





Citation: van Mansfeld R, de Been M, Paganelli F, Yang L, Bonten M, Willems R (2016) Within-Host Evolution of the Dutch High-Prevalent *Pseudomonas aeruginosa* Clone ST406 during Chronic Colonization of a Patient with Cystic Fibrosis. PLoS ONE 11(6): e0158106. doi:10.1371/journal.pone.0158106

Editor: Sanjay Haresh Chotirmall, Lee Kong Chian School of Medicine, SINGAPORE

Received: March 2, 2016
Accepted: June 12, 2016
Published: June 23, 2016

Copyright: © 2016 van Mansfeld et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Sequencing reads have been submitted to the European Nucleotide Archive under study accession PRJEB12885 (http://www.ebi.ac.uk/ena/data/view/PRJEB12885).

Funding: M. Bonten is supported by the Netherlands Organization for Scientific Research (VICI NWO Grant 918.76.611), http://www.nwo.nl/en. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

RESEARCH ARTICLE

Within-Host Evolution of the Dutch High-Prevalent *Pseudomonas aeruginosa* Clone ST406 during Chronic Colonization of a Patient with Cystic Fibrosis

Rosa van Mansfeld^{1*}, Mark de Been¹, Fernanda Paganelli¹, Lei Yang², Marc Bonten¹, Rob Willems¹

- 1 Department of medical microbiology, University Medical Centre Utrecht, Utrecht, The Netherlands, 2 Novo Nordisk Foundation center for Biosustainability, Technical University of Denmark, Copenhagen, Denmark
- * r.vanmansfeld@vumc.nl

Abstract

This study investigates adaptation of ST406, a prevalent P. aeruginosa clone, present in 15% of chronically infected cystic fibrosis (CF) patients in the Netherlands, in a newly infected CF patient during three years using whole genome sequencing (WGS), transcriptomics, and phenotypic assays, including biofilm formation. WGS-based phylogeny demonstrates that ST406 is genetically distinct from other reported CF related strains or epidemic clones. Comparative genomic analysis of the early (S1) and late (S2) isolate yielded 42 single nucleotide polymorphisms (SNPs) and 10 indels and a single 7 kb genomic fragment only found in S2. Most SNPs and differentially expressed genes encoded proteins involved in metabolism, secretion and signal transduction or transcription. SNPs were identified in regulator genes mexT and exsA and coincided with differential gene expression of mexE and mexF, encoding the MexE/F efflux pump, genes encoding the type six secretion system (T6SS) and type three secretion system (T3SS), which have also been previously implicated in adaptation of other P. aeruginosa strains during chronic infection of CF lungs. The observation that genetically different strains from different patients have accumulated similar genetic adaptations supports the concept of adaptive parallel evolution of P. aeruginosa in chronically infected CF patients. Phenotypically, there was loss of biofilm maturation coinciding with a significant lower level of transcription of both bfmR and bfmS during chronic colonization. These data suggest that the high-prevalent Dutch CF clone ST406 displays adaptation to the CF lung niche, which involves a limited number of mutations affecting regulators controlling biofilm formation and secretion and genes involved in metabolism. These genes could provide good targets for anti-pseudomonal therapy.



Competing Interests: The authors have declared that no competing interests exist.

Introduction

Pseudomonas aeruginosa is a versatile Gram-negative rod with a relatively large genome of more than 6 million base pairs that can thrive in many different niches[1]. P. aeruginosa chronically infects the lungs of patients with cystic fibrosis (CF) and lung deterioration due to inflammation is the major cause of death in these patients [2]. More than half of the CF patients become chronically infected during their lifetime and most chronically infected patients harbor the same *P. aeruginosa* strain for many years or even life-long[3-5]. Several studies have investigated adaptation of P. aeruginosa trying to understand what the essential factors are for colonization and persistence in the lungs of CF patients despite the host immune system and aggressive antibiotic treatment [2;6-13]. Genome analysis of *P. aeruginosa* demonstrated that genes involved in biofilm formation, motility, transmembrane transport, hemolysis, secretion systems and resistance to oxidative stress and antibiotics are associated with CF lung adaptation [6]. Adaptation is mediated through remodeling of regulatory networks involving central metabolism and pathogenicity[13]. Phenotypic adaptation often involves a reduced growth rate, loss of motility, loss of substantial catabolic activities and inactivation of important regulatory functions[14]. Several studies demonstrated that parallel evolutionary events in these adapted strains do not seem to be due to loss or gain of certain genes[15], but rather to pathoadaptive mutations [10;13;16]. However, some CF adapted strains share accessory genomic elements like the LES-prophage-1, LESGI-2, and LESGI-4, that may contribute to increased competitiveness in the CF lung[17]. Adapted strains can be transmitted between CF patients and some of these epidemic *P. aeruginosa* clones seem associated with worse clinical outcome[18-23]. In the Netherlands, an epidemic P. aeruginosa clone with MLST type ST406, MLVA type CC27 and AT-chip single nucleotide polymorphism (SNP)-type A418 or E418 was found in up to 50% of CF patients between 15 and 25 years of age in 2007[24-26]. This clone was not found in clinical cultures from non-CF patients and was genotypically different from the epidemic CF strains found in other countries [26;27]. SNP-typing of clinical and environmental P. aeruginosa strains from all over the world only detected this genotype in Dutch CF patients[28]. The fact that this clone is specifically linked to the CF lung niche suggests that this genotype is very well adapted to the CF lung. In contrast to some other epidemic clones an association with clinical deterioration could not be detected for ST406[29]. Another clone (ST497) was found in 5% of Dutch CF patients and was associated with older age (>25 years)[25]. Likewise this genotype has thus far only been isolated from Dutch CF patients.

Insights into the evolutionary dynamics during chronic colonization of the lungs of CF patients and genetic relatedness of epidemic CF clones that are specifically adapted to the cystic fibrosis lung niche might give further clues for new eradication therapies or preventive measures. The aims of this study were to infer the phylogenomic relatedness of two high-prevalent *P. aeruginosa* clones in the Netherlands, ST406 and ST497, to a collection of *P. aeruginosa* strains and to describe genotypic and phenotypic changes of ST406 during adaptation to the CF airway and compare this to other known CF clones. To investigate the traits that may have contributed to adaptation of the *P. aeruginosa* ST406 clone to the CF lung we indexed genomewide alterations by performing comparative genomics in combination with comparative transcriptomics and cataloging phenotypic differences between two ST406 strains obtained from the same patient, but separated by a timespan of three years. The first isolate was cultured one month after the patient became first colonized with *P. aeruginosa* ("early isolate" or S1) while the second isolate was recovered from the same patient after three years of chronic colonization ("late isolate" or S2).



Methods

Strains

The first, early, *P. aeruginosa* ST406 isolate (S1) was cultured from sputum from a 15 years old, Dutch CF patient in 2004, one month after the very first detection of *P. aeruginosa* in this patient. This patient became chronically infected after the first positive culture in 2004. The second, late, *P. aeruginosa* ST406 isolate (S2) was cultured in 2007 from sputum from the same CF patient. Both isolates were non-mucoid phenotype. ST497 isolate (S3) was cultured from another 26 year old, Dutch CF patient that had already been chronically infected with *P. aeruginosa* for more than seven years. Both patients were regularly visiting the University Medical Centre in Utrecht for their regular medical care. *P. aeruginosa* strain PA01 and *E. coli* strains that harbor *lasB-gfp* (MH155) or *rhlA-gfp* fusion for quorum sensing detection were kindly provided by Søren Molin from DTU, Denmark.

Whole genome sequencing and SNP analysis

Genomic DNA was extracted from *P. aeruginosa* isolates S1, S2 and S3 using a Wizard Genomic DNA Purification kit (Promega). Whole genome sequencing was performed on an Illumina MiSeq platform. Raw 2×250 bp paired-end Illumina reads were quality-filtered using Nesoni 0.109 [https://github.com/Victorian-Bioinformatics-Consortium/nesoni] (with following options: adaptor-clip (yes), adaptor-match (10), adaptor-match max error (1), clip ambiguous (yes), quality cutoff (10), minimum read length (150)). De novo assembly was done using SPAdes 2.5.1[30]: kmers used (21, 33, 55, 77, 99, 127), with "careful" option turned on and cutoffs for final assemblies: minimum contig/scaffold size = 500 bp, minimum contig/scaffold average Nt coverage = 10-fold. Gene prediction and annotation was done using Prokka[31] (default options; using a Pseudomonas-specific BLAST database, as provided in the Prokka package) and RAST[32]. Additional functional annotation of protein-coding genes was done using the COG database[33]. Protein sequences were queried against the COG myva database using BLAST 2.2.29+[34] and received the same COG as their best BLAST hit (E-cutoff ≤ 1e-10). Sequencing reads have been submitted to the European Nucleotide Archive under study PRJEB12885.

To compare differences between the ST406 S1 and S2 isolates and investigate within-patient adaptation an in-house read-mapping pipeline was used to detect SNPs and small indels. Nesoni-filtered reads of S2 were mapped against the S1 assembly using Bowtie2[35]. To filter for genomic repeats, we removed reads that mapped to multiple positions in the S1 assembly. SAMtools 0.1.18[36] was used to call SNPs and indels with following conditions: Qscore \geq 50, mapping quality \geq 30, and calls were required to be homozygous (under diploid model). Additional criteria included a mapping depth \geq 10 reads, a consensus of \geq 75% to support a call, and \geq 1 supporting read in each direction. To correct for potential assembly errors, we also performed the SNP/indel-calling procedure described above after mapping Nesoni-filtered reads of S1 against its own assembly. Genomic positions containing SNPs and indels in the S1 vs S1 comparison were ignored in the S2 vs S1 comparison. Identified SNPs and indels were linked to features (i.e. genes) and were inspected for synonymous vs non-synonymous mutations (in case of SNPs).

To compare gene content between S1 and S2 we used Inparanoid v4.129. All S1 genes that did not have a predicted orthologous relationship with an S2 gene and vice versa were considered as potentially unique for that given strain. To further verify these potential strain-specific genes, S2 reads were mapped against the S1 assembly and vice versa as described above, but without filtering for reads mapping to multiple genomic positions. Using the same criteria



described above for finding SNPs and small indels, genomic regions significantly covered by mapped reads were identified. Potentially strain-specific genes identified using Inparanoid were considered to be truly strain-specific if $\leq 10\%$ of their entire length was covered by reads.

Core genome alignment

To investigate phylogenetic relatedness of the three *P. aeruginosa* isolates from Dutch CF patients that represent two prevalent Dutch clones with an international collection of *P. aeruginosa* strains all publicly available *P. aeruginosa* whole genome sequences (WGSs) were downloaded from GenBank on 12 February 2014. The genome sequences of these publicly available strains and of strains S1, S2 and S3 were aligned using an in-house pipeline that makes use of the NUCmer v3.23 alignment algorithm[37]. All genomes were aligned against the completed genome sequence of *P. aeruginosa* PA01 (i.e. used as reference strain). Repetitively aligning genomic regions were removed from each pairwise alignment after which all pairwise alignments were merged into one *P. aeruginosa* core genome alignment. A phylogenetic tree was built from the polymorphic sites in the core genome alignment using RAxML[38]. The tree was built under the GTR model. Confidence was inferred by running 1000 bootstrap replicates under the same model. The tree was visualized using MEGA6[39]. Odds ratio's for calculating significance of associations between source (CF sputum isolate, clinical isolate, or environmental isolate) and clade assignment was done using MedCalc software (https://www.medcalc.org/calc/odds-ratio.php). The *P*-value threshold for statistical significance was set at 0.05.

Transcriptome analysis

Transcriptome analysis using GeneChip *P. aeruginosa* genome arrays (Affymetrix) was performed as described previously[14]. *P. aeruginosa* strains were grown aerobically in Luria-Bertani (LB) medium starting from OD600 = 0.01 and harvested at OD600 = 0.5. RNA was isolated using RNeasy Mini Purification kit (QIAGEN) and transcribed into cDNA using random primers (Invitrogen Life Technologies). Subsequently, cDNA was purified (QIAquick, QIAGEN), fragmented and labeled and hybridized on an Affymetrix *P. aeruginosa* PA01 gene chip. The probe arrays were scanned with a GeneChip Scanner 3000 and the raw data was obtained using the Affymetrix GeneChip Operating System 1.4. Microarray data analysis was performed as described before[24] using bioconductor in R environment (http://www.bioconductor.org). Normalization and expression index calculation was done with rma function[40]. The fold change was calculated using the average expression levels of three replicates. A cut off p-value < 0.05 and fold change > 2 between transcriptome levels of S1 and S2 was considered significant. The annotations and functional classes were assigned according to the Pseudomonas Genome Database[41].

Amplification of S2 specific genes

To verify differences in gene content between S1 and S2 (detected by comparing whole genome sequence data as described above) and analyze presence of these genes in other *Pseudomonas aeruginosa* isolates we performed PCR amplification in 29 non-ST406 clinical isolates (6 ICU, 6 community acquired, 6 CF, 6 other clinical isolates from a previous study[26] and 3 isolates of the Liverpool epidemic PA01 and the Midlands epidemic strain) (S1 Table). Isolates were cultured on tryptic soy agar (TSA)-blood plates (Becton, The Netherlands) overnight at 37°C, suspended in 20 μl lysis buffer (0.25% (wt/vol) SDS, 0.05 N NaOH) and incubated at 95°C for 20 min. The cell lysate was centrifuged and diluted with 180 μl buffer (10 mM Tris-HCl, pH 8.5). After thoroughly mixing, another centrifugation for 5 min at 16,000 x g was performed to remove cell debris. Supernatants were frozen at -20°C until further use. Two μl of the lysate



was used in a touchdown PCR using Hotstar Taq DNA polymerase (Qiagen Benelux B.V.), and 5 μ l Q-buffer (Qiagen Benelux B.V., Venlo, the Netherlands). The PCR was conducted as follows: 10 min at 96°C, then 10 cycles of 30 s at 95°C, 30 s at 65°C with 1°C less every cycle and 1 min at 72°C. This was followed by 25 cycles of 30 s at 95°C, 30 s at 55°C and 1 min at 72°C, after which the final step of 10 min at 72°C followed. Presence of PCR products was checked by electrophoresis on 1% (wt/vol) agarose gel. S2 was used as positive control. Primers used for PCR amplification of S2 specific genes and non-ST406 *P. aeruginosa* strains used to test presence of these genes are shown in S1 Table.

Phenotypic experiments

Phenotypic traits known to be important virulence factors that are involved in adaptation to chronic infection of the CF lung are growth rate, excreted proteases, motility, quorum sensing and biofilm formation were compared between S1 and S2. PA01 was used as control.

Motility assays. Swimming, swarming and twitching motility and protease assays were performed in triplicate. Bacteria were grown overnight on LB agar plates. Swimming, swarming and twitching motility assays were performed on respectively 0.3% (wt/vol), 0.6% (wt/vol) and 1.5% (wt/vol) agarose plates containing Agrobacterium (AB) minimal medium with glucose (0.5%(wt/vol)) and casamino acids (0.5% (wt/vol)). One bacterial colony was inoculated in, on and through the agar by sterile toothpick, respectively. Swimming ability was assessed after 24 h incubation at 30°C by measuring the maximum diameter of the zone of growth. Swarming and twitching was assessed after 48 h of incubation at 37°C by measuring zones of growth. Protease production in skim milk was measured by inoculating 100 μl of supernatant of overnight culture (after spinning at 7000 g for 5 min) on a LB-agar plate with 10% (wt/vol) skim milk concentration. Clear zones of protease production were measured after 48 h of incubation at 37°C.

Growth curves. Growth curves were made by measuring OD600 every 50 min, starting with an OD600 of 0.01. Doubling time was calculated in Microsoft Excel using the exponential trend line ($tD = (ln2/\mu)$).

Quorum sensing. Quorum sensing signal production was tested by streaking isolates and a positive and negative control near two *E. coli* monitor strains that harbor *lasB-gfp* (MH155) or *rhlA-gfp* fusion products on a plasmid[42] that can be induced by the Las (3-O-C12-HSL) and Rhl signal molecule (C4-HSL), respectively. GFP (green fluorescent protein) production was visualized after 24 hincubation at 37°C using a Zeiss Axion2 microscope with an integrated Coolsnap color cf camera at 2.5x magnification and 200ms exposure time.

Phenotype MicroArrays. To investigate difference between S1 and S2 in growth under conditions using different nutritional sources and other chemical compounds like toxins, phenotypic microarrays were performed. Quantitative measurements of 1821 phenotypes of S1, S2 and of *P. aeruginosa* strain PA01 were determined using phenotype MicroArrays (Biolog, Inc Hayward CA, USA). In this assay cell respiration, measured by color change caused by reducing a tetrazolium dye, is used as a universal reporter for both amplification and precise quantitation of phenotypes. The measured average height of two replicates of each phenotype was compared between S1 and S2. Phenotypes with more than two-fold difference were considered different phenotypes. Among the 1821 phenotypes tested by this method were 190 Carbonsources, 384 Nitrogen-sources, 60 Phosphor- sources, 192 phenotypes related to osmolytes and pH and 960 to chemical sensibility.

Antibiotic susceptibility and colony morphology. Susceptibilities were determined by disk diffusion on Muller Hinton plates using EUCAST breakpoints for ciprofloxacin, levofloxacin, piperacillin, ceftazidime, tobramycin, amikacin, meropenem, piperacillin/tazobactam, colistin trimethoprim, trimethoprim/sulfamethoxazole, chloramphenicol.



Colony morphology was checked under the plate microscope.

Biofilm experiments. A semi-static biofilm model was used to assess biofilm formation of S1 and S2, as described before [43]. Overnight bacterial cultures were diluted to an OD660 of 0.01 in 6 ml LB medium with 1%(wt/vol) glucose and added to a coverslip coated with poly-Llysine (0.45 μm; diameter, 12 mm; Becton Dickinson) inside a well from a six-well polystyrene plate (Corning Inc.). Biofilms were grown at 30°C for 48 hat 120 rpm. After 48 h, the coverslips were washed with 0.85% (wt/vol) NaCl and the biofilms were chemically fixed with 8% (vol/vol) glutaraldehyde (Merck) for 20 min. Subsequently, the biofilms were stained with 15 μg/ml propidium iodide (PI) in 0.85% (wt/vol) NaCl that was removed after 15 min. The coverslips were transferred to glass microscope slides and analyzed by a confocal laser scanning microscope (CLSM) (Leica SP5), equipped with an oil plan-Neofluar ×63/1.4 objective. PI was excited at 488 nm. Z stacks were taken with an interval of 0.42 μm. Pictures were analyzed with LAS AF software (Leica), and biofilm thickness and biomass were quantified using Comstat [44]/Matlab R2013b software (the MathWorks). The average thickness and biomass of the biofilms were measured at ten randomly chosen positions and statistical significance determined by unpaired two-tailed Student's t-test. Experiments were performed twice in duplicate.

Results and Discussion

Whole genome sequencing and phylogenetic inferences

To determine the phylogenetic relationship between Dutch prevalent clones, ST406 (S1 and S2) and ST497 (S3), and other epidemic clones whole genome sequencing was performed. Illumina sequencing and subsequent assembly of the three P. aeruginosa isolates S1, S2 and S3 yielded draft genomes with an average assembly size of 6.32, 6.32, and 6.28 Mbp, consisting of 45, 42, and 48 scaffolds with a nucleotide coverage of 56.4, 55.9 and 81.1, respectively. To assess the phylogenetic relatedness of the two Dutch epidemic clones, represented by the three P. aeruginosa isolates from Dutch CF patients S1, S2 and S3, with an international collection of P. aeruginosa strains, publicly available genome sequences of 37 P. aeruginosa strains were downloaded and aligned together (\$2 Table). This resulted in a core genome alignment of 3.27 Mb, containing 258,000 SNPs. An initial SNP-based phylogenetic tree showed that strains PA7 & VRFPA01 clustered far away from the other *P. aeruginosa* strains. To increase resolution, PA7 & VRFPA01 were left out and the remaining 38 strains were re-aligned. Identical strains were removed from the tree for clarity (MPA01P1, MPA01P2 and PA0579 represented by PA01 and PADK2CF510 represented by DK2). The final core genome alignment was 4.02 Mb and a phylogenetic tree including 34 strains was constructed using the 121,244 SNPs contained in this alignment (Fig 1).

The phylogenetic tree demonstrated that the Dutch high-prevalent P. aeruginosa CF strains, S1/S2 and S3 were evolutionary not closely related to other epidemic CF related clones. Also the Liverpool epidemic strain clustered separately, while the Manchester and Danish epidemic CF strain shared a common ancestry. Two distinct clades (A and B) were identified in the tree, which contain both environmental and clinical isolates as well as isolates from CF patients. However, clinical non-CF isolates were significantly associated with clade B (OR: 8.0; 95% CI: 1.5-43.7; p<0.01), while CF isolates were enriched in clade A, although this was not significant (OR: 7.4; 95% CI: 0.8-68; p = 0.08). Even though, there seems to be a trend towards significance and a larger number of isolates would be required to investigate the finding further and these results could suggests that clinical non-CF isolates may have an evolutionary background that is distinct from isolates colonizing CF patients. Of course, numbers of strains in our study are relatively small, and selection of strains that were chosen for WGS might contain a bias.



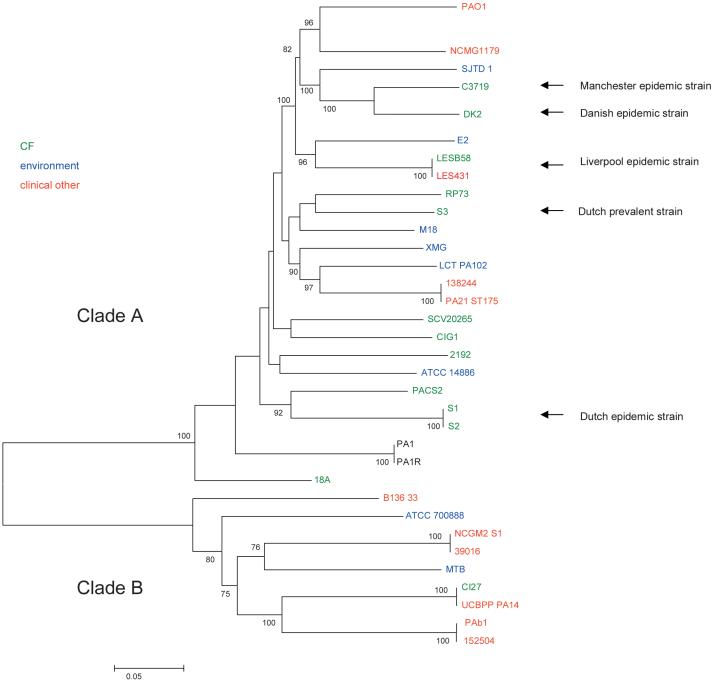


Fig 1. Phylogenetic tree including 34 P. aeruginosa strains built from an alignment of 121k core SNPs. Numbers along the branches indicate bootstrap values and only bootstrap values \geq 75% are shown. Branch lengths correspond to the scale bar, in units of changes/nucleotide position. Green = CF sputum isolate, red = clinical isolate, not CF, blue = environmental isolate. The Dutch high-prevalent CF strains and internationally known epidemic CF strains have been indicated.

In contrast to these observations, several studies investigating both clinical and environmental isolates have demonstrated that P. aeruginosa has a non-clonal epidemic population structure with no clear association between specific lineages and certain niches [23;28;45–50]. However, these studies indexed only a limited number of loci providing limited power for



inferring phylogenetic relationships. The fact that not all epidemic CF isolates group in one lineage indicates that these CF isolates do not form one monophyletic group, thus do not share a recent common ancestor. Clearly, the Manchester, Danish, Liverpool, and both Dutch epidemic CF strains form a polyphyletic group in which phylogenetically different sequences have converted into a successful epidemic phenotype in CF patients.

Comparative genomics and transcriptomics of the early (S1) and late (S2) ST406 isolate

To investigate genomic differences between S1 and S2, 6,284,869 bases (99.4–99.5% of total assembly size) of the two isolates were aligned and this yielded 42 high quality SNPs and 10 indels. Twelve SNPs were synonymous, while 24 were non-synonymous (<u>Table 1</u>). Five SNPs were located in intergenic regions and one occurred in an rRNA encoding gene. Most (29%) non-synonymous mutations were in genes coding for COG categories involved in metabolism, followed by non-synonymous SNPs in COG category transcription and signal transduction (21%) and category unknown (25%) (<u>Fig 2</u>).

Metabolic alterations and changes in regulatory functions have also been described for other CF adapted clones[13]. Whole genome sequencing of various isolates from different CF patients that harbored the DK2 clone for many years revealed loss of catabolic activities and inactivation of important regulatory systems[14]. Furthermore, analysis of non-synonymous SNPs in core genes of 32 CF related strains including LES (Liverpool epidemic strain) isolates revealed SNPs in genes encoding proteins involved in oxidoreductase activity, secretion and heterocycle metabolism[6].

Comparative transcriptomics between ST406 S1 and S2 yielded 179 genes with different transcription levels (p< 0.05), of which 110 genes had a more than 2-fold difference in gene expression (S3 Table). Gene ontology enrichment analysis revealed overrepresentation of differentially expressed genes involved in metabolism and posttranslational modification that were higher expressed in S1 and genes encoding membrane proteins or involved in membrane or cell wall biogenesis and proteins involved in secretion that were higher expressed in S2 (Fig 3). Approximately 92% of all the genes in strains S1 & S2 were covered by uniquely and perfectly matching probes on the Affymetrix *P. aeruginosa* PA01 GeneChip when using a cutoff of at least 4 probes per gene (~96% in PAO1). This indicates that only a small proportion of S1 and S2 genes was not analyzed for transcription differences.

All SNPs were located in genes that have been previously described in adaptation of *P. aeru-ginosa* to the lung of CF patients and of which some were designated pathoadaptive genes (<u>Table 1</u>). These SNPs were in the same genes but not identical to other published SNPs. SNPs and transcriptome differences between S1 and S2 in genes that were deemed important for observed phenotypic differences are described in more detail below.

Recent studies have shown that in one sputum sample of a CF patient a variety of *P. aeruginosa* isolates of the same genotype with variable phenotypes and based on WGS data several SNPs can be detected [7;13;51–53]. This suggests that in these niches *P. aeruginosa* behaves as a quasispecies, *i.e.* a large group or cloud of related genotypes that exist in an environment of high mutation rate [54]. This adaptive radiation can explain why in various longitudinal studies in CF patients, in addition to common parallel evolutionary events in chronically colonizing *P. aeruginosa* isolates, also different unique sets of putative adaptive SNPs and phenotypes have been identified.

ExsA and the type III secretion system (T3SS). A non-synonymous SNP (change from P156L in isolate S2) was identified in *exsA* (<u>Table 1</u>). Furthermore, one nucleotide insertion was detected upstream in the promoter region of this gene (<u>Table 2</u>). ExsA is a member of the



Table 1. 24 non-synonymous SNPs detected between early (S1) and late (S2) ST406 isolate.

ST 406 gene	SNP [aa change] ^a	amino acid position	PA01 annotation	gene	COG function classification	SNP in same gene in literature
3848	AAC[N] > ATC [I]	63	PA3545 ^b	algG	Inorganic ion transport and metabolism	[9], [13], [14], [53]
5027	AAC[N] > ATC [I]	178	PA4102	bfmS	Signal transduction mechanisms	[13], [14]
3581	AAG[K] > GAG[E]	297	PA1782		General function prediction only, Signal transduction mechanisms, Transcription, Replication, recombination and repair	[13]
5598	AAT[N] > GAT [D]	41	PA4276	secE	Intracellular trafficking, secretion, and vesicular transport	[13]
5091	ACC[T] > GCC[A]	460	PA4039		Cell motility	[13]
1522	ACG[T] > GCG[A]	1063	PA2727 + PA2728		Replication, recombination and repair	[13]
4626	AGC[S] > AGA[R]	224	PA2385 ^b	pvdQ	General function prediction only	[13]
4879	ATC[I] > AGC [S]	3973	PA0690		secretion	[13], [14]
0345	ATC[I] > GTC [V]	456	PA1147		Amino acid transport and metabolism	[13]
2241	ATG[M] > AAG[K]	86	PA4519	speC	Amino acid transport and metabolism	[9]
3651	CCG[P] > CTG[L]	156	PA1713	exsA	Transcription	[8],[9],[10], [13], [53]
5490	CTG[L] > ATG [M]	30	PA4163		Translation, ribosomal structure and biogenesis	[13]
0807	GCC[A] > ACC[T]	703	PA0794		Energy production and conversion	[13]
3973	GCC[A] > GGC[G]	61	PA3280	oprO	Inorganic ion transport and metabolism	[13]
5588	GCC[A] > GTC[V]	489	PA4266 ^b	fusA1	Translation, ribosomal structure and biogenesis	[10], [13], [53]
2194	GCG[A] > GTG[V]	15	PA4562		General function prediction only	[13]
0321	GGC[G] > GTC[V]	118	PA1171		Cell wall/membrane/envelope biogenesis	[13]
4666	GTC[V] > ATC [I]	204	PA2346		Lipid transport and metabolism	[13]
4610	GTC[V] > GCC[A]	1523	PA2402		Secondary metabolites biosynthesis, transport and catabolism	[13]
4794	GTC[V] > GCC[A]	80	PA2552		Lipid transport and metabolism	[13]
2089	TAG[stop] >TGG[W]	51	PA4661	pagL	no COG	[13], [14]
4856	TTC[F] > TTG [L]	34	PA2492 ^b	mexT	Transcription	[8], [13]
0194	TTG[L] > TTT [F]	68	PA1300		Transcription	[13]
5101	TTG[M] > CTG[L]	213	PA4029		Function unknown	[13]

^a A = Alanine, D = Aspartic acid, E = Glutamic acid, F = Phenylalanine, I = Isoleucine, K = Lysine, L = Ieucine, M = methionine, N = asparagine, P = proline, R = arginine, S = serine, T = treonine, V = valine, W = tryptophan. SNPs are indicated within their codon context, in the order: codon-S1 > codon-S2.

^b pathoadaptive gene by Marvig et al or Winstanley et al.[10;51]



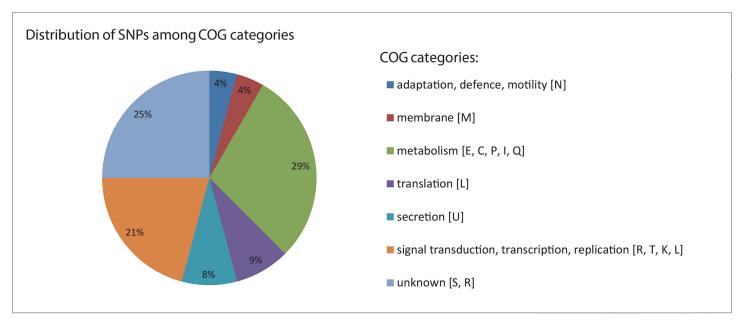
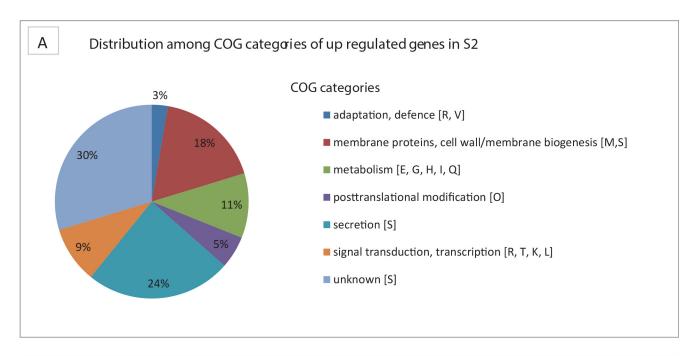


Fig 2. COG categories of non-synonymous SNP differences found between S1 (early isolate) and S2 (late isolate) isolated three years apart from the same patient. Percentage per category of total non-synonymous SNPs is indicated.

AraC family of transcriptional regulators and is the primary regulator of the type III secretion system (T3SS). ExsA controls expression of genes implicated in T3SS biogenesis by directly binding to promoter sequences of these genes leading to activation of transcription[55]. In correspondence with these findings, our transcriptome analysis indicated differential expression of T3SS genes, with a significantly lower level of expression in isolate S (S3 Table). Also *exoS*, encoding the ExoS toxin, which is secreted by T3SS was expressed at a lower level in S2. Type III secretion mediated translocation of exoenzymes is used by *P. aeruginosa* to deliver exoenzyme effector molecules, like ExoS, into the eukaryotic cell. Our findings of higher expression of genes encoding T3SS in S1 relative to S2 suggest that T3SS may play an important role in acute CF infection but not or to a lesser extent during chronic infection. This is consistent with previous findings that also demonstrated down regulation of ExoS in chronically infected CF patients[56]. On the other hand, a recent report also found up-regulation of *exoS* during chronic infections[57]. These conflicting data indicate that further studies are needed to elucidate the exact role of ExoS during acute and chronic infections.

MexT, the MexEF efflux pump and the type VI secretion system (T6SS). Another non-synonymous SNP (change from F34L in isolate S2) was located in *mexT*, which encodes the transcriptional regulator MexT that positively regulates the MexEF-OprN efflux pump and negatively regulates genes encoding T6SS[58–61] (Table 1). Transcriptome analysis revealed that the *mexEF* genes encoding the MexEF efflux pump were expressed at a higher level in the early S1 isolate, while genes encoding T6SS were expressed at higher levels in S (S3 Table). This suggests that the non-synonymous mutation in *mexT* affects some of the effector functions of MexT. The Mex E/F-oprN efflux operon also confers resistance to quinolones, chloramphenicol and trimethoprim. Down-regulation of *mexEF* genes in S2 coincided with less resistance against chloramphenicol and trimethoprim/ sulfa-methoxazole, while both isolates are resistant to quinolones and trimethoprim (S4 Table). It is interesting to note that *mexT* has been described as a mutational "hot spot," where mutations can contribute to global phenotypic changes in *P. aeruginosa* [62].





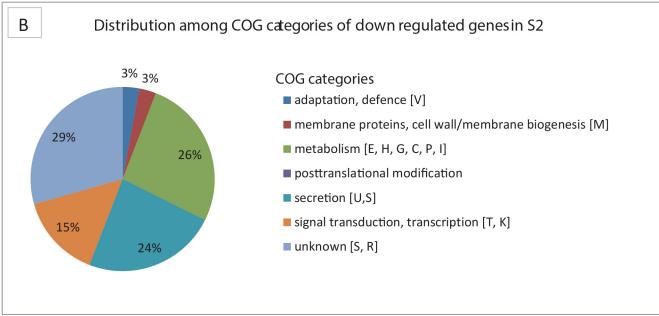


Fig 3. Functional categories of up- and down-regulated genes in S2 relative to S1 detected with Affymetrix transcriptome analysis. A, indicates the percentage of genes per COG category that were higher expressed in S2 compared to S1, while B, indicates the percentage of genes per COG category that were lower expressed in S2 compared to S1. COG category is indicated between brackets.

MigA, PagL and LPS. The early and late isolates displayed different colony morphology with a rougher colony morphology of S2 compared to S1 (Fig 4). Interestingly, a single nucleotide deletion in S2 relative to S1 and a non-synonymous SNP were detected in two genes, migA and pagL respectively, implicated in lipopolysaccharides (LPS) biogenesis. Furthermore, S2 contains a single nucleotide deletion just downstream of pagL, which may explain the observed differences in colony morphology (Tables 1 and 2). The migA gene encodes an Alpha-



Table 2. Indels in isolate S2 relative to isolate S1.

Indel in S2	PA01 annotation	gene	Gene product annotation
GC > GCC (+1C)	PA1302		putative heme utilization protein
GGC > GCGGTCCTGCAACTGC (+13CGGTCCTGCAACT)	PA4967	parE	topoisomerase IV subunit B
GCTGCGGCGC > GCTGCGGCGCGCCCTGCGGCGC (+15CTGCGGCGCCGCCGC)	PA2727		histidine kinase
TGGG > TGG (-1G)	PA4661	downstream of pagL	lipid A 3-O-deacylase
GGGT > G (-3GGT)	PA4379		methyltransferase domain-containing protein
TGGTAGGTA > TGGTA (-4GGTA)	PA3806		putative Fe-S-cluster redox enzyme
TGCCGGCCG > TGCCG (-4GCCG)	PA0347	downstream of glpQ	glycerophosphoryl diester phosphodiesterase
T > TG (+1G)	PA1713	upstream of exsA	AraC family of transcriptional regulators
GT > G (-1T)	PA2172		putative cellulase
CT > C (-1T)	PA0705	migA	alpha-1,6-rhamnosyltransferase

1,6-rhamnosyltransferase involved in producing uncapped core oligosaccharide[63] while pagL encodes a lipid A 3-O-deacylase, which recognizes either 3-OH C10 or 3-OH C14 moieties of the lipid A component of LPS and adjusts the structure of lipid A. Mutations in genes affecting LPS biogenesis have also previously been described in *P. aeruginosa* during CF lung adaptation[9;10;14]. In *P. aeruginosa* isolates from CF respiratory infections the O-antigen is not produced in high amounts. Since O-antigen is highly immunogenic this adaptation possibly facilitates chronic persistence[12]. PagL is overexpressed in *P. aeruginosa* isolates isolated from CF infants compared to those from acute infections or environmental isolates. However, it has also been shown before that in chronic infection with severe lung disease PagL function is lost[64]. Together with our findings this indicates that LPS is under selective pressure in the CF lung and that LPS modifications contribute to adaptation of *P. aeruginosa* to the CF airway.

Biofilm formation and comparison of other phenotypic differences

Analyzing biofilm formation of S1 and S2 demonstrated that, in a semi-static model of biofilm formation, S1 forms significantly thicker biofilms with more biomass than S2 (Fig 5). This coincides with lower expression levels of *bfmR* and *bfmS* genes (S3 Table) in S2. BfmS is part of the two-component regulatory systems named BfmR/S (PA4101/PA4102) that regulates biofilm maturation in *P. aeruginosa*. BfmR/S has been described as essential for biofilm maturation in *P. aeruginosa* by limiting bacteriophage-mediated lysis and thus, eDNA release [65;66]. Other research indicates that BfmS controls the Rhl quorum-sensing system by repressing BfmR, which directly reduces production of quorum-sensing molecules (C4-HSL) by suppressing expression of *rhlR*. Deletion of *bfmS* causes substantially reduced virulence in lettuce leaf, reduced cytotoxicity, enhanced invasion, and reduced bacterial survival during acute mouse lung infection [67]. However, BfmS is able to switch its function from repressor of BfmR to activator through point mutations found in CF-related strains such as the Danish epidemic clone (DK2)[67]. This suggests that under certain conditions reduced expression BfmS and BfmR may indeed coincide with the observed reduced biofilm maturation in S2 relative to S1.

At this moment, however, we can only speculate why loss of biofilm maturation is beneficial in the CF lung. Also, in vitro experiments with surface-attached biofilms may not reflect in the in vivo situation in CF lungs. It has been reported before that other "cooperative traits" like







Fig 4. Colony morphologies on sheep blood agar. S1 (left) has a smooth colony morphology, while S2 (right) has a more rough colony morphology.

secretion of pyoverdin, elastase, protease and some quorum-sensing molecules are also down-regulated in chronic CF isolates, indicating loss of social behavior of isolates that profit from products produced by cooperative strains [68]. In their study, Jiricny and co-workers did not observe that loss of social behavior also resulted in lower-level biofilm formation, but possibly reduced biofilm formation is advantageous at certain stages of chronic infection of the CF lung by saving energy or increasing transmissibility.

Directly adjacent to bfmRS is a cluster of 5 genes (PA4103-PA4107) that is, like bfmRS, also significantly lower expressed in S2 (Fig 6). It is known that BfmR directly binds to its own and to the promoters of PA4107, PA4103, and rhlR and that expression of the operons PA4103-Pa4104 and PA4105-PA4107 is upregulated in CF lungs due to BfmRS activation[67]. The function of these genes is largely unknown with the exception of PA4107, which was recently renamed efhP and found to be important for Ca²⁺ homeostasis and virulence[69]. Levels of Ca²⁺ are increased in CF lungs[70] and in the presence of high Ca²⁺ strains lacking functional EfhP were unable to produce pyocyanin, developed less biofilm, and had decreased resistance to oxidative stress (H2O2).

Besides differences in colony morphology and biofilm formation as described above, other phenotypic characteristics like swimming, swarming, twitching, production of quorum-sensing molecules and growth rate were comparable between S1 and S2 (Table 3). Protease production was slightly higher in S2. Both strains displayed a slow growing phenotype with lack of motility and lower protease production, compared to PA01. This is something that was also reported for other CF adapted strains[68;71;72]. Both S1 and S2 did not exhibit a mucoid phenotype, which is not unexpected. In general, chronically infected CF patients harbor mucoid *P. aeruginosa*, however, mucoidity can be lost by secondary site mutations and mucoid and non-mucoid isolates are often found to co-exist in chronic lung infections[73]. Known virulence-associated phenotypes that are important for *P. aeruginosa* colonization and infection, like motility and



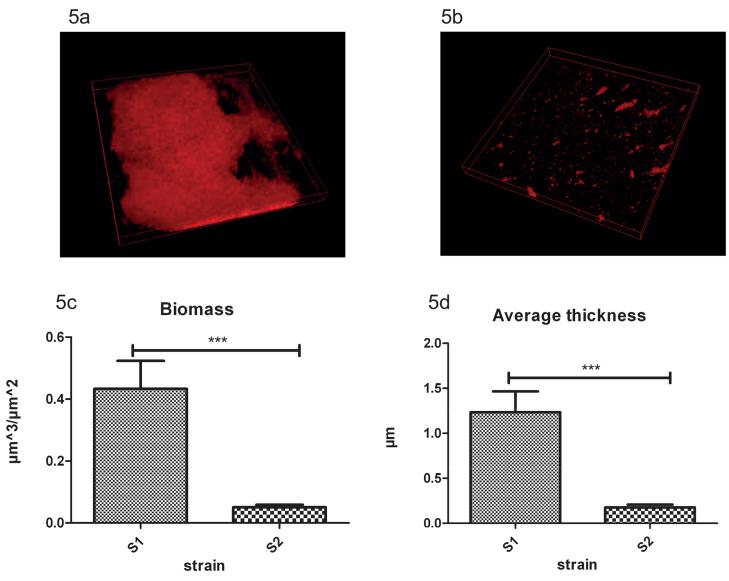


Fig 5. Biofilm formation. Biofilm formation in semi-static model of S1 (a) and S2 (b) after 48 h in LB+1% (wt/vol) glucose; Difference in total biomass (c) and average thickness (d) with Standard deviation error bars. Asterisks represent significant differences (*** P<0.005) as determined by an unpaired two-tailed Student's t-test.

quorum-sensing are not displayed in both the early and late ST406 isolate. Reduced virulence in CF adapted strains is also described in other studies [74;75].

To further quantitatively measure phenotypic traits of S1 and S2, growth rates of S1 and S2 were measured on different substrates using Phenotype MicroArrays. Relative to S1, S2 demonstrated enhanced metabolism in the presence of toxic substances like chloride and bromide detergents, and diverse oxidizing agents, while the S2 isolate displayed decreased growth capabilities on various nutritional supplements like several amino acids and different Carbon (1,2-Propanediol, Adenosine, D-Arabitol, D-Glucosaminic Acid, D-Glucuronic Acid, L-Alaninamide), Phosphor (Dithiophosphate, D-Mannose-6-Phosphate, O-Phospho-D-Serine,





Fig 6. Overview of the gene cluster encompassing genes PA4101-4107. Genes PA4101-4107 are depicted in blue with differences in transcription level between S1 and S2 and adjacent genes in white. Fold changes in expression level in S1 relative to S2 are indicated underneath. PA4103 is a hypothetical gene encoding a putative ferric reductase transmembrane component superfamily, PA4104 is a hypothetical gene encoding a polypeptide with similarities to the DoxX superfamily, PA4105 and PA4106 are hypothetical genes encoding DUF 2063 and DUF 692 superfamily protein, respectively. *EfhP* is the gene originally named PA4107. Arrows indicate direction of transcription.

O-Phospho-L-Serine) and Nitrogen sources (Fig 7 and S5 Table). This indicated that the transition from early to late isolate was accompanied with major metabolic changes. Higher expression of genes involved in metabolism in S1 from transcriptome data matches the phenotypic microarray data. Metabolic changes have also been described before in chronic CF strains[14].

Differences in gene content between the early (S1) and late (S2) isolate

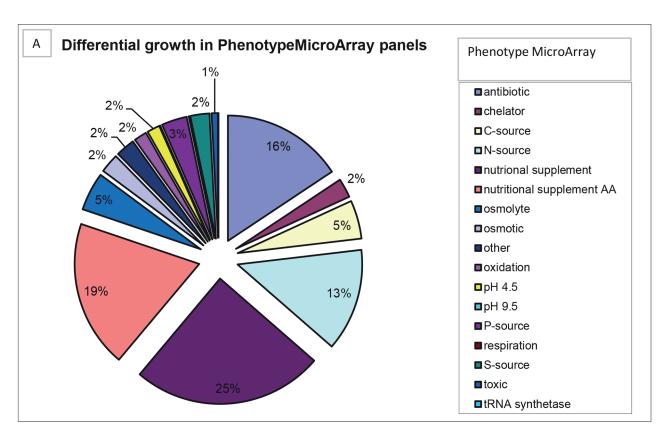
Annotation of the S1 and S2 genomes revealed 5766 protein-encoding genes in S1 and 5772 protein-encoding genes in S2. This cluster of seven genes in S2 was absent in S1. This means that either this cluster was acquired by S2 during chronic colonization or that this insert was already present in the *P. aeruginosa* (sub-) population at the early time point when strain S1 was sampled. These seven genes encode five hypothetical proteins, a putative replication initiation factor and a phage integrase, suggesting incorporation of phage-associated genes in S2. This seven-gene cluster is located next to a tRNA-SeC (p)-TCA gene and does not seem to disrupt a coding region or a predicted operon. These features may suggest that this seven-gene region represents a mobile genetic element, although we have no formal proof for this. On both sides of the 7 genes there is an identical stretch of 36 bp. This partly overlaps with the tRNA gene on the 3' side (Fig. 8). PCR performed on three targets of this 7 kb DNA islet respectively including part of predicted genes S2-1398/1399 (893 bp), S2-1401/1402 (426 bp) and S2-1404 (399 bp), in 29 non-ST406 clinical isolates (6 ICU, 6 community acquired, 6 CF, 6 other clinical isolates from a previous study [26] and 3 isolates of the Liverpool epidemic strain, PA01 and the Midlands epidemic strain) did not result in a positive PCR product, indicating that this insertion is specific for the ST406 S2 and, therefore, has probably no general role in adaptation of *P. aeruginosa* to the lung of CF patients.

Table 3. Phenotypic traits of S1 and S2 compared to PA01.

· · · · · · · · · · · · · · · · · · ·						
phenotypic assay	S1	S2	PA01			
swim	+	+	++			
swarm	-	-	++			
twitch	-	-	++			
protease	+	++	+++			
quorum sensing	-	-	+			
growth rate (doubling time in min)	50	46	29			

doi:10.1371/journal.pone.0158106.t003





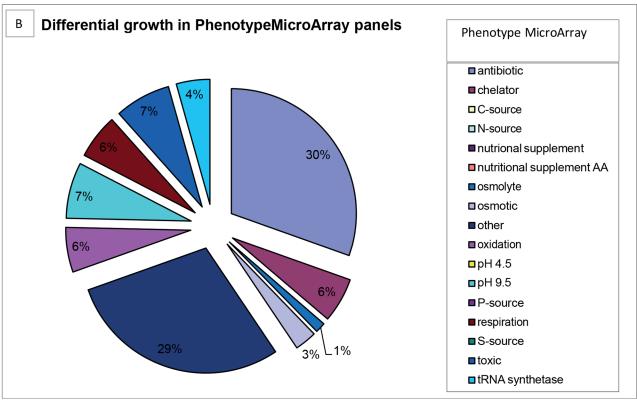




Fig 7. Phenotype MicroArray analysis of the early (S1) and late (S2) ST406 isolate. Indicated are the Phenotype MicroArray panels for which at least a two-fold difference in metabolism (= respiration) between S1 and S2 was detected. A, indicate the percentage of Phenotype MicroArray panels for which the growth rate of S2 was at least two-fold lower than that of S1, while B indicates the percentage of Phenotype MicroArray panels for which the growth rate of S2 was at least two-fold higher than that of S1. Phenotype MicroArray panel categories are color-coded and indicated on the right (C = carbon, S = Sulphur, N = Nitrogen, P = Phosphor).

Conclusion

This study characterized genome-wide and phenotypic changes that have occurred in the Dutch high-prevalent clone ST406 after three years of chronic carriage. The data demonstrate that within-host evolution of the Dutch ST406 clone within a CF patient during three years was not driven by major gene acquisition or gene loss, but mostly by point mutations that coincided with differences in transcriptome levels and phenotype, most notably with differences in biofilm formation. These adaptive evolutionary events in ST406 partly overlap with adaptations described for other chronic CF isolates. In general, clone ST406 seems to have a phenotypic profile of low virulence. Drawing firm conclusions on diversification over time are limited by the fact that in this study only one colony per sputum cultured was used and diversity within the P. aeruginosa population the CF lung at the same time point have been demonstrated[7;13;51-53]. Since the early isolate was recovered from a CF patient just one month after first time colonization with *P. aeruginosa*, diversification at that time was probably low, but the single late isolate might not capture the full diversity of *P. aeruginosa* isolates present at that time point capture the full diversity of *P. aeruginosa* isolates present at that time point [53]. The low virulence profile of the early isolate suggests that at the time of isolation it had already undergone a certain level of adaptation to the CF lung. Common adaptive changes that are repeatedly found in chronically infecting isolates may provide novel leads for targeted therapeutic interventions to combat chronic colonization with *P. aeruginosa* in CF patients.

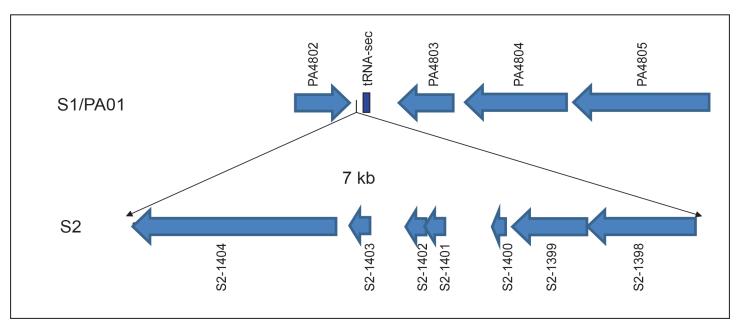


Fig 8. Overview of the seven gene cluster present in the S2 ST406 isolate, but not in the S1 isolate. The 7 kb insert next to tRNAsec and upstream of gene PA4802 detected only in S2 includes a phage integrase gene (S2-1398) and a gene encoding a putative replication initiation factor (S2-1399) and five genes encoding hypothetical proteins.

doi:10.1371/journal.pone.0158106.g008



Supporting Information

S1 Table. Primers and strains used for S2 specific gene amplification. (DOCX)

S2 Table. Published strains used in SNP-based phylogenetic tree. (DOCX)

S3 Table. Differential expressed genes between early and late isolate by Affymetrix (p<0.05 and >2fold change.

(DOCX)

S4 Table. Antibiotic susceptibility of S1 and S2.

(DOCX)

S5 Table. Phenotype MicroArray panels for which at least a two-fold difference in metabolism (= respiration) between S1 and S2 was detected. (XLSX)

Acknowledgments

We would like to thank Prof. S. Molin for his assistance in experimental design and generously providing us lab space and equipment to do transcriptomic and phenotypic experiments.

Author Contributions

Conceived and designed the experiments: RvM MdB FP LY MB RW. Performed the experiments: RvM MdB FP LY. Analyzed the data: RvM MdB FP MB RW. Contributed reagents/materials/analysis tools: MdB FP LY MB RW. Wrote the paper: RvM MdB FP LY MB RW.

References

- Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ, et al. Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. Nature 2000 Aug 31; 406 (6799):959–64. PMID: <u>10984043</u>
- Folkesson A, Jelsbak L, Yang L, Johansen HK, Ciofu O, Hoiby N, et al. Adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis airway: an evolutionary perspective. Nat Rev Microbiol 2012 Dec; 10 (12):841–51. doi: 10.1038/nrmicro2907 PMID: 23147702
- NCFS. Dutch cystic fibrosis registry report on the year 2012.
- Cramer N, Wiehlmann L, Ciofu O, Tamm S, Hoiby N, Tummler B. Molecular epidemiology of chronic Pseudomonas aeruginosa airway infections in cystic fibrosis. PLoS One 2012; 7(11):e50731. doi: 10.1371/journal.pone.0050731 PMID: 23209821
- 5. Cystic fibrosis foundation. Patient Registry annual data report 2012.
- Dettman JR, Rodrigue N, Aaron SD, Kassen R. Evolutionary genomics of epidemic and nonepidemic strains of *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A 2013 Dec 24; 110(52):21065–70. doi: 10.1073/pnas.1307862110 PMID: 24324153
- Mowat E, Paterson S, Fothergill JL, Wright EA, Ledson MJ, Walshaw MJ, et al. Pseudomonas aeruginosa population diversity and turnover in cystic fibrosis chronic infections. Am J Respir Crit Care Med 2011 Jun 15; 183(12):1674–9. doi: 10.1164/rccm.201009-1430OC PMID: 21297072
- Smith EE, Buckley DG, Wu Z, Saenphimmachak C, Hoffman LR, D'Argenio DA, et al. Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. Proc Natl Acad Sci U S A 2006 May 30; 103(22):8487–92. PMID: 16687478
- Cramer N, Klockgether J, Wrasman K, Schmidt M, Davenport CF, Tummler B. Microevolution of the major common *Pseudomonas aeruginosa* clones C and PA14 in cystic fibrosis lungs. Environ Microbiol 2011 Jul; 13(7):1690–704. doi: 10.1111/j.1462-2920.2011.02483.x PMID: 21492363



- 10. Marvig RL, Johansen HK, Molin S, Jelsbak L. Genome analysis of a transmissible lineage of *Pseudo-monas aeruginosa* reveals pathoadaptive mutations and distinct evolutionary paths of hypermutators. PLoS Genet 2013; 9(9):e1003741. doi: 10.1371/journal.pgen.1003741 PMID: 24039595
- Hogardt M, Heesemann J. Microevolution of Pseudomonas aeruginosa to a chronic pathogen of the cystic fibrosis lung. Curr Top Microbiol Immunol 2013; 358:91–118. doi: 10.1007/82 2011 199 PMID: 22311171
- Hauser AR, Jain M, Bar-Meir M, McColley SA. Clinical significance of microbial infection and adaptation in cystic fibrosis. Clin Microbiol Rev 2011 Jan; 24(1):29–70. doi: 10.1128/CMR.00036-10 PMID: 21233507
- Marvig RL, Sommer LM, Jelsbak L, Molin S, Johansen HK. Evolutionary insight from whole-genome sequencing of Pseudomonas aeruginosa from cystic fibrosis patients. Future Microbiol 2015; 10 (4):599–611. doi: 10.2217/fmb.15.3 PMID: 25865196
- Yang L, Jelsbak L, Marvig RL, Damkiaer S, Workman CT, Rau MH, et al. Evolutionary dynamics of bacteria in a human host environment. Proc Natl Acad Sci U S A 2011 May 3; 108(18):7481–6. doi: 10.1073/pnas.1018249108 PMID: 21518885
- Rau MH, Marvig RL, Ehrlich GD, Molin S, Jelsbak L. Deletion and acquisition of genomic content during early stage adaptation of *Pseudomonas aeruginosa* to a human host environment. Environ Microbiol 2012 Aug; 14(8):2200–11. doi: 10.1111/j.1462-2920.2012.02795.x PMID: 22672046
- 16. Klockgether J, Miethke N, Kubesch P, Bohn YS, Brockhausen I, Cramer N, et al. Intraclonal diversity of the *Pseudomonas aeruginosa* cystic fibrosis airway isolates TBCF10839 and TBCF121838: distinct signatures of transcriptome, proteome, metabolome, adherence and pathogenicity despite an almost identical genome sequence. Environ Microbiol 2013 Jan; 15(1):191–210. doi: 10.1111/j.1462-2920. 2012.02842.x PMID: 22882573
- 17. Winstanley C, Langille MG, Fothergill JL, Kukavica-Ibrulj I, Paradis-Bleau C, Sanschagrin F, et al. Newly introduced genomic prophage islands are critical determinants of in vivo competitiveness in the Liverpool Epidemic Strain of *Pseudomonas aeruginosa*. Genome Res 2009 Jan; 19(1):12–23. doi: 10.101/gr.086082.108 PMID: 19047519
- 18. Al-Aloul M, Crawley J, Winstanley C, Hart CA, Ledson MJ, Walshaw MJ. Increased morbidity associated with chronic infection by an epidemic *Pseudomonas aeruginosa* strain in CF patients. Thorax 2004 Apr; 59(4):334–6. PMID: 15047956
- Jones AM, Dodd ME, Doherty CJ, Govan JR, Webb AK. Increased treatment requirements of patients with cystic fibrosis who harbour a highly transmissible strain of *Pseudomonas aeruginosa*. Thorax 2002 Nov; 57(11):924–5. PMID: 12403871
- O'Carroll MR, Syrmis MW, Wainwright CE, Greer RM, Mitchell P, Coulter C, et al. Clonal strains of Pseudomonas aeruginosa in paediatric and adult cystic fibrosis units. Eur Respir J 2004 Jul; 24 (1):101–6. PMID: <u>15293611</u>
- Jones AM, Dodd ME, Morris J, Doherty C, Govan JR, Webb AK. Clinical outcome for cystic fibrosis patients infected with transmissible *Pseudomonas aeruginosa*: an 8-year prospective study. Chest 2010 Jun; 137(6):1405–9. doi: 10.1378/chest.09-2406 PMID: 20081099
- Tingpej P, Elkins M, Rose B, Hu H, Moriarty C, Manos J, et al. Clinical profile of adult cystic fibrosis patients with frequent epidemic clones of *Pseudomonas aeruginosa*. Respirology 2010 Aug; 15 (6):923–9. doi: 10.1111/j.1440-1843.2010.01792.x PMID: 20573059
- Fothergill JL, Walshaw MJ, Winstanley C. Transmissible strains of *Pseudomonas aeruginosa* in cystic fibrosis lung infections. Eur Respir J 2012 Jul; 40(1):227–38. doi: 10.1183/09031936.00204411 PMID: 22323572
- 24. Tramper-Stranders GA, van der Ent CK, Molin S, Yang L, Hansen SK, Rau MH, et al. Initial Pseudomonas aeruginosa infection in patients with cystic fibrosis: characteristics of eradicated and persistent isolates. Clin Microbiol Infect 2012 Jun; 18(6):567–74. doi: 10.1111/j.1469-0691.2011.03627.x PMID: 21883670
- 25. van Mansfeld R, Willems R, Brimicombe R, Heijerman H, van Berkhout FT, Wolfs T, et al. Pseudomonas aeruginosa genotype prevalence in Dutch cystic fibrosis patients and age dependency of colonization by various P. aeruginosa sequence types. J Clin Microbiol 2009 Dec; 47(12):4096–101. doi: 1128/JCM.01462-09 PMID: 19828746
- 26. van Mansfeld R, Jongerden I, Bootsma M, Buiting A, Bonten M, Willems R. The population genetics of Pseudomonas aeruginosa isolates from different patient populations exhibits high-level host specificity. PLoS One 2010; 5(10):e13482. doi: 10.1371/journal.pone.0013482 PMID: 20976062
- Tramper-Stranders GA, van der Ent CK, Wolfs TF, Kimpen JL, Fleer A, Johansen U, et al. Pseudomonas aeruginosa diversity in distinct paediatric patient groups. Clin Microbiol Infect 2008 Aug 26.
- Wiehlmann L, Wagner G, Cramer N, Siebert B, Gudowius P, Morales G, et al. Population structure of Pseudomonas aeruginosa. Proc Natl Acad Sci U S A 2007 May 8; 104(19):8101–6. PMID: <u>17468398</u>



- 29. de Vrankrijker AM, Brimicombe RW, Wolfs TF, Heijerman HG, van MR, van Berkhout FT, et al. Clinical impact of a highly prevalent *Pseudomonas aeruginosa* clone in Dutch cystic fibrosis patients. Clin Microbiol Infect 2011 Mar; 17(3):382–5. doi: 10.1111/j.1469-0691.2010.03295.x PMID: 20807225
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 2012 May; 19(5):455– 77. doi: 10.1089/cmb.2012.0021 PMID: 22506599
- Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics 2014 Jul 15; 30(14):2068–9. doi: 10.1093/bioinformatics/btu153 PMID: 24642063
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 2008; 9:75. doi: 10.1186/1471-2164-9-75 PMID: 18261238
- **33.** Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 2000 Jan 1; 28(1):33–6. PMID: 10592175
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997 Sep 1; 25 (17):3389–402. PMID: 9254694
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012 Apr; 9 (4):357–9. doi: 10.1038/nmeth.1923 PMID: 22388286
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009 Aug 15; 25(16):2078–9. doi: 10.1093/bioinformatics/btp352 PMID: 19505943
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. Genome Biol 2004; 5(2):R12. PMID: 14759262
- Stamatakis A, Ludwig T, Meier H. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics 2005 Feb 15; 21(4):456–63. PMID: 15608047
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 2013 Dec; 30(12):2725–9. doi: 10.1093/molbev/mst197 PMID: 24132122
- Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003 Jan 22; 19(2):185–93.
 PMID: 12538238
- 41. Winsor GL, Lam DK, Fleming L, Lo R, Whiteside MD, Yu NY, et al. Pseudomonas Genome Database: improved comparative analysis and population genomics capability for Pseudomonas genomes. Nucleic Acids Res 2011 Jan; 39(Database issue):D596–D600. doi: 10.1093/nar/gkq869 PMID: 20929876
- **42.** Hentzer M, Riedel K, Rasmussen TB, Heydorn A, Andersen JB, Parsek MR, et al. Inhibition of quorum sensing in *Pseudomonas aeruginosa* biofilm bacteria by a halogenated furanone compound. Microbiology 2002 Jan; 148(Pt 1):87–102. PMID: <u>11782502</u>
- 43. Paganelli FL, Willems RJ, Jansen P, Hendrickx A, Zhang X, Bonten MJ, et al. Enterococcus faecium biofilm formation: identification of major autolysin AtlAEfm, associated Acm surface localization, and AtlAEfm-independent extracellular DNA Release. MBio 2013; 4(2):e00154. doi: 10.1128/mBio.00154-13 PMID: 23592262
- 44. Heydorn A, Nielsen AT, Hentzer M, Sternberg C, Givskov M, Ersboll BK, et al. Quantification of biofilm structures by the novel computer program COMSTAT. Microbiology 2000 Oct; 146 (Pt 10):2395–407. PMID: 11021916
- **45.** Pirnay JP, De VD, Cochez C, Bilocq F, Vanderkelen A, Zizi M, et al. *Pseudomonas aeruginosa* displays an epidemic population structure. Environ Microbiol 2002 Dec; 4(12):898–911. PMID: <u>12534471</u>
- Pirnay JP, Bilocq F, Pot B, Cornelis P, Zizi M, Van EJ, et al. Pseudomonas aeruginosa population structure revisited. PLoS One 2009; 4(11):e7740. doi: 10.1371/journal.pone.0007740 PMID: 19936230
- **47.** Pirnay JP, Matthijs S, Colak H, Chablain P, Bilocq F, Van EJ, et al. Global *Pseudomonas aeruginosa* biodiversity as reflected in a Belgian river. Environ Microbiol 2005 Jul; 7(7):969–80. PMID: 15946293
- 48. Curran B, Jonas D, Grundmann H, Pitt T, Dowson CG. Development of a multilocus sequence typing scheme for the opportunistic pathogen *Pseudomonas aeruginosa*. J Clin Microbiol 2004 Dec; 42 (12):5644–9. PMID: 15583294
- 49. Maatallah M, Cheriaa J, Backhrouf A, Iversen A, Grundmann H, Do T, et al. Population structure of Pseudomonas aeruginosa from five Mediterranean countries: evidence for frequent recombination and epidemic occurrence of CC235. PLoS One 2011; 6(10):e25617. doi: 10.1371/journal.pone.0025617 PMID: 21984923



- 50. Grosso-Becerra MV, Santos-Medellin C, Gonzalez-Valdez A, Mendez JL, Delgado G, Morales-Espinosa R, et al. *Pseudomonas aeruginosa* clinical and environmental isolates constitute a single population with high phenotypic diversity. BMC Genomics 2014; 15:318. doi: 10.1186/1471-2164-15-318 PMID: 24773920
- Winstanley C, O'Brien S, Brockhurst MA. Pseudomonas aeruginosa Evolutionary Adaptation and Diversification in Cystic Fibrosis Chronic Lung Infections. Trends Microbiol 2016 Feb 10.
- Ashish A, Paterson S, Mowat E, Fothergill JL, Walshaw MJ, Winstanley C. Extensive diversification is a common feature of *Pseudomonas aeruginosa* populations during respiratory infections in cystic fibrosis. J Cyst Fibros 2013 Dec; 12(6):790–3. doi: 10.1016/j.jcf.2013.04.003 PMID: 23642644
- Chung JC, Becq J, Fraser L, Schulz-Trieglaff O, Bond NJ, Foweraker J, et al. Genomic variation among contemporary *Pseudomonas aeruginosa* isolates from chronically infected cystic fibrosis patients. J Bacteriol 2012 Sep; 194(18):4857–66. doi: 10.1128/JB.01050-12 PMID: 22753054
- 54. Biebricher CK, Eigen M. What is a quasispecies? Curr Top Microbiol Immunol 2006; 299:1–31. PMID: 16568894
- 55. Bleves S, Viarre V, Salacha R, Michel GP, Filloux A, Voulhoux R. Protein secretion systems in Pseudo-monas aeruginosa: A wealth of pathogenic weapons. Int J Med Microbiol 2010 Dec; 300(8):534–43. doi: 10.1016/j.ijmm.2010.08.005 PMID: 20947426
- 56. Lee VT, Smith RS, Tummler B, Lory S. Activities of Pseudomonas aeruginosa effectors secreted by the Type III secretion system in vitro and during infection. Infect Immun 2005 Mar; 73(3):1695–705. PMID: 15731070
- 57. Harmer C, Alnassafi K, Hu H, Elkins M, Bye P, Rose B, et al. Modulation of gene expression by Pseudomonas aeruginosa during chronic infection in the adult cystic fibrosis lung. Microbiology 2013 Nov; 159(Pt 11):2354–63. doi: 10.1099/mic.0.066985-0 PMID: 24014664
- Kohler T, Epp SF, Curty LK, Pechere JC. Characterization of MexT, the regulator of the MexE-MexF-OprN multidrug efflux system of *Pseudomonas aeruginosa*. J Bacteriol 1999 Oct; 181(20):6300–5. PMID: 10515918
- 59. Tian ZX, Fargier E, Mac AM, Adams C, Wang YP, O'Gara F. Transcriptome profiling defines a novel regulon modulated by the LysR-type transcriptional regulator MexT in *Pseudomonas aeruginosa*. Nucleic Acids Res 2009 Dec; 37(22):7546–59. doi: 10.1093/nar/gkp828 PMID: 19846594
- 60. Tian ZX, Mac AM, O'Connor HF, Fargier E, Mooij MJ, Adams C, et al. MexT modulates virulence determinants in *Pseudomonas aeruginosa* independent of the MexEF-OprN efflux pump. Microb Pathog 2009 Oct; 47(4):237–41. doi: 10.1016/j.micpath.2009.08.003 PMID: 19683048
- Olivares J, Alvarez-Ortega C, Linares JF, Rojo F, Kohler T, Martinez JL. Overproduction of the multidrug efflux pump MexEF-OprN does not impair *Pseudomonas aeruginosa* fitness in competition tests, but produces specific changes in bacterial regulatory networks. Environ Microbiol 2012 Aug; 14 (8):1968–81. doi: 10.1111/j.1462-2920.2012.02727.x PMID: 22417660
- 62. Luong PM, Shogan BD, Zaborin A, Belogortseva N, Shrout JD, Zaborina O, et al. Emergence of the P2 phenotype in *Pseudomonas aeruginosa* PAO1 strains involves various mutations in mexT or mexF. J Bacteriol 2014 Jan: 196(2):504–13. doi: 10.1128/JB.01050-13 PMID: 242444000
- Poon KK, Westman EL, Vinogradov E, Jin S, Lam JS. Functional characterization of MigA and WapR: putative rhamnosyltransferases involved in outer core oligosaccharide biosynthesis of *Pseudomonas* aeruginosa. J Bacteriol 2008 Mar; 190(6):1857–65. doi: 10.1128/JB.01546-07 PMID: 18178733
- 64. Ernst RK, Adams KN, Moskowitz SM, Kraig GM, Kawasaki K, Stead CM, et al. The *Pseudomonas aeru-ginosa* lipid A deacylase: selection for expression and loss within the cystic fibrosis airway. J Bacteriol 2006 Jan; 188(1):191–201. PMID: 16352835
- **65.** Petrova OE, Schurr JR, Schurr MJ, Sauer K. The novel *Pseudomonas aeruginosa* two-component regulator BfmR controls bacteriophage-mediated lysis and DNA release during biofilm development through PhdA. Mol Microbiol 2011 Aug; 81(3):767–83. doi: 10.1111/j.1365-2958.2011.07733.x PMID: 21696457
- 66. Petrova OE, Sauer K. A novel signaling network essential for regulating *Pseudomonas aeruginosa* bio-film development. PLoS Pathog 2009 Nov; 5(11):e1000668. doi: 10.1371/journal.ppat.1000668 PMID: 19936057
- 67. Cao Q, Wang Y, Chen F, Xia Y, Lou J, Zhang X, et al. A novel signal transduction pathway that modulates rhl quorum sensing and bacterial virulence in *Pseudomonas aeruginosa*. PLoS Pathog 2014 Aug; 10(8):e1004340. doi: 10.1371/journal.ppat.1004340 PMID: 25166864
- 68. Jiricny N, Molin S, Foster K, Diggle SP, Scanlan PD, Ghoul M, et al. Loss of social behaviours in populations of *Pseudomonas aeruginosa* infecting lungs of patients with cystic fibrosis. PLoS One 2014; 9(1): e83124. doi: 10.1371/journal.pone.0083124 PMID: 24454693



- 69. Sarkisova SA, Lotlikar SR, Guragain M, Kubat R, Cloud J, Franklin MJ, et al. A Pseudomonas aeruginosa EF-Hand Protein, EfhP (PA4107), Modulates Stress Responses and Virulence at High Calcium Concentration. PLoS One 2014; 9(2):e98985. doi: 10.1371/journal.pone.0098985 PMID: 24918783
- Halmerbauer G, Arri S, Schierl M, Strauch E, Koller DY. The relationship of eosinophil granule proteins to ions in the sputum of patients with cystic fibrosis. Clin Exp Allergy 2000 Dec; 30(12):1771–6. PMID: 11122216
- Jeukens J, Boyle B, Kukavica-Ibrulj I, Ouellet MM, Aaron SD, Charette SJ, et al. Comparative genomics
 of isolates of a *Pseudomonas aeruginosa* epidemic strain associated with chronic lung infections of cystic fibrosis patients. PLoS One 2014; 9(2):e87611. doi: 10.1371/journal.pone.0087611 PMID:
 24505294
- 72. Manos J, Hu H, Rose BR, Wainwright CE, Zablotska IB, Cheney J, et al. Virulence factor expression patterns in *Pseudomonas aeruginosa* strains from infants with cystic fibrosis. Eur J Clin Microbiol Infect Dis 2013 Dec; 32(12):1583–92. doi: 10.1007/s10096-013-1916-7 PMID: 23832143
- 73. Bragonzi A, Paroni M, Nonis A, Cramer N, Montanari S, Rejman J, et al. *Pseudomonas aeruginosa* microevolution during cystic fibrosis lung infection establishes clones with adapted virulence. Am J Respir Crit Care Med 2009 Jul 15; 180(2):138–45. doi: 10.1164/rccm.200812-1943OC PMID: 19423715
- 74. Lore NI, Cigana C, De F, I, Riva C, Juhas M, Schwager S, et al. Cystic fibrosis-niche adaptation of Pseudomonas aeruginosa reduces virulence in multiple infection hosts. PLoS One 2012; 7(4):e35648. doi: 10.1371/journal.pone.0035648 PMID: 22558188
- 75. Jelsbak L, Johansen HK, Frost AL, Thogersen R, Thomsen LE, Ciofu O, et al. Molecular epidemiology and dynamics of *Pseudomonas aeruginosa* populations in lungs of cystic fibrosis patients. Infect Immun 2007 May; 75(5):2214–24. PMID: 17261614