatics survey of pneumococcal strains.

Cross-induction between colonies

Pneumococci in the nasopharynx live in spatially structured colonies or biofilms. In order to determine if cross-induction could occur under these conditions, we examined interactions between neighboring colonies endogenously

219 secreting either cognate or non-cognate signals (Fig. 5). In control assays, we first demonstrated that colonies were
220 induced by exogenous addition of peptide to the plate surface; these results were concordant with those in Figure 4B in
221 14 of 15 combinations (Fig. S3B). Next, we measured expression of reporter strains when grown adjacent to wild-type
222 colonies that secreted BlpC peptides at endogenous levels (Fig. 5A). We observed a response in the reporter strains as
223 estimated by increased LacZ activity in 3 out of 6 strains, with 2 examples of induction by non-cognate BlpC signals.
224 Interestingly, when the reporter strain expressing the BlpH from Hermans-1012 was grown adjacent to its wild type
225 counterpart, there was no induction; instead this strain was induced by PMEN-14, which also produced the Charlie
226 signal. The same strain was also induced by PMEN-2, which produced the Foxtrot signal (which induces Hermans-
227 1012 at a lower concentration than with its cognate signal; Fig. 4C), and strain PMEN-18 (Golf/Hotel BlpH clade) was
228 induced by PMEN-14, which produced the Charlie signal (Fig. 5). This may suggest that in addition to differences in
229 the binding affinities of BlpC and BlpH, strains may also vary in the concentration of diffusible signal that they secrete,
230 at least under these experimental conditions. Consistent with our *in vitro* assays with synthesized peptides, these results
231 show that *blp* operon expression can be activated by crosstalk between neighboring competing colonies secreting
232 peptides at wild-type concentrations.
233
234 **Evolutionary consequences of eavesdropping genotypes**
235 Because the *blp* operon is auto-induced via a quorum dependent process, cross-induction can potentially
236 influence other strains by lowering the population density required for auto-induction. To examine this issue, we
237 developed a spatially explicit stochastic model to investigate conditions where genotypes with eavesdropping receptors
238 may be favored over strains only able to respond to a single peptide signal. We further varied the signal affinity to
239 eavesdropping receptors to determine how this altered the selective benefits of cross-induction. Simulations are initiated
240 with cells of four genotypes randomly spaced upon a plane. The four genotypes each release their own QS signal at
241 equal concentrations (Table S1). Cells bind these secreted signals in a concentration dependent manner, at which point
242 they are induced to produce bacteriocins that kill susceptible neighbor cells at the cost of reduced growth for the
243 producer (27). While two faithful-signaling genotypes are only able to respond to their own signals, the two other
244 eavesdropping genotypes can respond to multiple signals. Our results shown in Fig. 6 lead to two important conclusions.
245 First, we observe strong benefits to eavesdropping cells that depends on the degree of cross-sensitivity, or affinity, to
246 non-cognate signals. Specifically, we found that higher affinity to non-cognate signal provides stronger ecological
247 benefits. This results from earlier potential activation (Fig. S4) and secretion of bacteriocins in these cells, an effect that
248 increases with greater affinity to non-cognate signals. Second, we find that the benefits to eavesdropping are strongly
249 negative frequency-dependent, i.e. eavesdropping cells only gain benefits (in the form of earlier bacteriocin induction)
250 when surrounded by faithful-signaling cells. When eavesdropping cells are rare, they benefit through maximum

exposure to the alternative peptide, while after they increase they must rely solely on auto-induction. Because the benefits of eavesdropping are frequency dependent, these simple simulations thus suggest that promiscuous receptor mutants with increased affinity to non-cognate signals will be able to rapidly invade populations of cells that can only respond to a single signal. Interestingly, the simulations also clarify that the affinity to non-cognate signals can be extremely low — even at 10% of the affinity to cognate signals — to provide benefits (Fig. 6).

Discussion

Pneumococcal bacteriocins are believed to play a key role in mediating intraspecific competitive interactions. Here, we show that the QS system regulating *blp* bacteriocins is highly polymorphic, that there is widespread evidence that QS signals are cross-reactive (crosstalk), and that promiscuous receptors can detect and respond to non-cognate signals (eavesdropping). Assays between adjacent colonies revealed that both behaviors are manifest at endogenous concentrations of secreted peptides, while simulations revealed significant ecological benefits to strains that express promiscuous receptors. Together, these results suggest that social interactions influenced by QS signaling may have a strong influence on pneumococcal competition.

Previous surveys (19, 28) of BlpC and BlpH identified four BlpC signals: the Alpha, Charlie, Foxtrot, and Golf signals in our nomenclature, which together represent ~75% of the strains in our sample (Table 1). By expanding our survey to 4,096 strains, we identified several additional signal peptide families (Fig. 2): the Echo, Hotel, Delta, Bravo, and Kilo signals. Two of these (Bravo and Hotel) were found only in the Maela dataset, consistent with the idea of strong geographic structuring in this species; furthermore, because these signal variants appear in significant frequencies (3.8% and 8.1%, respectively, within the Maela data set), we infer that they are not actively selected against. The additional signals reported here suggest flexibility (and perhaps diversifying selection) in the mature signal peptide sequence, such as in the first two residues (which differ between the Alpha, Bravo, and Kilo signals, which share a co-occurring BlpH clade) and in signal residue 22, which differs between the Echo and Foxtrot signals despite the signals activating BlpH variants near identically in our experiments (Fig. 4B, Fig. 4C). Previous work suggested that differences in the electric charge of signal residue 14 is crucial for specificity (26); this residue is undoubtedly important because it differentiates the Alpha/Bravo/Kilo, Echo/Foxtrot, and Golf/Hotel signal groups. However, other signal residues are also likely to be important for BlpH binding, as signals that are identical at this site 14 (e.g. Golf/Hotel, and Alpha/Bravo/Kilo/Charlie) differentially bind/activate BlpH (Fig. 4). Although rarefaction analysis indicates that we have essentially saturated the diversity of signal and receptor types (Fig. S1B), more extensive sampling in other geographic locations, such as with the Global Pneumococcal Sequencing project (http://www.pneumogen.net/gps), is likely to uncover further rare variants.

The concordance between the phylogenies of *blpC* and *blpH* and the extensive co-occurrence in individual genomes suggest that these genes are co-evolving (Fig. 2, Fig. 3). At the same time, both *blpC* and *blpH* are changing rapidly, as indicated by their relatively high levels of nucleotide diversity (Fig. S2) and their high non-synonymous / synonymous substitution rates (Table 2), which is consistent with the idea that the genes are evolving non-neutrally. The co-variation between BlpC and BlpH also allows inferences on the key residues within each gene that mediate their binding. We used mutual information based upon this co-variation to identify residues in BlpH that are crucial for binding the QS signal, and these results can serve as a guide for experimental approaches to unravel the specificity of the BlpC/BlpH interaction (Fig. S5). Notably, our analysis of the BlpH receptor residues support previous findings that residues 17 and 119-124 are important for activation by BlpC signals (26), although additional residues co-vary with specific BlpC signals (Fig. S5).

While the correlation between *blpH* clade and co-occurring BlpC signal is high, in some clades the correlation drops to 75.5%, and BlpH / BlpC mismatches (Fig. 3) are widespread across the pneumococcal phylogeny; this can be compared to the exceptionally tight, > 99% correlation between the ComD QS receptor and CSP signal also in *S. pneumoniae* (25). There are at least two explanations for this difference. First, we do not have know if different BlpH variants are functionally distinct; all *blpH* alleles could, in principle, be most responsive to their co-occurring BlpC. This seems unlikely, given the high frequency (up to 36 signal:receptor pairs) of *blpH* clade / BlpC mismatches (Fig. 3). Second, weaker selection for a highly auto-inducing *blp* QS could explain the difference between the *blp* and *com* QS systems. After a recombination event that results in a sub-optimal BlpH/ BlpC pair for auto-induction, the BlpC signal may still be able to active the co-occurring BlpH variant through crosstalk, albeit at a higher concentration of BlpC (Fig. 4C). While auto-induction may be decreased, such a genotype would gain an eavesdropping receptor that can potentially detect signals of surrounding genotypes. For comparison, there is no eavesdropping between CSP pherotypes in the *com* QS system, and very rare signal/receptor mismatches (24, 25).

These signal/receptor mismatches can result in two outcomes for cell-to-cell communication. First, this can result in cells unable to detect the signal that they produce, rendering them unable to auto-induce. The lack of QS activation in strains producing the Delta signal (Hermans-33; Fig. 4) seemingly fits into this description; however, interestingly, this is not caused by signal / receptor mismatch because there is perfect concordance between the Delta signal and the Delta *blpH* clade, and no tested signal activated strains with Delta *blpH*. Instead, all 143 strains carrying the Delta signal have a frameshift in *blpR*, which suggests functional deterioration of the QS system in these strains, which has not yet led to deterioration of *blpH* and *blpC*. These Delta BlpC strains are not simply 'cheater' cells, as they continue to pay the cost of synthesizing BlpC. This suggests there may be weakened selection for functional *blp* QS.

The second outcome of signal/receptor mismatches for cell-to-cell communication is crosstalk and eavesdropping. We have ample evidence for crosstalk in the *blp* QS system, as all signal peptides except for the Alpha

315 signal activated QS receptors in genotypes that produce other QS signals (Fig. 4B, Fig. 4C). Similarly, BlpH receptors
316 (aside from the Alpha clade) were eavesdropping QS receptors able to detect more than one QS peptide signal (Fig. 4B,
317 Fig. 4C). Each of the receptors we tested (except for the signal-blind BlpH Delta clade) was maximally induced with a
318 single set of related signals and decreased to 0-71% with signals that the receptors were eavesdropping upon (Fig 4B).
319 This suggests that there are no 'generalist' receptors that are able to listen to multiple signals with equal affinity.
320 Previous research indicated that *blpH* alleles with more crosstalk were less sensitive to BlpC (26); however, reporter
321 strain PMEN-2, Hermans-1012, and PMEN-14 were all sensitive to their complementary signal (≤1.0 ng/ml) but show
322 extensive crosstalk (Fig. 4C). Crosstalk was seen in previous research, specifically induction from both the Charlie
323 (signal 6A in (26)) and Foxtrot signal (signal 184 in (26)) in strains with a BlpH Charlie clade allele (6A.3 in (26)), and
324 in induction from the Charlie and Golf (signal T4 in (26)) signals with the BlpH Golf clade (T4 in (26)). These two
325 cases do require 2-7.8 times more of their own signal for minimum activation, compared to our BlpH Alpha clade and
326 BlpH Foxtrot clade strains; this finding does supports a trade-off between signal sensitivity and specificity (26) but not
327 to the exclusion of crosstalk. Overall, our experimental evidence supports extensive crosstalk and eavesdropping in the
328 *blp* QS system.

329 What are the potential consequences of crosstalk and eavesdropping? The result of crosstalk could be to
330 manipulate other, non-clonal, strains into inducing their QS system at lower densities, thereby causing competing
331 strains to secrete bacteriocins and induced immunity proteins earlier. At present, it is unclear how such crosstalk would
332 be beneficial to cells producing cross-reactive signals, unless premature production of bacteriocins or immunity
333 introduces energetic or other costs to cells responding at sub-quorum densities. Similar benefits are thought to exist for
334 other bacterial public goods (8, 29). By contrast, it is easier to envision the potential benefits of eavesdropping, which
335 can both lead to earlier activation of bacteriocins (although this may have attendant costs) and earlier induction of cross-
336 reactive immunity. Importantly, our simulations suggest that this could be beneficial even if the affinity of promiscuous
337 receptors is only 10% of the affinity for their cognate signal (Fig. 6). This value falls within the range of responses we
338 measured experimentally (Fig. 4C). This level of responsiveness is also sufficient to induce the *blp* operon among
339 adjacent colonies secreting peptides at endogenous levels (Fig. 5).

340 Our results make clear that signaling interactions *in vitro* can lead to complex ecological outcomes that may
341 strongly influence competitive interactions between strains. As yet, however, it is unclear how these interactions will
342 play out in the complex within-host environment of the human nasopharynx. In addition, it remains unclear how these
343 interactions directly influence bacteriocin-mediated killing and immunity. Clearly, the heterogeneous conditions *in vivo*
344 differ markedly between liquid cultures or even agar surfaces. Diffusion is more limited, while population densities may
345 be strongly constrained overall and spatially. These factors, among others, may alter the level and dispersion of signal
346 peptides as well as the sensitivity of individual strains to these signals. It will remain an important aim for future work

to elucidate the influence of these real-world factors. More generally, our results reinforce the importance of social interactions among bacteria for mediating competitive dynamics. Many ecologically relevant bacterial traits are regulated by QS, and many of these systems, especially in Gram-positive peptide signaling systems, are polymorphic. While some of these systems (e.g. pneumococcal competence regulated by the *com* QS system) have only few signal types and show no cross-reactivity, many others (e.g. Agr in *S. aureus* (30) and ComX in *B. subtilis* (31)) do or have yet to be measured. It remains to be investigated which, if any, of these polymorphic QS signals have ecological effects and which factors (such as co-colonization or extensive intraspecific competition) result in the evolution of crosstalk and eavesdropping.

Materials and Methods

Genomic data

We used *S. pneumoniae* genomes from eight publicly available sets, six of which contain strains that were randomly sampled from either cases of disease or asymptomatic carriage: 3,017 genomes from refugees in Maela, Thailand (32); 616 genomes from Massachusetts carriage strains (33); 295 genomes from GenBank, which include 121 genomes from Atlanta, Georgia, The United States (34); 142 genomes from Rotterdam, the Netherlands (Hermans data set) (18, 35); and 26 PMEN (Pneumococcal Molecular Epidemiology Network) genomes (18, 36). The PMEN-1 (37) and Clonal Complex 3 (38) data sets, containing 240 and 82 genomes, respectively, were a result of targeted sampling for specific clonal complexes of *S. pneumoniae*; as such, these strains were excluded from analyses that assumed random sampling. In each of these genomes, we located the *blpC* and *blpH* alleles using a DNA reciprocal BLAST search as previous described (18).

Phylogenetic Analysis

We examined if the rates of transition in changing BlpC signals were independent from the rates of transition between bacteriocin groups along the whole-genome phylogeny from (18). We did so by measuring the likelihood while estimating both independent and dependent transitions using BayesTraits (39) for each signal and bacteriocin group separately. We used a log ratio test to measure P-values from the likelihoods.

To infer the evolutionary history of *blpC*, we aligned the 37 alleles of *blpC* from all genomes, removed nucleotide sites caused by insertion mutations in single alleles, and reconstructed the phylogeny using Geneious 7.1.9 (40) and MrBayes 3.2.2 (40–42) with an HKY+γ nucleotide substitution model as determined by jModelTest 2.1.7 (43).

Because *blpH* contained evidence of recombination, we modified our approach to infer its evolutionary history. We aligned the unique 161 full-length ($\geq$ 1341 bp) *blpH* alleles using Geneious 7.1.9 (40). We then used Gubbins 1.4.2

(44) to detect and place recombination events onto a phylogenetic tree. To reconstruct the phylogeny while measuring the confidence for each clade, we used the filtered polymorphic sites from Gubbins, which replaced recombination fragments with N's, as input for MrBayes 3.2.2 (40–42) using a GTR+γ nucleotide substitution model as determined by jModelTest 2.1.7 (43). We defined *blpH* based on genetic association with BlpC; in cases of ambiguity, such as grouping the clade of six *blpH* variants with the Charlie clade located at position 9 o'clock on Figure 2 within the Charlie *blpH* clade instead of within the Golf/Hotel *blpH* clade, we used parsimony with regard to BlpC to assign *blpH* clades.

Sequence Analysis

We measured the nucleotide diversity and d_N / d_S ratio of *blpH*, *blpC*, as well as the seven housekeeping genes used for MLST analysis (*aroE*, *ddlA*, *gdhA*, *glkA*, *lepB*, *recP*, and *xpt*) using DNAsp 5.10.01 (45). We separated *blpH* after nucleotide 687 to measure these values for the receptor and kinase domains separately (26).

We extracted the region 1500 bp before *blpT* to 100bp after *blpB* from the 4,096 randomly sampled genomes from a previous alignment with R6_uid57859 (25), which included sequences with interrupted BlpA and BlpB genes (46). We measured nucleotide entropy (measured as Shannon entropy by site) and recombination breakpoints (measured from previously calculated data from GeneConv 1.81a (25, 47) for these sequences. Using only pairwise recombination breakpoints with unique start and stop positions prevented overestimating signs of recombination events caused by the common descent of strains.

Mutual information was calculated using a custom Python script. For each signal, we grouped signals into signal groups, such as Alpha:Bravo:Kilo or Golf:Hotel. We grouped amino acids as either acidic, aliphatic, aromatic, basic, cyclic, or hydroxyl residues. We then normalized mutual information to correct for the entropy of signals and of amino acids at each site.

Transmembrane domains in BlpH were predicted using TOPCONS 2.0 (48).

Bacterial strains and growth conditions

S. pneumoniae strains were grown as liquid cultures in C+Y medium (49) at 37°C and transformed as described previously (21). For selection, *S. pneumoniae* was plated on Columbia agar plates supplemented with 2% defibrinated sheep blood (Johnny Rottier, Kloosterzande, Netherlands) and 1 µg/ml tetracycline, 100 µg/ml spectinomycin or 0.25 µg/ml erythromycin, when appropriate. *E. coli* was grown in LB medium with shaking at 37°C or plated on LA

407 containing 100 µg/ml ampicillin.

408

409 Strain construction

410 Strains and plasmids used in this study are listed in Table S2.

411

412 Constructs for expression of *blpSRH* from different strains in *S. pneumoniae* D39.

413 The *blpSRH* genes, including the constitutive *blpS*-promoter (21), was amplified from the genome of *S.*
414 *pneumoniae* strains D39, PMEN-18 and Hermans-1012 using primers blpS-F-ClaI-SphI and blpH-R-Hermans-1012-
415 NotI-SpeI, from PMEN-2, using primers blpS-up-F-PMEN2-SphI and blpS-down-R-PMEN2-SpeI-NotI from PMEN14
416 with primers BlpS-PMEN14-F-SphI and BlpH-R-Hermans-1012-NotI-SpeI and from Hermans-33 with primers blpS-F-
417 ClaI-SphI and blpH-Hermans33/35-R-NotI-SpeI. The PCR products were digested with SphI and NotI and ligated into
418 the corresponding sites of plasmid pJWV25 (between the *bgaA* homology regions) and transformed into *E. coli* DH5α.
419 The resulting plasmids were verified by PCR and sequencing. The plasmids were then transformed into *S. pneumoniae*.
420 Correct integration of the P_{blpS}-*blpSRH* constructs into the non-essential *bgaA*-locus was verified by PCR. Primers used
421 for these constructs are listed in Table S3.

422

423 Deletion of *blpSRHC*.

424 The native *blp*-regulatory genes (*blpS, blpR, blpH, blpC*) of *S. pneumoniae* D39 were deleted by replacement
425 with an erythromycin-resistance cassette as described previously (21).

426

427 Reporter constructs

428 Two different *blp* promoters were used to monitor *blp*-expression; P_{blpK}, controlling expression of the bacteriocin
429 *blpK* and P_{blpT}, controlling expression of the functionally uncharacterized gene *blpT*. The P_{blpK} and P_{blpT} promoters have
430 previously been shown to be co-regulated (21) and act as reporters for BlpR activation across the tested pneumococcal
431 strains (as there is almost no variation in the DNA-binding motif of different BlpR). The reporter constructs P_{blpK}-*luc-*
432 *gfp-lacZ* and P_{blpT}-*luc-gfp-lacZ* (a tripartite reporter cassette) integrated into the non-essential *cep*-locus of *S.*
433 *pneumoniae* D39, has been described previously (21).

434

435 Luciferase assays

436 Luciferase assays were performed essentially as described before (21, 50). Briefly, *S. pneumoniae* cultures pre-
437 grown to OD_{600} 0.4 were diluted 100-fold in C+Y medium (pH 6.8) with 340 µg/ml luciferin. Luc-activity was

438 measured in 96-well plates at 37°C, and OD_{600} and luminesencence (as relative luminescence units, RLU) were
439 recorded every 10 minutes using Tecan Infinite 200 PRO. Synthetic peptides (BlpCs) were purchased from Genscript
440 (Piscataway, NJ). Different concentrations of BlpCs were added to the culture wells after 100 min or in the beginning of
441 the experiment, depending on the experiment. The data was plotted as RLU/OD over time to analyze induction of *blp*
442 expression.
443
444 **LacZ assays on agar plates**
445 LacZ assays were performed on C+Y agar plates (pH 7.2) covered with 40 µl of 40 mg/ml solution X-gal (spread
446 on top of the plates). All strains were pre-grown to OD_{600} 0.4, before 2 µl of the wild-type strains (BlpC producers)
447 were spotted and allowed to dry. Then 2 µl of the different reporter strains were spotted next to the dried spot. The
448 plates were incubated at 37°C over-night.
449 For induction with synthetic BlpC, C+Y agar plates were covered with 40 µl of 40 mg/ml solution X-gal and 5
450 µl 1 mg/ml BlpC (spread on top of the plates), and different reporter strains were spotted on top. The plates were
451 incubated at 37°C over-night.
452
453 **Stochastic Model**
454 We built a spatial, stochastic model in which cells are modeled individually and interact in a grid. At each
455 discrete time point, each cell can divide with probability 80%, which produces an identical offspring in an empty,
456 randomly chosen adjacent position on the grid. Each cell also secretes one of four signaling molecules, which
457 accumulate and diffuse in the space around the cell (defined as “diffusion area”). When the amount of signaling
458 molecules within the sensitivity area (which is an area with half the radius of the diffusion area) of a cell reaches a
459 defined threshold, this cell becomes induced and starts producing bacteriocins. A bacteriocin-producing cell can kill up
460 to six neighboring cells depending on their genotype, as explained below. Induced cells, which produce bacteriocins,
461 grow 20% slower than uninduced cells. Every cell has a 0.1% probability of death at each time point.
462 We modeled four genotypes, which differ in the signaling molecule and bacteriocins that they produce as well
463 as in the number and identity of signals that they respond to (Table S1). Bacteriocins produced by genotypes 1 and 2
464 specifically could kill genotypes 3 and 4 and vice versa. Signals produced by genotype 1 could induce genotypes 1 and
465 2 and similarly, signals produced by genotype 3 could induce genotypes 3 and 4; we therefore classify genotypes 2 and
466 4 as “eavesdropping genotypes”. Genotypes 1 and 3 can only respond to their own signal, as “signal-faithful genotypes”.
467 All four genotypes have equivalent growth rates, which are only variable depending on if a cell is induced or uninduced.
468 Eavesdropping cells respond to signals that they do not produce with certain degrees of affinity. If we consider the

469 affinity of a cell to its own signal as 100%, we ranged the affinity to the other signals in the case of eavesdropping
470 genotypes as 0% - 90% for different simulations.
471 We used an initial grid of 400 by 400 positions, and we started with 10% of the grid randomly populated with
472 eavesdropping and signal-faithful phenotypes at a range of proportions. As cells grow and interact, the center of the grid
473 remains occupied, while the cell population can expand on the boundaries past the initial 400 by 400 grid. Simulations
474 were run for until the number of cells reached 110% of initial grid size. We calculated fitness as the difference between
475 relative frequencies of eavesdropping cells at the last time point and at the initial time point, i.e.:
476 $$(\text{Eavesdropping cells/Total cells})_{\text{Final Time Point}} - (\text{Eavesdropping cells/Total cells})_{\text{Initial Time Point}}$$
477

478 **Acknowledgements**

479 We would like to thank Frank Lake for technical assistance.
480

481 **Funding**

482 This work was supported by the Biotechnology and Biological Sciences Research Council (grant number
483 BB/J006009/1) to DER and ISR and by the Wellcome Trust (105610/Z/14/Z) to the University of Manchester. MA is
484 supported by the Biotechnology and Biological Sciences Research Council (grant number BB/M000281/1). Work in the
485 Veening lab is supported by the EMBO Young Investigator Program, a VIDI fellowship (864.12.001) from the
486 Netherlands Organisation for Scientific Research, Earth and Life Sciences (NWO-ALW) and ERC starting grant
487 337399-PneumoCell. MK is supported by a grant from The Research Council of Norway (250976/F20).
488
489

References

1. Miller MB, Bassler BL (2001) Quorum Sensing in Bacteria. *Annu Rev Microbiol* 55:165–199.

2. Waters CM, Bassler BL (2005) QUORUM SENSING: Cell-to-Cell Communication in Bacteria. *Annu Rev Cell Dev Biol* 21(1):319–346.

3. Redfield RJ (2002) Is quorum sensing a side effect of diffusion sensing? *Trends Microbiol* 10(8):365–370.

4. Bassler BL, Greenberg EP, Stevens AM (1997) Cross-species induction of luminescence in the quorum-sensing bacterium *Vibrio harveyi*. *J Bacteriol* 179(12):4043–4045.

5. Strassmann JE, Gilbert OM, Queller DC (2011) Kin Discrimination and Cooperation in Microbes. *Annu Rev Microbiol* 65(1):349–367.

6. Schluter J, Schoech AP, Foster KR, Mitri S (2016) The Evolution of Quorum Sensing as a Mechanism to Infer Kinship. *PLOS Comput Biol* 12(4):e1004848.

7. Atkinson S, Williams P (2009) Quorum sensing and social networking in the microbial world. *J R Soc Interface* 6(40):959–78.

8. Diggle SP, Griffin AS, Campbell GS, West SA (2007) Cooperation and conflict in quorum-sensing bacterial populations. *Nature* 450(7168):411–414.

9. Regev-Yochay G, et al. (2004) Nasopharyngeal carriage of *Streptococcus pneumoniae* by adults and children in community and family settings. *Clin Infect Dis* 38(5):632–639.

10. Wyllie AL, et al. (2014) *Streptococcus pneumoniae* in saliva of Dutch primary school children. *PLoS One* 9(7):1–8.

11. Sauver JS, et al. (2000) Risk factors for otitis media and carriage of multiple strains of *Haemophilus influenzae* and *Streptococcus pneumoniae*. *Emerg Infect Dis* 6(6):622–630.

12. García-Rodríguez JA, Fresnadillo Martínez MJ (2002) Dynamics of nasopharyngeal colonization by potential respiratory pathogens. *J Antimicrob Chemother* 50 Suppl S:59–73.

13. Brugger SD, Frey P, Aebi S, Hinds J, Muhlemann K (2010) Multiple colonization with *S. pneumoniae* before and after introduction of the seven-valent conjugated pneumococcal polysaccharide vaccine. *PLoS One* 5(7):e11638.

14. Meats E, et al. (2003) Stability of serotypes during nasopharyngeal carriage of *Streptococcus pneumoniae*. *J Clin Microbiol* 41(1):386–392.

15. Turner P, et al. (2012) A longitudinal study of *Streptococcus pneumoniae* carriage in a cohort of infants and their mothers on the Thailand-Myanmar border. *PLoS One* 7(5). doi:10.1371/journal.pone.0038271.

16. Dawid S, Roche AM, Weiser JN (2007) The *blp* bacteriocins of *Streptococcus pneumoniae* mediate intraspecies competition both *in vitro* and *in vivo*. *Infect Immun* 75(1):443–451.

17. Lux T, Nuhn M, Hakenbeck R, Reichmann P (2007) Diversity of bacteriocins and activity spectrum in *Streptococcus pneumoniae*. *J Bacteriol* 189(21):7741–7751.

18. Miller EL, Abrudan MI, Roberts IS, Rozen DE (2016) Diverse ecological strategies are encoded by *Streptococcus pneumoniae* bacteriocin-like peptides. *Genome Biol Evol* 8(4):1072–1090.

19. De Saizieu A, et al. (2000) Microarray-based identification of a novel *Streptococcus pneumoniae* regulon controlled by an autoinduced peptide. *J Bacteriol* 182(17):4696–4703.

20. Reichmann P, Hakenbeck R (2000) Allelic variation in a peptide-inducible two-component system of *Streptococcus pneumoniae*. *FEMS Microbiol Lett* 190(2):231–236.

21. Kjos M, et al. (2016) Expression of *Streptococcus pneumoniae* Bacteriocins Is Induced by Antibiotics via Regulatory Interplay with the Competence System. *PLOS Pathog* 12(2):e1005422.

22. Håvarstein LS, Diep DB, Nes IF (1995) A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. *Mol Microbiol* 16(2):229–240.

23. Wei-Yun W, Kochan TJ, Storck DN, Dawid S (2016) Coordinated Bacteriocin Expression and Competence in *Streptococcus pneumoniae* Contributes to Genetic Adaptation through Neighbor Predation. *PLoS Pathog*. doi:DOI: 10.1371/journal.ppat.1005413.

24. Iannelli F, Oggioni MR, Pozzi G (2005) Sensor domain of histidine kinase ComD confers competence pherotype specificity in *Streptoccoccus pneumoniae*. *FEMS Microbiol Lett* 252(2):321–326.

25. Miller EL, Evans BA, Cornejo OE, Roberts IS, Rozen D (2016) Pherotype polymorphism in *Streptococcus pneumoniae* and its effects on population structure and recombination. *bioRxiv*. Available at: http://biorxiv.org/content/early/2016/08/17/070011.abstract.

26. Pinchas MD, LaCross NC, Dawid S (2015) An electrostatic interaction between BlpC and BlpH dictates pheromone specificity in the control of bacteriocin production and immunity in *Streptococcus pneumoniae*. *J Bacteriol* 197(7):1236–1248.

27. Ruparell A, et al. (2016) The fitness burden imposed by synthesizing quorum sensing signals. *Sci Rep*

546 6(August):33101.

547 28. Reichmann P, Hakenbeck R (2000) Allelic variation in a peptide-inducible two-component system of
548 *Streptococcus pneumoniae*. *FEMS Microbiol Lett* 190(2):231–236.

549 29. West SA, Winzer K, Gardner A, Diggle SP (2012) Quorum sensing and the confusion about diffusion. *Trends*
550 *Microbiol* 20(12):586–594.

551 30. Ji G, Beavis R, Novick RP (1997) Bacterial interference caused by autoinducing peptide variants. *Science*
552 276(5321):2027–2030.

553 31. Stefanic P, et al. (2012) The quorum sensing diversity within and between ecotypes of *Bacillus subtilis*. *Environ*
554 *Microbiol* 14(6):1378–1389.

555 32. Chewapreecha C, et al. (2014) Dense genomic sampling identifies highways of pneumococcal recombination.
556 *Nat Genet* 46(3):305–9.

557 33. Croucher NJ, et al. (2013) Population genomics of post-vaccine changes in pneumococcal epidemiology. *Nat*
558 *Genet* 45(6):656–63.

559 34. Chancey ST, et al. (2015) Composite mobile genetic elements disseminating macrolide resistance in
560 *Streptococcus pneumoniae*. *Front Microbiol* 6(February):1–14.

561 35. Bogaert D, et al. (2001) Pneumococcal carriage in children in the Netherlands: A molecular epidemiological
562 study. *J Clin Microbiol* 39(9):3316–3320.

563 36. McGee L, et al. (2001) Nomenclature of major antimicrobial-resistant clones of *Streptococcus pneumoniae*
564 defined by the pneumococcal molecular epidemiology network. *J Clin Microbiol* 39(7):2565–2571.

565 37. Croucher NJ, et al. (2011) Rapid pneumococcal evolution in response to clinical interventions. *Science*
566 331(6016):430–434.

567 38. Croucher NJ, et al. (2013) Dominant role of nucleotide substitution in the diversification of serotype 3
568 pneumococci over decades and during a single infection. *PLoS Genet* 9(10). doi:10.1371/journal.pgen.1003868.

569 39. Barker D, Meade A, Pagel M (2007) Constrained models of evolution lead to improved prediction of functional
570 linkage from correlated gain and loss of genes. *Bioinformatics* 23(1):14-20.

571 40. Kearse M, et al. (2012) Geneious Basic: An integrated and extendable desktop software platform for the
572 organization and analysis of sequence data. *Bioinformatics* 28(12):1647–1649.

41. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17(8):754–755.

42. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(12):1572–1574.

43. Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9(8):772–772.

44. Croucher NJ, et al. (2015) Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 43(3):e15–e15.

45. Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25(11):1451–1452.

46. Son MR, et al. (2011) Conserved mutations in the pneumococcal bacteriocin transporter gene, *blpA*, result in a complex population consisting of producers and cheaters. *MBio* 2(5). doi:10.1128/mBio.00179-11.

47. Sawyer SA (1999) GENECONV: A computer package for the statistical detection of gene conversion.

48. Tsirigos KD, Peters C, Shu N, Kall L, Elofsson a. (2015) The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res* 43(May):1–7.

49. Martin B, Garcia P, Castanié M-P, Claverys J-P (1995) The *recA* gene of *Streptococcus pneumoniae* is part of a competence-induced operon and controls lysogenic induction. *Mol Microbiol* 15(2):367–379.

50. Slager J, Kjos M, Attaiech L, Veening J-W (2014) Antibiotic-induced replication stress triggers bacterial competence by increasing gene dosage near the origin. *Cell* 157(2):395–406.

Figure Legends

Figure 1. QS eavesdropping, crosstalk, and regulation. A) Eavesdropping occurs when a QS receptor of a cell is activated by a QS signal that the cell does not produce, such as activation of the blue QS receptor by both the cognate blue square signal and non-cognate green triangle signal. Crosstalk occurs when a QS signal activates more than one receptor, such as the green triangle signal activating both the cognate green QS receptor and the non-cognate blue QS receptor. B) *blp* QS regulation. External BlpC signal binds to histidine kinase receptor BlpH. This activates response regulator BlpR through phosphorylation, which increases transcription of *blpABC*, *blpT*, and the *blp* bacteriocins (including *blpK*) and immunity genes. Pre-BlpC is processed and transported out of the cell by ABC transporters ComAB and BlpAB.

Figure 2. Bayesian unrooted phylogenetic tree of *blpC*. Taxa are colored by mature BlpC signal with the signal designation followed by the number of genomes containing the allele. Internal nodes show the posterior probabilities of clades; we collapsed clades with less than 0.75 posterior probability.

Figure 3. Bayesian unrooted phylogenetic tree of *blpH* alleles. The outer ring shows the number of 4,096 genomes with each *blpH* allele, color-coded by their co-occurring BlpC signal and on a log scale. The inner ring denotes the *blpH* clade type, and recombination events within *blpH* are shown as solid green lines. Mismatches between *blpH* clade and BlpC signal are indicated by dashed lines. Internal nodes show the posterior probabilities of clades; we collapsed clades with less than 0.75 posterior probability.

Figure 4. A) Proportion of each BlpC signal within genomes containing each *blpH* clade. The phylograms are simplified versions of Fig. 1 and Fig. 2. B) The relative maximal expression levels of *luc* following addition of 1 µg/ml of synthesized BlpC signal peptide. The maximum expression level for each reporter strain was set to 1. Raw data is found in Fig. S6. C) The minimum concentration of synthesized BlpC signal peptide required for *luc* induction in reporter strains with different BlpH. Example of raw data is provided in Fig. S6. The Bravo, Kilo, and Hotel signal peptides were not synthesized and are denoted with slashes.

Figure 5. LacZ induction by neighboring colonies on agar plates. A) The wild-type strains were spotted next to the reporter strains (see box), and induction of *blp* expression by the wild-type produced BlpC is shown as faint blue colonies. The experiment was repeated three times with the same result, and a representative photo of the plates is shown. B) Summary of the results from B. Squares in white indicate no induction of the reporter strain for colony pairs, while black and blue indicate induction by complementary and on-complementary BlpCs, respectively.

Figure 6. Average fitness of eavesdropping genotypes that produce bacteriocins in response to multiple signals in a spatially explicit, stochastic model. Simulations were started with five proportions of eavesdropping genotypes mixed with signal-faithful genotypes, as indicated on the x-axis. Absolute fitness values on the y-axis above 1.0 indicate that the genotype can increase in frequency in the population. Affinity to other genotypes' signals are a percentage of affinity to a genotype's own signal for eavesdropping genotypes. Error bars link the 25% and 75% quantiles for the final eavesdropping genotypes' fitness across 100 simulations.

Table 1. Predicted *blpC* mature peptide signals and their frequencies in each data set

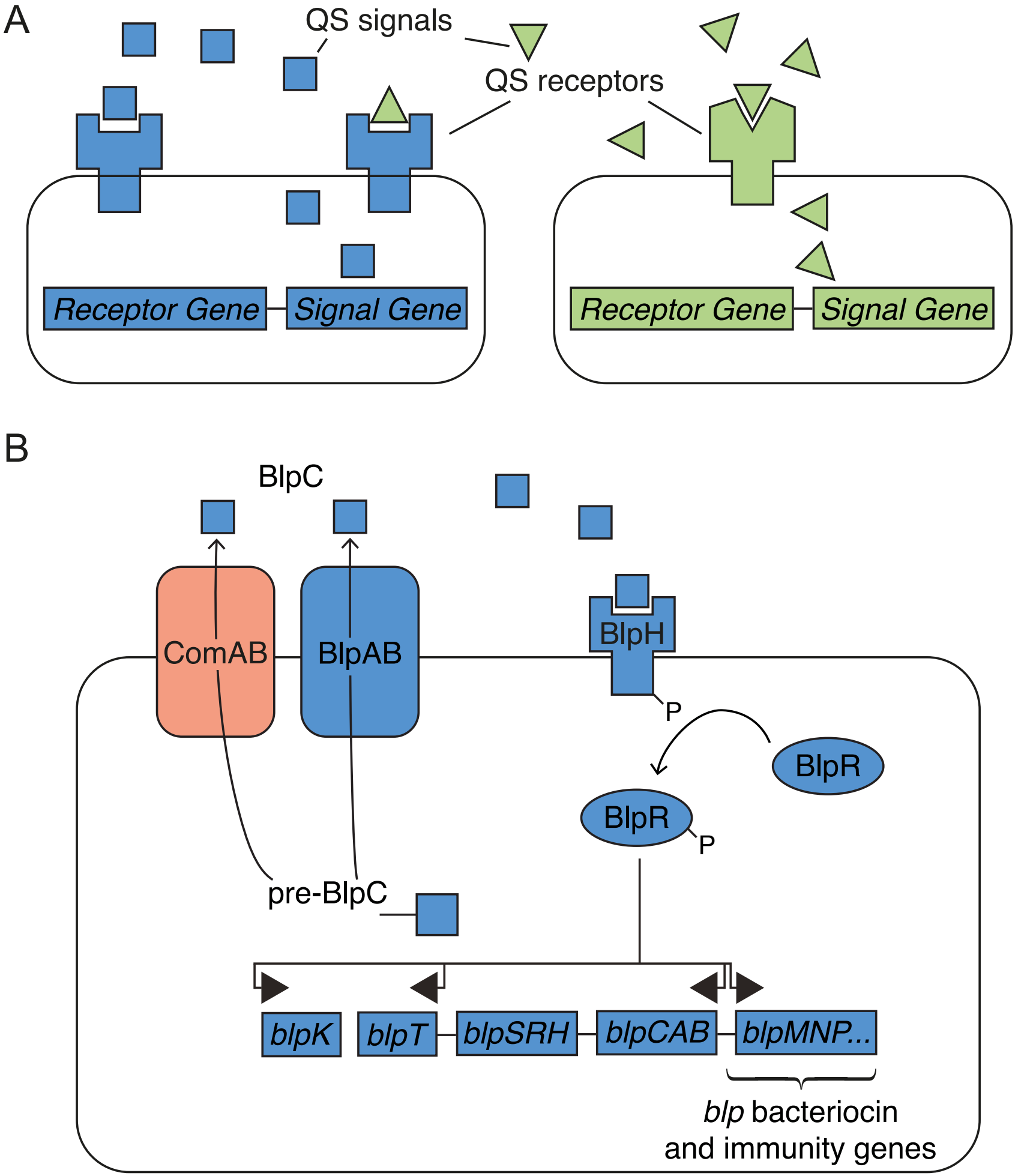
				Frequencies								
				Unbiased Sets								
Signal	Mature AA Sequence[a]	Genomes Used with Mature Signal	Reference Genomes with Mature Signal	All Unbiased Sets	Maela	Mass. Asymptomatic	GenBank	Hermans	Georgia GenBank	PMEN	Complex 3	PMEN-1
Charlie	**G**L**WE**DI**L**YSLNIIKHNN**TK**GLHH**PIQL**	Hermans-1012, PMEN-14	SPNA45, INV200	0.291	0.302	0.235	0.282	0.289	0.314	0.308	0.024	0.004
Golf	**G**L**WE**DL**L**YNINRYAHYI**T**	PMEN-18	TIGR4, INV104	0.223	0.185	0.42	0.19	0.197	0.256	0.154		0.938
Alpha	**G**W**WE**EL**L**HETILSKFKI**TK**ALEL**PIQL**	D39	D39, R6	0.125	0.1	0.175	0.184	0.275	0.215	0.308		
Foxtrot	**G**W**WE**DF**L**YRFNIIEQKN**TK**GFYQ**PIQL**	PMEN-2	P164	0.125	0.13	0.091	0.195	0.092	0.099	0.192		
Echo	**G**W**WE**DF**L**YRFNIIEQKN**TK**GFHQ**PIQL**			0.087	0.099	0.039	0.069	0.092	0.041	0.038		
Hotel	**G**L**WE**DL**L**YNINRYAHYI**T**QELHH**PIQL**			0.06	0.081							
Delta	**G**W**W**KDL**L**HRFNVIEQNN**TK**GFNQ**PIQL**	Hermans-33	ECC_3510	0.035	0.042	0.016	0.011	0.028				
Bravo	**G**L**WE**EL**L**HETILSKFKI**TK**ALEL**PIQL**			0.028	0.038							
Kilo	EW**WE**EL**L**HETILSKFKI**TK**ALEL**PIQL**		OXC141	0.007	0.002	0.018	0.052	0.007	0.017		0.976	
Other, or no 'GG' cleaveage site				0.009	0.012	0.002		0.021	0.05			0.058
No *blpC*				0.011	0.01	0.003	0.017		0.008			
Total Strains				4096	3017	616	174	142	121	26	82	240

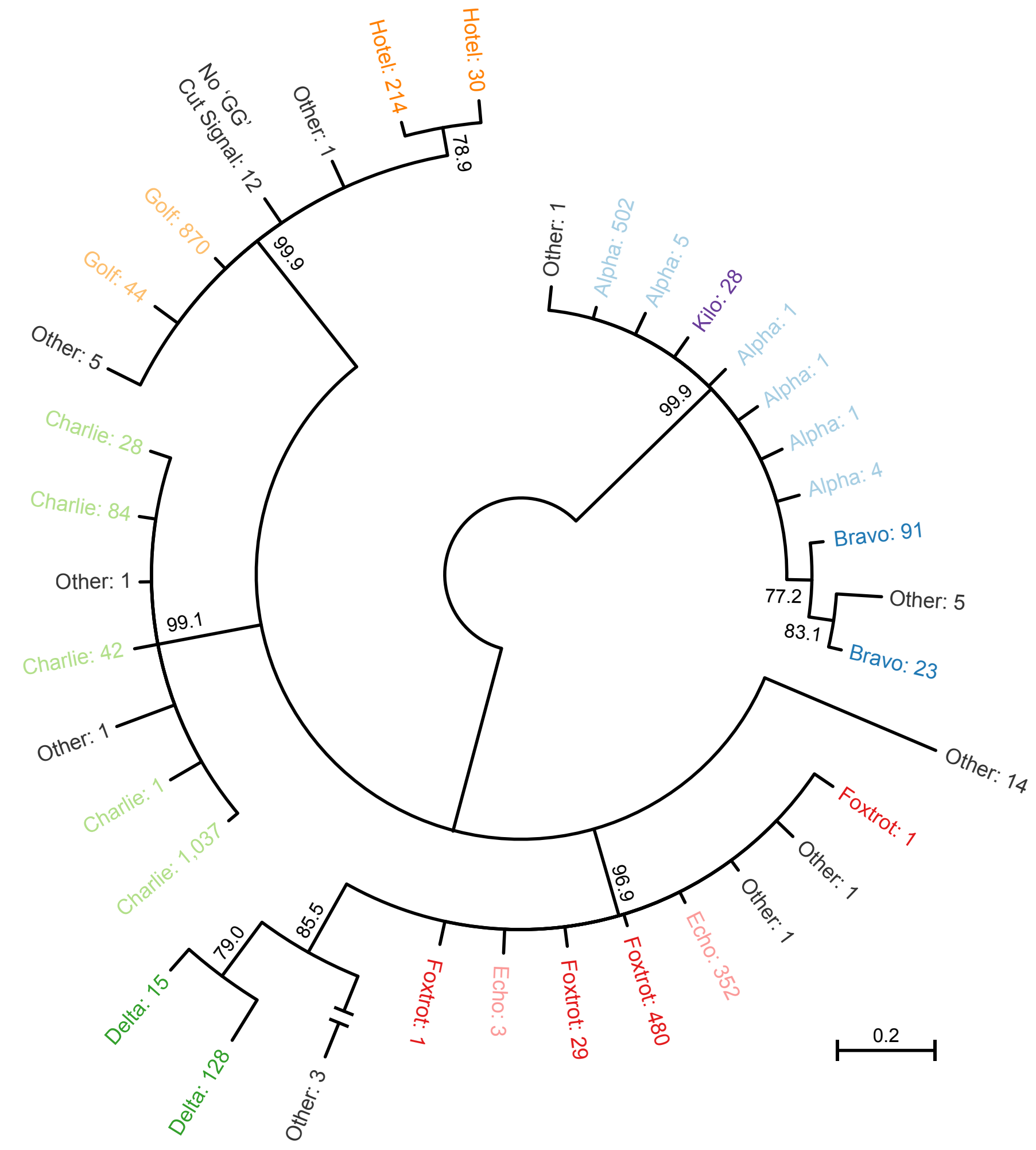
[a] Conserved residues are in bold

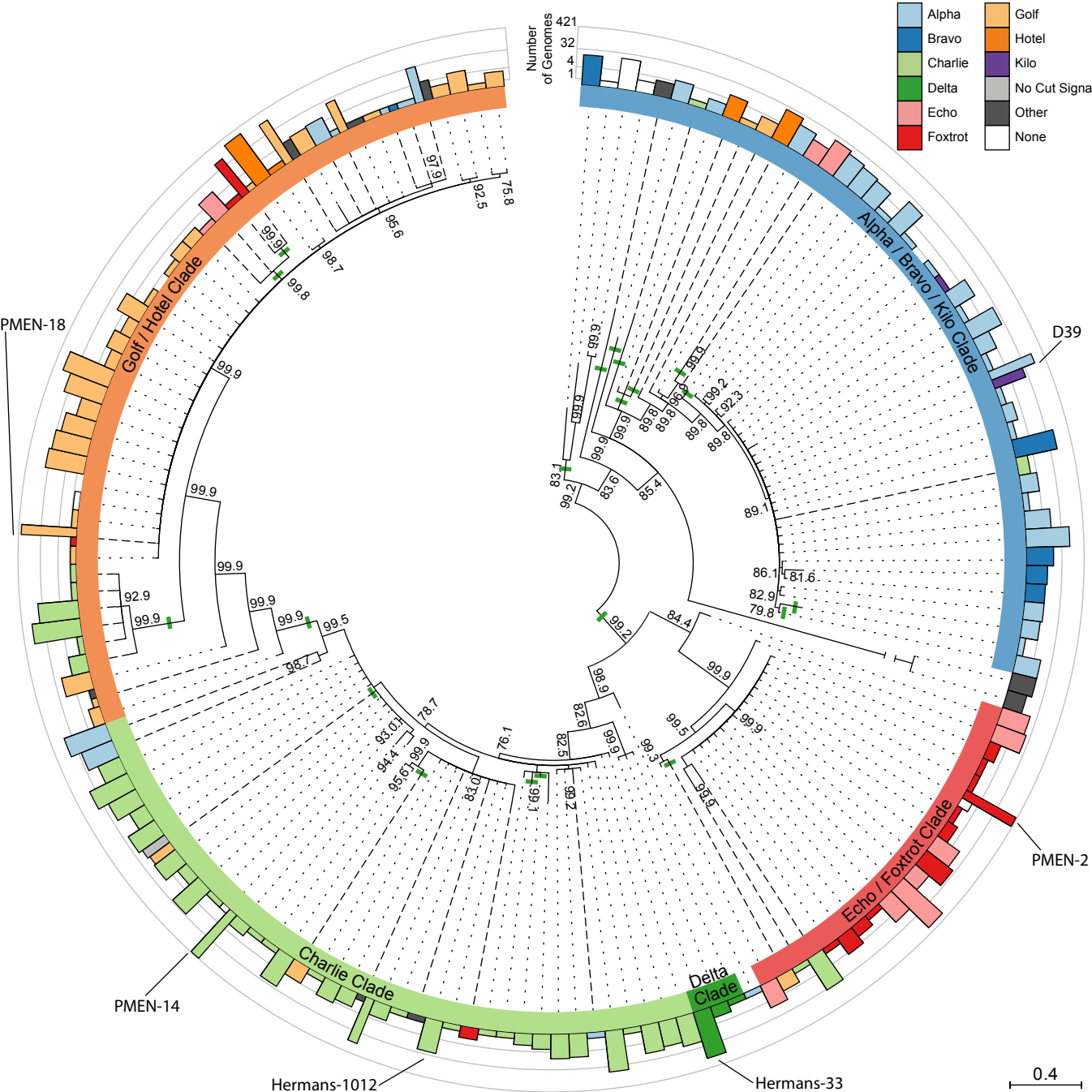
636 **Table 2.** Nucleotide diversity and d_N / d_S for selected genes

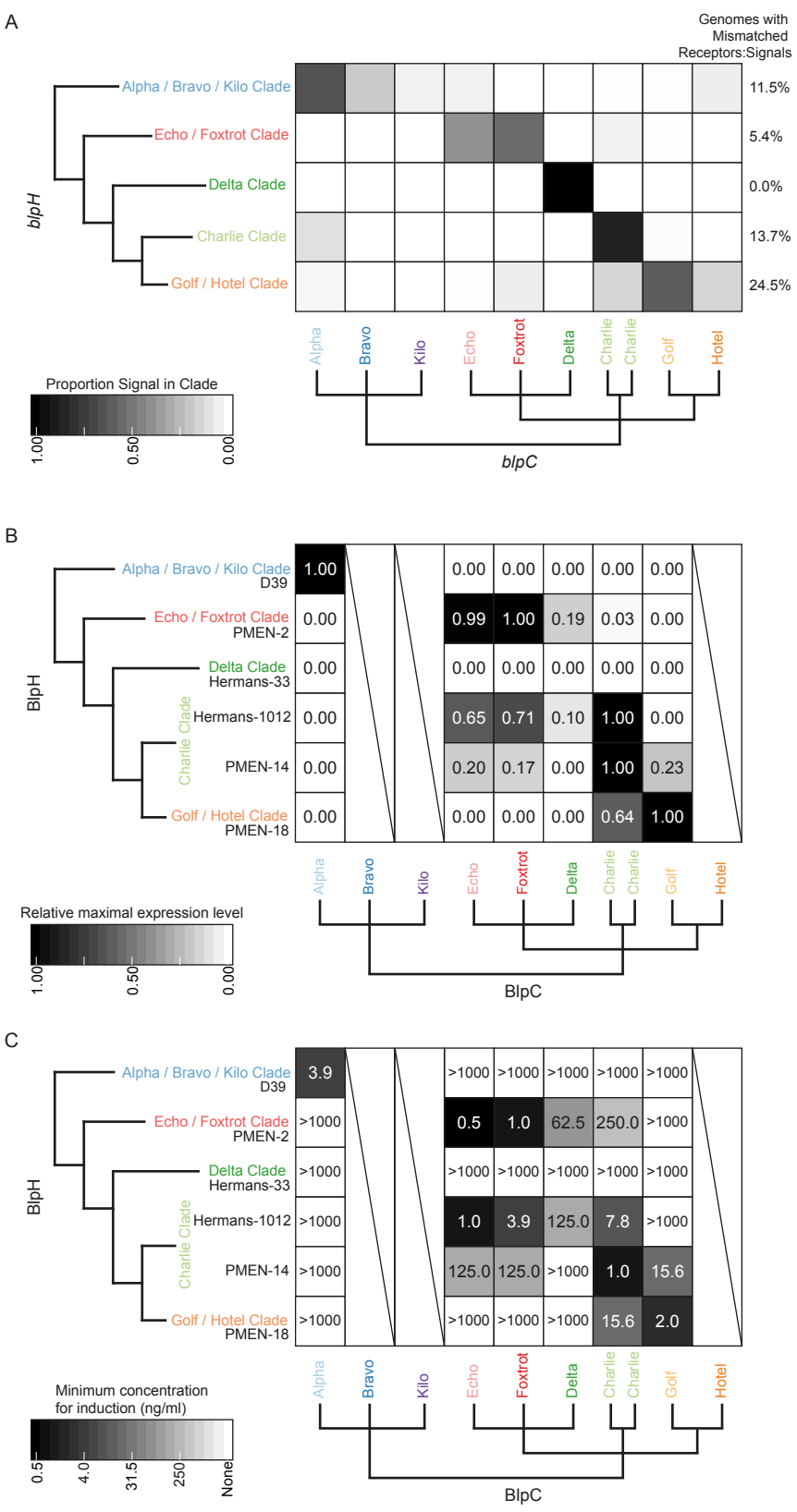
Gene	Function	Nucleotide Diversity (π)	d_N / d_S	Number of Sites	Number of Unique Sequences
blpC	QS Signal	0.124	0.965	126	34
blpH (receptor)	Histidine Kinase	0.110	0.377	687	98
blpH (entire)	Histidine Kinase	0.090	0.266	1353	161
blpH (kinase)	Histidine Kinase	0.069	0.146	666	103
ddlA	Housekeeping Gene	0.044	0.064	1041	170
lepB	Housekeeping Gene	0.027	0.026	612	90
aroE	Housekeeping Gene	0.022	0.146	852	123
glkA	Housekeeping Gene	0.022	0.061	975	143
xpt	Housekeeping Gene	0.020	0.083	579	111
gdhA	Housekeeping Gene	0.015	0.085	1485	185
recP	Housekeeping Gene	0.011	0.091	1974	231

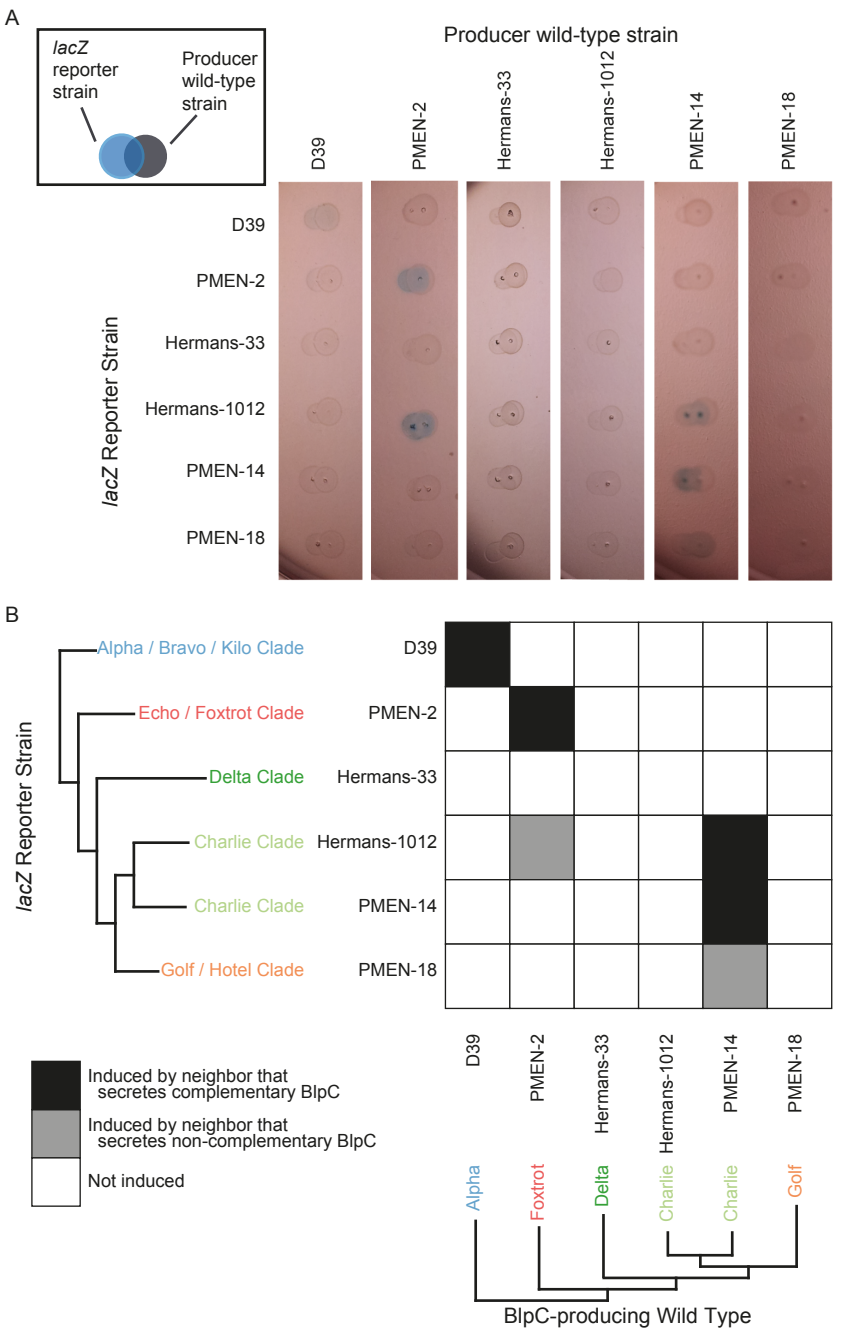
637

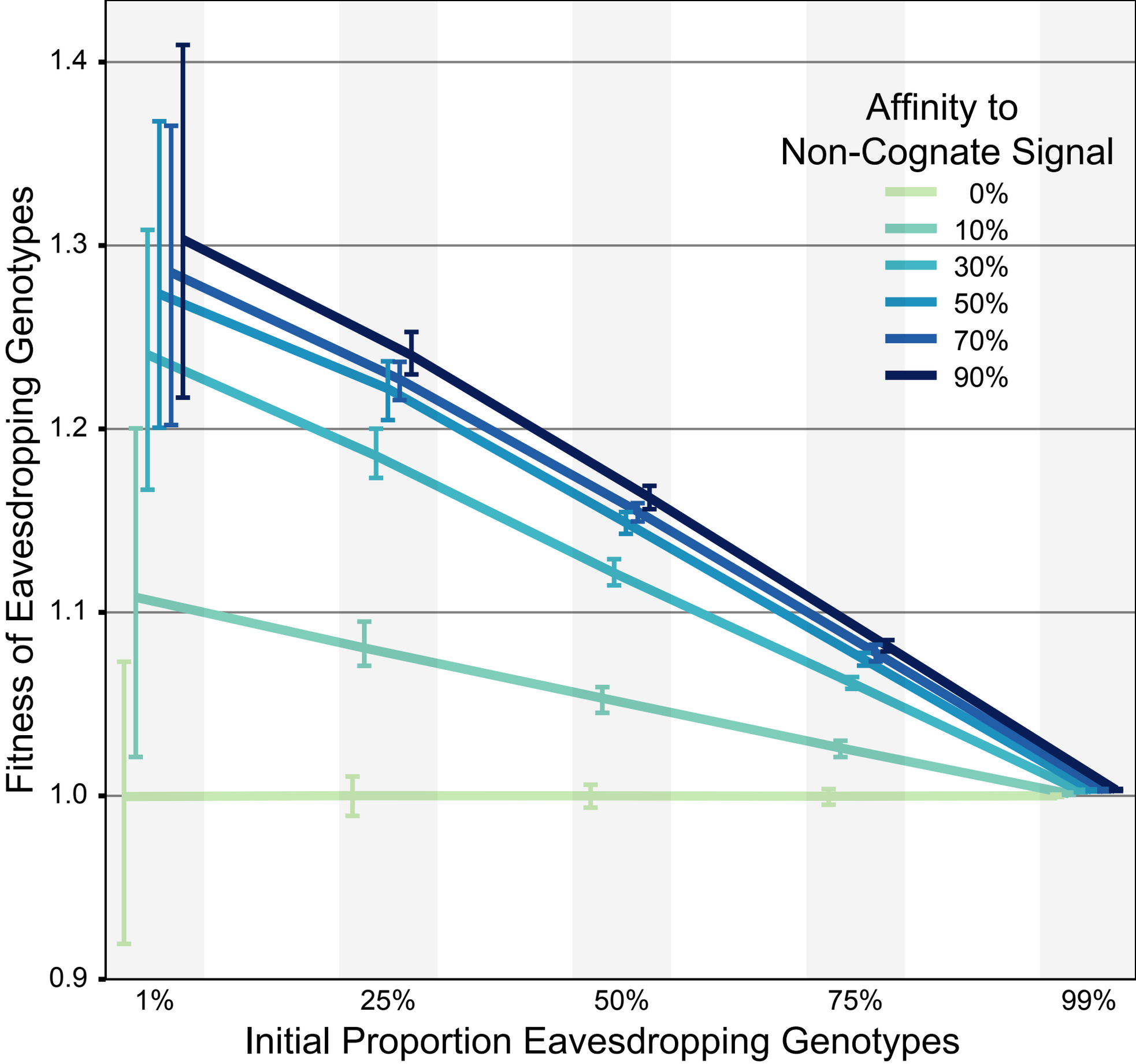

A
QS signals
QS receptors
Receptor Gene
Signal Gene
Receptor Gene
Signal Gene
B
BlpC
ComAB
BlpAB
BlpH
P
BlpR
BlpR
P
pre-BlpC
blpK
blpT
blpSRH
blpCAB
blpMNP...
blp bacteriocin
and immunity genes


Hotel: 30
Hotel: 214
Other: 1
No 'GG' Cut Signal: 12
Golf: 870
Golf: 44
Other: 5
Charlie: 28
Charlie: 84
Other: 1
Charlie: 42
Other: 1
Charlie: 1
Charlie: 1,037
Delta: 15
Delta: 128
Other: 3
Foxtrot: 1
Echo: 3
Foxtrot: 29
Foxtrot: 480
Echo: 352
Other: 1
Other: 1
Foxtrot: 1
Other: 14
Bravo: 23
Bravo: 91
Other: 5
Alpha: 4
Alpha: 1
Alpha: 1
Alpha: 1
Kilo: 28
Alpha: 5
Alpha: 502
Other: 1
78.9
99.9
99.9
99.1
77.2
83.1
96.9
85.5
79.0
0.2


Number of Genomes
421
32
4
1
Alpha
Bravo
Charlie
Delta
Echo
Foxtrot
Golf
Hotel
Kilo
No Cut Signa
Other
None
Alpha / Bravo / Kilo Clade
Golf / Hotel Clade
Echo / Foxtrot Clade
Charlie Clade
Delta Clade
PMEN-18
D39
PMEN-2
PMEN-14
Hermans-1012
Hermans-33
0.4


A
Genomes with Mismatched Receptors:Signals
Alpha / Bravo / Kilo Clade
Echo / Foxtrot Clade
Delta Clade
Charlie Clade
Golf / Hotel Clade
blpH
11.5%
5.4%
0.0%
13.7%
24.5%
Alpha
Bravo
Kilo
Echo
Foxtrot
Delta
Charlie
Charlie
Golf
Hotel
blpC
Proportion Signal in Clade
1.00
0.50
0.00
B
Alpha / Bravo / Kilo Clade
D39
Echo / Foxtrot Clade
PMEN-2
Delta Clade
Hermans-33
Charlie Clade
Hermans-1012
PMEN-14
Golf / Hotel Clade
PMEN-18
BlpH
1.00 0.00 0.00 0.00 0.00 0.00
0.00 0.99 1.00 0.19 0.03 0.00
0.00 0.00 0.00 0.00 0.00 0.00
0.00 0.65 0.71 0.10 1.00 0.00
0.00 0.20 0.17 0.00 1.00 0.23
0.00 0.00 0.00 0.00 0.64 1.00
Alpha
Bravo
Kilo
Echo
Foxtrot
Delta
Charlie
Charlie
Golf
Hotel
BlpC
Relative maximal expression level
1.00
0.50
0.00
C
Alpha / Bravo / Kilo Clade
D39
Echo / Foxtrot Clade
PMEN-2
Delta Clade
Hermans-33
Charlie Clade
Hermans-1012
PMEN-14
Golf / Hotel Clade
PMEN-18
BlpH
3.9 >1000 >1000 >1000 >1000 >1000
>1000 0.5 1.0 62.5 250.0 >1000
>1000 >1000 >1000 >1000 >1000 >1000
>1000 1.0 3.9 125.0 7.8 >1000
>1000 125.0 125.0 >1000 1.0 15.6
>1000 >1000 >1000 >1000 15.6 2.0
Alpha
Bravo
Kilo
Echo
Foxtrot
Delta
Charlie
Charlie
Golf
Hotel
BlpC
Minimum concentration for induction (ng/ml)
0.5
4.0
31.5
250
None


A
lacZ reporter strain
Producer wild-type strain
Producer wild-type strain
D39
PMEN-2
Hermans-33
Hermans-1012
PMEN-14
PMEN-18
lacZ Reporter Strain
D39
PMEN-2
Hermans-33
Hermans-1012
PMEN-14
PMEN-18
B
lacZ Reporter Strain
Alpha / Bravo / Kilo Clade
Echo / Foxtrot Clade
Delta Clade
Charlie Clade
Charlie Clade
Golf / Hotel Clade
D39
PMEN-2
Hermans-33
Hermans-1012
PMEN-14
PMEN-18
Induced by neighbor that secretes complementary BlpC
Induced by neighbor that secretes non-complementary BlpC
Not induced
D39
PMEN-2
Hermans-33
Hermans-1012
PMEN-14
PMEN-18
Alpha
Foxtrot
Delta
Charlie
Charlie
Golf
BlpC-producing Wild Type


Affinity to
Non-Cognate Signal
0%
10%
30%
50%
70%
90%
Fitness of Eavesdropping Genotypes
1.4
1.3
1.2
1.1
1.0
0.9
1%
25%
50%
75%
99%
Initial Proportion Eavesdropping Genotypes