Supplementary Material

Supplement S1: Synonymous codon usage

In this study, we found that cooperative genes have significantly higher polymorphism than private genes. Importantly, this elevated polymorphism was at both synonymous and non-synonymous sites. This result is initially puzzling, since we expect cooperative genes to accumulate polymorphism due to increased equilibrium frequency of deleterious mutations, and so we might expect to see elevated polymorphism only at non-synonymous sites. Such an expectation is however built upon the assumption that synonymous variation is neutral. This assumption is increasingly seeming unlikely, particularly in bacteria (1, 2). In particular, it is now well-documented that selection acts on synonymous codon usage across bacteria (1, 3). Certain codons may be preferred due to metabolic costs (4), a general preference for GC as opposed to AT (5), or due to relative availability of tRNAs (6). In this section, we investigate whether the elevated synonymous polymorphism we observe in cooperative genes can be explained by observable differences in codon usage, and whether any differences fit our hypothesis of relaxed selection based on what we know about codon usage bias in *P. aeruginosa*.

Codon usage bias analysis

There are many established methods for measuring codon usage bias (7, 8). They can be broadly split into two types: 1) those which make comparisons to equal use of each codon within an amino acid; 2) those which make comparisons to a set of highly expressed genes (under the assumption that highly expressed genes have stronger selection for codon usage). The methods that we use to test for codon usage are shown in Table S1.1, alongside our predictions if selection is relaxed on cooperative genes. For each metric, we can make a prediction of whether it should be greater in cooperative or private genes if selection is relaxed on cooperative genes.

We also look at GC content, as GC rich codons tend to be favored (9). *P. aeruginosa* has a genome-wide GC content of 67%, despite having the AT mutational bias characteristic of most bacteria (10), leading us to hypothesize that relaxed selection on cooperative genes may lead to a weaker preference for GC rich codons than cooperative genes.

Table S1.1: Codon usage bias metrics, and their predictions if selection is relaxed

| Metric                      | Description                                                                 | Prediction if selection relaxed                                      |
|-----------------------------|-----------------------------------------------------------------------------|-----------------------------------------------------------------------|
| RSCU                        | Relative use of a codon compared to null expectation of each codon for an    | Cooperative genes use relatively more of AT rich codons               |
| Relative Synonymous Codon   | amino acid being used equally (RSCU=1)                                      |                                                                       |
| Usage (7)                   |                                                                            |                                                                       |
| CAI                         | Relative use of a codon compared to that in highly expressed genes (as       | Codon adaptation index higher in private genes                         |
| Codon Adaptation Index (7)  | determined by RA³)                                                         |                                                                       |
| FOP                         | Uses data on highly expressed genes to define favourite(s) codon for each   | Frequency of optimal codon(s) greater in private genes                |
| Frequency of Optimal Codon  | amino acid, and calculate relative usage of favourite vs. non-favourite     |                                                                       |
| (8)                         | codon(s) for each gene                                                     |                                                                       |
RA (Relative Adaptiveness of codons (7)) Uses data on highly expressed genes to create a reference table of how often each codon is used in genes where selection on codon usage is presumed to be strongest.

For each metric, we calculated the value for each gene in our set of cooperative and private QS genes using the R package seqinR (11). For metrics where we needed to define a set of essential highly expressed genes, we used the dataset of 321 genes from (12).

Results

First, we look at relative synonymous codon usage (RSCU), as this requires no assumption about the difference in selection between highly expressed genes and our chosen genes. The comparison between cooperative and private genes is shown in Figure S1.1. The figure shows that any differences in usage of individual codons are minor (with no one codon showing significantly different RSCU from t-tests with FDR adjusted p-values). If there was strong evidence for relaxed selection, we might expect to see cooperative genes being equally likely to use all codons within an amino acid (RSCU = 1), or (more likely) RSCU greater for the codons which are AT rich, since this is where the mutational bias is.

Figure S1.1: Codon usage bias (RSCU) in cooperative (blue) vs private (yellow) QS genes. Each amino acid is represented by a contrasting grey column, with each point and error bar representing the mean ± 1 s.d. for that codon. The colour bar shows the GC content of each codon, with red=3, yellow=2, green=1, blue=0. The colour in the label of each amino acid represents the biosynthesis cost of each amino acid (as determined in E. coli (4), with blue the lowest cost, yellow intermediate cost, and red high cost).

For the other codon usage bias metric, results are presented in Table S1.2. It is important to note that we know there are some codon usage differences between cooperative and private genes, because polymorphism at synonymous sites is significantly higher in cooperative genes (see main text). The minor differences we observe necessarily sum to the significant differences in polymorphism, since both are calculated from exactly the same data. The question, therefore,
is whether these differences fit the pattern expected if selection is relaxed on cooperative traits.
There is a general trend noticeable, and it is in the direction predicted. Cooperative genes appear to use slightly fewer ‘adaptive’ codons (relative to highly expressed genes) than private genes, and have slightly lower GC content. There is however no difference in the frequency of optimal codon (FOC). The slightly lower GC content of cooperative genes could be evidence of relaxed selection, given that it is likely there is selection favoring high GC content (5).

Table S1.2: Codon usage bias statistics for cooperative vs. private genes

| Metric   | Cooperative genes | Private genes |
|----------|-------------------|---------------|
| CAI      | 0.676             | 0.681         |
| FOC      | 0.45              | 0.45          |
| GC content % | 66.6            | 67.1          |

Previous studies have shown that CAI is high throughout the genome in P. aeruginosa (13, 14), rather than being strongly linked to gene expression as it is in other species (15). This has been suggested to be linked to the need for P. aeruginosa to have optimal gene expression across the range of metabolic and environmental conditions that they experience (13). This may explain why we don’t detect a strong signature of relaxed selection in the codon usage data presented here. Given the known stronger relationship between gene expression and codon usage bias in other species (particularly fast-growing bacteria) (16), this method of detecting signatures of relaxed selection may prove more fruitful in further studies in other species. It is also worth noting that selection on codon usage is by no means restricted to microbes (17).
Supplement S2: Clinical Strains analysis

To complement our analysis from the main text on environmental isolates, we conducted an analysis on clinical isolates. We randomly sampled 41 clinical isolates from our dataset of strains for which raw sequence read data is publicly available, and called SNPs and conducted molecular population genetics using the exact same methods as in the main text. The strains used and relevant meta-information is in Table S2.1.

| Strain                  | Accession | Location | Source | Disease               | SRA         |
|-------------------------|-----------|----------|--------|-----------------------|-------------|
| LES_CF_sputum_CF03      | SAMEA2732173 | UK       | sputum | Cystic Fibrosis (CF)  | ERR953499   |
| AUS717                  | SAMN10478459 | Australia | sputum | CF                    | SRR8612814  |
| AUS074                  | SAMN10478455 | Australia | sputum | CF                    | SRR8612818  |
| Zw31                    | SAMN10478533 | Austria  | throat | CF                    | SRR8612774  |
| WH-SGI-V-07256          | SAMN04128707 | Belgium  | Hospital | -                   | SRR2939468  |
| WH-SGI-V-07263          | SAMN04128712 | Belgium  | Burn   | -                     | SRR2939473  |
| WH-SGI-V-07310          | SAMN04128742 | Bulgaria | Burn   | -                     | SRR2939503  |
| WH-SGI-V-07322          | SAMN04128751 | Colombia | blood  | -                     | SRR2939512  |
| WH-SGI-V-07302          | SAMN04128738 | Czech Republic | Hospital | -                   | SRR2939499  |
| WH-SGI-V-07291          | SAMN04128731 | DRC      | blood  | -                     | SRR2939492  |
| WH-SGI-V-07406          | SAMN04128580 | France   | Hospital | -                   | SRR2939341  |
| PAL1.1                  | SAMN09205884 | France   | lung   | Pneumonia            | SRR7632685  |
| WH-SGI-V-07249          | SAMN04128700 | Germany  | Burn   | -                     | SRR2939461  |
| Zw73_2                  | SAMN10478563 | Germany  | throat | CF                    | SRR8612786  |
| H2                      | SAMN10478529 | Germany  | Catheter | -                   | SRR8612893  |
| HJ2                     | SAMN10478528 | Germany  | sputum | CF                    | SRR8612898  |
| BK5                     | SAMN04128575 | India    | cornea | Keratitis            | SRR2063980  |
| MCF430                  | SAMN10478518 | Israel   | throat | CF                    | SRR8612873  |
| Zw49                    | SAMN10478534 | Italy    | throat | CF                    | SRR8612773  |
| CF16053                 | SAMN12635929 | Mexico   | blood  | Bacteremia            | SRR10093385 |
| WH-SGI-V-07306          | SAMN04128740 | Netherlands | Burn   | -                     | SRR2939501  |
| PA_185                  | SAMN10248972 | Pakistan | Intensive care unit | -          | SRR10094501 |
| WH-SGI-V-07279          | SAMN04128722 | Portugal | sputum | CF                    | SRR2939483  |
| WH-SGI-V-07282          | SAMN04128725 | Portugal | sputum | CF                    | SRR2939486  |
### Results

In general, the pattern observed in environmental strains is the same as in environmental strains. Here, we present the results of the main analysis.

**Cooperative QS genes are more polymorphic than private QS genes**

Cooperative genes are significantly more polymorphic than private genes, as measured by average pairwise nucleotide diversity $\pi$ (ANOVA $F_{1,2352}=12.0$, $p<0.001$. Tukey’s HSD $p=0.01$) (Figure S2.1). The increased polymorphism is manifest at both synonymous and non-synonymous sites (synonymous: ANOVA $F_{1,2351}=29.9$, $p<10^{-7}$. Tukey’s HSD $p=0.004$) (non-synonymous: Kruskall-Wallis $X^2(2) = 22.0$, $p < 10^{-4}$. Dunn Test $p=0.04$) (Figure S2.2).

Due to elevated polymorphism at both non-synonymous and synonymous sites, the ratio between non-synonymous polymorphism and synonymous polymorphism is not significantly different between cooperative and private genes (ANOVA $F_{1,2339}=32.4$, $p<10^{-7}$. Tukey’s HSD $p=0.63$), but the ratio is significantly higher in private QS genes vs. private genes across the genome (Tukey’s HSD $p<10^{-4}$) (Figure S2.3).
Figure S2.1: Nucleotide diversity per site for Private QS (yellow) and Cooperative QS (blue) genes. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.

Figure S2.2: Nucleotide diversity per site for (A) synonymous (B) non-synonymous sites. Private QS (yellow) and Cooperative QS (blue) genes. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.
Figure S2.3: Ratio between non-synonymous and synonymous nucleotide diversity per site for Private QS (yellow) and Cooperative QS (blue) genes. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.

Cooperative genes are more divergent than private genes

As divergence is measured by substitutions in comparison to an outgroup (see methods), the results for divergence don’t differ between this set of strains and the set of strains used in the main text, where we found that divergence is significantly greater in cooperative genes compared to private genes.

Cooperative genes aren’t more likely to be evolving under balancing selection or with selective sweeps compared to private genes

Using a binomial test to assess significant over- or under-representation of cooperative genes in the set of genes which show significant deviations from the null expectation for various measures of the allele frequency spectrum (see methods) we find no evidence that cooperative genes are experiencing a different form of selection to private genes (Table S2.2). If cooperative genes were more polymorphic than private genes due to being more likely under balancing selection (a reasonable and often cited hypothesis), then we would expect to detect this using these measures.

Table S2.2: Significance tests for representation of cooperative genes within the group of genes with significant deviation from neutrality in allele frequency spectrum measures

| Test          | Genes with Significant deviation | Cooperative genes with significant deviation | p-value |
|---------------|---------------------------------|---------------------------------------------|---------|
| Tajima’s D    | 1514                            | 16                                         | 0.100   |
| Fu & Li’s D*  | 282                             | 4                                          | 1.000   |
| Fu & Li’s F*  | 298                             | 4                                          | 1.000   |
Cooperative genes aren’t more likely to be evolving under positive selection than private genes. If cooperative genes were more likely to be evolving under positive selection than private genes, then we would expect to detect that in various population genetic measures that use a comparison of polymorphism and divergence for a gene to assess deviations from neutral expectations.

We find no significant difference in neutrality index ANOVA $F_{1,1447}=4.26$, $p=0.01$. Tukey’s HSD $p=0.36$ (Figure S4.4) or Direction of Selection statistic (ANOVA $F_{1,1084}=0.163$, $p=0.687$. Tukey’s HSD $p=0.672$) (Figure S2.5) for cooperative and private genes.

We also find no significant over- or under-representation of cooperative genes in the group of genes for which the Mcdonald-Krietman test shows a significant deviation from the null hypothesis (for either positive or balancing selection) (7/245 genes extracellular, binomial test $p=0.12$).

**Figure S2.4:** -log-transformed neutrality index for Private QS (yellow) and Cooperative QS (blue) genes. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.
Deleterious Mutations

To test the hypothesis that cooperative genes should be overrepresented in genes with segregating deleterious mutations, we looked for two types of mutations in our dataset: mutations that generate stop codons and frameshift mutations. Of the 356 variants that we find, 8 are in cooperative genes (extracellular). This does not represent a significant overrepresentation of cooperative genes (binomial test $p=0.289$). This doesn’t change if we consider number of genes with at least one mutation, rather than total mutations ($7/288$ extracellular, binomial test $p=0.236$). We note however that the gene with the most of this type of mutation is the QS signal receptor lasR. Strains with these mutations in this gene act as social ‘cheats’, which don’t respond to the QS signal and are often seen in clinical infections (18).
Cooperative genes aren’t more likely to be evolving under balancing selection or with selective sweeps compared to private genes

Population genetic measures which use data from the allele frequency spectrum are powerful tools for detecting selection by comparing the distribution of allele frequencies with the neutral expectation. They are however best interpreted by looking at the subset of genes which show a statistically significant deviation from the neutral expectation.

Tajima’s D compares pairwise differences with the number of segregating sites, and is commonly used to detect balancing selection (53). Demographic fluctuations such as population bottlenecks and fluctuations, or strong population structure can change the frequency of rare alleles and therefore the ability of the test to detect balancing selection (54, 55).

Using a binomial test to assess significant over- or under-representation of cooperative genes in the set of genes which show significant deviations from the null expectation (see methods) we find no evidence that cooperative genes are experiencing a different form of selection to private genes (Table S3.1). If cooperative genes were more polymorphic than private genes due to being more likely under balancing selection (a reasonable and often cited hypothesis), then we would expect to detect this using these measures.

**Table S3.1**: Significance tests for representation of cooperative genes within the group of genes with significant deviation from neutrality in allele frequency spectrum measures

| Test         | Genes with Significant deviation | Cooperative genes with significant deviation | p-value |
|--------------|----------------------------------|---------------------------------------------|---------|
| Tajima’s D   | 484                              | 4                                           | 0.206   |
| Fu & Li’s D* | 282                              | 4                                           | 1.000   |
| Fu & Li’s F* | 298                              | 4                                           | 1.000   |
Supplement S4: Positive selection

Cooperative genes aren’t more likely to be evolving under positive selection compared to private genes

The McDonald-Kreitman (MK) test is a powerful method for detecting selection by comparing synonymous and non-synonymous polymorphism and divergence. A p-value is obtained for each gene by comparing the ratio of non-synonymous to synonymous divergence to the ratio of non-synonymous to synonymous polymorphism using Fisher’s exact test. The MK test is considered robust to synonymous mutations not being selectively equivalent (19) and to demographic changes (20). If synonymous sites are under selection, as appears to be the case here, the MK test is likely to be biased towards finding a signature of positive selection (21), so we supplement our analysis below with other indices such as the Direction of Selection indices.

In order to make an assessment about different selection that may be acting on cooperative vs. private genes, we can look at the subset of genes for which the MK test shows a statistically significant deviation from the neutral expectation – which in this case is 132 genes. It is important to note that the MK test can be significant either because of a significant overrepresentation of non-synonymous fixed differences (indicative of positive selection), or because of an overrepresentation of non-synonymous polymorphism (which is indicative of strong balancing selection). Of the 123 genes with evidence for positive selection and known protein localization, 3 are cooperative, which is not a significant over- or under-representation (binomial test P=0.452). Of the 11 genes with evidence for strong balancing selection and known protein localization, 1 is cooperative, which is not a significant over-representation (binomial test P=0.162). Because of the small sample size, this test doesn’t allow us to determine if this is an underrepresentation. If cooperative genes are experiencing relaxed selection, we would have predicted that one signature of this would be an under-representation in genes evolving under positive selection, but we don’t find this here.

There are other ways of using the same data and underlying theory to assess selection. We looked at neutrality index summary statistic (log-transformed such that +ve values suggest positive selection) which shows a trend for cooperative genes to be under weaker positive selection (Figure S4.1), which is in the direction that we would expect if selection is indeed relaxed on cooperative genes. We also looked at $\alpha$, which is meant to represent the proportion of non-synonymous mutations fixed by positive selection. When calculated on a gene-by-gene basis, many of these values will be negative so the interpretation as a proportion breaks down. As such, we calculated $\alpha$ based on the average of each parameter for a class of genes (cooperative or private), giving a single value for cooperative genes and for private genes. For cooperative genes, $\alpha=0.33$ (95% CI: 0.17-0.49), whereas for private genes, $\alpha=0.30$ (95% CI: 0.20-0.37). This implies that the proportion of positively selected substitutions is not different in cooperative vs. private genes. Whilst we may have expected $\alpha$ to be lower in cooperative genes if selection is relaxed, the finding that $\alpha$ isn’t higher provides evidence against the positive selection interpretation of the greater divergence of cooperative genes.
Figure S4.1: -log-transformed neutrality index for Private QS (yellow) and Cooperative QS (blue) genes. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.

Direction of selection is another statistic that uses the same information to try and produce a quantitative measure that can be used to compare between classes of genes for different signatures of selection – with the ratio of non-synonymous divergence to total divergence compared to the ratio of non-synonymous polymorphism to total polymorphism. There is no difference evident between cooperative and private genes from the measure as calculated for each gene individually (Figure S4.2). When we used the same approach as before and calculated a single value for each class of gene (by averaging parameters), bootstrapping (by sampling with replacement) to calculate confidence intervals (5 and 95 percentiles of bootstrap parameter – true parameter for n=10000 iterations), we found that DoS for cooperative genes is 0.0855 (95% CI: 0.041-0.145), and private 0.0705 (95% CI: 0.042-0.145). Together, these findings from comparing polymorphism to divergence data indicate that there are no fundamental differences in the form of selection experienced by cooperative and private genes. We don’t find evidence in support of the hypotheses that cooperative genes may have greater divergence due to being more likely to be experiencing positive selection, or greater polymorphism due to being more likely under balancing selection. Only by looking at polymorphism and divergence data both separately and together (in tests like MKT and DoS), alongside allele frequency measures can we build a more complete picture of what selection genes are under. Looking at any one of these classes of evidence in isolation could have easily led us to make plausible conclusions that aren’t supported by the whole evidence.
Figure S4.2: Direction of Selection (DoS) statistic for Private QS (yellow) and Cooperative QS (blue) genes. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.
Supplement S5: Comparisons between cooperative QS genes and private genes

Cooperative QS genes were compared to private QS genes for the main comparison, and we also made comparisons to a background set of private genes in the rest of the genome, for which we used genes whose proteins localise to the cytoplasm, as these are the class of gene least likely to have a cooperative function. Indeed, cytoplasmic genes are known to be over-represented with essential genes (22), which suggest an overrepresentation of genes with functions such as central metabolism and replication.

Reported stats are for post-hoc comparisons between cooperative genes and background private genes. Where an ANOVA was used for the main test, Tukey’s HSD is used. Where the non-parametric Kruskall-Wallis test was used for the main text, Dunn’s test is used to perform the post-hoc comparison.

Supplementary Table S5.1: Statistical analysis of comparison between cooperative and background private genes for seven parameters. For all parameters, an omnibus test was conducted (either ANOVA or Kruskall-Wallis, see main text) and the relevant post-hoc comparison is reported here for the cooperative genes in comparison to background private genes.

| Parameter                                      | Post-hoc test | p-value |
|------------------------------------------------|---------------|---------|
| Nucleotide diversity                           | Tukey’s HSD   | p<0.01  |
| Synonymous nucleotide diversity                | Tukey’s HSD   | p<0.03  |
| Non-synonymous nucleotide diversity            | Dunn’s Test    | p<0.001 |
| Non-synonymous nucleotide diversity / synonymous nucleotide diversity | Tukey’s HSD | p<0.03  |
| Synonymous divergence                           | Tukey’s HSD   | p<0.003 |
| Non-synonymous divergence                      | Dunn’s Test    | p<0.02  |
| Non-synonymous divergence / synonymous divergence | Dunn’s Test   | p<0.003 |
Supplement S6: Secondary comparisons of private vs. cooperative genes

In the main text we present an analysis of the overall patterns in private vs. cooperative genes by conducting a paired analysis where each of our six comparisons have one datapoint for the mean of cooperative genes, and one for the mean of private genes. These six pairs of means were then compared using a paired t-test. As an alternative, here we combine all of these six trait types into one analysis, where we compare the mean of cooperative genes of all six trait types with the mean of private genes of all six trait types. We find that cooperative genes have significantly higher polymorphism and divergence than cooperative genes (Supplementary Table S6.1)

Supplementary Table S6.1: Statistical analysis of comparison between cooperative and private genes for five parameters. For all parameters, mean values of the variable in question are compared using Welch’s two-sample t-test, with all tests performed on log transformed variables.

| Parameter                      | t    | df | p-value   |
|--------------------------------|------|----|-----------|
| Nucleotide diversity          | 3.920| 175| 0.00013   |
| Synonymous nucleotide diversity| 3.859| 184| 0.00016   |
| Non-synonymous nucleotide diversity | 2.843| 150| 0.0051    |
| Non-synonymous divergence     | 4.353| 147| 0.000025  |
| Synonymous divergence         | 4.273| 149| 0.000034  |
Supplement S7: Analysis of whether gene length explains differences between private and cooperative genes

To see if the known effects of gene length on molecular population genetic parameters such as polymorphism and divergence can explain the differences between private and cooperative genes observed, we re-analyzed the data after removing the smallest quartile of genes (those with <188 amino acids). We find that this makes no difference to the qualitative results (Supplementary Table S7.1) with cooperative genes consistently more polymorphic and divergent than private genes.

Supplementary Table S7.1: Statistical analysis of comparison between cooperative and private genes for five parameters when the shortest quartile of genes has been removed. For all parameters, either ANOVAS with post-hoc Tukey’s test, or non-parametric Kruskal-Wallis tests are conducted.

| Parameter                                | Test                  | Test statistic | df  | Test p-value | Post-hoc test p-value |
|------------------------------------------|-----------------------|----------------|-----|--------------|-----------------------|
| Nucleotide diversity                     | ANOVA                 | \( F = 4.69 \) | 1,2105 | <0.01       | 0.006                 |
| Synonymous nucleotide diversity          | ANOVA                 | \( F = 6.63 \) | 1,2105 | <0.01       | 0.001                 |
| Non-synonymous nucleotide diversity      | Kruskal-Wallis        | \( \chi^2 = 9.80 \) | 2   | <0.01       | 0.049                 |
| Non-synonymous divergence                | Kruskal-Wallis        | \( \chi^2 = 11.49 \) | 2   | 0.03        | 0.025                 |
| Synonymous divergence                     | ANOVA                 | \( F = 3.68 \) | 1,1944 | 0.025       | 0.027                 |
Supplement S8: Comparisons to other studies

Many studies have measured population genetic parameters in *Pseudomonas aeruginosa*. These studies often use clinical isolates, frequently from a specific collection (such as a nationwide collection of clinical isolates (Muthukumarasmy 2020), or from multiple locations (Kordes 2019) or timepoints (Andersen 2018) within a single patient). The aims of these studies are very different from our aims here, but we can compare the average values of several molecular population genetic parameters.

A recent study by Muthukumarasmy *et al.* (2020) examined 99 clinical isolates from hospitals in Germany, finding that the median rate of protein evolution ($d_N/d_S$) of core genes was 0.14. Another study that used 181 mostly clinical isolates found a very similar median $d_N/d_S$ of 0.1 (Mosquera-Rendón 2016). In our study, we found a median $d_N/d_S = 0.13$, which is very similar to the results from these studies.

A study by Dotsch *et al.* (2010) found median sequence variation from a set of 36 clinical isolates was 0.47%. The equivalent measure from our set of isolates is 0.76%, which is slightly higher but still similar – and importantly much lower than estimates for many other gammaproteobacteria (Spencer 2003).

Studies have shown that environmental and clinical isolates of *P. aeruginosa* are very similar (Wolfgang 2003, Grosso-Becerra 2014), but certain genotypes are overrepresented in environmental compared to clinical isolates (Rutherford 2018), and population structure is more related to clinical vs. environmental differences than geographic separation (Ozer 2019).
Supplement S9: Genes regulated by both QS systems

In the main analysis, we use a set of 316 QS-regulated genes from Schuster et al. (2003). This gene set was calculated by comparing gene expression between the wildtype and a mutant lacking the receptors for both major QS systems (LasR & RhlR). As such, some of these genes are controlled by just one of the QS systems, and some of the genes are controlled by both. This may mean that there is substantial variation in the conditional expression of the genes in our gene set.

There is a continuum of how strongly any given gene responds to each QS signal, so we follow the system outlined in Schuster & Greenberg (2006) to categorize genes as either specific to one QS signal, or responsive to both. This gives 238 genes (76% of all QS-regulated) that are regulated by both QS systems. In this supplement, we conduct an analysis in which we only use this smaller set of QS-controlled genes.

Polymorphism

Cooperative genes are significantly more polymorphic that private genes, as measured by average pairwise nucleotide diversity $\pi$ (ANOVA $F_{2,2050}=3.51$, $p<0.05$. Tukey’s HSD $p=0.03$) (Figure S9.1). The increased polymorphism is manifest at both synonymous and non-synonymous sites (synonymous: ANOVA $F_{2,2050}=3.86$, $p<0.05$. Tukey’s HSD $p=0.02$) (non-synonymous: Kruskall-Wallis $X^2(2) = 10.4$, $p < 0.01$. Dunn Test $p=0.03$) (Figure S9.2).

**Figure S9.1**: Nucleotide diversity per site for private QS (yellow) and cooperative QS (blue) genes. Each dot represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome. Genes for cooperative traits showed significantly higher polymorphism than genes for private traits.
Figure S9.2: Nucleotide diversity per site for (A) synonymous (B) non-synonymous sites. Private QS (yellow) and Cooperative QS (blue) genes. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.

Divergence

We found that genes regulating cooperative traits had significantly higher divergence than genes regulating private traits (Figure S9.3). The difference was significant when examining both non-synonymous (Figure 3A; Kruskal-Wallis $X^2(2) = 13.8, p < 0.02$. Dunn Test p=0.016) and synonymous sites (Figure 3B; ANOVA $F_{2,1902}=4.73$, p=0.09. Tukey’s HSD p<0.01).

Figure S9.3: Divergence at synonymous (A) and non-synonymous (B) sites, measured as rates of protein evolution (e.g. non-synonymous substitutions per non-synonymous site) for Private QS (yellow) and Cooperative QS (blue) genes. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome. Genes for cooperative traits showed significantly higher divergence than genes for private traits.
**Direction of Selection**

The Direction of Selection statistic (Stoletzki & Eyre-Walker 2011) is a quantitative measure that allows us to compare substitution relative to polymorphism, allowing us to compare if a certain set of genes are experience a different form of selection from another. Positive values indicate positive selection, and negative values indicate balancing selection.

Cooperative genes do not significantly differ in form of selection from private genes (ANOVA $F_{2,1521} = 0.08$, $p = 0.92$. Tukey’s HSD $p = 0.97$) (Figure S9.4).

**Figure S9.4:** Direction of Selection (DoS) statistic for Private QS (yellow) and Cooperative QS (blue) genes. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.

**Genes specific to one QS system**

There are 29 rhl-specific genes in our dataset, and 26 las-specific genes. Within these sets, there are 3 cooperative rhl-specific genes, and 8 cooperative las-specific genes. These smaller sample sizes make statistical comparisons difficult, but we note that polymorphism is on average higher in the cooperative genes for both of these gene sets.
Supplement S10: Conditional expression

In the main text we are interested in assessing the effect of sociality on signatures of selection. In order to do this, we had to carefully control for the effect of conditional expression, as this is predicted to have a similar effect as sociality in causing selection to effectively be relaxed (van Dyken et al. 2010).

In this supplement, we conduct the opposite comparison, aiming to control for the effect of sociality and see if conditional expression predicts the degree to which selection is relaxed. We control for the effect of sociality by analyzing extracellular genes, as these are the class most likely to be social. We use genes that are QS-regulated vs. those that are not as our comparison of conditional expression vs. more continuous expression. This analysis is predicated on the assumptions that 1) QS genes are expressed at a lower rate than other genes throughout the genome, 2) extracellular genes are likely to be social.

We find that extracellular QS-regulated genes are not significantly more polymorphic or more divergent than extracellular genes that aren’t QS regulated (ANOVA $F_{1,5781}=0.185$, $p=0.831$, Tukey’s HSD $p=0.994$). Furthermore, we find that there is no difference in the ratio of non-synonymous to synonymous polymorphism (ANOVA $F_{1,5778}=7.09$, $p<0.001$, Tukey’s HSD $p=0.812$), and no difference in polymorphism if we use our set of clinical isolates instead of the environmental ones (ANOVA $F_{1,5785}=0.131$, $p=0.877$, Tukey’s HSD $p=0.986$).

It seems likely that many of the extracellular genes that are not regulated by QS also have conditional expression, such as upon encountering other strains or a host’s immune system. A full test of the conditional expression hypothesis is beyond the aims of this study, but would likely require careful selection of two sets of genes that differed in their expression level but not their sociality.
In order to see whether cooperative genes were more likely to be transferred horizontally we conducted an analysis using pangenome data from panX (23) (downloaded from pangenome.org).

Specifically, we tested whether cooperative genes are overrepresented in the accessory genome (present in <90% of strains) compared to the core genome (present in ≥90% of strains). This analysis is built on the assumption that genes in the accessory genome are more likely to have been transferred than genes in the core genome (24, 25).

13.2% of the genes in the reference strain PAO1 are part of the accessory genome, with 86.8% of genes in the core genome. 40 of our 41 cooperative genes are part of the core genome, with the only gene that isn’t present in 90% of strain but duplicated in one.

This analysis suggests that cooperative genes were not more likely to be transferred horizontally than private genes, which would have make population genetic analysis more difficult.
In our main analysis we compared levels of polymorphism between cooperative and private genes. In theory, we can use this data to estimate the relatedness $r$ in the population, with a number of caveats and assumptions.

The work of Linksvayer & Wade (2009) showed that the degree to which selection is relaxed for cooperative traits is inversely proportional to $r$. This result arises because $r$ is the relationship between the genotype of the individual conducting the social action and the genotype of the individual who receives the fitness effect. A simple population genetics model such as those presented in Linksvayer & Wade (2009) shows that the equilibrium frequency of a slightly deleterious allele with a cooperative effect will be inversely proportional to $r$. Given standard assumptions about weak selection, large population size, and ignoring higher order terms, results for equilibrium allele frequencies of deleterious alleles can be translated to predictions for relative polymorphism, leading to the prediction that relative polymorphism of social and direct effect genes will be (inversely) proportional to $r$.

There are some important caveats and assumptions of this analysis. The first is that as $r$ becomes small we don’t expect to see polymorphism continue to increase proportional to $r$, because there is a ceiling imposed by the neutral expectation. Put another way, polymorphism cannot be infinitely higher for social genes when $r$ is extremely small because the neutral expectation defines maximal polymorphism. As we are in the parameter space where $r$ is not small, that is not a problem for this analysis. A further caveat is that the theory compares two genes that differ only in mode of selection (direct vs. social), but have equal (and weak) strength of selection. Here, we are comparing two classes of gene, so we have to assume that the average strength of selection is the same between the two classes, and that the distribution of fitness effects is also similar.

With these caveats in mind, we can make an estimate of $r$ from our data. We choose levels of polymorphism, as this is the parameter that most closely aligns with the predictions of the basic model, and we focus only on the set of genes regulated by QS, as they are likely to have similar strength of selection acting on them, and are expressed at similar levels. Cooperative genes have a median polymorphism of 0.00895, and private genes have a median polymorphism of 0.00757. This leads us to estimate $r = 0.84$. We obtain a very similar estimate when using the data from clinical strains ($r = 0.85$). Furthermore, when we removed the divergent strain that had a mutation in mutL (see Supplement S13), we observed overall lower estimates of polymorphism, but we obtain the same estimate for relatedness ($r = 0.84$).
In our main analysis we measured population genetic parameters on a set of environmental isolates. One of those isolates, WH-SGI-V-07287, has an in-frame deletion in the mutL gene, a mismatch repair gene which is associated with large chromosomal deletion (Shen et al. 2018). This strain also exhibits the most polymorphism relative to the other strains in our analysis. In order to check whether our central result of higher polymorphism in cooperative genes was driven by this highly divergent strain, we conducted a robustness analysis by removing this strain and re-doing the main analysis.

We found that although overall polymorphism was lower when we removed this strain (as expected), cooperative genes still have significantly greater levels of polymorphism than private genes (ANOVA $F_{1,2107}=4.14$, $p=0.02$; Tukey’s HSD $p=0.01$ Supplementary Figure S13.1), and the greater levels of polymorphism are manifest at both synonymous and non-synonymous sites (Synonymous: ANOVA $F_{1,2107}=5.28$, $p<0.01$; Tukey’s HSD $p<0.01$ Supplementary Figure S13.2. Non-synonymous: Kruskal-Wallis $X^2(2) = 7.72, p = 0.02$; Dunn Test $p=0.04$ Supplementary Figure S13.3). These results indicate that our results weren’t driven by one strain, which makes logical sense as it seems unlikely that a mutator element would lead to elevated variation in only cooperative genes but not private genes.

**Supplementary Figure S13.1**: Nucleotide diversity per site for private QS (yellow) and cooperative QS (blue) genes. Each dot represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.
Supplementary Figure S13.2: Synonymous nucleotide diversity per site for private QS (yellow) and cooperative QS (blue) genes. Each dot represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.

Supplementary Figure S13.3: Non-synonymous nucleotide diversity per site for private QS (yellow) and cooperative QS (blue) genes. Each dot represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.
Supplement S14: Population structure

It is plausible that the environmental isolates we are using in this study come from isolated subpopulations. In this case, there could be different mutations segregating in each subpopulation, which could alter the ability of tests such as Tajima’s D to detect balancing selection.

Whilst balancing selection acting in subpopulations is plausible, we would expect to see 1) a signature of balancing selection in Tajima’s D, and 2) low divergence. We don’t find either of these two predictions in any of the traits or sets of isolates analysed. Nonetheless, we present here a brief analysis on population structure in the environmental isolates we analysed.

When using principal component analysis and FST to investigate population structure in our isolates, we find that the first two components separate out 3 isolates from the remaining 38. The third component splits the 41 isolates into a group of 15 and a group of 26. The FST value based on this grouping is 0.036, which implies little genetic differentiation. (Hartl & Clark 1997) Because of this, alongside our sampling procedure, we feel that we don’t have strong population structure in our dataset.
S15: Supplementary Methods

SNP Calling

We downloaded raw sequencing reads for each of the 41 strains plus the outgroup PA7 (SRA: SRR9418201) from the European Bioinformatics Institute's European Nucleotide archive (www.ebi.ac.uk/ena) – see Supplementary Table 6 for the relevant ID of each sequencing run. We trimmed reads for each strain to remove adapters and low quality reads using Trimmomatic (26). We removed leading and trailing reads with a quality score <3, and also removed reads if average quality in a four base sliding window was below 20. The resulting reads were quality-control checked using FastQC (27).

Next, we mapped reads for each strain, and aligned to the reference strain PAO1 (Accession: SAMN02603714) using BWA (28). We sorted and converted the resulting SAM files to BAM files using SAMtools (29). We then removed PCR duplicates using Picard tools (30).

We called variants on all strains using BCFtools (31), and converted to a VCF file for analysis. Next, we filtered variants to removed INDELs, and further quality filtering conducted using the default settings of the vcfutils python script in SAMtools (29) to filter for minimum mapping quality (=10), minimum read depth (=2), and minimum p-value for strand bias (=0.001).

We used the featureCounts tool in Subread (32) to assess coverage of each gene in each strain, removing any strains with <2 reads in >50% of genes (which in this case was no strains). We used the coverageBed tool in BEDtools (33) to analyse what proportion of each gene’s length had been mapped – so that we could adjust per-site population genetic measures to the mapped length of a gene, rather than the length of the gene in the reference genome.

We removed any site in the genome which hadn’t been called in >80% of strains. This meant that each site had a call in at least 33 strains. We conducted a brief power analysis by removing 8 strains from the VCF file to ensure that downstream population genetic measures would not substantially altered by this lowering of sample size. After filtering, we had a VCF file with a total of 391,770 SNPs among the 41 environmental strains (not including the outgroup).
We conducted the majority of the molecular population genetics analysis using the PopGenome package (version 2.7.5) in R (34). Specifically, we calculated the parameters nucleotide diversity $\pi$, Tajima’s $D$, Fu and Li’s $F^*$, Fu and Li’s $D^*$, Mcdonald Kreitman test, Neutrality index, alpha, and the Direction of Selection statistic using the PopGenome package. For statistic where an outgroup is necessary, we used PA7, a known taxonomic outlier commonly used (35). Where necessary, to obtain per-site measures, we calculated parameters separately for synonymous and non-synonymous sites and scaled to the relevant mapped length. Genes with mapped length $<50\%$ were removed from the analysis at this stage – leaving a final set of 5234 genes.

We calculated rates of protein evolution ($k_a/k_s$) by comparison of the reference strain PAO1 to a known taxonomic outlier, PA7 (35). Next, we extracted SNPs for PA7 from the VCF file, and inserted them into the sequence of PAO1 using the ‘FastaAlternateReferenceMaker’ tool in the GATK suite (36). We compared this pseudo-genome sequence to the sequence of PAO1 using the seqinR package in R (11) to determine $k_a$, $k_s$, and $k_a/k_s$ for each gene. Genes which weren’t aligned between the two strains return were removed from this analysis.

For some tests, we conducted further analysis by analyzing whether cooperative genes were overrepresented in the subset of genes which had a statistically significant result for a given parameter. Some tests such as Mcdonald-Kreitman are designed to test the null hypothesis for an individual gene. We used various measures that use the same information as the MKT to allow comparisons across genes (e.g. neutrality index, alpha, direction of selection statistic), and we also extracted the set of genes for which the test is significant (meaning an excess of either non-synonymous substitutions or nonsynonymous polymorphism). For statistics that use data on the site frequency distribution (Tajima’s $D$, Fu & Li’s $D^*$, Fu & Li’s $F^*$), we also extracted the genes with a significant value. For Tajima’s $D$ this was conducted using the beta distribution test (37) conducted in the R package Pegas (38). For Fu & Li’s $D^*/F^*$ statistics we used the critical value’s from the original paper (39) for $n=100$ genes to test significance at the $\alpha=0.025$ level. Although we have many more genes than 100, the critical value for these tests will be proportional to $\ln(n)$ so this is a reasonable approximation. After extracting the subset
of genes which are significant for a given test, we test for whether cooperative genes (see below) are over- or under-represented in this class using a binomial test.

One signature of relaxed selection on sociality genes is an increase in deleterious mutations, such as those which have large disruptive effects on the function of a gene. For this analysis, we annotated variants with SNPeff (40) and counted mutations that generate premature stop codons. We included INDELs at this stage so that we could also count frameshift mutations.

To test whether cooperative genes are over- or under-represented in a set of genes, it is necessary to use a proxy for cooperative genes, because our designation of cooperative genes is not a systematic genome-wide assignment and so we cannot confidently say if any number is an overrepresentation since we don’t know how many ‘cooperative’ genes there are in the genome. We used extracellular proteins as a proxy for cooperative genes, which has been used several times before (41, 42) and can be systematically calculated for a whole genome using PSORTb (43). Although it is evident that not all cooperative genes are extracellular and not all extracellular proteins are cooperative, if there was a strong signature of sociality captured by measures such as Tajima’s D, we would expect to see an effect with this proxy.
**Supplementary Figures**

**Supplementary Figure 1**: Nucleotide diversity per site for (A) synonymous (B) non-synonymous sites. Private QS (yellow) and Cooperative QS (blue) genes. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.

**Supplementary Figure 2**: Ratio between non-synonymous and synonymous nucleotide diversity per site for Private QS (yellow) and Cooperative QS (blue) genes. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.
**Supplementary Figure 3:** Ratio between divergence at non-synonymous and synonymous sites, measured as rates of protein evolution for Private QS (yellow) and Cooperative QS (blue) genes. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.

**Supplementary Figure 4:** Nucleotide diversity per site for functional categories of QS controlled genes, as determined by eggNOG. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.
**Supplementary Figure 5**: Divergence at non-synonymous (A) and synonymous (B) sites, measured as rates of protein evolution (e.g. non-synonymous substitutions per non-synonymous site) for functional categories of QS controlled genes, as determined by eggNOG. Each point represents a gene, and the horizontal line represents the median for each group. The grey dotted line represents the median for private genes across the genome.

**Supplementary Figure 6**: Cooperative vs. private comparison for six trait types for synonymous polymorphism (nucleotide diversity)
Supplementary Figure 7: Cooperative vs. private comparison for six trait types for non-synonymous polymorphism (nucleotide diversity)

Supplementary Figure 8: Cooperative vs. private comparison for six trait types for synonymous divergence (synonymous substitutions per non-synonymous site)
## Supplementary Table 1: Cooperative QS induced genes

| Gene ID (PA) | Name | Function | PSORTb | Reference |
|--------------|------|----------|--------|-----------|
| 0122         | rahU | Modulates innate immunity and inflammation in host cells | Extracellular | (44) |
| 0355         | pflp | Protease | Cytoplasmic | (45) |
| 0572         | impA | Immunomodulating metalloprotease | Outer Membrane | (46) |
| 0852         | cpbD | Chitin-binding protein | Extracellular | (47) |
| 0996         | pqsA | Quinolone signal system (PQS) synthesis | Cytoplasmic Membrane | (18) |
| 0997         | pqsB | Quinolone signal system (PQS) synthesis | Cytoplasmic | (18) |
| 0998         | pqsC | Quinolone signal system (PQS) synthesis | Cytoplasmic | (18) |
| 0999         | pqsD | Quinolone signal system (PQS) synthesis | Cytoplasmic | (18) |
| 1000         | psqE | Thioesterase – alternative ligand to RhlI | Cytoplasmic | (48) |
| 1130         | rhlC | Enzyme involved in production of rhamnolipid public good | Unknown | (49) |
| 1246         | aprD | Alkaline protease biosynthesis | Cytoplasmic Membrane | (50) |
| 1247         | aprE | Alkaline protease biosynthesis | Cytoplasmic Membrane | (50) |
| 1248         | aprF | Alkaline protease biosynthesis | Outer Membrane | (50) |
| 1249         | aprA | Alkaline protease | Extracellular | (50) |
| 1432         | lasI | Autoinducer synthesis protein | Cytoplasmic | (18) |
| 1869         | acp1 | Fatty acid synthesis | Cytoplasmic | (51) |
| 1871         | lasA | Protease, enhances activity of lasB | Extracellular | (52) |
| 2193         | hcnA | Hydrogen cyanide | Unknown | (53) |
| 2194         | hcnB | Hydrogen cyanide | Cytoplasmic | (53) |
| 2195         | hcnC | Hydrogen cyanide | Cytoplasmic | (53) |
| 2300         | chiC | Chitinase | Extracellular | (54) |
| 2302         | ambD | Toxin biosynthesis | Cytoplasmic | (55) |
| 2303         | ambC | Toxin biosynthesis | Cytoplasmic | (55) |
| 2304         | ambB | Toxin biosynthesis | Cytoplasmic | (55) |
| 2305         | ambA | Toxin biosynthesis | Unknown | (55) |
| 2570         | lecA | Lectin – suppress host immunity | Cytoplasmic | (56) |
| 2587         | pqsH | Quinolone signal system (PQS) synthesis | Cytoplasmic | (18) |
| 2939         | pepB | Aminopeptidase PaAP essential for cooperative proteolytic growth | Extracellular | (57) |
| 3226         | clpP2 | Peptidase required to degrade MucA to produce alginate | Cytoplasmic | (58) |
| 3361         | lecB | Lectin – suppress host immunity | Unknown | (56) |
| 3476         | rhlI | Autoinducer of signal-receptor pathway | Unknown | (48) |
| 3477         | rhlR | Quorum-sensing receptor | Cytoplasmic | (48) |
| 3478         | rhlB | Involved in production of rhamnolipid | Cytoplasmic | (59) |
| 3479         | rhlA | Rhamnolipid biosynthesis | Cytoplasmic | (18) |
| 3535         | eprS | Serine protease – induces host inflammatory responses | Outer Membrane | (60) |
| 3724         | lasB | Extracellular protease elastase | Extracellular | (18) |
| 4175         | piv | Protease that disrupts innate mucosal defences | Extracellular | (61) |
| 4190         | pqsL | Quinolone signal system (PQS) synthesis | Cytoplasmic | (18) |
| 5161         | rmlB | Polysaccharide biosynthesis | Cytoplasmic | (62) |
| 5162         | rmlD | Polysaccharide biosynthesis | Cytoplasmic | (62) |
| 5164         | rmlC | Polysaccharide biosynthesis | Cytoplasmic | (62) |
## Supplementary Table 2: Cooperative and private genes in pyoverdine and pyochelin pathways (63, 64)

| Gene ID (PA) | Name | Function | Sociality |
|-------------|------|----------|-----------|
| 2385        | *pvdQ* | Periplasmic Biosynthesis | Cooperative |
| 2386        | *pvdA* | Cytoplasmic Biosynthesis | Cooperative |
| 2387        | *fpvI* | Receptor Regulation | Private |
| 2388        | *fpvR* | Uptake and disassociation | Private |
| 2389        | *pvdR* | Secretion | Cooperative |
| 2390        | *pvdT* | Secretion | Cooperative |
| 2391        | *opmQ* | Secretion | Cooperative |
| 2392        | *pvdP* | Periplasmic Biosynthesis | Cooperative |
| 2393        | *pvdM* | Periplasmic Biosynthesis | Cooperative |
| 2394        | *pvdN* | Periplasmic Biosynthesis | Cooperative |
| 2395        | *pvdO* | Periplasmic Biosynthesis | Cooperative |
| 2396        | *pvdF* | Cytoplasmic Biosynthesis | Cooperative |
| 2397        | *pvdE* | Export into periplasm | Cooperative |
| 2398        | *fpvA* | Receptor | Private |
| 2399        | *pvdD* | Cytoplasmic Biosynthesis | Cooperative |
| 2400        | *pvdI* | Cytoplasmic Biosynthesis | Cooperative |
| 2402        | *fpvG* | Uptake and disassociation | Private |
| 2404        | *fpvH* | Uptake and disassociation | Private |
| 2405        | *fpvJ* | Uptake and disassociation | Private |
| 2406        | *fpvK* | Uptake and disassociation | Private |
| 2407        | *fpvC* | Uptake and disassociation | Private |
| 2408        | *fpvD* | Uptake and disassociation | Private |
| 2409        | *fpvE* | Uptake and disassociation | Private |
| 2410        | *fpvF* | Uptake and disassociation | Private |
| 2412        | *-* | Cytoplasmic Biosynthesis | Cooperative |
| 2413        | *pvdH* | Cytoplasmic Biosynthesis | Cooperative |
| 2424        | *pvdL* | Cytoplasmic Biosynthesis | Cooperative |
| 2425        | *pvdG* | Cytoplasmic Biosynthesis | Cooperative |
| 2426        | *pvdS* | Regulation | Cooperative |
| 2428        | *pvdX* | Cytoplasmic Biosynthesis | Cooperative |
| 2531        | *ptaA* | Periplasmic Biosynthesis | Cooperative |
| 5531        | *tonB1* | Uptake and disassociation | Private |
| 4168        | *fpvB* | Uptake and disassociation | Private |

### PYOCHELIN

| Gene ID (PA) | Name | Function | Sociality |
|-------------|------|----------|-----------|
| 3753        | *fpB* | Transport | Private |
| 4218        | *fpx* | Transporter | Private |
| 4220        | *fptb* | Receptor | Private |
| 4221        | *fptA* | Receptor | Private |
| 4222        | *pchI* | Export | Cooperative |
| 4223        | *pchH* | Export | Cooperative |
| 4224        | *pchG* | Biosynthesis | Cooperative |
| 4225        | *pchF* | Biosynthesis | Cooperative |
| 4226        | *pchE* | Biosynthesis | Cooperative |
| 4227        | *pchR* | Regulation | Cooperative |
| 4228        | *pchD* | Biosynthesis | Cooperative |
| 4229        | *pchC* | Biosynthesis | Cooperative |
**Supplementary Table 3:** Cooperative and private genes involved in AMR (65–68)

| Gene ID (PA) | Name                  | Function                          | Sociality  |
|--------------|-----------------------|-----------------------------------|------------|
| 4109         | ampR                  | Beta-lactamase                    | Cooperative|
| 4110         | ampC                  | Beta-lactamase                    | Cooperative|
| 4522         | ampD                  | Beta-lactamase                    | Cooperative|
| 2315         | -                     | Putative Beta-lactamase           | Cooperative|
| 5542         | PIB-1                 | Beta-lactamase                    | Cooperative|
| 5514         | blaOXA-50             | Putative Beta-lactamase           | Cooperative|
| 4119         | aph                   | aminoglycoside phosphotransferase | Cooperative|

**Efflux Pumps**

| Gene ID (PA) | Name                  | Function |
|--------------|-----------------------|----------|
| 0424-0427    | MexR/A/B-OprM         | Efflux pump | Private |
| 1435-1436    | MexM/N                | Efflux pump | Private |
| 2018, 2020   | MexY/Z                | Efflux pump | Private |
| 2491-2495    | MexS/T/E/F-OprN       | Efflux pump | Private |
| 3521-3523    | MexQ/P-OpmE           | Efflux pump | Private |
| 3676-3678    | MexK/J/L              | Efflux pump | Private |
| 4205-4208    | MexG/H/I-OpmD         | Efflux pump | Private |
| 4374-4375    | MexV/W                | Efflux pump | Private |
| 4597-4599    | MexD/C-OprJ           | Efflux pump | Private |
| 2525-2528    | MuxC/B/A-OpmB         | Efflux pump | Private |
| 4974         | OpmH                  | Efflux pump | Private |

**Porins**

| Gene ID (PA) | Name | Function |
|--------------|------|----------|
| 1777         | oprF | porin    | Private |
| 3186         | oprB | porin    | Private |
| 2291         | oprB2| porin    | Private |
| 4067         | oprG | porin    | Private |
| 1178         | oprH | porin    | Private |
| 3279         | oprP | porin    | Private |
| 3280         | oprO | porin    | Private |
| 0958         | oprD | porin    | Private |
| 0162         | oprC | porin    | Private |
| 4503         | oprP | porin    | Private |
| 2505         | oprT | porin    | Private |
| 0189         | oprI | porin    | Private |
| 2760         | oprQ | porin    | Private |
| 2700         | oprB | porin    | Private |
| 2420         | oprJ | porin    | Private |
| 4898         | oprK | porin    | Private |
| 0240         | oprF | porin    | Private |
| 2113         | oprO | porin    | Private |
| 4137         | oprL | porin    | Private |
| 0755         | oprH | porin    | Private |
| 3038         | oprQ | porin    | Private |
| 1025         | oprD | porin    | Private |
| 0291         | oprE | porin    | Private |
| 2213         | oprG | porin    | Private |
| 4179         | oprN | porin    | Private |
### Supplementary Table 4: Cooperative and private genes involved in competition (69–71)

| Gene ID (PA) | Name | Function | Sociality |
|--------------|------|----------|-----------|
| **R & F Pyocins** |
| 0609         | trpE | Pyocin   | Cooperative |
| 0610         | prtN | Pyocin   | Cooperative |
| 0611         | prtR | Pyocin   | Cooperative |
| 0649         | trpG | Pyocin   | Cooperative |
| 0616-0628    | -    | Pyocin   | Cooperative |
| 0633-0648    | -    | Pyocin   | Cooperative |
| **Type VI Secretion system** |
| 0070-0095    | HSI-I| T6SS     | Private   |
| 1654-1671    | HIS-II| T6SS     | Private   |
| 2359-2373    | HIS-III| T6SS    | Private   |
| 2685         | vgrG4| T6SS     | Private   |
| 0262         | vgrG2b| T6SS     | Private   |
| 3486         | vgrG4b| T6SS     | Private   |
| 5090         | vgrG5| T6SS     | Private   |
| 1511         | vgrG2a| T6SS     | Private   |
| 3294         | vgrG4a| T6SS     | Private   |
| 5266         | vgrG6| T6SS     | Private   |

### Supplementary Table 5: Cooperative and private genes involved in sticking and stickiness (72–75)

| Gene ID (PA) | Name | Function | Sociality |
|--------------|------|----------|-----------|
| **Extracellular Polysaccharides & Rhamnolipids** |
| 2231         | pslA | Extracellular Polysaccharide | Cooperative |
| 2245         | pslO | Extracellular Polysaccharide | Cooperative |
| 3058         | pelG | Extracellular Polysaccharide | Cooperative |
| 3064         | pelA | Extracellular Polysaccharide | Cooperative |
| 3478         | rhlB | Rhamnolipid | Cooperative |
| 3479         | rhlA | Rhamnolipid | Cooperative |
| 1130         | rhlC | Rhamnolipid | Cooperative |
| **Type IV Pili & Flagella** |
| 395-396      | pII/T/U | Flagella | Private   |
| 408-412      | pII/G/H/I | Flagella | Private   |
| 2960         | pIIZ | Flagella | Private   |
| 3805         | pIIF | Flagella | Private   |
| 4525-4526    | pIIA/B | Flagella | Private   |
| 4528         | pIID | Flagella | Private   |
| 4549-4550    | fimII/U | Flagella | Private   |
| 4551-4556    | pIIV/W/X/Y1/Y2/E | Flagella | Private   |
| 5040-5044    | pIIQ/P/O/N/M | Flagella | Private   |
| 1452-1453    | flhA/F | Flagella | Private   |
| 1454         | fleN | Flagella | Private   |
| Strain Name | Accession | Location | Source | SRA     |
|-------------|-----------|----------|--------|---------|
| M8A.2       | SAMN02360740 | Colombia | environmental isolate | SRR1014512 |
| WW          | SAMN12784948 | South Africa | water | SRR10257272 |
| LV          | SAMN08930812 | Brazil | orange fruit with canker citrus | SRR13065837 |
| DR1         | SAMN03160602 | India | rice rhizosphere soil | SRR1639604 |
| WH-SGI-V-07166 | SAMN04128504 | Switzerland | Environment water | SRR2939265 |
| WH-SGI-V-07284 | SAMN04128727 | Romania | Water | SRR2939488 |
| WH-SGI-V-07286 | SAMN04128728 | Puerto Rico | Well water | SRR2939489 |
| WH-SGI-V-07287 | SAMN04128729 | Puerto Rico | Tree | SRR2939490 |
| WH-SGI-V-07297 | SAMN04128736 | Pakistan | River | SRR2939497 |
| WH-SGI-V-07305 | SAMN04128739 | Germany | Drinking water | SRR2939500 |
| Hex1T       | SAMN03735017 | Argentina | contaminated soil | SRR2993251 |
| N002        | SAMN00996515 | India | crude oil-polluted soil | SRR502988 |
| 4014        | SAMN06718840 | France | agricultural soil | SRR5513015 |
| 11987-2-5   | SAMN09270742 | UK | Sea water | SRR7230067 |
| GOM1        | SAMN10839112 | Mexico | sea water | SRR8513685 |
| JYH21       | SAMN10478512 | Japan | River | SRR8612775 |
| AUS176      | SAMN10478443 | Australia | Pool | SRR8612785 |
| JYH13       | SAMN10478508 | Japan | River | SRR8612821 |

**Supplementary Table 6:** Strains used in whole-genome analysis
| Code    | Accession   | Country     | Location     | Description   | SRR     |
|---------|-------------|-------------|--------------|---------------|---------|
| W5Aug28 | SAMN10478487| Belgium     | River        |               | SRR8612823|
| PT6     | SAMN10478536| Germany     | River/sewer  |               | SRR8612836|
| PT12    | SAMN10478537| Germany     | River/sewer  |               | SRR8612837|
| AUS277  | SAMN10478423| Australia   | Tank water   |               | SRR8612849|
| JYH17   | SAMN10478510| Japan       | River        |               | SRR8612867|
| JYH25   | SAMN10478514| Japan       | River        |               | SRR8612874|
| AUS500  | SAMN10478447| Australia   | River        |               | SRR8612885|
| JYH7    | SAMN10478505| Japan       | River        |               | SRR8612887|
| W15Apr4 | SAMN10478482| Belgium     | River        |               | SRR8612888|
| AUS258  | SAMN10478418| Australia   | Pool         |               | SRR8612901|
| Jp100   | SAMN10478473| Japan       | Sea water (open ocean) |     | SRR8612921|
| Jp241   | SAMN10478476| Japan       | Sea water (coastal) |           | SRR8612923|
| Jp222   | SAMN10478474| Japan       | Sea water (open ocean) |       | SRR8612925|
| m183    | SAMN08014105| India       | stem         |               | SRR8835649|
| 62      | SAMN02360656| USA         | soil         |               | SRR1014018|
| C20     | SAMN02360744| Unknown     | environmental isolate |      | SRX366136|
| C52     | SAMN02360750| Unknown     | environmental isolate |     | SRR1014414|
| E2      | SAMN02360657| USA         | tomato       |               | SRR1014494|
| M9A.1   | SAMN02360743| Colombia    | environmental isolate |     | SRR1014533|
| MSH10   | SAMN02360659| USA         | water        |               | SRR1014539|
| MSH3    | SAMN02360658| USA         | water        |               | SRR1014548|
| PS42    | SAMN02360753| Venezuela   | environmental isolate |     | SRR1019995|
| PS50    | SAMN02360754| Venezuela   | environmental isolate |     | SRR1019997|
1. F. Hildebrand, A. Meyer, A. Eyre-Walker, Evidence of selection upon genomic GC-content in bacteria. *PLoS Genet.* **6** (2010).

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

3. P. M. Sharp, E. Bailes, R. J. Grocock, J. F. Peden, R. E. Sackett, Variation in the strength of selected codon usage bias among bacteria. *Nucleic Acids Res.* **33**, 1141–1153 (2005).

4. H. Akashi, T. Gojobori, Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 3695–3700 (2002).

5. R. Raghavan, Y. D. Kelkar, H. Ochman, A selective force favoring increased G+C content in bacterial genes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 14504–14507 (2012).

6. E. P. C. Rocha, Codon usage bias from tRNA’s point of view: Redundancy, specialization, and efficient decoding for translation optimization. *Genome Res.* **14**, 2279–2286 (2004).

7. P. M. Sharp, W. H. Li, The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **15**, 1281–1295 (1987).

8. T. Ikemura, Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. *J. Mol. Biol.* **151**, 389–409 (1981).

9. S. E. H. West, B. H. Iglewski, Codon usage in *Pseudomonas aeruginosa*. *Nucleic Acids Res.* **16**, 9323–9335 (1988).

10. J. R. Dettman, J. L. Sztepanacz, R. Kassen, The properties of spontaneous mutations in the opportunistic pathogen *Pseudomonas aeruginosa*. *BMC Genomics* **17**, 1–14 (2016).

11. D. Charif, J. R. Lobry, “SeqinR 1.0-2: A Contributed Package to the R Project for Statistical Computing Devoted to Biological Sequences Retrieval and Analysis” in (2007) https://doi.org/10.1007/978-3-540-35306-5_10.

12. B. E. Poulsen, *et al.*, Defining the core essential genome of *Pseudomonas aeruginosa.* *Proc. Natl. Acad. Sci. U. S. A.* **116**, 10072–10080 (2019).

13. A. Dötsch, *et al.*, Evolutionary conservation of essential and highly expressed genes in *Pseudomonas aeruginosa*. *BMC Genomics* **11** (2010).

14. C. Kiewitz, B. Tümmler, Sequence Diversity of *Pseudomonas aeruginosa*: Impact on Population Structure and Genome Evolution. *J. Bacteriol.* **182**, 3125–3135 (2000).

15. E. P. C. Rocha, A. Danchin, An Analysis of Determinants of Amino Acids Substitution Rates in Bacterial Proteins. *Mol. Biol. Evol.* **21**, 108–116 (2004).

16. S. Vieira-Silva, E. P. C. Rocha, The systemic imprint of growth and its uses in ecological (meta)genomics. *PLoS Genet.* **6** (2010).

17. S. K. Behura, D. W. Severson, Comparative analysis of Codon usage bias and Codon context patterns between dipteran and hymenopteran sequenced genomes. *PLoS One* **7** (2012).

18. R. Chen, E. Déziel, M. C. Groletau, A. L. Schaefer, E. P. Greenberg, Social cheating in a *Pseudomonas aeruginosa* quorum-sensing variant. *Proc. Natl. Acad. Sci. U. S. A.* (2019) https://doi.org/10.1073/pnas.1819801116.

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

20. R. Nielsen, Molecular signatures of natural selection. *Annu. Rev. Genet.* **39**, 197–218 (2005).
21. G. A. Bazykin, A. S. Kondrashov, Major role of positive selection in the evolution of conservative segments of Drosophila proteins. *Proc. R. Soc. B Biol. Sci.* **279**, 3409–3417 (2012).

22. C. Peng, F. Gao, Protein localization analysis of essential genes in prokaryotes. *Sci. Rep.* (2014) https://doi.org/10.1038/srep06001.

23. W. Ding, F. Baumdicker, R. A. Neher, panX: pan-genome analysis and exploration. *Nucl. Acids Res.* **46**, e5 (2018).

24. T. J. Treangen, E. P. C. Rocha, Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS Genet.* **7** (2011).

25. J. O. McInerney, A. McNally, M. J. O’Connell, Why prokaryotes have pangenomes. *Nat. Microbiol.* **2** (2017).

26. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* (2014) https://doi.org/10.1093/bioinformatics/btu170.

27. S. Andrews, FastQC. Babraham Bioinforma. (2010).

28. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* (2009) https://doi.org/10.1093/bioinformatics/btp324.

29. H. Li, et al., The Sequence Alignment/Map format and SAMtools. *Bioinformatics* (2009) https://doi.org/10.1093/bioinformatics/btp352.

30. Broad Institute, Picard toolkit. *Broad Institute, GitHub Repos*. (2019).

31. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* (2011) https://doi.org/10.1093/bioinformatics/btr509.

32. Y. Liao, G. K. Smyth, W. Shi, FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* (2014) https://doi.org/10.1093/bioinformatics/btt656.

33. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* (2010) https://doi.org/10.1093/bioinformatics/btq033.

34. B. Pfeifer, U. Wittelsbürger, S. E. Ramos-Onsins, M. J. Lercher, PopGenome: An efficient swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* **31**, 1929–1936 (2014).

35. P. H. Roy, et al., Complete genome sequence of the multiresistant taxonomic outlier Pseudomonas aeruginosa PA7. *PLoS One* **5**, 1–10 (2010).

36. A. McKenna, et al., The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* (2010) https://doi.org/10.1101/gr.107524.110.

37. F. Tajima, Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* (1989).

38. E. Paradis, Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* **26**, 419–420 (2010).

39. Y. X. Fu, W. H. Li, Statistical tests of neutrality of mutations. *Genetics* (1993) https://doi.org/10.1093/genetics/133.3.693.

40. P. Cingolani, et al., A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)*. (2012) https://doi.org/10.4161/fly.19695.

41. M. Garcia-Garcera, E. P. C. Rocha, Community diversity and habitat structure shape the repertoire of extracellular proteins in bacteria. *Nat. Commun.* **11**, 1–11 (2020).

42. T. Nogueira, M. Touchon, E. P. C. Rocha, Rapid Evolution of the Sequences and Gene Repertoires of Secreted Proteins in Bacteria. *PLoS One* **7**, 1–10 (2012).

43. N. Y. Yu, et al., PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes.
1053 *Bioinformatics* **26**, 1608–1615 (2010).

1054 44. J. Rao, et al., RahU: An inducible and functionally pleiotropic protein in *Pseudomonas aeruginosa* modulates innate immunity and inflammation in host cells. *Cell. Immunol.* **270**, 103–113 (2011).

1055 45. L. Fernández, E. B. M. Breidenstein, D. Song, R. E. W. Hancock, Role of intracellular proteases in the antibiotic resistance, motility, and biofilm formation of *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **56**, 1128–1132 (2012).

1056 46. B. W. Bardoel, et al., Identification of an immunomodulating metalloprotease of *Pseudomonas aeruginosa* Carry Multiple Post-Translational Modifications on Their Lysine Residues. *J. Proteome Res.* **18**, 923–933 (2019).

1057 47. S. Mukherjee, et al., The PqsE and RhlR proteins are an autoinducer synthase–receptor pair that control virulence and biofilm development in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E9411–E9418 (2018).

1058 48. K. E. Boyle, et al., Metabolism and the evolution of social behavior. *Mol. Biol. Evol.* **34**, 2367–2379 (2017).

1059 49. B. W. Bardoel, K. P. M. Van Kessel, J. A. G. Van Strijp, F. J. Milder, Inhibition of *pseudomonas aeruginosa* virulence: Characterization of the AprA-AprI interface and species selectivity. *J. Mol. Biol.* **415**, 573–583 (2012).

1060 50. J. C. Ma, Y. Q. Wu, D. Cao, W. Bin Zhang, H. H. Wang, Only Acyl carrier protein 1 (AcpP1) functions in *Pseudomonas aeruginosa* fatty acid synthesis. *Front. Microbiol.* **8**, 1–14 (2017).

1061 51. B. A. Cowell, S. S. Twining, J. A. Hobden, M. S. F. Kwong, S. M. J. Fleischig, Mutation of lasA and lasB reduces *Pseudomonas aeruginosa* invasion of epithelial cells. *Microbiology* **149**, 2291–2299 (2003).

1062 52. M. Wang, A. L. Schaefer, A. A. Dandekar, E. P. Greenberg, Quorum sensing and policing of *Pseudomonas aeruginosa* social cheaters. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 2187–2191 (2015).

1063 53. J. Folders, et al., Characterization of *pseudomonas aeruginosa* chitinase, a gradually secreted protein. *J. Bacteriol.* **183**, 7044–7052 (2001).

1064 54. N. R. Murcia, et al., The *Pseudomonas aeruginosa* antimetabolite L-2-amino-4-methoxy-trans-3-butenoi acid (AMB) is made from glutamate and two alanine residues via a thiotemplate-linked tripeptide precursor. *Front. Microbiol.* **6**, 1–13 (2015).

1065 55. D. Passos da Silva, et al., The *Pseudomonas aeruginosa* lectin LecB binds to the exopolysaccharide Psl and stabilizes the biofilm matrix. *Nat. Commun.* **10**, 1–11 (2019).

1066 56. T. Robinson, P. Smith, E. R. Alberts, M. Colussi-pelaez, M. Schuster, crosssm Aminopeptidase in the *Pseudomonas aeruginosa* RpoS Response. *Biocatalysis & Biotransformation* **11**, 1–20 (2020).

1067 57. D. Qiu, V. M. Eisinger, N. E. Head, G. B. Pier, H. D. Yu, ClpXP proteases positively regulate Alginate overexpression and mucoid conversion in *Pseudomonas aeruginosa*. *Microbiology* **154**, 2119–2130 (2008).

1068 58. A. Bazire, A. Dufour, The *Pseudomonas aeruginosa* rhlG and rhlAB genes are inversely regulated and RhlG is not required for rhamnolipid synthesis. *BMC Microbiol.* **14**, 1–9 (2014).

1069 59. Y. Kida, J. Taira, K. Kuwano, EprS, an autotransporter serine protease, plays an important role in various pathogenic phenotypes of *Pseudomonas aeruginosa*. *Microbiol. (United Kingdom)* **162**, 318–329 (2016).

1070 60. J. L. Bradshaw, et al., *Pseudomonas aeruginosa* Protease IV Exacerbatates
Pneumococcal Pneumonia and Systemic Disease. *mSphere* 3, 1–10 (2018).

62. M. S. Byrd, *et al.*, Genetic and biochemical analyses of the Pseudomonas aeruginosa Psl exopolysaccharide reveal overlapping roles for polysaccharide synthesis enzymes in Psl and LPS production. *Mol. Microbiol.* (2009) https://doi.org/10.1111/j.1365-2958.2009.06795.x.

63. L. Michel, A. Bachelard, C. Reimmann, Ferricyochelin uptake genes are involved in pyochelin-mediated signalling in Pseudomonas aeruginosa. *Microbiology* 153, 1508–1518 (2007).

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

65. E. Amanatidou, *et al.*, Biofilms facilitate cheating and social exploitation of β-lactam resistance in Escherichia coli. *npj Biofilms Microbiomes* 5, 1–10 (2019).

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

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

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

69. M. Ghoul, *et al.*, Bacteriocin-mediated competition in cystic fibrosis lung infections. *Proc. R. Soc. B Biol. Sci.* 282 (2015).

70. J. D. Mougous, *et al.*, A virulence locus of Pseudomonas aeruginosa encodes a protein secretion apparatus. *Science (80-. ).* 312, 1526–1530 (2006).

71. R. D. Hood, *et al.*, a Toxin to Bacteria. *Cell* 7, 25–37 (2011).

72. Y. Irie, *et al.*, The Pseudomonas aeruginosa PSL polysaccharide is a social but noncheatable trait in biofilms. *MBio* (2017) https://doi.org/10.1128/mBio.00374-17.

73. J. B. Xavier, W. Kim, K. R. Foster, A molecular mechanism that stabilizes cooperative secretions in Pseudomonas aeruginosa. *Mol. Microbiol.* 79, 166–179 (2011).

74. N. Dasgupta, *et al.*, A four-tiered transcriptional regulatory circuit controls flagellar biogenesis in Pseudomonas aeruginosa. *Mol. Microbiol.* (2003) https://doi.org/10.1046/j.1365-2958.2003.03740.x.

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