Kin selection for cooperation in natural bacterial populations

Laurence J. Belcher[a,1], Anna E. Dewar[a], Melanie Ghoul[a,2], and Stuart A. West[a,2]

[a]Department of Zoology, University of Oxford, Oxford OX1 3SZ, United Kingdom

Bacteria produce a range of molecules that are secreted from the cell and can provide a benefit to the local population of cells. Laboratory experiments have suggested that these "public goods" molecules represent a form of cooperation, favored because they benefit closely related cells (kin selection). However, there is a relative lack of data demonstrating kin selection for cooperation in natural populations of bacteria. We used molecular population genetics to test for signatures of kin selection at the genomic level in natural populations of the opportunistic pathogen *Pseudomonas aeruginosa*. We found consistent evidence from multiple traits that genes controlling putatively cooperative traits have higher polymorphism and greater divergence and are more likely to harbor deleterious mutations relative to genes controlling putatively private traits, which are expressed at similar rates. These patterns suggest that cooperative traits are controlled by kin selection, and we estimate that the relatedness for social interactions in *P. aeruginosa* is $r = 0.84$. More generally, our results demonstrate how molecular population genetics can be used to study the evolution of cooperation in natural populations.

relatedness | public goods | population genetics | inclusive fitness

The growth and success of many bacteria appear to depend upon a stunning array of cooperative behaviors (1–3). Cells produce and secrete a range of factors that benefit the local group of cells and so, act as cooperative "public goods." Examples include molecules to scavenge iron (siderophores) (4), enzymes that break down proteins (proteases) (5), and molecules to aid cell movement (rhamnolipids) (6).

The potential problem with such cooperation is that it can be exploited by noncooperators ("cheats") that do not produce public goods but can still benefit from those produced by others (7). A likely solution to this problem in bacteria is that clonal growth keeps close relatives together, and limited diffusion keeps public goods close to producers (8). Consequently, the benefits of cooperation tend to be shared with related cells that share the gene for cooperation, and so, cooperation is favored by kin selection (9).

However, most evidence for cooperation and kin selection in bacteria has come from laboratory experiments (10–18). To what extent are test-tube cultures, often utilizing extreme gene knockouts, representative of natural populations (1, 12). A problem here is that while bacteria and other microorganisms offer many advantages for laboratory experiments, they can be very difficult to study in their natural environment.

Population genetics offers a way to study natural populations because kin selection can leave signatures ("footprints") of selection at the genomic level (10–12, 15, 19–28). In a clonal population, where the relatedness (*r*) between interacting cells is $r = 1$, the benefits of cooperating will always be passed onto other individuals that carry the gene for cooperation. In contrast, as relatedness decreases ($r < 1$), the benefits of cooperation will increasingly be passed onto individuals that do not carry the gene for cooperation (Fig. 1*A*). This reduces (dilutes) the kin-selected benefit of cooperation, making beneficial mutations less likely to fix and deleterious mutations more likely to fix (Fig. 1*B*) (9, 25).

Population genetic theory, therefore, predicts that, in nonclonal populations ($r < 1$), cooperative traits favored by kin selection will show increased polymorphism and divergence relative to traits that provide private benefits (Fig. 1 *C* and *D*) (15, 23, 25). Nonclonal populations appear to be very common in bacteria. At the scale of the social interaction, groups often contain multiple species, let alone multiple lineages of the same species (17, 29, 30). In addition, molecular and genomic studies have demonstrated selection for noncooperative cheats that exploit the cooperation of others as well as a diversity of mechanisms for attacking nonrelatives (14, 16, 31). Clonal interactions seem to be limited to extreme cases, such as cyanobacteria filaments (30).

We tested for genomic signatures of kin selection for cooperation in the opportunistic pathogen *Pseudomonas aeruginosa*. Laboratory experiments have suggested that *P. aeruginosa* produces a range of cooperative public goods that facilitate both growth and virulence (4, 32, 33). A potential problem with genomic analyses is that they can be confounded by conditional gene expression. If a gene is only occasionally expressed, in certain conditions, this can also lead to relaxed selection, making beneficial mutations less likely to fix and deleterious mutations more likely to fix (10, 22). We controlled for this influence of conditional gene expression by making targeted comparisons between cooperative and private traits that are likely to be expressed at similar rates.

Results and Discussion

We compared genetic variation in traits that are hypothesized to be cooperative with traits that are hypothesized to be private (Fig. 1). The predicted results from the population

Significance

Bacteria secrete many molecules outside the cell, where they provide benefits to other cells. One potential reason for producing these "public goods" is that they benefit closely related cells that share the gene for cooperation (kin selection). While many laboratory studies have supported this hypothesis, there is a lack of evidence that kin selection favors cooperation in natural populations. We examined bacterial genomes from the environment and used population genetics theory to analyze the DNA sequences. Our analyses suggest that public goods cooperation has indeed been favored by kin selection in natural populations.

[1]To whom correspondence may be addressed. Email: laurence.belcher@zoo.ox.ac.uk.

[2]M.G. and S.A.W. contributed equally to this work.

EVOLUTION
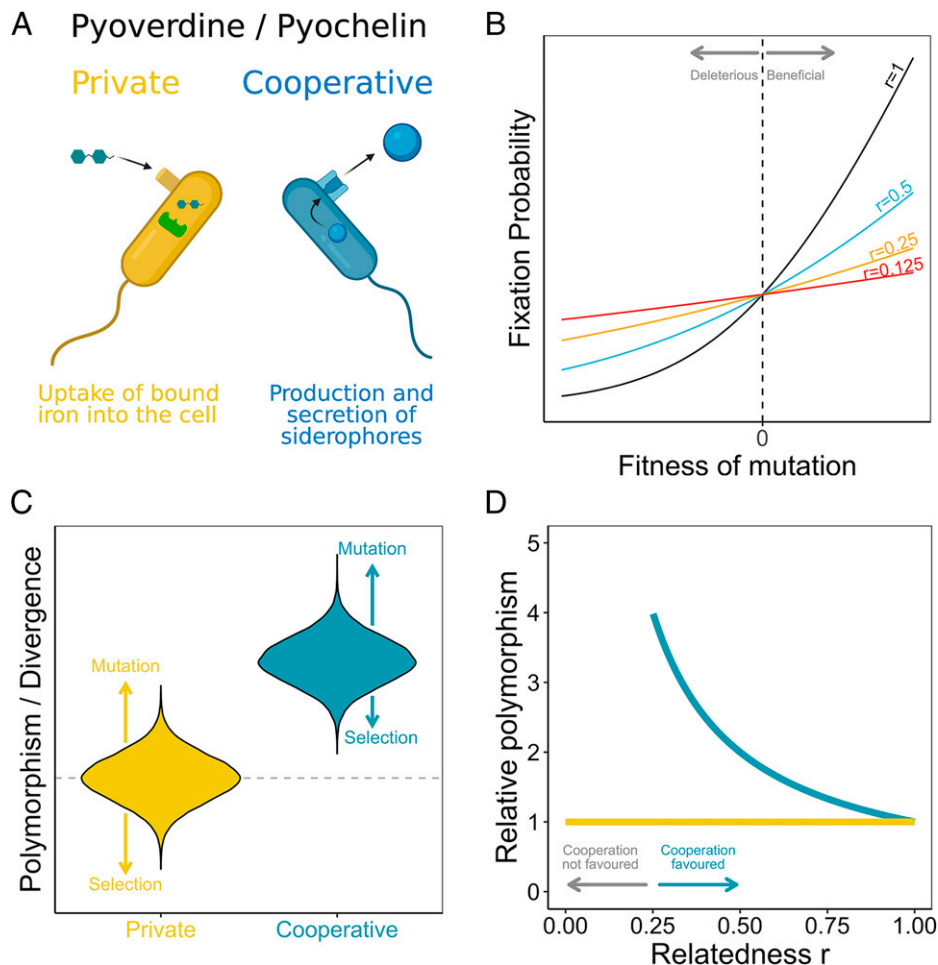
Fig. 1. Population genetic theory for cooperative traits (15, 25). (*A*) Representation of how traits are categorized as private or cooperative. Cooperative traits are those involving the production and secretion of molecules where the fitness benefits can potentially be shared with other cells in the local group. Private traits are those where the fitness benefits are only felt by the individual expressing the gene (e.g., internal metabolism). (*B*) Probability of fixation for deleterious or beneficial mutations of varying effect (*x* axis) for mutations influencing private (black line) and cooperative (social; all lines) traits. In clonal populations, where the relatedness (*r*) between interacting individuals $r = 1$, the prediction is the same for mutations influencing private and cooperative traits (black line). As relatedness decreases, the prediction changes for mutations influencing cooperation, with beneficial mutations becoming less likely to fix and deleterious mutations becoming more likely to fix. Consequently, in nonclonal populations, there is relaxed selection on genes controlling cooperative traits relative to those controlling private traits. Adapted from ref. 15. (*C*) Prediction for relative polymorphism and divergence for cooperative (blue) relative to private (yellow) genes assuming a fixed $r < 1$. Due to the increased fixation likelihood of deleterious mutations and decreased fixation likelihood of beneficial mutation, genes for cooperative traits should have relatively greater levels of polymorphism and divergence. (*D*) Predicted polymorphism of private (yellow) and cooperative (blue) genes as relatedness varies for a trait where cooperation is favored when $r > 0.25$. For private traits, polymorphism is independent of relatedness. For cooperative traits, expected polymorphism relative to a private trait is inversely proportional to *r* when cooperation is favored. When $r = 1$, there is no difference in polymorphism between cooperative and private traits. When $r < 0.25$, cooperation is not favored, so relatedness no longer predicts the level of polymorphism observed.

genetic analysis for kin selection and other competing hypotheses are shown in Table 1. As no single measure can separate the different possible forms of selection, it is important to consider all of these measures together. We examined 41 genomes of *P. aeruginosa* environmental isolates, focusing our analyses on six groups of traits where the cooperative and private traits were likely to be expressed at relatively similar rates (*SI Appendix*, Table 6).

Table 1. Predicted results from population genetics analysis for four different forms of selection

Selection type	Divergence	Polymorphism	Deleterious mutations	Tajima's D	McDonald–Krietman
Positive/directional	High	Low	—	<<0	P < 0.05
Kin	High	High	High	~0	n.s.
Balancing	Low	High	—	>>0	P < 0.05
Purifying	Low	Low	—	~0	n.s.

Levels of divergence, polymorphism, and frequency of deleterious mutations are shown as values for cooperative genes relative to private genes. Tajima's *D* uses information about the frequency of polymorphism, and predictions are shown as the absolute value, with extreme values indicative of positive or balancing selection. The McDonald–Krietman test compares levels of polymorphism and divergence, and a significant result is indicative of either positive or balancing selection. From refs. 15 & 25. n.s., not significant.

Quorum Sensing. We started by examining genes induced by the quorum sensing (QS) signaling system (34, 35). This system regulates gene expression in response to the density of a diffusible signal molecule produced by cells. As cell density increases, the concentration of the signal molecule also increases, leading to the up-regulation of many genes. In *P. aeruginosa*, the QS network regulates several hundred genes, which comprise ~6% of the genome (36).

There are four advantages to examining the QS system. First, it regulates a number of traits that are hypothesized to be cooperative as well as a number of traits that have only private benefits (Fig. 1*A*) (37, 38): for example, the secretion of enzymes to digest proteins outside the cell (cooperative) vs. the production of enzymes to metabolize molecules within the cell (private). Second, control by the shared QS network means that the genes coding for these different traits are likely to be expressed at relatively similar rates on average (34, 35). This allows us to control for the potentially confounding influence that expression rates may have on patterns of genetic variation (22). Third, coregulation of genes acts as a control for mutations in noncoding regulatory and promoter regions that could affect the production of public goods. Fourth, the large size of the network means that there are sufficient genes for a meaningful comparison (*Methods*).

We used a combination of gene annotations and experimental data to assign genes as controlling either cooperative or private traits (*Methods*). For example, we categorized the extracellular elastase Las*B* as cooperative because it has been shown to be an exploitable public good in laboratory experiments (39). We also included several other extracellular proteases controlled by QS signaling, such as Protease IV (PIV) and PepB, which can provide benefits to the local group of cells and are known virulence factors (40, 41). Private traits include genes encoding proteins such as Nuh, an intracellular enzyme that allows cells to metabolize adenosine within the cell (5). The set of cooperative genes and their function are given in *SI Appendix*, Table 1. Our set of genes contains some that respond specifically to only one of the two major QS signals, so we checked the robustness of our results by restricting the analysis to only genes that respond to both QS signals in *SI Appendix*, section S9.
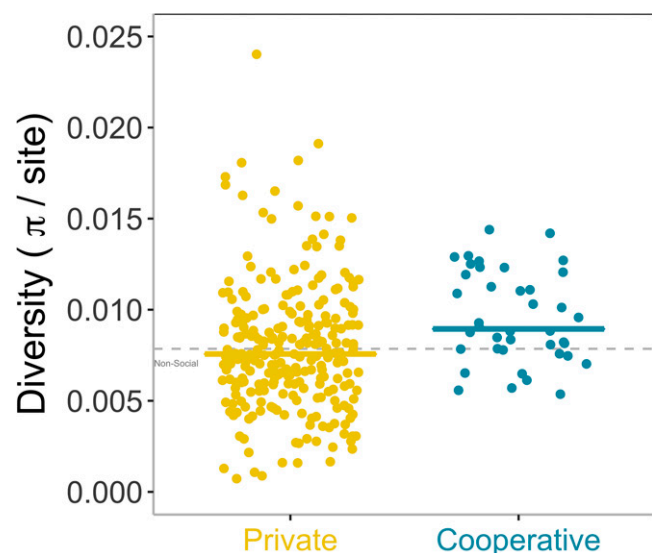
Fig. 2. Nucleotide diversity per site for private QS (yellow) and cooperative QS (blue) genes. Each dot represents a gene, and the horizontal lines represent the median for each group. The gray dotted line represents the median for private genes across the genome. Genes for cooperative traits showed significantly higher polymorphism than genes for private traits.
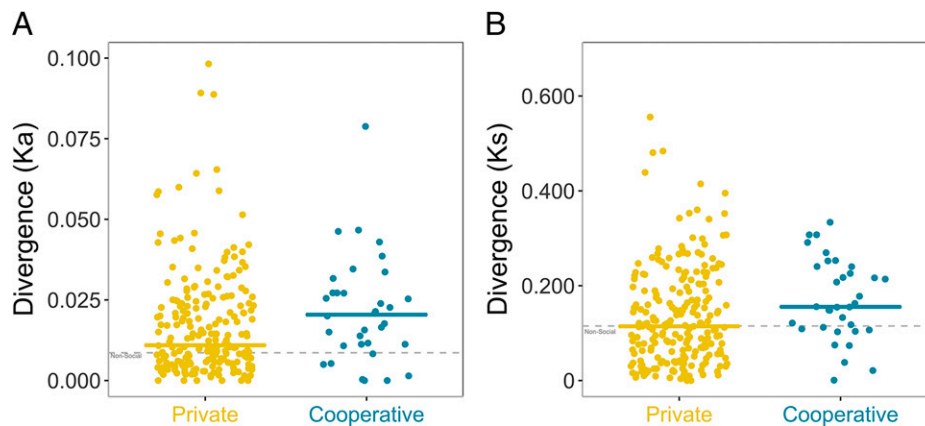
QS: Polymorphism. We found that genes regulating cooperative traits had significantly higher levels of polymorphism than genes regulating private traits (Fig. 2) (ANOVA $F_{1,2351} = 12.0$, $P < 0.01$; Tukey's honest significance test [HSD] $P = 0.009$). This difference was also significant when examining synonymous and nonsynonymous sites separately (synonymous: ANOVA $F_{1,2350} = 30.0$, $P < 10^{-7}$; Tukey's HSD $P = 0.004$; nonsynonymous: Kruskal–Wallis $X^2(2) = 22.7$, $P < 10^{-4}$; Dunn test $P = 0.04$) (*SI Appendix*, Fig. 1). In all cases, the average pairwise nucleotide diversity per site (π) was significantly higher in cooperative genes compared with private genes. We discuss possible reasons for increased polymorphism being manifest at synonymous sites as well as nonsynonymous sites in the following section.

We also found the same pattern of elevated polymorphism in cooperative genes when comparing with a background set of 2,459 private genes not involved in the QS system (*SI Appendix*, section S5). This background set was made up of genes whose proteins localize to the cytoplasm since these are the class of gene least likely to have a cooperative function. However, some cytoplasmic genes will be critical to the process of producing and secreting public goods, particularly in complex public goods such as pyoverdine that require several biosynthesis steps (42). Examining QS-controlled genes, the ratio of nonsynonymous to synonymous polymorphism did not differ significantly between genes controlling cooperative vs. private traits (ANOVA $F_{1,2338} = 32.4$, $P < 10^{-7}$; Tukey's HSD $P = 0.963$). However, QS-regulated private genes had a significantly higher ratio than the background set of private genes (Tukey's HSD $P < 0.03$) (*SI Appendix*, Fig. 2). This result reflects the finding that polymorphism is increased at both nonsynonymous and synonymous sites in cooperative compared with private genes and that QS-regulated genes may be under overall stronger selection than the background set of private genes. This could be because QS-regulated genes include many virulence factors and genes with large fitness effects, such as those involved in biofilms, social motility, and obtaining nutrients (38).

QS: Divergence. We found that genes regulating cooperative traits had significantly higher divergence than genes regulating private traits (Fig. 3). We measured divergence as the rate of protein evolution, quantified as the number of substitutions per site when comparing the reference genome PAO1 with the known taxonomic outlier PA7 (43). The difference was significant when examining both nonsynonymous (Fig. 3*A*) (Kruskal–Wallis $X^2(2) = 25.5$, $P < 10^{-5}$; Dunn test $P = 0.045$) and synonymous sites (Fig. 3*B*) (ANOVA $F_{1,2118} = 0.08$, $P = 0.771$; Tukey's HSD $P = 0.03$).

Divergence was significantly elevated at both nonsynonymous and synonymous sites in cooperative genes, and the ratio of nonsynonymous to synonymous divergence does not differ between the two classes of gene (Kruskal–Wallis $X^2(2) = 37.8$, $P < 10^{-8}$; Dunn test $P = 0.40$). However, both cooperative and private QS genes have a significantly higher ratio than the background private genes (Tukey's HSD cooperative $P < 10^{-2}$, private $P < 10^{-6}$), and cooperative genes have a slightly higher median ratio than private genes (*SI Appendix*, Fig. 3). We found that the baseline levels of polymorphism and divergence that we observed were consistent with previous studies (*SI Appendix*, section S8).

Our finding that cooperative genes have significantly elevated polymorphism at both synonymous and nonsynonymous sites suggests that mutations at synonymous sites are under selection and not evolving neutrally. In microbes, there is substantial evidence that synonymous mutations have fitness effects (44), such as increasing antibiotic resistance (45) and generating public goods cheats in viruses (46). Synonymous mutations in pyoverdine biosynthesis genes repeatedly occur in

Belcher et al.
Kin selection for cooperation in natural bacterial populations

PNAS | 3 of 10
https://doi.org/10.1073/pnas.2119070119

Fig. 3. Divergence at nonsynonymous (*A*) and synonymous (*B*) sites measured as rates of protein evolution (e.g., nonsynonymous substitutions per nonsynonymous site) for private QS (yellow) and cooperative QS (blue) genes. Each point represents a gene, and the horizontal lines represent the median for each group. The gray dotted lines represent the median for private genes across the genome. Genes for cooperative traits showed significantly higher divergence than genes for private traits.

experimental evolution of *P. aeruginosa* biofilms (47), and synonymous mutations in QS genes of *Vibrio campbellii* are associated with intermediate QS phenotypes (48). Similar patterns of elevated polymorphism at both nonsynonymous and synonymous sites were also found in the social microbe *Dictyostelium discoideum* (10). We did not find evidence for systematic differences in codon usage that could explain the synonymous variation that we see (*SI Appendix*, section S1).

QS: Deleterious Mutations. Population genetic theory also predicts that deleterious mutations are more likely be observed in genes controlling cooperative traits that are maintained by kin selection (10, 25). This prediction is a result of relaxed selection making deleterious mutations less likely to be removed by selection. We tested this prediction by looking at the overrepresentation of a subset of loss-of-function mutations that are easily identifiable. Specifically, we looked for 1) mutations that generate stop codons and 2) frameshift mutations. Our previous designation of cooperative genes was based on searching the literature for QS-regulated genes that have been demonstrated to be cooperative in the laboratory. Because of this, we do not know how many other "cooperative" genes there are that were not included in our previous dataset. Therefore, to test whether genes with deleterious mutations were more likely to be cooperative, we needed to use a proxy of cooperative genes that examined all genes in the genome. We used the production of extracellular proteins as a proxy for cooperation, as has been done previously (49, 50), since this can be systematically calculated for the whole genome using the protein subcellular localization prediction tool PSORTb (51).

We found that deleterious mutations were more common in genes controlling the production of extracellular proteins, which were, therefore, more likely to be cooperative. Of the 359 genes that have known protein localization and at least one deleterious mutation, 12 code for extracellular proteins (3.3%). Genes coding for extracellular proteins make up 1.6% of all genes with known protein localization but 3.3% of genes with deleterious mutations, which represents a significant overrepresentation of genes coding for extracellular proteins in genes containing deleterious mutations (binomial test, $P < 0.05$). Additionally, this increased to 4.4% (19 of 431 mutations) when we counted the total number of deleterious mutations in genes coding for extracellular proteins (rather than the number of genes with at least one mutation), suggesting that extracellular proteins are also likely to contain multiple mutations per gene. Interestingly, we observed a particularly high rate of

deleterious mutations in LasR, the master regulator of the QS system. While LasR is not an extracellular protein, LasR mutants are common in generating "cheaters" in clinical isolates (5, 52), and we show here that they also appear to be common in environmental isolates.

QS: Robustness and Competing Hypotheses. Our conclusion that kin selection favors cooperation was further supported by five further analyses that eliminated alternative explanations for the patterns that we observed. First, genes for cooperative traits could alternatively have significantly greater polymorphism than genes for private traits if they were more likely to be under balancing selection: for example, due to frequency-dependent selection between cooperators and cheats (11, 12, 25, 53, 54). However, we found no evidence that genes for cooperative traits are overrepresented in genes evolving under balancing selection and no evidence that balancing selection explained the elevated polymorphism we observed (*SI Appendix*, section S3).

Second, genes for cooperative traits could have significantly greater divergence than genes for private traits because they are more likely to be under positive selection and therefore, have fixed adaptive differences (24, 25). However, we found no evidence that genes for cooperative traits are overrepresented in genes evolving under positive selection and no evidence that positive selection explains the elevated divergence we observe (*SI Appendix*, section S4). The population genetic parameters that we analyzed are designed to test deviation from neutral expectations and therefore, have various underlying assumptions. Neutral theory (55) is based on the idea that polymorphisms are added by mutation, and their fate is largely determined by drift (56). This means that populations are at mutation–drift equilibrium, and we can make predictions about the level of polymorphism we expect in a population. We can then use tests like Tajima's *D* or the McDonald–Krietman test to test for deviations from the predictions of the standard neutral model. While we cannot completely rule out problems in interpreting these tests due to issues such as selection acting at different sites in subpopulations (*SI Appendix*, section S14), no alternative hypotheses can explain the patterns we see across multiple sets of isolates and across multiple traits.

Third, our findings could reflect some other shared aspect of cooperative genes rather than being cooperative per se. We performed a functional annotation of all the QS-controlled genes using the eggNOG database (57), which splits genes into functional categories, such as "metabolism," "cellular processes and signaling," and "information storage and processing." We

found that while genes for cooperative traits are overrepresented in genes annotated as metabolism and underrepresented in genes annotated as information storage and processing, there was no difference in polymorphism between these two functional categories (*SI Appendix*, Fig. 4). While we did find a difference for divergence (*SI Appendix*, Fig. 5), it is information storage and processing genes that have higher divergence. Overall, it appears that there is no other shared function of genes for cooperation that explains greater divergence and polymorphism.

Fourth, cooperative genes could appear more polymorphic and divergent than private genes because of differences in gene length. In human genomes at least, shorter genes tend to have higher expression (58) and greater divergence (59) than longer genes. If cooperative genes tend to be much shorter than private genes, this could bias our results, even though we control for gene length by using polymorphism measures calculated per site and control for variation in expression by analyzing QS-controlled genes, which should have similar average expression. However, cooperative genes did not differ in length compared with private genes (t test $t = 0.448$, $P = 0.657$). Further, when considering all genes, there is no significant correlation between gene length and polymorphism (Pearson's correlation $t = -0.650$, $P = 0.516$). We checked the robustness of this analysis by removing the bottom quartile of genes (<188 amino acids) from our analysis and found that this makes no difference to the qualitative results (*SI Appendix*, section S7).

Fifth, if cooperative and private genes differed in their likelihood of being transferred horizontally over their evolutionary history, that could affect comparisons due to the inherent problems that horizontal gene transfer raises in population genetics (60). We conducted an analysis using pangenome data (*SI Appendix*, section S11) showing that the cooperative genes we used either are part of the core genome or present in most strains with rare duplications. More generally, recent work has shown that, across bacteria, cooperative genes are not more likely to be on plasmids (and therefore, transferred) than chromosomes, including in *P. aeruginosa* (61).

Other Forms of Cooperation. Our analyses on QS provided support for cooperation being favored by kin selection. We then tested the robustness of this conclusion by examining five other cases where we could compare genes for cooperative and private traits that were likely to be expressed at similar rates: 1) iron-scavenging siderophore pyoverdine, 2) iron-scavenging siderophore pyochelin, 3) antimicrobial resistance, 4) toxins, and 5) adhesion and movement (Fig. 4 and Table 2). As each comparison considers traits with the same or similar fitness components, the strength of selection is expected to be similar between the "private" and cooperative genes, aiding comparisons with theory (25). We have focused on cooperation because we are examining genes for cooperative traits, but if $r < 1$, then we could also expect selection for conflict and exploitation, as has been examined in the slime mold *D. discoideum* (11, 62).

Examining across these different cases, we consistently found that genes coding for relatively cooperative traits were more polymorphic and showed greater divergence than genes coding for relatively private traits. Comparing across all six cases, including QS, the average level of polymorphism was consistently greater (six of six cases) in genes coding for cooperative traits (Fig. 5) (Wilcoxon signed rank exact test, V = 21, $P = 0.03$). We found analogous patterns when analyzing synonymous and nonsynonymous sites separately (synonymous: six of six cases, Wilcoxon signed rank exact test, V = 21, $P = 0.03$ [*SI Appendix*, Fig. 6]; nonsynonymous: five of six cases, Wilcoxon signed rank exact test, V = 20, $P = 0.06$ [*SI Appendix*, Fig. 7]).

Comparing across all six cases, the average level of nonsynonymous divergence was consistently greater (six of six cases) in genes coding for cooperative traits (Fig. 6) (Wilcoxon signed rank exact test, V = 21, $P = 0.03$), with divergence also higher when analyzing synonymous divergence separately (*SI Appendix*, Fig. 8) (six of six cases, Wilcoxon signed rank exact test, V = 21, $P = 0.03$).

In the above analysis, we examined whether there was a consistent pattern across different types of trait, taking each trait type as a single data point ($n = 6$). One reason that we have taken this relatively conservative approach is that the six traits differ in their power to test between cooperative and private traits. For example, with toxins as well as adhesion and movement, we are comparing relatively private traits that are likely to still have some cooperative benefit compared with relatively more cooperative traits (Table 2). With antibiotic resistance, private and cooperative traits can also involve resistance to difference antibiotics (Table 2). Nonetheless, while some of these other five comparisons could have had less power than our analysis of QS, we found the same consistent pattern across all cases (Figs. 5 and 6).

As an alternative analysis, we also combined all genes from all traits into a single dataset ($n = 92$ cooperative genes, $n = 405$ private genes). In this case, we also found the same pattern: that genes for cooperative traits showed significantly greater polymorphism and divergence (nucleotide diversity: $t_{175} = 3.920$, $P < 0.001$; nonsynonymous divergence: $t_{147} = 4.353$, $P < 0.0001$) (*SI Appendix*, section S6). We did not analyze the patterns within each type of trait separately because the sample size in some groups was too low. For example, we were only able to analyze three private pyochelin genes and four cooperative conflict genes (R and F pyocin bacteriocins).

Clinical Isolates. The robustness of our results was also supported when we analyzed whole genomes from 41 clinical isolates. While most clinical strains are often acquired from the environment (63), it is generally thought that they are not transmitted back to the environment (64). We, therefore, focused on environmental isolates because they are more likely to represent natural populations. Furthermore, certain environmental conditions, such as treatment with antibiotics, may affect diversity at some genes (e.g., those involved in immune escape) but not others, so we decided not to analyze clinical and environmental isolates together. Nonetheless, when analyzing clinical isolates, we found the same qualitative patterns, with genes for cooperative traits showing increased polymorphism consistent with relaxed selection (*SI Appendix*, section S2). Further, our results for polymorphism and divergence are in line with previous studies in this species (*SI Appendix*, section S8).

Relatedness. Given that the predicted degree to which selection is relaxed on cooperative traits is inversely proportional to relatedness between producers and recipients of cooperative traits, we can use our data to estimate relatedness. We do this by comparing the relative level of polymorphism between cooperative and private QS-regulated genes, as we can make direct predictions of relative polymorphism from a simple population genetic model with some assumptions (*SI Appendix*, section S12) (15, 25). In particular, because the theory is about comparing one cooperative gene with one private gene under equal strength of selection, we have to assume that the magnitude and distribution of selection coefficients on cooperative and private traits are on average the same.

We estimate that relatedness is $r = 0.84$ for the natural isolates and $r = 0.85$ for the clinical isolates (*SI Appendix*, section S12). This method allows us to estimate relatedness in natural populations, when it would otherwise be problematic to estimate directly. In order to estimate relatedness directly, it would be necessary to both genotype cells and identify the spatial scale over which social interactions take place. This is possible
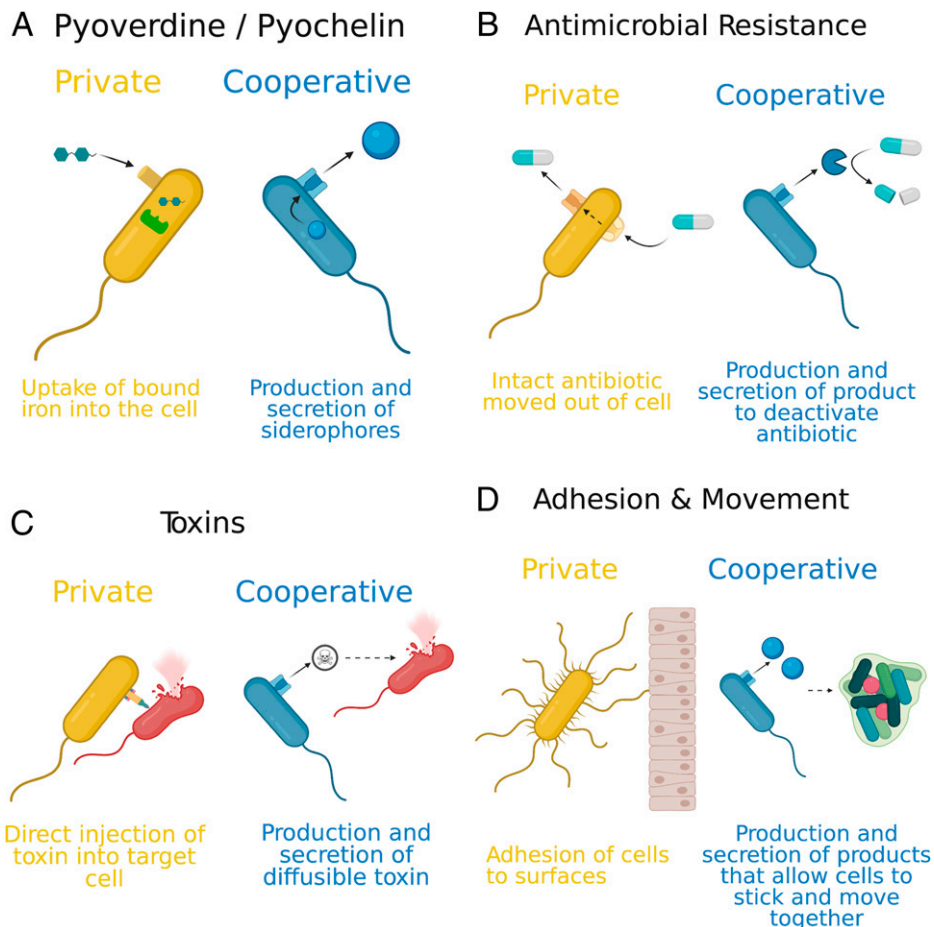
Fig. 4. Secondary comparisons of cooperative vs. private traits. (*A*) Pyoverdine and pyochelin siderophores. (*B*) Antimicrobial resistance. (*C*) Toxins. (*D*) Cell adhesion and movement. More details are in *Methods* and Table 2. Gene lists for each comparison are in *SI Appendix*, Tables 2–4.

in cases where interactions take place in a defined social group, such as a fruiting body (18). In contrast, things get much more difficult with public goods, especially as cells live and grow in a range of different environments and produce a variety of public goods (65). Indeed, laboratory data could even lead to very misleading estimates. In contrast, by using an indirect population genetics approach, we are effectively letting natural selection work out the spatial scale of interaction for us (66). Natural

Table 2. Additional comparisons of cooperative vs. private genes

Comparison	Relatively private genes	Relatively cooperative genes
Pyoverdine	Genes involved in the uptake and use of iron-bound pyoverdine in the cell	Genes involved in the biosynthesis and export of pyoverdine into the extracellular space
Pyochelin	Genes involved in the uptake and use of iron-bound pyochelin in the cell	Genes involved in biosynthesis and export of pyochelin into the extracellular milieu
Antimicrobial resistance	Genes that control efflux pumps, which expel unaltered antibiotics back into the environment, and outer porins, which alter resistance through traits such as membrane stability	Genes where the antibiotic is modified and all cells in the local population benefit; this includes the production of beta-lactamases and enzymes that deactivate aminoglycoside antibiotic
Toxins	Genes that control mechanisms to eliminate competitors via direct contact and the injection of toxins, such as the T6SS; this may still provide an indirect benefit to other cells but relatively less than diffusible toxins	Genes involved in the production of bacteriocins to eliminate competitors, such as R and F pyocins, which diffuse through the environment
Adhesion and movement	Genes that allow cells to stick to and move across surfaces, such as flagella and pili	Genes producing EPSs and rhamnolipids that allow cells to stick and move together

We examined five scenarios. In the first two of these, we compared genes for the same trait with either private (uptake) or cooperative (production and export) fitness consequence: pyoverdine and pyochelin. For the other three, we compared genes for traits with similar functions but where traits varied in the extent to which they were relatively private or relatively cooperative: antimicrobial resistance, toxins, and adhesion/movement.
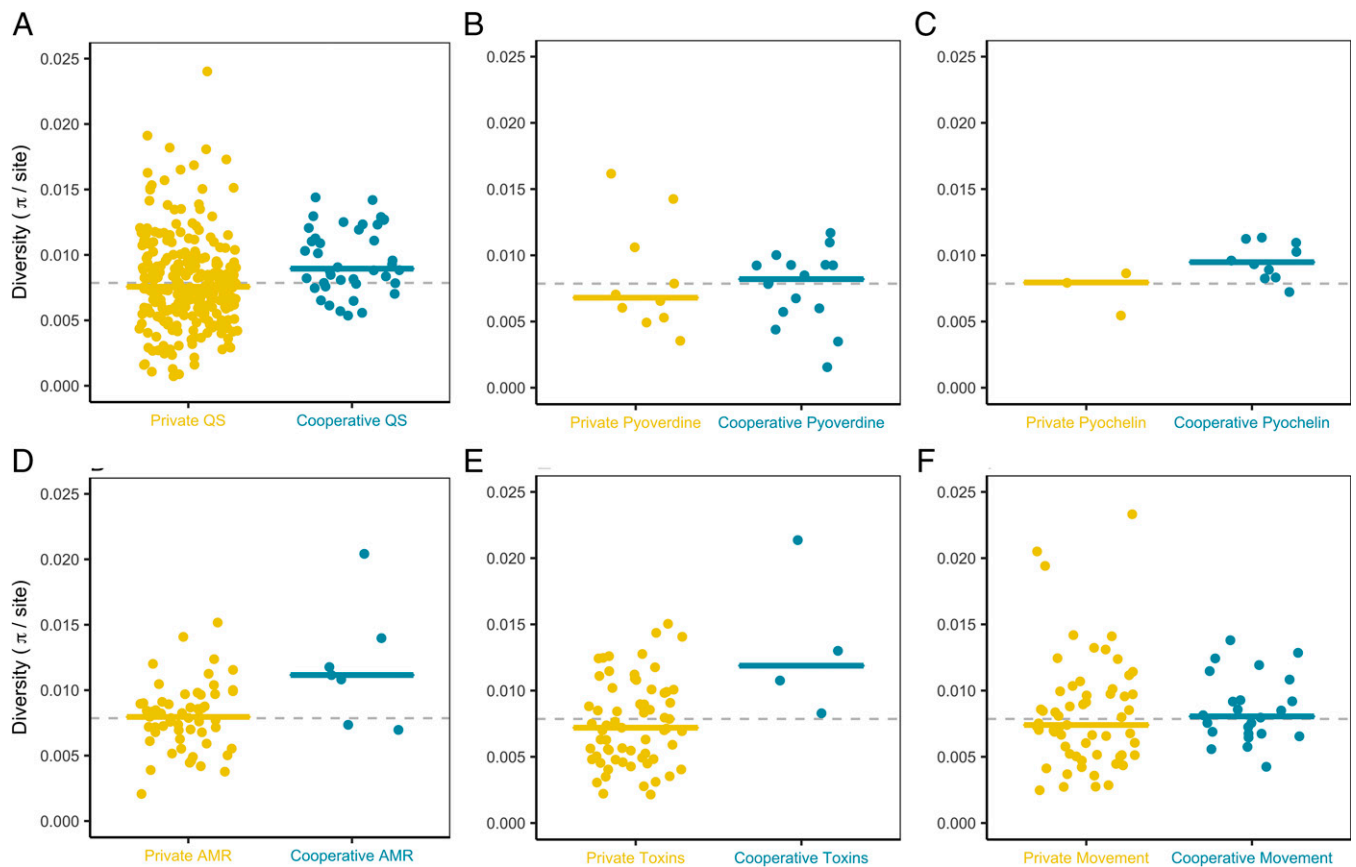
Fig. 5. Private vs. cooperative comparisons for six trait types for polymorphism (nucleotide diversity). *A* is the private vs. cooperative comparison for QS genes from the main analysis (Fig. 2) shown for comparison. *B–F* show the private vs. cooperative comparison for five other traits (see *Methods*). Across different traits, genes for cooperative traits showed a consistent trend toward higher polymorphism than genes for private traits.

selection will respond to the average relatedness, which will depend upon all the factors that would be hard or impossible for us to directly estimate.

Other Species. Our results build upon previous studies to show how cooperative social behaviors can be favored by kin selection in an analogous way across the natural world. Van Dyken and Wade's (15) groundbreaking analysis on QS genes across seven bacterial species found similarly increased polymorphism and divergence but did not have sufficient information at the time to distinguish between private and cooperative QS-controlled traits or control for expression rates (22). Population genetic analyses on the slime mold *D. discoideum* have examined both social conflict and relaxed selection (10–12, 67). In social insects, genes for cooperative traits diverge and evolve faster (28, 68), yet they experience lower rates of adaptive evolution (19). Furthermore, selection on cooperative worker traits appears to be relaxed with increased mating frequency, when relatedness is lower (20, 21, 26). Social insects have the advantage that genes can be readily separated by gene expression data into worker traits, which are presumably cooperative (because workers are largely sterile), and queen traits, which are likely to evolve under direct fitness effects (25, 69–71).

Conclusions

Molecular population genetics offers a powerful tool to study how selection acts in natural populations (56). In combination with theory, this type of analysis can determine the extent to which microbial traits are cooperative and how important this cooperation is in microbes (10–12, 15, 22–25). These results

add to the growing evidence that cooperation plays an important role in natural populations of bacteria and other microorganisms. Experiments carried out in hosts have shown that natural populations engage in cooperation (72–75) and can be exploited by noncooperative cheats (13, 14, 16, 76, 77). Looking across species, comparative studies have found higher levels of cooperation in species where the relatedness between interacting individuals is higher (17, 30). We have shown here that molecular population genetics can also provide evidence for the role of cooperation in natural populations.

Methods

Controlling for Levels of Expression. The central predictions of elevated divergence and polymorphism are characteristic of relaxed selection, but there are factors alongside kin selection (indirect fitness) that can lead to relaxed selection. Notably, conditional expression can also produce the same effect via the same mechanism of weakening the association between possessing a genotype and producing a phenotype that can be seen by selection. Specifically, if a gene is expressed by only a fraction of individuals or by all individuals but in only a fraction of generations, selection is relaxed (22).

In order to control for the effect of conditional expression, we restricted our primary analysis to the subset of genes coinduced by the QS signaling system. QS is a mechanism for coordinating gene expression whereby diffusible signals accumulate as cell density increases, eventually reaching a threshold where the receptor is activated and expression of a set of genes is triggered. In *P. aeruginosa*, there are several hundred genes whose expression is controlled by QS signaling, of which there is an overrepresentation of proposed cooperative traits as well as many private traits (37, 38). We, therefore, compare cooperative genes with private genes within this set of QS-controlled genes, allowing us to control for the effect of conditional expression. In a separate analysis, we assess whether conditional expression itself predicts levels of polymorphism and divergence (*SI Appendix*, section S10).

EVOLUTION
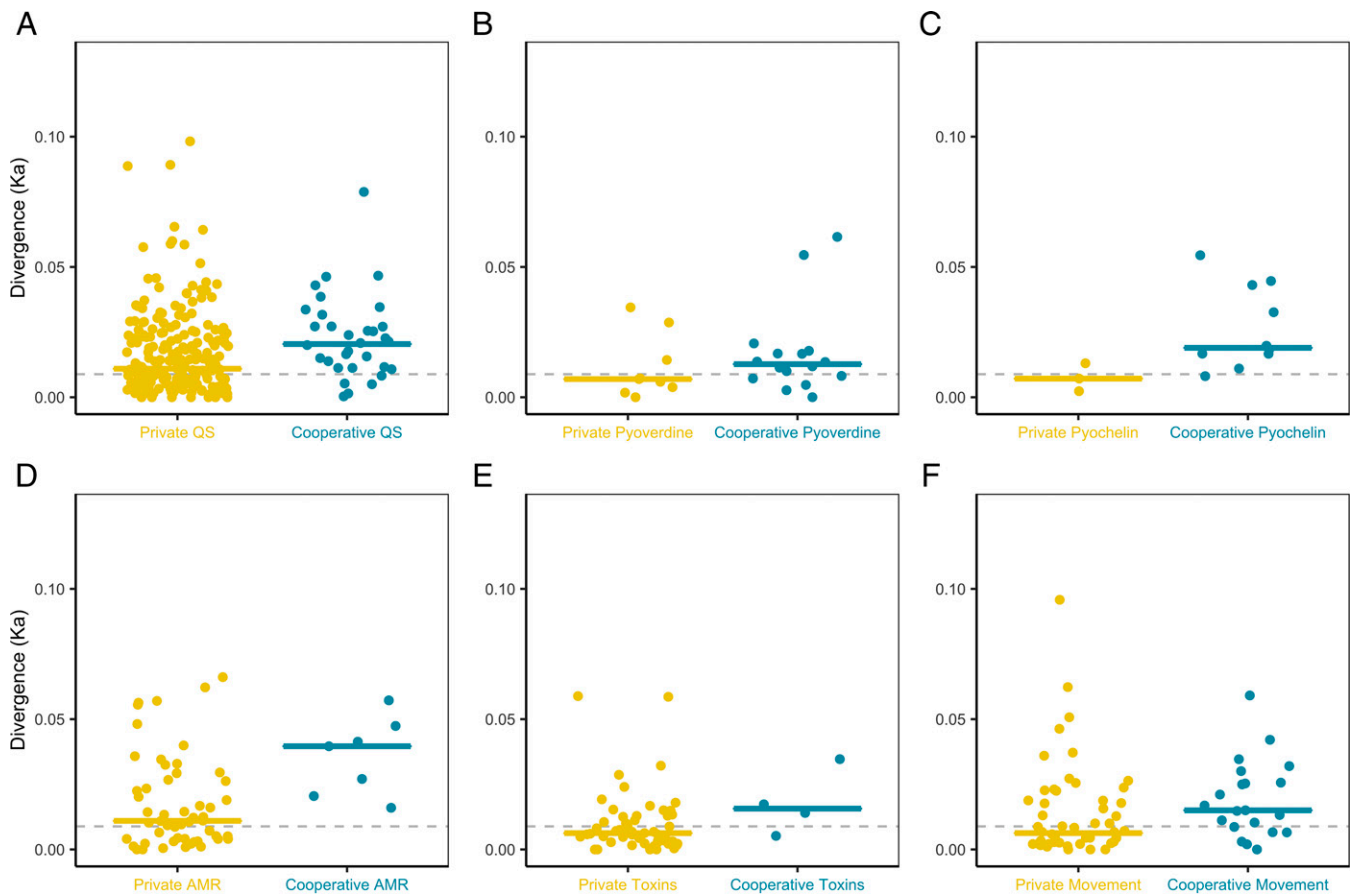
Fig. 6. Private vs. cooperative comparisons for six trait types for divergence (nonsynonymous). *A* is the private vs. cooperative comparison for QS genes from the main analysis (Fig. 3) shown for comparison. *B–F* show the private vs. cooperative comparison for five other traits (see *Methods*). Across different traits, genes for cooperative traits showed a consistent trend toward higher divergence than genes for private traits.

Categorization of Genes. For the main analysis, we focus on genes induced by QS signaling in the *P. aeruginosa* reference strain PAO1, for which we use the set of genes described in ref. 36. Within these 315 genes, we selected a set of genes that are putatively cooperative by manually assessing gene function from annotation in the *Pseudomonas* Genome Database (78) as well as a literature search for any experiments demonstrating a cooperative fitness effect. This was determined by looking for studies that show the basic prediction for a public good (producers outperform nonproducers clonally, but nonproducers outperform producers in groups). The set of cooperative genes and their functions are shown in *SI Appendix*, Table 1. This is not intended to be a fully comprehensive list of genes with any cooperative effect; indeed, there are several QS-induced genes of unknown/predicted function that are plausibly cooperative and several categories of genes that may have at least some cooperative component. We compared cooperative QS with private QS genes for the main comparison and made further comparisons with a background set of private genes in the rest of the genome. For this set of background genes, we used proteins localized to the cytoplasm, as these are the class of gene least likely to have a cooperative function. Such cytoplasmic genes are known to be overrepresented with essential genes (79), which suggests an overrepresentation of genes with functions such as central metabolism and replication.

For some analyses where we needed a set of cooperative genes across the whole genome, we followed the approach of previous studies, which have used extracellular localization as a proxy for sociality (49, 80). Extracellular localization can be reliably and systematically calculated using PSORTb (51). While it is evident that not all cooperative genes are extracellular and not all extracellular proteins are cooperative, any strong effect of sociality is very likely to be captured by this proxy. For further investigation into properties that may differ between cooperative and private genes, we used eggNOG functional annotations (57).

Secondary Comparisons. In our secondary analysis, we examined five comparisons of cooperative vs. private genes (Table 2). First, we used pyoverdine, an iron-scavenging siderophore that is extremely well studied for sociality (4, 32,

81). We separated the genes involved in the pyoverdine pathway into cooperative and private components, which is possible thanks to good knowledge of the function and localization of the genes involved (42). We classified genes involved in the biosynthesis and export of pyoverdine into the extracellular milieu as cooperative and genes involved in the uptake and disassociation of iron-bound pyoverdine in the cell as private (*SI Appendix*, Table 2). Pyochelin, the secondary siderophore of *P. aeruginosa*, was separated into cooperative and private components using the same principles, forming our next secondary comparison.

For the two iron-scavenging comparisons, we separated a single trait into cooperative and private functions, whereas for the other comparisons here, we used separate traits for the private vs. cooperative comparison while making an effort to ensure that the traits are directly comparable.

Antimicrobial resistance is a broad feature that has been well studied in *P. aeruginosa* for its social fitness effects. There are many ways in which cells can express resistance. One such mechanism is through the production of beta-lactamases, which detoxify the environment and therefore, can provide cooperative benefits to the local population (82). Aminoglycoside resistance can also be a cooperative trait; the antibiotic is modified, and therefore, the environment is detoxified (83). This is in contrast to efflux pumps, which expel unaltered antibiotics back into the environment (84) and therefore, have private fitness effects. Outer porins are another private mechanism (85) that alters resistance through traits such as membrane stability (86). The genes used in this analysis are shown in *SI Appendix*, Table 3.

Toxin production is another aspect of bacterial life that can be separated into relatively cooperative and private components. In *P. aeruginosa*, there are various mechanisms by which strains compete with and kill each other, which can again be separated into cooperative and private components. Type VI secretion systems (T6SSs) involve direct contact with competitors and the use of a needle to inject toxins (87), therefore having a private fitness effect. By contrast, bacteriocins, such as R and F pyocins, do not require direct contact and diffuse through the environment (33), which allows cooperative fitness effects on other cells. Elimination of competitors via direct contact can still

have a cooperative social benefit, and so, our comparison here is between relatively cooperative and relatively private. The gene list for R and F pyocins comes from Ghoul et al. (33). Note that we only use the R and F pyocins and not the S pyocins. R and F pyocins are made up of many genes, which form a structure that resembles a bacteriophage tail (88). S pyocins, however, consist only of killing and immunity genes (33) and so, are less comparable with T6SS. The T6SS gene list comes from the set of genes in the known three distinct T6SS loci in *P. aeruginosa* (87) alongside the *vgr* genes (89) (*SI Appendix*, Table 4).

The final cooperative vs. private comparison we used was a broad distinction between extracellular polysaccharides (EPSs) and rhamnolipids, which allow cells to stick and move together and are presumed cooperative, and flagella and pili, which allow cells to stick to and move across surfaces. For EPSs, we used the genes for two of the major *P. aeruginosa* polysaccharides Psl and Pel (90) but not the third polysaccharide EPS alginate, which is only a major component of EPS production in clinical settings (91). For rhamnolipids, a known cooperative trait (6), we used the three biosynthesis genes. For flagella, we used the gene list in Dasgupta et al. (92). For pili, we used the gene list in the review by Burrows (93). This category lumps together some different functions and represents our most tentative grouping (*SI Appendix*, Table 5).

We used paired samples Wilcoxon tests to test if cooperative genes differ significantly from private genes for each population genetic parameter, with the cooperative and private comparison for each trait type forming a pair. We chose the nonparametric form of a paired *t* test because the sample size is quite low for the cooperative genes in some comparisons; so, differences were rarely normally distributed, and means were strongly affected by extreme values. We calculated two-sided *P* values using the wilcox.test function in R.

Sequences. *P. aeruginosa* is an opportunistic pathogen, with most clinical strains also widely spread environmentally (63). To avoid complications from the selection faced in clinical settings, we focused our primary analysis on environmental isolates. It is generally thought that while clinical infections are acquired from the environment, clinical isolates generally are not transmitted back to the environment (64). We chose strains from a list of *P. aeruginosa* strains on the *Pseudomonas* Genome Database accessed at https://pseudomonas.com/ (78). We gathered all available metadata on isolation sources and locations and first filtered for strains for which the raw sequence read data were publicly available (in the form of a sequence read archive [SRA]); then, we further filtered for strains that were unambiguously environmental (by first removing any strains for which the metadata mentioned "human," "clinical," or the name of a disease and then, further removing any records for which it was not possible to ascertain their source). This gave a list of 96 possible strains at the time of analysis. This strain list had heavy

representation of multiple strain collections from the same locality or environment type, so we took a smaller sample of 41 strains by sampling randomly while ensuring that no country was featured more than five times. We also screened the isolates to make sure no strain had a known mutator element, such as *mutS*, that could increase diversity and affect comparisons. While one strain had an in-frame deletion mutation in the mismatch repair gene *mutL*, removing this strain makes no difference to our conclusions (*SI Appendix*, section S13). The 41 strains used are shown in *SI Appendix*, Table 6.

Statistical Analysis. We used R (94) for all statistical analyses and graph plotting. For the main analysis comparing cooperative QS genes with private QS genes, we used a background set of genes for comparison, which was composed of all genes in the genome localized as cytoplasmic by PSORTb. This created a large set of genes; some may, of course, be cooperative, but arguably, this is the group least likely to be cooperative.

Where possible, we used an ANOVA to analyze whether there were any significant differences between our three classes of genes (cooperative, private, and background). Data were transformed using the Box–Cox transformation (95), which finds a value of λ such that the transformation $\frac{y^\lambda - 1}{\lambda}$ gives the best approximation of a normal distribution. Transformed variables were checked for normality with the Kolmogorov–Smirnov test. For some variables, the Box–Cox transformation was not appropriate (as the formulation used does not allow zeros), so a transformation of the form $\log(y + c)$ was used, where c is a constant. After transformation and checking the assumptions of ANOVA tests, we conducted the omnibus ANOVA in R and used Tukey's HSD for post hoc comparisons. Where data transformation was not sufficient to meet the assumptions of an ANOVA, we used the nonparametric Kruskal–Wallis test, which compared medians in a ranked order approach. The Dunn test was used for post hoc comparisons of Kruskal–Wallis tests and was only performed where the omnibus test was significant.

Figures. Results figures were all produced using the ggplot2 package in R (96). Conceptual figures were created with BioRender.

1. J. E. Strassmann, O. M. Gilbert, D. C. Queller, Kin discrimination and cooperation in microbes. *Annu. Rev. Microbiol.* **65**, 349–367 (2011).
2. S. A. West, A. S. Griffin, A. Gardner, S. P. Diggle, Social evolution theory for microorganisms. *Nat. Rev. Microbiol.* **4**, 597–607 (2006).
3. L. McNally, M. Viana, S. P. Brown, Cooperative secretions facilitate host range expansion in bacteria. *Nat. Commun.* **5**, 1–8 (2014).
4. A. S. Griffin, S. A. West, A. Buckling, Cooperation and competition in pathogenic bacteria. *Nature* **430**, 1024–1027 (2004).
5. K. M. Sandoz, S. M. Mitzimberg, M. Schuster, Social cheating in *Pseudomonas aeruginosa* quorum sensing. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15876–15881 (2007).
6. J. B. Xavier, W. Kim, K. R. Foster, A molecular mechanism that stabilizes cooperative secretions in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **79**, 166–179 (2011).
7. M. Ghoul, A. S. Griffin, S. A. West, Toward an evolutionary definition of cheating. *Evolution* **68**, 318–331 (2014).
8. R. Kümmerli, A. S. Griffin, S. A. West, A. Buckling, F. Harrison, Viscous medium promotes cooperation in the pathogenic bacterium *Pseudomonas aeruginosa*. *Proc. Biol. Sci.* **276**, 3531–3538 (2009).
9. W. D. Hamilton, The genetical evolution of social behaviour. I. *J. Theor. Biol.* **7**, 1–16 (1964).
10. J. L. de Oliveira et al., Conditional expression explains molecular evolution of social genes in a microbe. *Nat. Commun.* **10**, 3284 (2019).
11. S. Noh, K. S. Geist, X. Tian, J. E. Strassmann, D. C. Queller, Genetic signatures of microbial altruism and cheating in social amoebas in the wild. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 3096–3101 (2018).
12. E. A. Ostrowski et al., Genomic signatures of cooperation and conflict in the social amoeba. *Curr. Biol.* **25**, 1661–1665 (2015).
13. J. B. Bruce, G. A. Cooper, H. Chabas, S. A. West, A. S. Griffin, Cheating and resistance to cheating in natural populations of the bacterium *Pseudomonas fluorescens*. *Evolution* **71**, 2484–2495 (2017).
14. E. Butaitė, M. Baumgartner, S. Wyder, R. Kümmerli, Siderophore cheating and cheating resistance shape competition for iron in soil and freshwater *Pseudomonas* communities. *Nat. Commun.* **8**, 1–12 (2017).
15. J. D. Van Dyken, M. J. Wade, Detecting the molecular signature of social conflict: Theory and a test with bacterial quorum sensing genes. *Am. Nat.* **179**, 436–450 (2012).
16. O. X. Cordero, L. A. Ventouras, E. F. DeLong, M. F. Polz, Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 20059–20064 (2012).
17. C. Simonet, L. McNally, Kin selection explains the evolution of cooperation in the gut microbiota. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2016046118 (2021).
18. O. M. Gilbert, K. R. Foster, N. J. Mehdiabadi, J. E. Strassmann, D. C. Queller, High relatedness maintains multicellular cooperation in a social amoeba by controlling cheater mutants. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 8913–8917 (2007).
19. M. R. Warner, A. S. Mikheyev, T. A. Linksvayer, Genomic signature of kin selection in an ant with obligately sterile workers. *Mol. Biol. Evol.* **34**, 1780–1787 (2017).
20. B. G. Hunt et al., Relaxed selection is a precursor to the evolution of phenotypic plasticity. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15936–15941 (2011).
21. B. G. Hunt et al., Sociality is linked to rates of protein evolution in a highly social insect. *Mol. Biol. Evol.* **27**, 497–500 (2010).
22. J. D. Van Dyken, M. J. Wade, The genetic signature of conditional expression. *Genetics* **184**, 557–570 (2010).
23. J. D. Van Dyken, T. A. Linksvayer, M. J. Wade, Kin selection-mutation balance: A model for the origin, maintenance, and consequences of social cheating. *Am. Nat.* **177**, 288–300 (2011).
24. T. A. Linksvayer, M. J. Wade, Theoretical predictions for sociogenomic data: The effects of kin selection and sex-limited expression on the evolution of social insect genomes. *Front. Ecol. Evol.* **4**, 1–10 (2016).
25. T. A. Linksvayer, M. J. Wade, Genes with social effects are expected to harbor more sequence variation within and between species. *Evolution* **63**, 1685–1696 (2009).
26. D. W. Hall, M. A. D. Goodisman, The effects of kin selection on rates of molecular evolution in social insects. *Evolution* **66**, 2080–2093 (2012).
27. D. W. Hall, S. V. Yi, M. A. D. Goodisman, Kin selection, genomics and caste-antagonistic pleiotropy. *Biol. Lett.* **9**, 1–4 (2013).
28. S. Vojvodic et al., The transcriptomic and evolutionary signature of social interactions regulating honey bee caste development. *Ecol. Evol.* **5**, 4795–4807 (2015).

Belcher et al.
Kin selection for cooperation in natural bacterial populations

PNAS | 9 of 10
https://doi.org/10.1073/pnas.2119070119

EVOLUTION

29. C. D. Nadell, J. B. Xavier, K. R. Foster, The sociobiology of biofilms. *FEMS Microbiol. Rev.* **33**, 206–224 (2009).

30. R. M. Fisher, C. K. Cornwallis, S. A. West, Group formation, relatedness, and the evolution of multicellularity. *Curr. Biol.* **23**, 1120–1125 (2013).

31. E. T. Granato, T. A. Meiller-Legrand, K. R. Foster, The evolution and ecology of bacterial warfare. *Curr. Biol.* **29**, R521–R537 (2019).

32. R. Kümmerli et al., Co-evolutionary dynamics between public good producers and cheats in the bacterium *Pseudomonas aeruginosa. J. Evol. Biol.* **28**, 2264–2274 (2015).

33. M. Ghoul et al., Bacteriocin-mediated competition in cystic fibrosis lung infections. *Proc. Biol. Sci.* **282**, 20150972 (2015).

34. S. Azimi, A. D. Klementiev, M. Whiteley, S. P. Diggle, Bacterial quorum sensing during infection. *Annu. Rev. Microbiol.* **74**, 201–219 (2020).

35. S. T. Rutherford, B. L. Bassler, Bacterial quorum sensing: Its role in virulence and possibilities for its control. *Cold Spring Harb. Perspect. Med.* **2**, a012427 (2012).

36. M. Schuster, C. P. Lostroh, T. Ogi, E. P. Greenberg, Identification, timing, and signal specificity of *Pseudomonas aeruginosa* quorum-controlled genes: A transcriptome analysis. *J. Bacteriol.* **185**, 2066–2079 (2003).

37. K. B. Gilbert, T. H. Kim, R. Gupta, E. P. Greenberg, M. Schuster, Global position analysis of the *Pseudomonas aeruginosa* quorum-sensing transcription factor LasR. *Mol. Microbiol.* **73**, 1072–1085 (2009).

38. M. Schuster, D. J. Sexton, B. A. Hense, Why quorum sensing controls private goods. *Front. Microbiol.* **8**, 885 (2017).

39. R. Chen, E. Déziel, M. C. Groleau, A. L. Schaefer, E. P. Greenberg, Social cheating in a *Pseudomonas aeruginosa* quorum-sensing variant. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7021–7026 (2019).

40. J. L. Bradshaw et al., *Pseudomonas aeruginosa* Protease IV exacerbates pneumococcal pneumonia and systemic disease. *MSphere* **3**, e00212-18 (2018).

41. T. Robinson, P. Smith, E. R. Alberts, M. Colussi-Pelaez, M. Schuster, Cooperation and cheating through a secreted aminopeptidase in the *Pseudomonas aeruginosa* RpoS response. *mBio* **11**, e03090-19 (2020).

42. M. T. Ringel, T. Brüser, The biosynthesis of pyoverdines. *Microb. Cell* **5**, 424–437 (2018).

43. P. H. Roy et al., Complete genome sequence of the multiresistant taxonomic outlier *Pseudomonas aeruginosa* PA7. *PLoS One* **5**, e8842 (2010).

44. E. Lebeuf-Taylor, N. McCloskey, S. F. Bailey, A. Hinz, R. Kassen, The distribution of fitness effects among synonymous mutations in a gene under selection. *Elife* **8**, e45952 (2019).

45. M. P. Zwart et al., Unraveling the causes of adaptive benefits of synonymous mutations in TEM-1 β-lactamase. *Heredity* **121**, 406–421 (2018).

46. M. Meir et al., Competition between social cheater viruses is driven by mechanistically different cheating strategies. *Sci. Adv.* **6**, eabb7990 (2020).

47. S. Azimi et al., Allelic polymorphism shapes community function in evolving *Pseudomonas aeruginosa* populations. *ISME J.* **14**, 1929–1942. (2020).

48. E. L. Bruger, D. J. Synder, V. S. Cooper, C. M. Waters, Quorum sensing provides a molecular mechanism for evolution to tune and maintain investment in cooperation. *ISME J.* **15**, 1236–1247 (2020).

49. M. Garcia-Garcera, E. P. C. Rocha, Community diversity and habitat structure shape the repertoire of extracellular proteins in bacteria. *Nat. Commun.* **11**, 758 (2020).

50. T. Nogueira, M. Touchon, E. P. C. Rocha, Rapid evolution of the sequences and gene repertoires of secreted proteins in bacteria. *PLoS One* **7**, e49403 (2012).

51. N. Y. Yu et al., PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**, 1608–1615 (2010).

52. L. R. Hoffman et al., *Pseudomonas aeruginosa* lasR mutants are associated with cystic fibrosis lung disease progression. *J. Cyst. Fibros.* **8**, 66–70 (2009).

53. J. Gore, H. Youk, A. van Oudenaarden, Snowdrift game dynamics and facultative cheating in yeast. *Nature* **459**, 253–256 (2009).

54. A. Ross-Gillespie, A. Gardner, S. A. West, A. S. Griffin, Frequency dependence and cooperation: Theory and a test with bacteria. *Am. Nat.* **170**, 331–342 (2007).

55. M. Kimura, Evolutionary rate at the molecular level. *Nature* **217**, 624–626 (1968).

56. M. W. Hahn, *Molecular Population Genetics* (OUP USA, 2018).

57. J. Huerta-Cepas et al., eggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314 (2019).

58. A. O. Urrutia, L. D. Hurst, The signature of selection mediated by expression on human genes. *Genome Res.* **13**, 2260–2264 (2003).

59. I. Lopes, G. Altab, P. Raina, J. P. de Magalhães, Gene size matters: An analysis of gene length in the human genome. *Front. Genet.* **12**, 559998 (2021).

60. E. P. C. Rocha, Neutral theory, microbial practice: Challenges in bacterial population genetics. *Mol. Biol. Evol.* **35**, 1338–1347 (2018).

61. A. E. Dewar et al., Plasmids do not consistently stabilize cooperation across bacteria but may promote broad pathogen host-range. *Nat. Ecol. Evol.* **5**, 1624–1636 (2021).

62. S. A. West, M. Ghoul, Conflict within cooperation. *Curr. Biol.* **29**, R425–R426 (2019).

63. J. P. Pirnay et al., *Pseudomonas aeruginosa* population structure revisited. *PLoS One* **4**, e7740 (2009).

64. E. A. Ozer, E. Nnah, X. Didelot, R. J. Whitaker, A. R. Hauser, The population structure of *Pseudomonas aeruginosa* is characterized by genetic isolation of exoU+ and exoS+ lineages. *Genome Biol. Evol.* **11**, 1780–1796 (2019).

65. P. Smith, M. Schuster, Public goods and cheating in microbes. *Curr. Biol.* **29**, R442–R447 (2019).

66. S. Nee, S. A. West, A. F. Read, Inbreeding and parasite sex ratios. *Proc. Biol. Sci.* **269**, 755–760 (2002).

67. E. A. Ostrowski, Enforcing cooperation in the social amoebae. *Curr. Biol.* **29**, R474–R484 (2019).

68. C. Tong, G. M. Najm, N. Pinter-Wollman, J. N. Pruitt, T. A. Linksvayer, Comparative genomics identifies putative signatures of sociality in spiders. *Genome Biol. Evol.* **12**, 122–133 (2020).

69. S. Patalano et al., Molecular signatures of plastic phenotypes in two eusocial insect species with simple societies. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13970–13975 (2015).

70. P. G. Ferreira et al., Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. *Genome Biol.* **14**, R20 (2013).

71. B. A. Taylor, A. Cini, C. D. R. Wyatt, M. Reuter, S. Sumner, The molecular basis of socially mediated phenotypic plasticity in a eusocial paper wasp. *Nat. Commun.* **12**, 775 (2021).

72. K. P. Rumbaugh et al., Quorum sensing and the social evolution of bacterial virulence. *Curr. Biol.* **19**, 341–345 (2009).

73. E. J. G. Pollitt, S. A. West, S. A. Crusz, M. N. Burton-Chellew, S. P. Diggle, Cooperation, quorum sensing, and evolution of virulence in *Staphylococcus aureus. Infect. Immun.* **82**, 1045–1051 (2014).

74. L. Zhou, L. Slamti, C. Nielsen-LeRoux, D. Lereclus, B. Raymond, The social biology of quorum sensing in a naturalistic host pathogen system. *Curr. Biol.* **24**, 2417–2422 (2014).

75. S. Gu et al., Competition for iron drives phytopathogen control by natural rhizosphere microbiomes. *Nat. Microbiol.* **5**, 1002–1010 (2020).

76. S. B. Andersen, R. L. Marvig, S. Molin, H. Krogh Johansen, A. S. Griffin, Long-term social dynamics drive loss of function in pathogenic bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10756–10761 (2015).

77. J. L. Sachs, M. O. Ehinger, E. L. Simms, Origins of cheating and loss of symbiosis in wild Bradyrhizobium. *J. Evol. Biol.* **23**, 1075–1089 (2010).

78. G. L. Winsor et al., Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res.* **44**, D646–D653 (2016).

79. C. Peng, F. Gao, Protein localization analysis of essential genes in prokaryotes. *Sci. Rep.* **4**, 6001 (2014).

80. T. Nogueira et al., Horizontal gene transfer of the secretome drives the evolution of bacterial cooperation and virulence. *Curr. Biol.* **19**, 1683–1691 (2009).

81. S. O'Brien, A. M. Luján, S. Paterson, M. A. Cant, A. Buckling, Adaptation to public goods cheats in *Pseudomonas aeruginosa. Proc. Biol. Sci.* **284**, 20171089 (2017).

82. E. Amanatidou et al., Biofilms facilitate cheating and social exploitation of β-lactam resistance in *Escherichia coli. NPJ Biofilms Microbiomes* **5**, 36 (2019).

83. K. Poole, Aminoglycoside resistance in *Pseudomonas aeruginosa. Antimicrob. Agents Chemother.* **49**, 479–487 (2005).

84. K. Poole, *Pseudomonas aeruginosa*: Resistance to the max. *Front. Microbiol.* **2**, 1–13 (2011).

85. S. Chevalier et al., Structure, function and regulation of *Pseudomonas aeruginosa* porins. *FEMS Microbiol. Rev.* **41**, 698–722 (2017).

86. R. E. W. Hancock, F. S. L. Brinkman, Function of pseudomonas porins in uptake and efflux. *Annu. Rev. Microbiol.* **56**, 17–38 (2002).

87. J. D. Mougous et al., A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science* **312**, 1526–1530 (2006).

88. Y. Michel-Briand, C. Baysse, The pyocins of *Pseudomonas aeruginosa. Biochimie* **84**, 499–510 (2002).

89. R. D. Hood et al., A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell Host Microbe* **7**, 25–37 (2010).

90. Y. Irie et al., The *Pseudomonas aeruginosa* PSL polysaccharide is a social but non-cheatable trait in biofilms. *MBio* **8**, e00374-17 (2017).

91. D. J. Wozniak et al., Alginate is not a significant component of the extracellular polysaccharide matrix of PA14 and PAO1 *Pseudomonas aeruginosa* biofilms. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7907–7912 (2003).

92. N. Dasgupta et al., A four-tiered transcriptional regulatory circuit controls flagellar biogenesis in *Pseudomonas aeruginosa. Mol. Microbiol.* **50**, 809–824 (2003).

93. L. L. Burrows, *Pseudomonas aeruginosa* twitching motility: Type IV pili in action. *Annu. Rev. Microbiol.* **66**, 493–520 (2012).

94. R version 4.0.5, R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, Austria, 2020).

95. G. E. P. Box, D. R. Cox, An analysis of transformations. *J. R. Stat. Soc. B* **26**, 211–243 (1964).

96. H. Wickham, ggplot2 Elegant Graphics for Data Analysis. https://ggplot2.tidyverse.org. Accessed 1 January 2021.

97. L. J. Belcher, A. E. Dewar, M. Ghoul, S. A. West, Kin selection for cooperation in natural bacterial populations. Oxford University Research Archive. https://doi.org/10.5287/bodleian:7J4Md7bA2. Deposited 12 January 2022.