Natural polymorphisms in *Pseudomonas aeruginosa* populations reveal a dual role for DNA gyrase in fluoroquinolone resistance and persistence

Supplementary text

S1 On the panel of *Pseudomonas aeruginosa* isolates

The collection of *P. aeruginosa* isolates consist of a wide array of natural isolates. For statistical analyses, two groups were considered based on their origin: clinical strains and environmental strains. Environmental isolates were obtained from water sources, such as oceans and rivers, and plants. Clinical isolates include isolates from humans but also from different animal species. The human isolates can be further divided in strains isolated either from patients suffering from cystic fibrosis (CF) or from non-CF patients. While strains were isolated from diverse geographic origins, statistical analyses based on the geographic origin of the isolates were not included in this study since isolates from Europa (67.2%) dominate the strain collection and isolates from Africa (1.9%) and South America (1.4%) are underrepresented. For strains of which only a sampling period is known, the median of that period was taken as representative date in the analyses and graphical representations.

Most strains (93.9%) belonged to two major clades (group A and B), while the remaining strains belonged to three smaller clades (group C, D and E). Group A comprised 30.3% of the strains and included the reference strain PA14. The reference strain PAO1 belonged to group B which is the largest phylogroup with 63.6% of the strains. Group C was the most genetically distant phylogroup and consisted of only four isolates, including the reference strain PA7.

S2 On the antibiotic resistance and persistence phenotypes

We treated a selection of *P. aeruginosa* strains with various concentrations ciprofloxacin (Extended Data Fig. 2) for extended time periods and analyzed the data for biphasic killing dynamics, a hallmark of the presence of drug-tolerant persisters that are killed at a slower rate compared to the sensitive regular cells[1]. Strains with variable resistance levels were tested and the applied antibiotic concentrations were adjusted each time to agree with a multifold of the respective MIC (10x, 20x and 100x MIC). From a treatment dose of 10-fold the MIC, killing dynamics were highly similar. No substantial differences were observed between treatment doses which confirms that persistence is independent of the applied antibiotic concentration as soon as the concentration strongly exceeds the MIC[2]. Furthermore, after 5 hours of treatment, all strains reached a persister plateau. Therefore, we followed a treatment scheme treating stationary phase cells for 5 hours with a concentration of 20 times the MIC in the assessment of persistence for the entire collection of strains.

Treating at 20 times the MIC also circumvents potential confounding effects of antibiotic resistance on survival. The high dose ensures that no clones can emerge that are resistant to such dose and result in survival because of this resistance and wrongfully are being counted as persisters, since the mutant prevention concentration of ciprofloxacin for *P. aeruginosa* strains is reported to be generally below 16µg/ml[3,4]. Scaling the antibiotic dose to the MIC of the strain also further tries to circumvent the confounding effect of resistance in our killing assays. While killing curves in function of concentration seems to agree with such scaling, it does not exclude a bona fide interaction between resistance and persistence levels.

S3 On the GWAS analyses

As clinical origin and sampling year affect both phenotypes, they were included as fixed effect covariates in the final mixed effect model. In addition, given the positive association between resistance and persistence, persistence levels were included as covariate in the GWAS for resistance and vice versa, aiming to disentangle both phenotypes and exclude potential confounding effects. As described in the methods, multidimensional scaling components were derived from the pairwise distance matrix and included as covariate in the linear mixed model to account for lineage effects. Furthermore, we corrected for population structure by a kinship matrix as a random covariate. When comparing various combinations of the herein discussed effects, our global outcomes and conclusions remain highly similar and significant (not shown) and a dominant effect of mutations in *gyrA* remained prominent. Additionally to GWAS on SNP genetic variation, we performed similar analyses on presence or absence of genes and unitig variation. While some additional hits came out of these association studies (see Supplementary Fig 3), also *gyrA* containing unitigs were picked up which is what we focus on in this paper.

In addition to GWAS on our own strain collection and phenotypes, we have also performed GWAS using a publicly available dataset of clinical *P. aeruginosa* isolate[5,6]. In this study, they examined persistence in 391 isolates obtained from 39 CF patients. In contrast to our quantitative measures, persistence levels were determined using a qualitative measure. Based on regrowth of spots after ciprofloxacin treatment, strains were classified as hip or lop, respectively the fourth quartile, containing isolates with the 25% highest and the first quartile, containing isolates with the 25% lowest persistence levels. A pangenome analyses revealed that the genetic diversity of this strain collection is lower compared to ours as only 18,924 orthologs were predicted in the entire pangenome. A pyseer analysis on SNPs showed that the *gyrA* T83I mutation was only significantly associated with ciprofloxacin resistance (FDR < 0.05). A possible explanation for the absence of the association between *gyrA* and persistence is the strict definition of persistence as a binary phenotype. Instead, a

Fisher's exact test on this data showed that the proportion of isolates with a mutant *gyrA* allele is significantly higher in the hip group compared to the lop group (Extended Data Fig. 4a). The *parC* S87L mutation was not observed in any of the clinical isolates.

Given its strong association with both ciprofloxacin resistance and persistence, subsequent analyses focused on *gyrA*, and more specifically on the effect of the amino acid change at position 83. Mutations in the QRDR of *gyrA* are highly prevalent and occur in 75.4% to 98.1% of the *P. aeruginosa* isolates resistant to ciprofloxacin[7]. These *gyrA* mutations are often observed together with a mutation in *parC,* more specifically an amino acid change at position 87, which further decreases ciprofloxacin susceptibility[7].

S4 On the causality of *gyrA* alleles in natural strains

Since genomic engineering of wild *P. aeruginosa* isolates can be challenging, we followed an alternative approach using *in vitro* evolution experiments which also introduces *gyrA* mutations in *P. aeruginosa* isolates[8–12]. Indeed, continuous exposure to sub-MIC concentrations results in increased resistance by the enrichment of pre-existing resistance mutants or the selection of *de novo* mutants[13–19]. Since we were interested in the effect of single mutations with minimal fitness costs, a low selection pressure was applied with a fixed concentration of ciprofloxacin[19]. Previous research has already demonstrated that under these conditions, at fixed sub-MIC concentrations, clinically relevant mutations conferring fluoroquinolone resistance emerge[17,13,16,15]. Nine independent lineages of each strain were serially passaged in MHB medium containing ciprofloxacin, and a control population was evolved without ciprofloxacin. The evolution experiment was stopped when the MIC of all populations experiencing antibiotic exposure exceeded the EUCAST cutoff of 0.5 µg/ml[20]. The MIC of the untreated control did not increase more than twofold throughout the experiment. The highest fold change in MIC levels was observed in the first 2 days of the evolution experiment (Supplementary Fig. 3).

Subsequently, we conducted whole-genome sequencing on a subset of clones. Their genome sequences were compared to the respective ancestor background. In addition to mutations in DNA gyrase, mutations in efflux related systems were also found. Indeed, the presence of (inactivating) mutations in one or multiple of repressor genes of such systems are also known to significantly increase MIC levels. When studying the resistome per parental background, we observed that the combination of mutations in genes encoding for DNA gyrases and negative regulators confers higher levels of resistance, which agrees with previous work[21]. No mutations in *parC* or *parE* were observed in the evolved clones. A mutation in *parC*, in combination with *gyrA* mutations, is often associated with higher levels of resistance [21–23]. On the other hand, *parE* mutations are less prevalent in fluoroquinolone-resistant strains [24]. Mutations in other efflux transporters such as MFS transporters and ABC

transporters were also identified, but do not have a dominant role in resistance in *P. aeruginosa* [25]

Remarkably, not only mutations in resistance genes were identified. Most notably, the LMG 14083-1 clone had a deletion of 16 bp in *mutS,* encoding for a DNA repair protein, which results in a mutator phenotype and thus explains the high number of mutations[26] (Supplementary Table 4). However, other clones also contained mutations in genes that are not directly related to fluoroquinolone resistance. All sequenced clones carried at least one mutation in *gyrA* or *gyrB*, or in one of the transcriptional regulators of multidrug efflux pumps (Supplementary Table 3). NalC, NalD and MexR are the transcriptional repressors of the MexAB-OprM efflux pump, while NfxB and MexS control the expression of the MexCD-OprJ and MexEF-OprN efflux pump, respectively[27].

For statistical analyses in Fig. 3b, persistence levels of the evolved clones were estimated based on the biphasic killing curves and expressed relatively compared to the parameters of the respective ancestral strains. While most clones show significantly increased persistence, CPHL 11450–2 and J66UH5 F21-1 deviate from this finding as they are resistant but do no show increase persistence (Figure 3b, Extended Data Fig. 5). Strikingly, these strains are not mutated in genes encoding for DNA gyrases (Supplementary Table 3). For example, CPHL 11450-2 possess a 12bp-deletion in *nfxB* and a SNP in *mexS*. We have shown that a single point mutation in *nfxB* or a knockout in the gene does not significantly impact persistence levels (not shown). However, PN119w-4 demonstrated increased persistence levels compared to its ancestor and the only resistance-associated mutation in this clone is in the *nfxB* gene, namely G180A. Further research is required to verify whether this high-persistence phenotype is caused by this specific *nfxB* mutation or other mutations in the PN199w-4 clone. Based on the current results, it seems that any mutation in *nfxB* confers resistance while only specific mutations also confer persistence.

Less extreme increased survival levels were also observed in PN119w–4 and TA03–2 which carried wild-type *gyrA* and *gyrB* alleles (Figure 3b, Supplementary Table 3). Moreover, the strongest increase in survival levels was observed in strains with the *gyrA* T83I mutation. Indeed, the LiA18/2003–2, PN199w–2 and TA03–4 clones possess the *gyrA* T83I mutation. The clone LMG 14083-2 also carried the *gyrA* T83I mutation but did not show increased survival levels compared to the other two clones. However, in contrast to other clones with the *gyrA* T83I mutation, LMG 14083-2 had a second mutation in *gyrA* (A51T) which might impact persistence levels as well. This mutation at position 51 is located outside the QRDR of *gyrA,* which is commonly associated with fluoroquinolone resistance. Moreover, mutations in *gyrA* at position 51 and 325 in strain LiA83/2003–1 and LMG 14083-2, and LMG14083–3, respectively, were not observed in any of the *P. aeruginosa* isolates of our

collection. It would be interesting for future research to verify the effect of these single mutations on resistance and persistence.

To fully grasp the importance of either DNA gyrase mutations and resistance on persistence, we performed regression modelling. Simple linear models with either only resistance or *gyrAB* mutations as explanatory factors were constructed and compared. Adjusted R^2 values were 0.54 and 0.95 respectively, indicating that 95% of the variation in tolerance is explained by *gyrAB* mutations while only 54% of the variation is explained by resistance. A more complex model with both factors is superior compared to the model with only *gyrAB* information (based on Likelihood ratio test, AIC and BIC) and explains 97.3% or 98.2% of the variation in persistence, depending on whether it is an additive model or a model that also includes interaction terms respectively.

While the model with interaction terms is statistically superior (Likelihood ratio test, AIC and BIC) it suffers from strong multicollinearity. The partial R^2 values (*rsq* package) of the additive model show that resistance, after accounting for *gyrAB* information, uniquely explains 48.5% of the variance in persistence. In contrast, *gyrAB* information, after accounting for resistance, uniquely explains 94.7% of the variance in persistence. The relative importance of *gyrAB* information is estimated to be 70% compared to only 30% for the resistance levels (*relaimpo* package). Thus, while resistance levels generally went up (*t*-test, $p \leq 0.0001$) and evolved strains generally showed increased persistence (intercept is significantly different from 0, $p \leq 0.0001$), it is the presence or absence of mutations in *gyrAB* that most strongly determines the persistence level. Tukey post-hoc test show significant increases in all groups (with/without DNA gyrase mutations) compared to the ancestors and showed a significant difference when comparing between all groups of DNA gyrase mutations ($p \leq 0.0001$). In fact, when adding information of *gyrAB* mutations to the model, the effect of resistance on persistence become significantly negative, i.e. increasing resistance decreases persister levels.

Ridge regression models were constructed as they can handle multicollinearity by regularization and penalization of the coefficient sizes while still including an interaction term. These models resulted in highly similar conclusions about the impact of resistance versus that of mutations in *gyrAB* on explaining persistence (see previous). Due to the correction for the multicollinearity, the effect of resistance is strongly diminished and even becomes statistically insignificant. This is actually logic from a methodological point of view as well, since we corrected for increased resistance levels by increasing the antibiotic concentration used in the persistence assay. The ridge regressions and a simple linear regression with only resistance is plotted in Fig. 3b.

To verify whether mutations in *gyrAB* explain persistence more than resistance, we compared models where each of the two phenotypes are response variable and *gyrAB* mutations are the explanatory variable. After normalization, the adjusted R^2 is 0.95 for tolerance while only 0.38 for resistance, showing that mutations in *gyrAB* better explain the variation in persistence than the variation in resistance. Beta coefficients are, however, not significantly different (estimated from bootstrapped 95% CIs of the beta coefficients) and coefficients from multivariate models are not significantly different.

S5 On the effect of *gyrA* alleles on subsequent resistance evolution.

Previously, both the level of resistance and the type of resistance mutations acquired during experimental evolution was shown to be different for strains with variable backgrounds[15,28]. Here, we focused on resistance evolution and the impact of mutations in *gyrA* given their predominant role in both resistance and persistence. To do so, we opted to describe the evolutionary trajectories to three important parameters: MIC_{end}, the time needed to reach the maximum MIC and the AUC. Although the AUC summarizes the evolutionary dynamics, it does not fully grasp resistance evolution. For example, low AUCs might point to either slow resistance evolution to high MIC levels or rapid evolution to low MIC levels. For this reason, we discuss resistance evolution using all three parameters to provide a general view of the rate and extent of evolution.

In addition to the isogenic strain pair PAO1 and PAO1 *gyrA* T83I, we selected ten isolates from our collection based on the presence of resistance-conferring mutations and variable resistance levels. Three isolates carry the *gyrA* T83I mutation of which two were considered as sensitive to ciprofloxacin (Lo53 and PSAE1656, PSAE1649). Given the previously demonstrated role of gyrB in resistance and persistence, we also selected one resistant isolate carrying the *gyrB* S466F mutation (10BR1). The other six isolates contained wild-type *gyrA* and *gyrB* alleles (Br764, MC110, A14, Br670, PHLS08916 and MC116) and displayed variable MIC and persister levels (Supplementary Fig. 6). Throughout the text, we refer to isolates carrying mutant *gyrA* or *gyrB* alleles as 'mutant' isolates, while isolates carrying wild-type *gyrA* and *gyrB* alleles will be called 'wild-type' isolates. Of note, none of the selected isolates have mutations in the QRDR of *parC* and *parE*.

We evolved six replica populations for each strain. Notably, some populations of Br670, PHLS08916, and MC110 (all expressing wild-type DNA gyrase) went to extinction. To verify the effect of mutations in DNA gyrase on the evolutionary process, we performed multivariate modelling. Here, we accounted for the differences in strains with a random effect. Furthermore, we accounted for possible confounding effects of start resistance and persistence levels by adding these as covariates. The resulting Bayesian generalized linear multivariate mixed regression showed that the presence of mutations in DNA gyrase result

in a highly probable increase in end MICs and AUCs while the time to maximum MIC is highly probable to decrease (99.26%, 100% and 99.99% with factors of 0.74, 10.13 and 3.08 respectively). While we did correct for the possible confounding effect of MIC at the start and used this model for the statements in the main text and figures, the brms model that included start MIC as a covariate was not statistically better based on Leave-One-Out Cross-Validation (LOO) comparisons. Higher start MICs did result in highly probable lower end MICs, AUCs and times to maximum MIC (99.34%, 92.53% and 99.49% with factors of -0.36, -1.05 and -0.79 respectively).

In our strain set, there is significant positive correlation between persister fraction at the start and the presence of mutations in DNA gyrase (as also shown in the previous sections of this paper). A model that also includes persister fraction at the start should be analyzed with caution because of the resulting multicollinearity as variance cannot fully be contributed to either DNA gyrase mutations or the initial persister fraction. Such more complex model was also not statistically better based on Leave-One-Out Cross-Validation (LOO) comparisons. Including the persister fraction at start decreased the significant contribution of DNA gyrase mutations to resistance development somewhat as credibility intervals widened. However, DNA gyrase mutations still significantly increased AUC and decreased the time to the maximum MIC (99.99% and 99.9% with factors of 7.61 and 3 respectively). Furthermore, the persister fraction at the start significantly increased resistance through the AUC and MIC_{end} (99.01% and 96.76% with factors 0.26 and 1.82 respectively) as we previously demonstrated in *E. coli* isolates [27].

S6 On the mechanism of gyrA-dependent high tolerance

We measured the intracellular concentration of PAO1 isogenic mutants after 1h treatment with ciprofloxacin at the previously used scaled concentration of 20x MIC. We tested isogenic mutants to minimize the effect of the genetic background on antibiotic uptake. PAO1 was treated with a concentration of 1 µg/ml, PAO1 *gyrA* T83I with a concentration of 20 µg/ml and PAO1 *gyrA* T83I *parC* S87L with a concentration of 160 µg/ml.