Omics-based tracking of *Pseudomonas aeruginosa* persistence in 'eradicated' CF patients

SUPPLEMENTAL METHODS

Cystic fibrosis (CF) diagnosis. Patients are defined as having CF if they are homo- or heterozygous with one or more of the CF causing mutations of the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) gene. *Pseudomonas aeruginosa* infected and non-infected patients are segregated into separate wards and seen on different days in the outpatient clinic according to bacteriology in the lower respiratory tract [33].

Culture and identification of PA*.* PA is isolated from the airways using standard microbiological methods [34]. Briefly, samples are plated on a Sabouraud plate, a 7% NaCl plate, a *Burkholderia cepacia* plate containing colistin and gentamicin, a 'blue plate' (modified Conradi Drigalski's medium) selective for Gram-negative rods, a 5% Danish blood agar plate, and a chocolate agar plate (SSI Diagnostica®, Hillerød, Denmark). Isolated bacteria are identified as described previously [34]. Before 2011, biochemical profiling of PA was based on API 20NE (bioMérieux), and from 2011 on MALDI-TOF mass spectrometry (Bruker, Germany).

Cohort evaluation. Our base cohort of 80 CF patients was determined by excluding patients with their first PA positive culture before January 2002 or after July 2018, and those that between January 2002 to December 2018 moved treatment centers, were given a lung transplant, or were deceased. We then examined continuity of treatment by evaluating the dates of all patient microbiology samples by both visualization of longitudinal sampling and calculating various metrics of sample gaps (periods of time where no sample results, pathogen positive or normal, were reported in our database). We further excluded patients with any sample gaps greater than one year (n=5, retained one patient with a year long sample gap which ended 4 years before first positive PA culture). Patients with sample gaps between 0.5 and 1 year in length or at least 3 instances of a sample gap greater than 0.25 years were manually evaluated to ensure that sample gaps did not overlap with first or maximum eradication periods or otherwise indicate routinely poor clinic attendance by the patient. Patients with 4 or more sample gaps greater than 0.25 years were predominantly

diagnosed as chronic during the study period and therefore none were excluded from the study; these patients often attend the clinic less regularly as monitoring for new PA infection is no longer a primary goal in their care (n=4 of 5).

In the resulting complete-record cohort, a median of 6 PA samples (SD: 32.7, range: 1-238) per patient were obtained over a median of 7 years (SD: 5.4, range: 0-16.7 years). A comparison of samples versus sequenced samples over time is shown in **Figure S2A**, highlighting a higher ratio of sequence to sample in early colonization, including both the 72 patients with sequenced samples and the remaining 8 patients with no sequenced isolates. Because we do not have all PA isolates sequenced, it is important to be aware of how sequenced isolates span both colonization time and number of samples (**Figure S2B)**. Patients with many sequenced isolates spanning a large proportion of the colonization of the given patient are large, green circles, while patients with few or no sequenced samples, that thus covers limited or no time of the entire colonization, are small yellow circles. Examples of the latter group would be patients who have been diagnosed with chronic PA infection and shown frequent or continuous positive culture since that chronic diagnosis date but with few or no samples sequenced after their diagnosis as chronically infected under the assumption that the same clone type is present. In summary, while we have sequence coverage variation across patients, our sequenced patients cover almost the entire spectrum of colonization lengths (that is, the green medium-large circles are present along the entire x-axis).

33 Høiby N, Pedersen S. Estimated risk of cross-infection with *Pseudomonas aeruginosa* in Danish cystic fibrosis patients. *Acta Paediatr Scand* 1989; 78: 395–404.
34 Johansen HK, Moskowitz SM, Ciofu O, *et al*. Spread of colistin resistant non-mucoid *Pseudomonas aeruginosa* among chronically infected Danish cystic fibrosis patients. *J Cyst Fibros* 2008; 7: 391–397.

Figure S1. *Summary of sampled data for all patients with more than 1 positive Pseudomonas aeruginosa culture.* Data from C2 to C4 is based on sequenced samples only. Both number of samples and sample range increase from C2 to C4 as completeness of patient records is improved through filtering (as described in Figure S2). A. Number of samples per cohort. B. Range of sampling periods per cohort in years. Boxplots show median and quartiles.

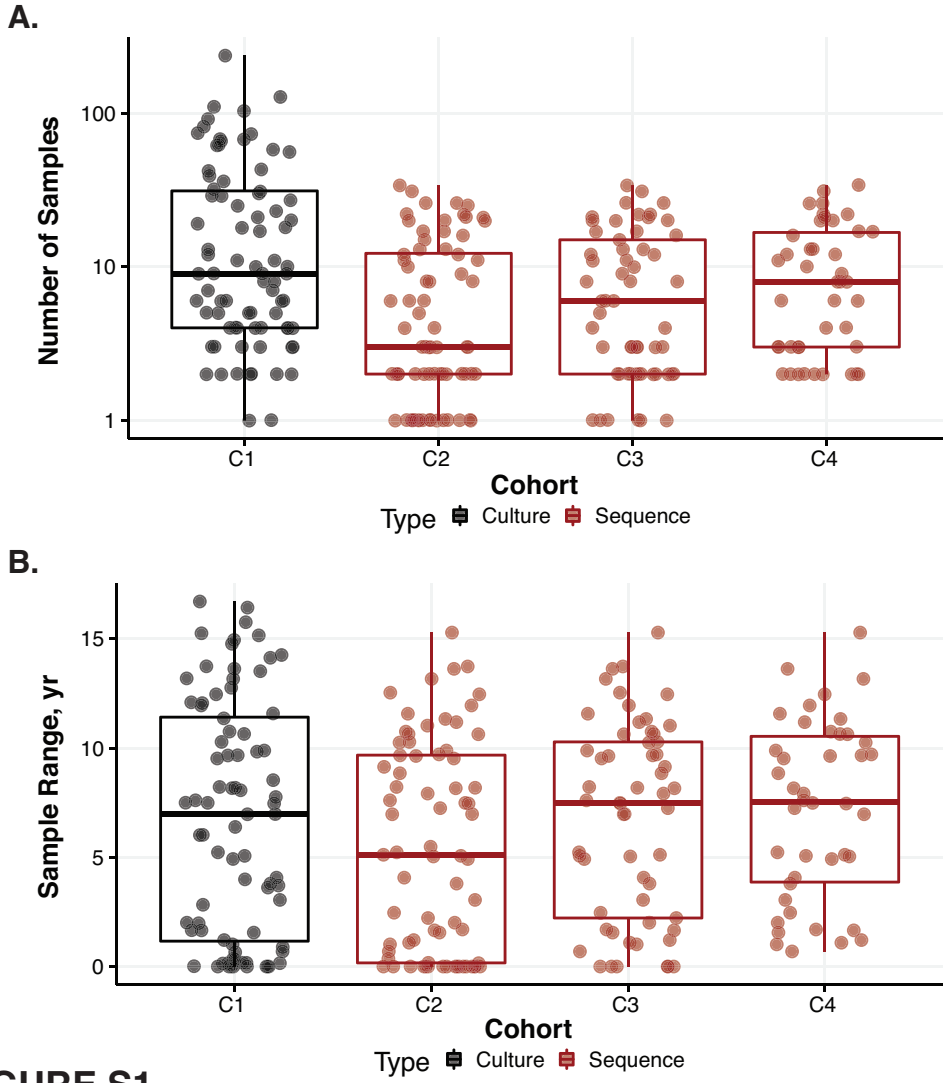

A.
Number of Samples
100
10
1
C1
C2
C3
C4
Cohort
Type
Culture
Sequence
B.
Sample Range, yr
15
10
5
0
C1
C2
C3
C4
Cohort
Type
Culture
Sequence

FIGURE S1

Figure S2. *Bacterial sampling from 80 young CF patients.* **(A)** Samples of *Pseudomonas aeruginosa* cultured over the monitoring period from patients with complete *P. aeruginosa* culture records (C1). Proportion of samples sequenced is indicated in red. **(B)** Sequence coverage of individual patients (circles) contextualized by total length of colonization (x axis), cohort membership (circle outline color), fraction of samples sequenced (size), and temporal range of sequenced samples (circle fill color).

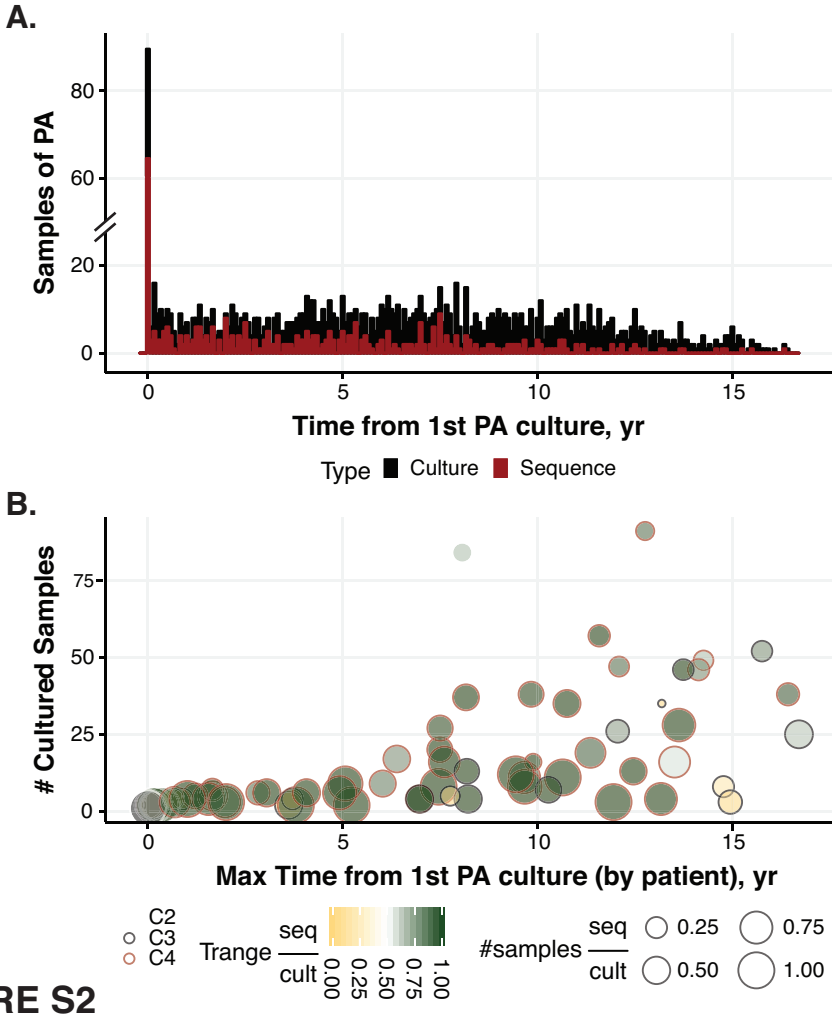

A.
Samples of PA
80
60
20
0
0
5
10
15
Time from 1st PA culture, yr
Type
Culture
Sequence
B.
Cultured Samples
75
50
25
0
0
5
10
15
Max Time from 1st PA culture (by patient), yr
C2
C3
C4
Trange
seq
cult
0.00
0.25
0.50
0.75
1.00
#samples
seq
cult
0.25
0.50
0.75
1.00

FIGURE S2

Figure S3. Maximum eradication period (MEP) length per patient versus length of colonization (max time from 1st *Pseudomonas aer*uginosa culture) for C3, highlighting patients with clone types that persist versus switch over the MEP (C4).

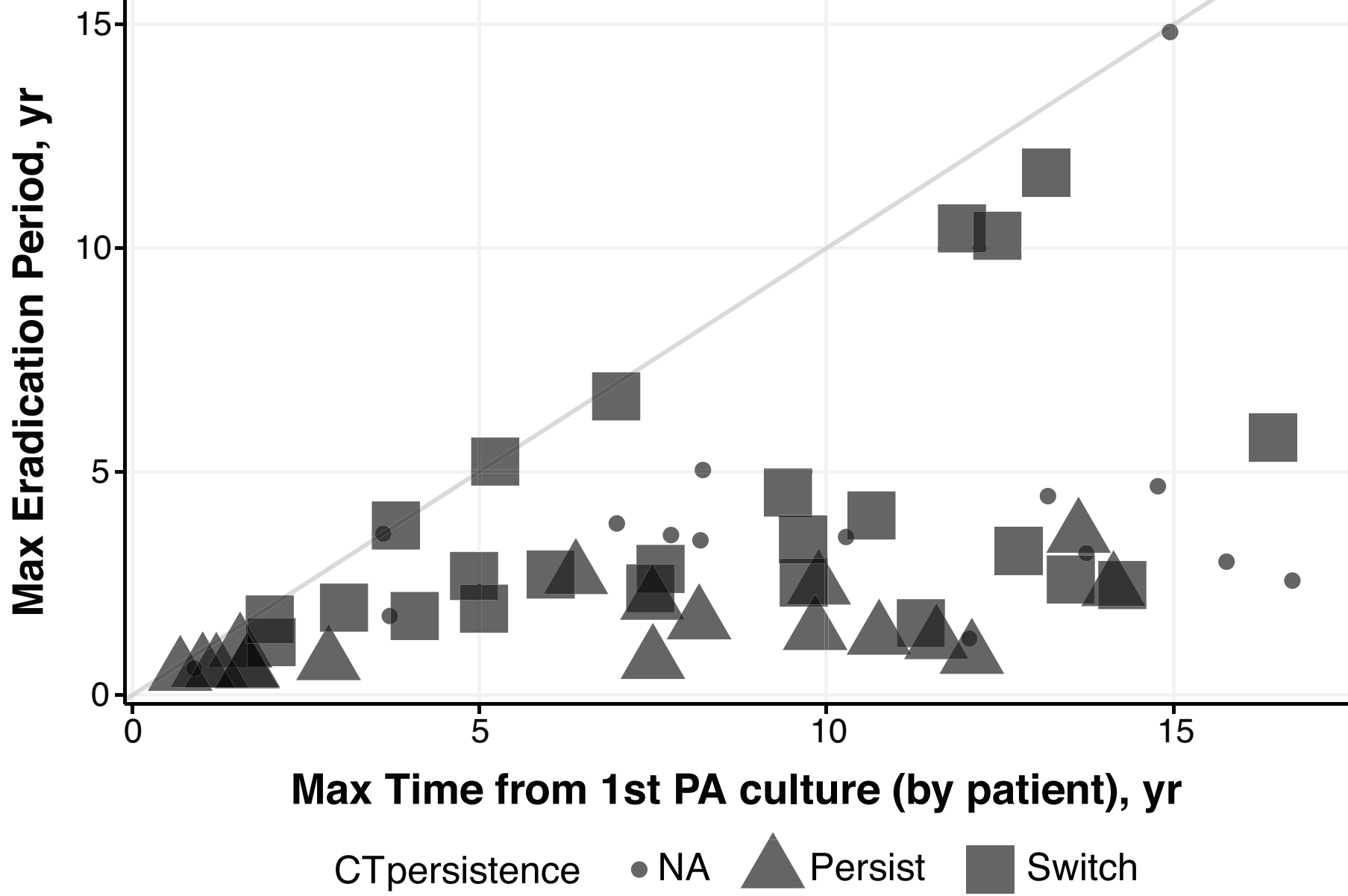
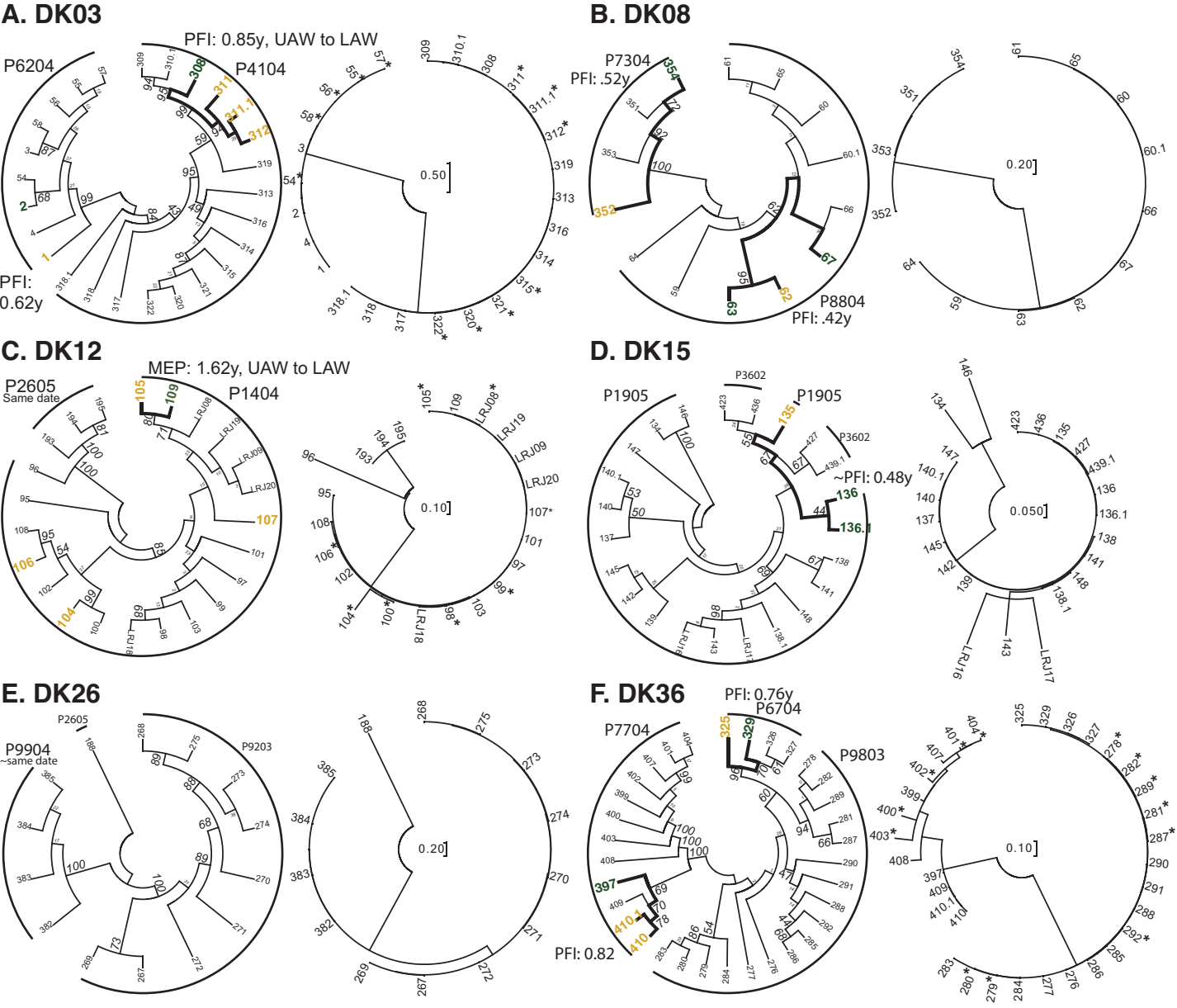

Max Eradication Period, yr
15
10
5
0
0
5
10
15
Max Time from 1st PA culture (by patient), yr
CTpersistence
NA
Persist
Switch

FIGURE S3

Figure S4. *Maximum likelihood cladograms and phylogenies of clone types found in multiple patients.* Bootstrap values >= 40 are shown with large text size in the cladogram where patients are also indicated with lines and patient IDs. In the phylogeny to the right the isolate IDs are indicated, which can be traced back to genotypes in Marvig *et al*[17] and phenotypes in Bartell *et al*[24]. The log likelihoods of trees are: (A) DK03: -2665.62, (B) DK08: -853.02, (C) DK12: -2567.09, (D) DK15: -2931.71, (E) DK26: -800.92, (F) DK36: -6641.36. The likelihood of mutation for the maximum likelihood phylogenies to the right are indicated with a bar and likelihood in the center of the phylogenies. In the phylogenies to the right, instances of eradication periods (either the MEP or the longest PFI as noted) are highlighted such that isolates at period start are yellow and isolates at period end are green. Length of eradication period and relevant context is included in text. The closest phylogenetic linkage between start and end isolates is also in bold. In the cladograms on the left, isolates derived from the upper airway (UAW) are highlighted with (*).



Isolates bridging longest MEP or PFI within CT phylogeny denoted as yellow (start) and green (end), * = UAW

FIGURE S4

Figure S5. *Archetypal analysis of isolates associated with MEP or longest PFI.* In addition to two patients for which we had data to model phenotypic adaptation over their MEP, seventeen patients had phenotypic data available to model isolates (points) at the start (yellow) or end (green) of an eradication period (MEP – circles, longest PFI – triangles). Here, we show all 19 patients, separated into clone types that persisted (panel A, 8 strains) versus switched over the eradication period (panel B, 11 strains).

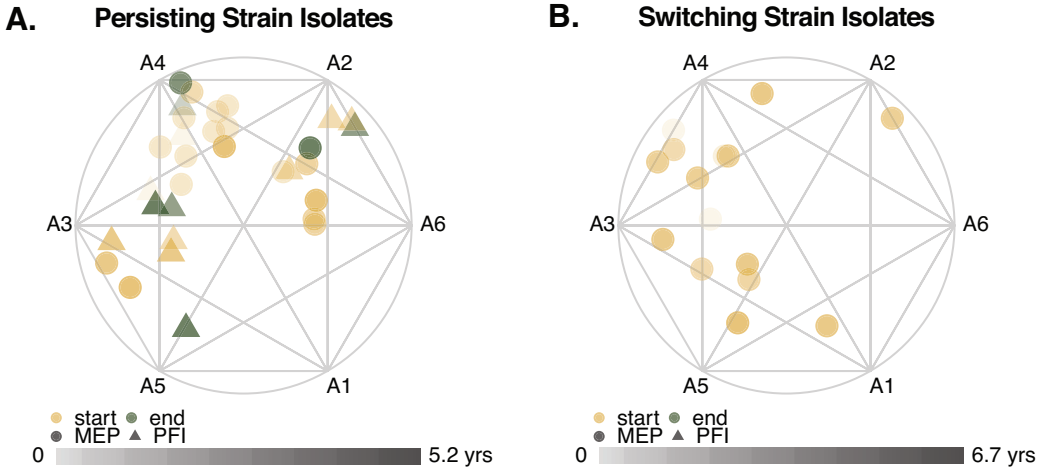

A.
Persisting Strain Isolates
A4
A2
A3
A6
A5
A1
start
end
MEP
PFI
0
5.2 yrs
B.
Switching Strain Isolates
A4
A2
A3
A6
A5
A1
start
end
MEP
PFI
0
6.7 yrs

Figure S6. *Archetypal analysis of isolates associated with MEP or longest PFI.* The 4 drug classes prescribed to the most patients in Panel A (aminoglycosides – amg, fluoroquinolones – flq, macrolides – mcl, and polymyxins – pmx, prevalence indicated by y-axis boxplot width) were evaluated with respect to usage combinations versus number of 'persist' or 'switch' patients.

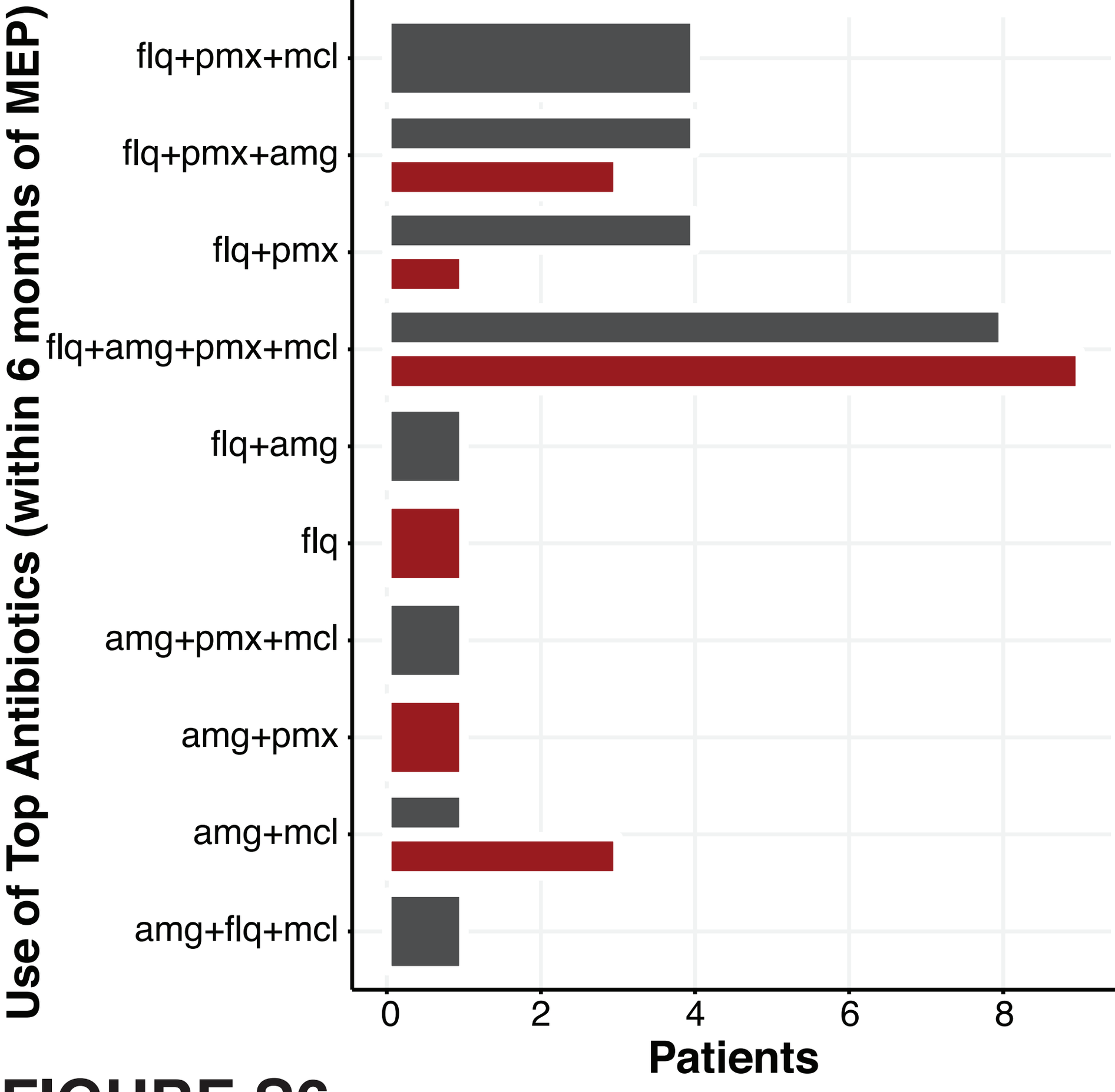

Use of Top Antibiotics (within 6 months of MEP)
flq+pmx+mcl
flq+pmx+amg
flq+pmx
flq+amg+pmx+mcl
flq+amg
flq
amg+pmx+mcl
amg+pmx
amg+mcl
amg+flq+mcl
0
2
4
6
8
Patients
Cohort
persist
switch

FIGURE S6