Pseudomonas aeruginosa host-adaptation in cystic fibrosis patients

Rau, Martin Holm

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Pseudomonas aeruginosa host-adaptation in cystic fibrosis patients

Ph.D. Thesis
by
Martin Holm Rau

Center for Systems Microbiology
Department of Systems Biology
Technical University of Denmark
March 2012
Contents

1 Preface ............................................................................................................................ v

2 Summary ............................................................................................................................ vii

3 Dansk resumé ................................................................................................................... ix

4 Publication list .................................................................................................................. xi

5 Introduction ......................................................................................................................... 1
  5.1 Cystic fibrosis ................................................................................................................. 1
    5.1.1 Cystic fibrosis lower airway environment ................................................................. 3
    5.1.2 Environment of the cystic fibrosis paranasal sinuses ................................................. 5
    5.1.3 P. aeruginosa infection progression in CF patients ................................................. 6
  5.2 Pseudomonas aeruginosa ............................................................................................... 8
    5.2.1 Virulence factors .......................................................................................................... 9
    5.2.2 Antibiotic resistance ................................................................................................... 10
    5.2.3 Quorum sensing ......................................................................................................... 11
    5.2.4 The mucoid phenotype .............................................................................................. 12
  5.3 The P. aeruginosa genome – a dynamic genome ........................................................... 15
    5.3.1 The accessory genome ............................................................................................... 15
    5.3.2 Genetic adaptation during CF infection ...................................................................... 17
    5.3.3 Evolutionary dynamics of infecting lineages ............................................................. 20
    5.3.4 Horizontal gene transfer ............................................................................................ 22
  5.4 Bacterial host-restriction adaptation ............................................................................ 24

6 Concluding remarks ......................................................................................................... 29

7 References ......................................................................................................................... 31

Paper 1: Early adaptive developments of Pseudomonas aeruginosa after the transition from life in the environment to persistent colonization in the airways of human cystic fibrosis hosts.

Paper 2: Evolution and diversification of Pseudomonas aeruginosa in the paranasal sinuses of cystic fibrosis children have implications for chronic lung infection

Paper 3: Deletion and acquisition of genomic content during early stage adaptation of Pseudomonas aeruginosa to a human host environment.
1 Preface

This thesis is submitted as a partial fulfilment of the requirements to obtain the Ph.D. degree at the Technical University of Denmark (DTU). The work was carried out in the Infection Microbiology Group, Center for Systems Microbiology, Department of Systems Biology at DTU under the supervision of Prof. Søren Molin. The work was performed from January 2009 to March 2012. This Ph.D. project was co-financed by DTU and the Lundbeck Foundation.

First of all I would like to express my gratitude to my supervisor Søren Molin for being a true inspiration, always having a positive mind-set and for excellent guidance throughout. The working environment has been outstanding not least because of the many international input but also the friendly atmosphere within the group. I would also like to thank especially Susse K. Hansen for your time and patience, for the many talks and for teaching me many important scientific skills. Further I would like to thank Lars Jelsbak for a friendly attitude, rewarding discussion and for showing an interest in my research. Thank you to Anders Folkesson and Claus Sternberg too for good discussions and helpful advice.

Of course also thank you to all the other IMG Ph.D. students, previous and present. To Lei Yang for being a pleasant conference travelling partner and for all the talks late in the day. To Rasmus Marvig for sharing bioinformatic frustrations and for great suggestions and to Vinoth, Nicholas, Søren D, Martin W, Liang, Yang, Maria, Trine and Juliane for being helpful and friendly colleagues. Importantly, also a big thanks to all the foreign guests coming and going (and coming and ...) throughout my stay. Fatima, thank you for all the chatting and time spent together and of course for being salsa champions. Marilo, it was fun times at various alternative places, come by DK more often. Thank you also to Adela, Inna, Jeff, Anna, Cristina, Julie and Gerdien for the parties and joyful company.

During my research stay at Stanford University I was lucky to work with Sünje J. Pamp. I would like to thank you for your kindness but also for scientific inspiration. Thanks to you and Janus for trips and get togethers as well. A further thank you to the students I was blessed with the supervision of, especially my favourite of course, Carlotta, for your positive attitude, great ideas and for teaching me the most important Italian words. Further thanks to Helle K.
Johansen, Niels Høiby and Oana Ciofu for helpful discussion and for enabling us all to carry out this exciting research and to Tove Johansen, Lisser St. Clair Norton, Mette Munk and Susanne Koefoed for helping hands when in need.

Last but not least I would like to thank my family and friends for support during the three years of ups and downs.

Martin Holm Rau
Kgs. Lyngby
March 2012
2 Summary

*Pseudomonas aeruginosa* is an opportunistic pathogen capable of transition from an environmental lifestyle to a host-associated lifestyle, as exemplified in the life-long airway infection of cystic fibrosis (CF) patients. Long-term infection is associated with extensive genetic adaptation of *P. aeruginosa* towards the CF airway environment generating variants with markedly altered phenotypes. Gaining insight into this adaptation process has great clinical relevance but simultaneously has the potential to increase our understanding of bacterial adaptation to a host environment. This has been the framework upon which this thesis is based.

Early *P. aeruginosa* colonization of the CF airways is the period in which the outcome of infection is determined, i.e. if the bacteria are eventually eradicated or persist. In three patient cases the evolutionary events from initiation of infection were explored to unravel the early adaptive processes possibly securing bacterial persistence. In this early stage, clinical isolates displayed few adaptive events however these included phenotypes often observed in late chronic infection isolates including the conversion to a mucoid phenotype and increased antibiotic resistance. Detailed characterization of the mucoid phenotype revealed profound pleiotropic effects such as reduction of virulence factors and the Rhl quorum sensing system. Hence the mucoid phenotype in itself encompasses many of the traits associated with chronic infection isolates.

In many CF patients intermittent colonization with *P. aeruginosa* is characterized by the recurrent detection of the same genotype with several months apart. Comparing patient sinus and lung isolates it was found that for some patients the paranasal sinuses constitute an important niche for *P. aeruginosa* during intermittent colonization. In fact, evolution within the sinus population gives rise to variants known to be adaptive to the CF lung environment including mucoid, antibiotic resistant and small colony variants. The likely downwards direction of migration between sinuses and lower airways at this stage suggests that the paranasal sinus population in these patients has the potential for seeding the lungs with pre-adapted variants and thus increase the risk of progression to chronic lung infection.

Long-term infection of *P. aeruginosa* in CF patients can be compared to the early stage of host-restriction adaptation. One example of CF long-term infection is the dominant clone, the DK2 clone type, infecting CF patients at the Copenhagen CF clinic for more than 35 years. Characterization of the genome evolution of DK2 isolates during host-adaptation revealed
substantial genome reduction; in one isolate 8 % of the genome was deleted. Thus large portions of the *P. aeruginosa* genome is dispensable for proliferation within the CF lung. The deletions were mediated by illegitimate and homologous recombination but were not insertion sequence (IS) element mediated as previously proposed for early stage host adaptation. The putative entry of a prophage in one isolate was the only sign of DNA uptake during infection revealing little importance of horizontal gene transfer (HGT) in the adaptation towards the CF lung and further demonstrating the general signal of loss of function compared to gain of function during adaptation to the CF lung.

The transition of *P. aeruginosa* from its natural aquatic and terrestrial habitats into the very different CF lung environment containing multiple stresses usually not encountered by the bacteria should pose a number of challenges. Yet, *P. aeruginosa* displays an inherent capability for successful colonization of the CF lung with persistence seemingly mediated by only a few adaptive events.
3 Dansk resumé

_Pseudomonas aeruginosa_ er en opportunistisk patogen som kan skifte fra naturen til en værts-associoteret levevis, eksemplificeret ved livslang infektion af luftvejene i cystisk fibrose (CF) patienter. Langvarig infektion er associateret med udstrakt genetisk adaptation af _P. aeruginosa_ til CF luftvejs-miljøet, hvilket generer varianter med markant forandrede fænotyper. At opnå indsigt i denne adaptations proces har stor klinisk relevans, men ligeledes potentialen til at forøge vores forståelse af bakteriel adaptation til et værtsmiljø. Dette er baggrunden for denne afhandling.

Den tidlige kolonisering med _P. aeruginosa_ i CF Luftvejene udgør den periode, som afgør hvordan infektionen falder ud, dvs. udryddelse eller persistering. De evolutionære hændelser i tre patienter blev undersøgt fra infektionens påbegyndelse, for at udrede de tidlige adaptive processer som potentielte sikrer bakteriel persistering. De kliniske isolater udviste i dette tidlige stadie få adaptive hændelser, til gengæld inkluderede disse, fænotyper ofte observeret i isolater fra sen kronisk infektion, inklusive konvertering til en mukoide fænotype samt forøget antibiotika resistens. En detaljeret karakterisering af den mukoide fænotype afslørede omfangsrigte pleiotropiske effekter, såsom reduktion af virulens faktorer samt Rhl quorum sensing systemet. Følgelig omfatter den mukoide fænotype i sig selv adskillige af de karaktertræk, som normalt er associateret med isolater fra kronisk infektion.

For mange CF patienter er intermitterende kolonisering med _P. aeruginosa_ karakteriseret ved gentagen detektering af den samme genotype, tidsmæssigt adskilt af mange måneder. Sammenligning af bihule og lunge isolate afslørede, at bihulerne for visse patienter udfør en vigtig niche for _P. aeruginosa_ i løbet af intermitterende kolonisering. Faktisk medfører evolution blandt bihule populationen dannelse af varianter kendt som værende adaptive overfor CF lungve miljøet, inklusive mukoide, antibiotika resistente samt små koloni-varianter. Den sandsynlige nedadgående migrering fra bihuler til lunger i dette infektionsstadie, antyder at bihule populationen i disse patienter potentielt kan tilføre lungerne præ-adapterede varianter og dermed forøge risikoen for udvikling af kronisk lunge infektion.

Langvarig _P. aeruginosa_ infektion i CF patienter kan sammenlignes med det tidlige stadie af adaptation til bakteriel restriktion til et værtsmiljø. Den dominante klon, DK2 klon typen, med en infektions historik på mere end 35 år udfør et eksempel på langvarig CF infektion. Karakterisering af evolutionen af genomet i DK2 isolater i løbet af deres værts
adaptering afslørede betragtelig reducering af genomet; i et isolat var 8 % af genomet forsvundet. Store bestanddele af *P. aeruginosa* genomet er dermed overflødigt med hensyn til vækst i CF lungen. Deleterings mekanismen bestod af illegitim samt homolog rekombination, men ikke af insertion sequence (IS) element rekombination som tidligere foreslået for tidlig værts adaptation. Det eneste tegn på DNA optag under infektionen var et potentielt optag af en pro-fag i et enkelt isolat, hvilket påviser en begrænset betydning af horizontal genoverførsel i adaptationen til CF ungemiljøet. Endvidere tjener det som demonstration på den generelle observation af tab af funktion, modsat erhverv af ny funktion i adaptationen til CF ungemiljøet for *P. aeruginosa*.

Overgangen fra dets naturlige habitat mod det særdeles anderledes CF ungemiljø indeholdende adskillige stress påvirkninger som *P. aeruginosa* normalt ikke møder, må formodes at præsentere væsentlige udfordringer for bakterien. Dog ses det at *P. aeruginosa* udviser en immanent egnerheit til succesfuld kolonisering af CF ungemiljøet, hvor persistens umiddelbart kan opnås via ganske få adaptive hændelser.
4 Publication list


Not included in this thesis


5 Introduction

5.1 Cystic fibrosis

Cystic fibrosis is the most common lethal autosomal recessive disease in Caucasian populations affecting 1 in 2500 newborns. It is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene (Riordan et al., 1989) encoding a chloride channel regulating the transport of ions and water in epithelial cells. The protein is expressed in respiratory, gastrointestinal, hepatobiliary and reproductive systems (Lyczak et al., 2002) and CFTR mutation consequently causes abnormalities in these organs such as poor nutrient absorption, male infertility and airway infections, the latter causing a median age of survival around 37 years for CF patients. The most frequent mutation (70-80% of all CF patients) is an in-frame deletion at position 508 (ΔF508) which concurrently is one of the most severe mutation types in terms of clinical manifestation (Koch and Hoiby, 1993). A defective CFTR channel leads to thick, sticky mucus caused by hyper-absorption of Na⁺ and Cl⁻ ions and subsequent water uptake by osmosis (Figure 1) (Rowe et al., 2005).

![Figure 1: Model of airway epithelia. (A) Normal airway epithelia with hydrated mucus and functional mucociliary clearance. (B) Dehydration of CF airway epithelia leads to](image-url)
malfunctional mucociliary clearance and (C) thick dehydrated mucus (Worlitzsch et al., 2002).

Thick mucus impairs the defense mechanism of mucociliary clearance of microbes enabling these to establish lifelong infections in the lungs. This leads to excessive inflammation and together these are the factors responsible for the morbidity and mortality of cystic fibrosis disease. Microbes frequently isolated from CF patients include *P. aeruginosa, Staphylococcus aureus* and *Haemophilus influenzae* (Figure 2).

![Figure 2: Age specific prevalence of bacterial species in cystic fibrosis patients (Cystic Fibrosis Foundation annual report 2009).](image-url)

*P. aeruginosa* is the main pathogen of CF patients and infection is a poor prognostic indicator and associated with lung function decline (Nixon et al., 2001; Li et al., 2005). One reason is its capability of proliferation in the CF lung - up to $10^9$ cells per ml in CF sputum samples are found.
5.1.1 Cystic fibrosis lower airway environment

The cystic fibrosis lung environment comprises a complex heterogeneous environment for microorganisms to adapt to. Apart from spatial heterogeneity it presents variation in nutrients, oxygen concentration, immune system defences and antibiotics and the considerable phenotypic diversity exhibited by P. aeruginosa CF clinical isolates (Govan and Deretic, 1996) provides a possible reflection of this heterogeneity. The lower airways can be partitioned into two distinct compartments, the conductive zone and the respiratory zone (Figure 3).

![Figure 3: The conductive and respiratory zones of the lungs (Høiby et al., 2010).](image-url)
The conductive zone constitutes about 5% of the lung volume and includes trachea, bronchi and terminal bronchioles. It contains cilia for mucociliary clearing and goblet cells that together with submucosal glands are responsible for mucus production. It is possible to reach high concentrations of antibiotics in the sputum of the conductive zone by nebulization (Høiby et al., 2010). The respiratory zone constitutes 95% of lung volume and contains respiratory bronchioles and alveolar ducts but not cilia or mucus producing cells. Only a thin barrier separates the blood from the air and intravenous antibiotic administration can provide high concentrations in the tissue of the respiratory zone (Høiby et al., 2010). Aerobic conditions are present in the respiratory zone unless mucus plugs restricts air passage creating microaerobic conditions.

Microbial growth in the conductive zone is usually restricted to the thick mucus containing gradients of oxygen ranging from aerobic through hypoxic to anaerobic (Worlitzsch et al., 2002). The anaerobic condition affects the antibiotic efficacy negatively and can sustain microbial growth through the presence of sufficient concentrations of nitrate utilized by P. aeruginosa as an electron acceptor (Palmer et al., 2007). Importantly, the CF lung is a nutrient rich environment with a nutritional composition of sputum containing high concentrations of free amino acids as well as lactic acid and glucose (Barth and Pitt, 1996; Palmer et al., 2007). Due to the low water activity in CF mucus the bacteria will also experience a high osmotic pressure. The immune defence in the lungs consists of non-inflammatory mechanisms such as mucociliary clearance, alveolar macrophages, defensins and secretory IgA while the inflammatory mechanisms include polymorphonuclear leukocytes (PMNs), IgG and complement (Chmiel and Davis, 2003). Bacterial presence within the CF lung leads to high levels of the inflammatory response consisting mainly of an influx of PMNs (up to 1,500 times normal levels) releasing proteases and reactive oxygen species (ROS) (Konstan et al., 1994). The release of these substances in turn leads to extensive tissue damage, reduced lung function ultimately with a lethal outcome. Bacteria proliferating in the CF lower airway environment consequently have to endure and adapt towards several stresses, e.g. antibiotic, oxidative and osmotic stress.
5.1.2 Environment of the cystic fibrosis paranasal sinuses

The respiratory system consists of both upper and lower airways. Since these are connected any bacteria entering e.g. the lower airways could also potentially enter more distal parts of the upper airways and not just the larynx and pharynx. For CF patients this is also the case as a large percentage (36% in one study) are suffering from chronic sinusitis and symptoms usually commence early in life (5-14 years) (Coste et al., 1995; Gysin et al., 2000; Wang et al., 2005). The paranasal sinuses comprise several small air-filled spaces surrounding the nasal cavity and draining directly or indirectly into this. Their biological function includes weight reduction of the skull and voice resonance and consist of the ethmoidal, frontal, maxillary and sphenoidal sinuses (Figure 4).

Figure 4: Anatomical overview of the location of the paranasal sinuses (Dorland, 2000).

The sinus and lower airway environment is in many ways similar. Epithelial cells contain the CFTR chloride channel and therefore similar effects of CFTR mutation is observed in both environments. In addition they both contain the same mucous membrane lining with the presence of cilia as well. Access of air to the CF paranasal sinuses can be restricted by blockade of the sinus ostia and the sinuses are also easily filled with thickened mucus and hypoxic or anaerobic conditions are known to exist on the mucosa (Aanaes et al., 2011). Thus, invading bacteria will encounter and have to adapt to similar conditions however a few important differences exist between the sinuses and lower airways. The immune response in
the sinuses is less challenging and of a different character with lower abundance of PMNs but higher levels of IgA (Johansen et al., 2009). Additionally the antibiotic load is lower in the sinuses due to e.g. obstruction of sinus cavities. Taken together these characteristics lower the protective barrier against e.g. *P. aeruginosa* colonization compared to the lower airways.

5.1.3 *P. aeruginosa* infection progression in CF patients

The presence of *P. aeruginosa* in the lungs of CF patients usually follows a certain pattern. Initially, after birth and a range of years thereafter, *P. aeruginosa* is not detected in lung samples. Subsequently, upon the first isolation of *P. aeruginosa* from the lungs follows a period of so-called intermittent infection in which *P. aeruginosa* is only occasionally detected in samples. This period will often be ensued by the period of chronic infection in which *P. aeruginosa* persists in the lungs and is continuously detected (Koch, 2002). In order to distinguish these phases several definitions using different criteria have been suggested (Ballmann et al., 1998; Lee et al., 2003; Proesmans et al., 2006; Pressler, 2011).

One of the most widely used definitions is that originating from the Copenhagen CF clinic. In this, the phase of intermittent colonization is defined as a culture of *P. aeruginosa* at least once and the presence of normal levels of precipitating antibodies against *P. aeruginosa*. The chronic infection phase is defined as the persistent presence of *P. aeruginosa* for at least 6 consecutive months, or less when combined with the presence of two or more *P. aeruginosa* precipitating antibodies (Hoiby et al., 1977; Pressler, 2011). The time point of the first acquisition (detection) of *P. aeruginosa* in CF patients is usually early in life. This time point is quite variable and different studies have yielded different median ages, some stating 1-2 years (Demko et al., 1995; Burns et al., 2001; Li et al., 2005) while others have found it to be 7-8 years (Kerem et al., 1989; Maselli et al., 2003). Factors that could influence these numbers are the frequency of sampling and possibly transmission from patients already carrying *P. aeruginosa* through co-localization at the same clinic (Hoiby and Pedersen, 1989). The duration of the period of intermittent colonization is also variable, one reason being the differences in definitions from clinic to clinic but the frequency of sampling can also be a limiting factor for accurate estimations. Furthermore the duration of this period will be heavily influenced by the antibiotic treatment regime of the clinic (Johansen and Hoiby, 1992; Koch, 2002). Aggressive antibiotic treatment can both end the period through
eradication of the bacteria but can importantly postpone the transition to a chronic infection (Koch, 2002). At the Copenhagen CF center an aggressive antibiotic treatment approach is followed during intermittent infection consisting of 3 weeks to 3 months of inhaled colistin and oral ciprofloxacin. During chronic infection two-week intravenous courses of e.g. a β-lactam and aminoglycoside every three months is initiated together with daily administering of ciprofloxacin and colistin (Hoiby et al., 2005).

The genotypes of intermittent colonization isolates have been found to be highly variable between patients and even within patients indicating that any P. aeruginosa genotype has the potential for infection (Burns et al., 2001; Munck et al., 2001; Jelsbak et al., 2007; Hansen et al., 2012). This observed level of genotype variation can however be confounded by cross-contamination of P. aeruginosa within the CF patient population by patient-to-patient contact. As observed in the Copenhagen CF clinic for older patients this can create dominant clone types since these are capable of outcompeting other invaders or clone types already present in the patients (Jelsbak et al., 2007; Yang et al., 2011b) (Jelsbak, Yang 2011). The DK2 clone type present in the Copenhagen CF clinic provides such an example, infecting more than forty individuals over a period of 35 years (Rau et al., in preparation).

In several cases of intermittent colonization (approximately 25 %) a period of several months without P. aeruginosa detection is followed by re-colonization with the same preceding genotype (Munck et al., 2001; Gibson et al., 2003; Doring et al., 2006; Hansen et al., 2012). Considering the normal genotype diversity of colonization it indicates the presence of a reservoir seeding the lungs of the patient having either an external environmental origin or an internal origin e.g. the paranasal sinuses. An association between sinus and lower airway colonization has also been established. Investigating this in young CF patients (Muhlebach et al., 2006) found that phenotypically identical micro-organisms, based on plate assaying, were shared between sinuses and lower airways in 12 out of 31 patients, while 10 of these also had identical genotypes based on Pulsed Field Gel Electrophoresis analysis. In a larger study covering 182 patients of all ages, 23 of the 24 patients carrying P. aeruginosa had identical genotypes in sinus and lower airways (Mainz et al., 2009). Although these studies establish an association between sinus and lung colonization nothing is revealed about any transfer between these two sites, let alone its direction. However a study on post lung-transplant CF patients showed a significant correlation between P. aeruginosa presence
in the sinuses and in the transplanted lung as well as the reverse; absence of *P. aeruginosa* in the sinuses was correlated with absence in the transplanted lung. Furthermore post-transplant sinus surgery was shown to decrease pneumonia (Holzmann et al., 2004). Consequently several studies point towards the transfer of bacteria between the two compartments.

This was finally confirmed by Hansen et al. (2012) finding not only a concordance between *P. aeruginosa* genotypes of sinus and lung isolates of intermittently colonized children but also distinct subpopulations in each side of the sinuses. The subsequent isolation of phenotypically identical strains from the lungs provided strong indications of transfer. The study further revealed that the likely directionality of transfer in this stage of infection is downwards from sinuses to lungs and additionally that evolution of *P. aeruginosa* occurs in the sinuses from which mucoid, small colony variants and antibiotic resistant clones were isolated. Thus the sinuses of intermittently CF patients can constitute a reservoir of variants pre-adapted to the CF lung environment revealing that infection of the paranasal sinuses could easily be a risk factor for development of chronic lung infection.

5.2 *Pseudomonas aeruginosa*

*Pseudomonas aeruginosa* is a rod-shaped Gram-negative bacterium displaying a high metabolic diversity. It is thought to be ubiquitous in nature colonizing aquatic and terrestrial habitats. Even though it is a successful environmental bacterium it is in addition an opportunistic pathogen capable of infecting amoeba, plants and animals (Rahme et al., 2000; Pukatzki et al., 2002). Human infection is generally restricted to immuno-compromised individuals causing systemic infections in burn patients, cancer and HIV patients, respiratory system infections including cystic fibrosis patients (Spencer, 1996; Brennan and Geddes, 2002; Sadikot et al., 2005; Church et al., 2006). It is capable of survival and growth in low nutrient environments and is known to persist in moist environments and water reservoirs such as bathrooms, sinks and kitchens providing opportunity for human contact. It is motile through the action of a flagellum and pili while having an optimum growth temperature of 37° but capable of survival up to 42°. The genome is one of the largest among bacteria being 5.5-7.0 Mbp with so far five assembled genomes (Stover et al., 2000; Lee et al., 2006;
having a high proportion of regulatory genes likely reflecting its versatility.

5.2.1 Virulence factors

Reflecting its potential as an opportunistic pathogen, *P. aeruginosa* has the potential to produce a multitude of virulence factors some of which are secreted while others are cell-associated (Figure 5).

![Figure 5: Pseudomonas aeruginosa virulence factors](image)

Secreted virulence factors include siderophores, pyocyanin, proteases, hydrogen cyanide and rhamnolipids. The two siderophores, pyoverdin and pyochelin, are produced and secreted during iron-limited conditions and taken up through specific cell-surface receptors (Beare et al., 2003). They can favourably compete with human transferrin for iron. Pyocyanin is a redox-active blue pigment with antimicrobial activity that upon reaction with oxygen can produce superoxide and hydrogen peroxide inhibiting mammalian cell respiration (Wilson et al., 1987). The functions of rhamnolipids are multifaceted. Their amphiphilic nature, mediated by its structure containing rhamnose and a fatty acid residue, provides detergent-like properties used for e.g. swarming motility. They are believed to dissolve lung surfactant phospholipids and protect against PMNs (Liu, 1974; Kohler et al., 2000; Van Gennip et al.,
2009). Hydrogen cyanide, for which *P. aeruginosa* has inherent resistance mechanisms, is known to cause inhibition of mitochondrial respiration and is a key factor in *Caenorhabditis elegans* killing (Gallagher and Manoil, 2001). In addition to these secreted virulence factors *P. aeruginosa* expresses the Type III secretion system involved in the secretion of cytotoxic effector proteins (ExoS, ExoT, ExoU and ExoY) into the cytoplasm of eukaryotic cells (Hauser et al., 2002). Cell-associated virulence factors include flagella, type IV pili and LPS. Apart from twitching motility, pili are involved in the attachment of *P. aeruginosa* to both abiotic and biotic surfaces (de Bentzmann et al., 1996) while flagella responsible for swimming motility are highly immunogenic (Mizel et al., 2003). The endotoxin LPS consists of lipid A, core oligosaccharide and repeated O-linked polysaccharide of either A-band or B-band form. The B-band is variable between the different O-band serotypes and is required for *P. aeruginosa* virulence (Cryz et al., 1984; Rocchetta et al., 1999).

### 5.2.2 Antibiotic resistance

A high intrinsic tolerance to many antibiotics is characteristic of *P. aeruginosa* due to e.g. a vast array of efflux pumps, and especially a potential for development of increased resistance through mutation affecting the expression of these. Usually the mutations are located in their respective regulator genes (Masuda et al., 1996; Vogne et al., 2004; Rau et al., 2010) (Vogne 2004, Masuda 1996, Rau 2010). Other mechanisms include enzymatic degradation of the antibiotic or modification of its target. Five families of efflux pumps, most of those in *P. aeruginosa* are of the Resistance-Nodulation-Division (RND) family of which it has 12 different members. These three-component systems differ in substrate spectrum and antibiotic class affinity with some overlap between them. The system of MexAB-OprM encompasses a broad substrate spectrum and is strongly associated with the high intrinsic resistance of *P. aeruginosa* (Poole et al., 1993) while MexXY-OprM is involved in resistance towards aminoglycosides (Aires et al., 1999). A third system, MexCD-OprJ is especially active against fluoroquinolones (Jalal et al., 2000).

In addition *P. aeruginosa* contains a chromosomally encoded gene *ampC* that confers β-lactam resistance and mutation of its regulator genes are known to cause high-level resistance (Campbell et al., 1997; Juan et al., 2005). The third mechanism of target modification is observed in the resistance towards cationic antimicrobial peptides. Resistance
towards these can be increased by mechanism of lipid A modification through the up-regulation of the *arn* operon (Moskowitz et al., 2004).

### 5.2.3 Quorum sensing

Bacterial quorum sensing is a means of communication between cells by the mechanism of signal molecule secretion and detection. Usually, upon reaching a critical concentration of autoinducer molecule in the surrounding environment a signal molecule sensor molecule activates the corresponding quorum sensing (QS) regulon. *Pseudomonas aeruginosa* possesses three QS systems, the *Las*, *Rhl* and PQS systems. The *Las* and *Rhl* systems consist of the autoinducer molecules 3-oxo-C12-HSL (*Las*) and C4-HSL (*Rhl*) and their cognate sensors, LasR and RhIR (Figure 6).

![Figure 6](image)

**Figure 6: The Las and Rhl quorum sensing system.** Included are the response regulators, LasR and RhIR, and their respective signal molecules. Upon signal molecule binding the response regulator activates transcription of target genes (Jimenez et al., 2012).
The QS regulatory network in *P. aeruginosa* is fairly complex involving multiple other regulatory proteins but in general the QS systems are organized hierarchically with the *Las* system capable of regulating the *Rhl* system while the *PQS* system can be regulated by both (de Kievit and Iglewski, 2000; Deziel et al., 2004; Schuster and Greenberg, 2006). The *P. aeruginosa* QS systems are key regulators of virulence factors. The *Las* system is known to regulate several proteases and be important for biofilm formation (Gambello and Iglewski, 1991; Toder et al., 1991) while numerous virulence factors are regulated by the *Rhl* system, e.g. rhamnolipid, proteases, hydrogen cyanide, pyocyanin and siderophores (Brint and Ohman, 1995; Latifi et al., 1995; Diggle et al., 2002). Consequently, the QS systems are global regulators of several functions normally related to pathogenesis.

### 5.2.4 The mucoid phenotype

Another characteristic of *P. aeruginosa* is its potential for excessive production of the polysaccharide alginate by mutation, as is often seen in infection of CF patients (Deretic et al., 1995; Govan and Deretic, 1996; Li et al., 2005; Rau et al., 2010; Yang et al., 2011b). Alginate is a negatively charged linear copolymer of partially O-acetylated β-1,4-linked D-mannuronic acid and its C5 epimer, α-L-guluronic acid (Linker and Jones, 1966) synthesized by the 12 gene alginate operon (*algD – algA*) and a 13th gene *algC* also involved in rhamnolipid biosynthesis. The precursor for alginate biosynthesis is fructose 6-phosphate which is converted into GDP-mannuronic acid that is polymerized and further processed into the two constituent sugar residues while being transported across the cell membrane (Ramsey and Wozniak, 2005) (Figure 7).

Synthesis is regulated mainly through the *algD* promoter which is under heavy regulation of several regulators, e.g. AlgT (AlgU, σ22), AlgR, AlgB, RpoN and AmrZ (Govan and Deretic, 1996; Baynham et al., 1999; Boucher et al., 2000). The alternative sigma factor AlgT is the key regulator during conversion to the mucoid phenotype and its corresponding gene is located in the *algT* operon consisting of *algT* and *mucABCD* genes (Figure 8).
The \textit{algT} operon is autoregulated by AlgT while the Muc proteins all function as negative regulators of AlgT and alginate production to varying extent (Boucher et al., 1996; Boucher et al., 1997b; Rowen and Deretic, 2000). MucA is the primary negative regulator and mutation of the gene is also the most frequent cause of conversion to the mucoid phenotype in CF isolates (Govan and Deretic, 1996; Ciofu et al., 2008) although \textit{mucB} and \textit{mucD} mutations were also identified (Boucher et al., 1997a). Mutation in numerous positions in \textit{mucA} have been observed although certain mutated alleles are predominant especially \textit{ΔG430} and \textit{T349C} constituting mutational hotspots (Bragonzi et al., 2006; Ciofu et al., 2008). MucA is the corresponding anti-sigma factor of AlgT and normally binds and sequesters AlgT at the cytoplasmic membrane hindering its transcription initiation (Schurr et al., 1996). Upon mutation MucA is cleaved from the membrane by AlgW and MucP and subsequently

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{alginate_biosynthesis.png}
\caption{Overview of alginate biosynthesis in \textit{P. aeruginosa}. Colored and lettered objects indicate alginate biosynthesis enzymes. O.M. and I.M. correspond to outer and inner membrane respectively (Ramsey and Wozniak, 2005).}
\end{figure}
degraded by ClpXP liberating AlgT activating *algT* operon transcription causing further synthesis of de-regulated AlgT (Qiu et al., 2007; Qiu et al., 2008).

**Figure 8:** AlgT regulation of the alginate biosynthesis operon and downstream regulators. AlgT can activate transcription of a number of genes, e.g. the *algD* operon, *algC*, *algR*, *algB* while being regulated by proteins such as MucA, MucB, AlgW. Arrows indicate positive regulation while blunt-ended lines indicate negative regulation (Sautter et al., 2012).

De-regulating AlgT has profound pleiotropic effects on the cell as its regulon is extensive. This is largely helped by its direct regulation of several regulatory genes such as *algR*, *algB* and *amrZ* (Goldberg and Dahnke, 1992; Wozniak and Ohman, 1994; Wozniak et al., 2003). The most renowned effect of AlgT activation is the increased transcription of the *algD* promoter leading to alginate overproduction however recently it was evidenced by transcriptome and phenotype analysis that the AlgT regulon also includes genes encoding virulence factors, the Rhl quorum sensing system, motility and possibly osmotic stress (Rau et al., 2010). Examining a mucoid clinical isolate, mucA mutation was shown to cause down-regulation of pyoverdin, hydrogen cyanide, rhamnolipid, the Type III secretion system, flagella and pili while putative osmotic stress genes were up-regulated (Rau et al., 2010). Although a previous transcriptome study (Firoved and Deretic, 2003) on the pleiotropic
effects of mucA mutation yielded different results, the above is corroborated by several experimental findings (Garrett et al., 1999; Lizewski et al., 2004; Wu et al., 2004b; Tart et al., 2006; Morici et al., 2007; Jones et al., 2010).

Apart from the direct regulation by AlgT as observed for the alginate biosynthesis operon and several regulatory proteins most of the pleiotropic effects are likely indirect through e.g. AlgR, AlgB and AmrZ. AlgR is e.g. known to have binding sites upstream of rhlI, part of the rhl QS system by which many of the mentioned virulence factors are regulated (Morici et al., 2007). AlgT has high homology to σE of Escherichia coli and Salmonella Typhimurium (DeVries and Ohman, 1994) a sigma factor known to be important for pathogenesis and various environmental stresses (Rowley et al., 2006). For persistent P. aeruginosa infection of CF patients this is especially true. The linkage of virulence factors, quorum sensing, oxidative stress response and possibly osmotic stress response to one regulon explains why this regulon is so frequently targeted for activation in the CF lung.

5.3 The P. aeruginosa genome – a dynamic genome

The genome of P. aeruginosa as a species is best described as a mosaic genome with a conserved core present in all P. aeruginosa strains and an accessory genome consisting of strain-specific elements from other species or genera. These elements are dispersed in the genome at a limited number of loci, often described as regions of genomic plasticity (Mathee et al., 2008; Klockgether et al., 2011).

5.3.1 The accessory genome

The accessory genome of P. aeruginosa consists of different element types, such as integrative conjugative elements (ICEs), replacement islands, prophages and transposons. The group of ICEs are self-transmissible genomic islands transferred between strains by conjugation while integration into the recipient genome is phage integrase mediated (Burrus et al., 2002). These elements can further exist as circular extra-chromosomal elements and are composed of a conserved backbone and individual cargo genes (Wurdemann and Tummler, 2007). Sizes range between 81 to 108 kbp (Kung et al., 2010). One family of ICEs are the pKLC102-related family comprising a highly prevalent family among P. aeruginosa
strains (Klockgether et al., 2011). Apart from pKLC102 these include e.g. PAPI-1 and PAGI-5 conferring pathogenic traits (Figure 9) (He et al., 2004; Battle et al., 2009).

Figure 9: Comparison of three ICEs belonging to the pKLC102 family. The ICEs share modules involved in different processes, e.g. integration (I), maintenance (M) and transfer (T). Grey lines indicate similar sequence (Kung et al., 2010).

Genes encoding several virulence factors, i.e. LPS O-antigen, pyoverdine, pili and flagella, are termed replacement islands as these loci have been horizontally acquired and display high sequence diversity between P. aeruginosa strains (Smith et al., 2006). The evolutionary rationale behind the high diversity is likely related to the insurance hypothesis for the general P. aeruginosa population to evade host immune systems and to resist phage predation. It is believed that most P. aeruginosa strains can be lysogenized by at least one temperate phage, of which there exists more than sixty members capable of P. aeruginosa infection (Akhverdian et al., 1984; Kung et al., 2010). Prophages can bestow the recipient organism with beneficial functions, e.g. prophage D3 causes sero-conversion of the recipient by O-antigen alteration, putatively providing advantageous in a host setting (Vaca-Pacheco et al., 1999; Newton et al., 2001). Likewise P. aeruginosa transposons can provide beneficial functions, among these the insertion of antibiotic resistance cassettes (Kung et al., 2010). Consequently, accessory elements play pivotal roles for the adaptation of P. aeruginosa to existing habitats but are equally important for niche expansion. In CF infection their role is
less explored however some findings have revealed their importance. The sequenced CF isolate *P. aeruginosa* LESB58 carries several prophage clusters and genomic islands of which genes from four prophage regions and one genomic island conferred increased fitness in a chronic rat lung infection model (Winstanley et al., 2009).

5.3.2 Genetic adaptation during CF infection

*Pseudomonas aeruginosa* is thought to be ubiquitous to aquatic and terrestrial environments and is therefore highly adapted to life in these environments. Even though its versatility allows it to be an opportunistic pathogen capable of causing a range of human infections it is not a host environment specialist. Consequently, a host associated life style spanning up to several years would be expected to be associated with genetic adaptation towards becoming a specialist in the host environment. Before the advent of readily accessible genome sequencing, phenotypic studies of *P. aeruginosa* CF isolates also revealed a multitude of changes at the phenotypic level during infection and in recent years whole genome sequencing of CF isolates have provided a more comprehensive overview of the level of genetic adaptation occurring in the transition from an environmental life style to host association. In order to understand this process it is important to be aware of what the adaptation is towards, i.e. the immune response, extensive antibiotic administrations in addition to a different nutritional composition and much adaptation would therefore be expected to be directed towards these components.

Some of the earliest observations of adaptation are the frequent loss of virulence factors. Although these have great importance for enabling acute infections the reverse seems to be the case for proliferation in the CF lung. Indeed, in this environment they seem to be selected against. This concerns especially the trait of motility mediated by loss of function of pilin and flagellar systems (Mahenthiralingam et al., 1994; Yang et al., 2011a), alteration of LPS by loss of B band O-antigen (Goldberg and Pler, 1996) as well as the loss of virulence factors such as proteases, phenazines, pyocins, siderophores and Type III secretion systems (Luzar and Montie, 1985; Jain et al., 2004; Hogardt et al., 2007; Yang et al., 2011a). Often this is mediated by inactivation of QS systems especially by *lasR* mutation (Smith et al., 2006) which additionally provides a growth advantage on certain amino acids (D’Argenio et al., 2007). The external nature of the virulence factors renders them highly immunogenic and their loss was therefore proposed as an adaptation mechanism for immune system evasion.
Another mechanism of this, is the excessive production of alginate, a hallmark of CF lung adaptation. Alginate overproduction can confer increased protection against the immune system in several ways: By masking complement and antibodies on the bacterial cell surface thereby preventing their subsequent phagocytosis, avoiding non-opsonic phagocytosis by neutrophils and macrophages and by functioning as an efficient scavenger of reactive oxygen intermediates released by PMNs (Cabral et al., 1987; Govan and Deretic, 1996; Mathee et al., 1999).

The selective pressure from frequent antibiotic administration is reflected in the increased antibiotic resistance observed in persisting CF isolates. These mechanisms include the mutational overexpression of efflux pumps such as MexXY for aminoglycoside resistance (Vogne et al., 2004), MexCD-OprJ for fluoroquinolone resistance (Jalal et al., 2000; Jeannot et al., 2008) or altered lipid A composition for cationic antimicrobial peptide resistance (Ernst et al., 1999; Moskowitz et al., 2004), gyrase mutations for fluoroquinolone resistance (Kureishi et al., 1994; Yang et al., 2011b) and increased ampC expression for β-lactam resistance (Giwereman et al., 1990). Most frequently the molecular mechanism facilitating increased resistance is by regulator mutations, a common theme in _P. aeruginosa_ adaptation to the CF lung. Increased resistance towards the different classes of antibiotics administered to CF patients is part of the inherent genetic potential of _P. aeruginosa_ and one of the major reasons it is the main cause of morbidity and mortality for CF patients. Most other organisms do not possess a comparable resistance repertoire and indeed before commencing an aggressive antibiotic treatment strategy the main pathogen of CF patients was _Staphylococcus aureus_.

_Pseudomonas aeruginosa_ infection of CF patients is also associated with development of the so-called small colony variant (SCV) phenotype characterized by an increased ability for biofilm formation, an important trait in CF infection (Haasby et al., 2010). As implied by the name these variants have a certain (small) colony morphology and have been found for a multitude of species isolated from various infections (Proctor et al., 2006). Unlike some species, this term comprises a somewhat heterogeneous collection of characteristics among _P. aeruginosa_ isolates. Consistent characteristics include an autoaggregative behaviour coupled with increased biofilm formation (Haassler, 2004) while increased motility by hyperpiliation and increased antibiotic resistance are frequent as well (Haassler et al., 2003; Starkey et al., 2009). The molecular mechanism is thought to occur via overexpression of _cup_
genes encoding fimbrial adhesin leading to autoaggregation (Vallet et al., 2001) while overexpression of *pel* and *psl* polysaccharides have been implicated in the biofilm formation of the special rugose SCV variant (Figure 10) (Starkey et al., 2009).

![Image](image.png)

**Figure 10:** Example of auto-aggregation and subsequent clumping of a rugose SCV on the left. The reaming tubes are rugose SCV psl gene mutants and consequently less aggregative (Starkey et al., 2009).

The presence of SCVs are correlated with poor lung function (Haussler et al., 1999) and their presence is likely due to increased biofilm formation capability providing protection towards the immune system and antibiotics.

Compared to aquatic or terrestrial environments the CF lung provides abundant nutrients. High concentrations of free amino acids, glucose and lactic acid have been reported although amino acids were preferentially catabolized over lactic acid and glucose (Barth and Pitt, 1996; Palmer et al., 2007). The significance of amino acid catabolism in the CF lung is also demonstrated by amino acid auxotrophy among clinical isolates (Thomas et al., 2000) and the up-regulation of the *atu* operon mediating an increased growth rate on leucine and likely other branched chain amino acids. In addition up-regulation of the glyoxylate shunt, observed in several clinical isolates (Lindsey et al., 2008; Yang et al., 2011a), seems to be adaptive, possibly by mediating increased hydrogen cyanide production (Hagins et al., 2009). Although metabolic adaptation ought to be important when entering a new environment knowledge on this area of *P. aeruginosa* adaptation to the CF lung is limited. One explanation is that most studies have focused on the two major selective forces, the
immune system and antibiotics, and studies focused on metabolic adaptation solely could provide more clues. Yet, increased growth rate is unlikely to be a key parameter in the adaptation since clinical isolates are commonly observed to have decreased in vitro growth rates (Head and Yu, 2004; Yang et al., 2008) a likely secondary effect of adaptation towards e.g. the immune system and antibiotics. Consequently, any metabolic changes observed for CF isolates need not confer advantages in nutrient utilization but could also be selected by providing improved tolerance towards the other stress factors present in the CF lung.

One frequently evolved trait of CF lung isolates is the trait of hypermutability. Considering the major divergent selection pressures existing in the CF lung and the extensive genetic adaptation the persisting bacteria are subject to, this finding might not be surprising. The mechanism of increased mutation rate is most frequently by inactivation of the mismatch repair (MMR) system conferring 100-1000 fold increases in mutation rates (Hogardt and Heesemann, 2012) while inactivation of the guanine 8-oxo-2-deoxyguanine (GO) system confers a weak mutator phenotype and is only occasionally observed (Mandsberg et al., 2009; Ciofu et al., 2010). Mutators have been associated with an increased rate of especially antibiotic resistance generation but also faster generation of other types of adaptive mutations (Oliver et al., 2000; Hogardt et al., 2007; Mena et al., 2008). Apart from the effect of selection pressures mutator phenotypes are also believed to arise through the oxidative stress conditions present in the CF lung by the liberation of reactive oxygen species by PMNs (Ciofu et al., 2005).

5.3.3 Evolutionary dynamics of infecting lineages

Global phenotypic and genetic analyses on clonal longitudinal isolates have provided valuable insight into the dynamics of evolution during CF lung infection. Transcriptional analysis of longitudinal isolates of three infecting lineages found evidence of parallel evolution between these as 24 genes were differentially expressed in the same direction in the three lineages while parallelism of gene expression between any two lineages was 86 genes (Huse et al., 2010). The majority of total differential gene expression (684 genes) was however dissimilar (Figure 11) and although many phenotypes are frequent among CF isolates, vast evolutionary diversity is also seems to be present when comparing the results of other global studies on
CF isolate evolution (Smith et al., 2006; Hoboth et al., 2009; Rau et al., 2010; Cramer et al., 2011; Yang et al., 2011b).

Figure 11: Common and distinct gene expression changes between three CF lineages, A, B and Cb. + and – indicate up or downregulation respectively (Huse et al., 2010).

One generally interesting aspect of *P. aeruginosa* CF infection is the temporal dynamics of the rate of evolution. Mutation rates can provide accurate measurements of this at the genome level however what is at least equally interesting is the rate on a phenotypic level. Analyzing transcriptome and global catabolic function over time of a long-term infecting lineage, the DK2 clone, revealed that the most substantial phenotypic changes occur initially (a few years) with limited phenotypic change thereafter (Yang et al., 2011b). Conversely, the rate of genomic evolution of the DK2 clone, based on single nucleotide polymorphism (SNP) analysis, remained quite constant and approximated linearity during infection. The limited phenotypic change was a likely effect of negative selection, based on the ratio of synonymous to non-synonymous substitutions, observed in the later stages but not initially. Consequently, genetic drift was the main mechanism of evolution in the later stage correlating well with the observed limited phenotypic change.
5.3.4 Horizontal gene transfer

Niche adaptation for *P. aeruginosa* can be facilitated by mechanism of HGT by acquiring accessory elements enhancing fitness in vacant niches. A sequenced *P. aeruginosa* strain e.g. contained a genomic island encoding enzymes for the degradation of terpenoids produced by trees likely conferring increased fitness in a very specific niche (Mathee 2008). Additionally, accessory elements of *P. aeruginosa* are known to confer increased resistance towards e.g. heavy metals (Aguilar-Barajas et al., 2010). Likely these processes contribute to the high versatility observed for *P. aeruginosa*, its proliferation in diverse habitats and ubiquitous presence in the environment however their relevance for *P. aeruginosa* adaptation to the CF lung has been less studied.

One likely mechanism of HGT within the CF lung would be spread of genetic material between microbes through transduction. Phages can exert significant influence on bacterial communities through e.g. spread of fitness enhancing genetic elements or through predation, also known as top-down control (Rodriguez-Valera et al., 2009). Systematic studies of phage influence and presence in the CF lung are limited although recent studies have shed more light on this subject. One study found that the CF lung contains an active phage population with a slightly lower diversity in CF patients compared to non-CF patients (Willner et al., 2009). The phage community of CF patients was enriched for sequence involved in especially carbohydrate and amino acid metabolism, alginate biosynthesis and virulence factors leading to suggestions of phages participating in the microbial diversification in the CF lung (Willner and Furlan, 2010). Investigating the spatial heterogeneity of phage in explant lungs and post-mortem acquired lungs from two patients further revealed phage sequence encoding the SCV phenotype and antibiotic resistance in post-mortem lungs. However, explanted lungs contained no bacteriophages in the highly diseased regions of the lungs indicating escape from phage predation by the microbial population here (Willner et al., 2012). As CF patients from Willner et al. (2009) had severe airway obstruction, indicating severe chronic infection, infection stage differences do not seem to influence the observed differences in diversity. Escape from phage predation was further indicated in the evolution of the DK2 lineage over 35 years in several patients, without the apparent detection of any prophage entry (one putative entry was detected) (Rau et al., *in preparation*). Consequently, the importance of HGT for *P. aeruginosa* adaptation to the CF lung seems limited as acquisition of novel DNA in the highly successful
DK2 lineage was non-existent or at least very limited during CF infection (Rau et al., in preparation). The intrinsic potential for increased fitness towards the stresses of the CF lung could explain the apparent insignificance of HGT adaptation in this niche or could be a reflection of e.g. the lack of external beneficial DNA or increased *P. aeruginosa* phage resistance.

Two processes of bacterium-phage co-evolution are known to occur, fluctuating selection and the arms race process (Figure 12).

![Figure 12: Model of fluctuating selection during bacterium-phage co-evolution.](image)

Arms race is characterized by increased phage infectivity and increased bacterial resistance over time while fluctuating selection is characterized by fluctuation in the fitness of different bacterial or phage genotypes, e.g. bacteria are more resistant to the contemporary phage genotype than to previous or future phage genotypes. The apparent lack of phage predation in explant lungs observed by Willner et al. (2012) indicates an arms race process in this specific scenario. Increased bacterial resistance in an arms race theoretically leads to increased costs while costs in fluctuating selection should be smaller. Actual fitness costs seem to depend on condition as any increase in resistance in a soil environment came at a relatively high cost leading to fluctuating selection (Gomez and Buckling, 2011) while no cost was observed in nutrient-rich laboratory media enabling an
arms race in this condition (Brockhurst et al., 2007). The nutrient rich conditions in the CF lung might therefore facilitate the possibility of an arms race between *P. aeruginosa* and the resident phage population. Many phage receptors, e.g. flagella, pili and LPS are altered or reduced during infection while the mucoid phenotype was selected upon co-evolution of *Pseudomonas fluorescence* and phage phi2 (Scanlan and Buckling, 2011). Consequently, these phenotypes might be adaptive towards evasion of both the immune system and phage predation.

### 5.4 Bacterial host-restriction adaptation

Long-term adaptation of *P. aeruginosa* to the CF lung frequently produces variants lacking e.g. motility, virulence factors, QS functionality in addition to slow growth and overproduction of alginate. These variants should possess inferior fitness in the natural aquatic and terrestrial environments. Hence, highly adapted variants can be considered as host restricted, i.e. their main habitat has been restricted to the host environment. Consequently, *P. aeruginosa* infection in the CF lung constitutes an example of the very early stage of host-restriction adaptation. Several more advanced-stage examples of adaptation to host restriction exist among strictly host-associated bacteria such as *Burkholderia mallei*, *Mycobacterium leprae*, and *Rickettsia* and *Buchnera* species (Andersson et al., 1998; Shigenobu et al., 2000; Gomez-Valero et al., 2007; Song et al., 2010). Investigating the mechanisms and dynamics of genome changes during such adaptation is usually carried out by comparison of present-day isolates to distant theoretical ancestors providing limited resolution of events during adaptation, however certain characteristics are observed among strains at different stages of host adaptation including changes in genome size, number of pseudogenes and quantity of selfish genetic elements (Figure 13).
Figure 13: Genome characteristics in different stages of host-restriction adaptation (Ochman and Davalos, 2006).

The reduction of genome size is a continuous process occurring in both early and late stages. Two general mechanisms are responsible for the observed nucleotide loss: Large deletions in addition to gene inactivation followed by gradual erosion of the inactivated gene (Moran and Plague, 2004). The latter was proposed to be the main mechanism in the early stage of adaptation based on comparison of Buchnera APS, Mycobacterium leprae, Shigella flexneri and Salmonella Typhi to theoretical ancestors (Silva et al., 2001; Dagan et al., 2006). However analyzing the same Buchnera genome researchers reached the conclusion that the main mechanism of genome reduction in the early stage was through large deletions of contiguous sequence (Moran and Mira, 2001). Several other studies have supported this to be a prominent feature of the early stage adaptation, e.g. adaptation of P. aeruginosa to the cystic fibrosis lung revealed substantial genome reduction (up to 8% in one lineage) during infection mediated by the loss of large regions of contiguous sequence (Rau et al., in preparation). Roughly double the amount of sequence was deleted compared to the quantity
of gene sequence affected by pseudogene formation (Rau, unpublished). Early stage in vitro experimental reductive evolution further showed the loss of large blocks of sequence, up to 200 kbp (Nilsson et al., 2005). The alternate conclusions reached by Dagan et al. (2006) and Silva and Mira (2001) could be affected by low resolution and the choice of analysis approach. Nonetheless, both the extent of reductive evolution and the proportion of pseudogene formation versus large deletions is likely to be influenced by the potential for nucleotide loss of the host-adapting strains yielding different ratios. Another feature of early adaptation to host-restriction is the proliferation of IS elements as exemplified by an increased density of these in the genomes of bacteria with a recent obligate host-association (Figure 14).

![Figure 14: Density of IS elements in the genomes of bacteria with either a free-living (grey), recent host-restricted (red) or ancient host-restricted (black) lifestyle (Moran and Plague, 2004).](image)

The mechanism of deletion can be through either illegitimate or homologous recombination. In the early stage adaptation of *Burkholderia mallei*, *Bordetella parapertussis* and *Bordetella pertussis* an association between end-regions of deletions and IS elements strongly indicated that the deletion mechanism occurred through IS element mediated homologous recombination (Parkhill et al., 2003; Song et al., 2010). This was the case for most deletion end-regions. However in the very early stage of adaptation towards host-restriction as exemplified by infection of CF patients with the *P. aeruginosa* DK2 clone,
deletion of up to 8 % of the genome occurred without any involvement of IS elements (Rau et al., in preparation). One reason was the low proliferation of IS elements keeping a low density of these throughout infection and consequently also low potential for IS mediated deletions. Yet, the findings demonstrate that a significant portion of the genome can be deleted without IS element involvement. The time scales between these studies are quite different (35 years for P. aeruginosa infection vs. 1-3 million years) likely affecting the outcomes and given more time, IS element mediated deletions might occur in the infecting P. aeruginosa strains. In fact, IS mediated inversions have been observed in CF isolates before (Kresse et al., 2003; Kresse et al., 2006). The IS element copy numbers within these isolates were also higher, 14 to 25 copies, indicating a higher proliferation rate of the elements. The mechanisms of deletion observed by Rau et al., (in preparation) instead consisted of illegitimate recombination and non-IS element mediated homologous recombination as also found during in vitro reductive evolution (Nilsson et al., 2005). Hence these could be the main deletion mechanisms in the beginning of host-restriction adaptation. A predominance towards illegitimate recombination was present in both studies likely reflecting the higher density of such regions within the genome.

Although not observed in the DK2 clone adaptation to the CF lung, host-restriction adaptation can occur by HGT albeit at a lower frequency compared to free-living bacteria (Toft and Andersson, 2010). Examples include the uptake of genes encoding secretion systems and eukaryotic-like proteins likely transferred by means of phage as the genes are located next to prophage sequence (Wu et al., 2004a; Berglund et al., 2009).

As mentioned host-restriction adaptation displays certain characteristics such as changes in genome size, number of pseudogenes and quantity of selfish genetic elements. There are a number of probable reasons to the observed characteristics during host-restriction adaptation. One is the potential small population sizes leading to decreased efficiency of purifying selection and thus fixation of deletions and IS element insertions through genetic drift. Inactivation of otherwise beneficial genes might occur in this process. Likewise loss of superfluous genes such as certain metabolic genes in a nutrient-rich host environment can be caused by reduced intensity of purifying selection, i.e. many genes are not selected for (Moran and Plague, 2004). Both processes should be present during P. aeruginosa adaptation to the CF lung. The P. aeruginosa population in the CF lung is not very
small however it is likely subject to severe bottlenecks especially during antibiotic treatment, creating the foundation for genetic drift. Likewise numerous P. aeruginosa genes should be neutral and some even deleterious in the CF environment creating further potential for reductive evolution.
6 Concluding remarks

The study of *Pseudomonas aeruginosa* adaptation to the cystic fibrosis lung has great merit both from a clinical and an evolutionary point of view. Knowledge on the process of adaptation can be of great benefit by providing novel targets for treatment of infection as well as uncovering the dynamics of *in vivo* microbial adaptation towards a markedly different environment. In this work one focus has been the early stage of *P. aeruginosa* colonization of CF airways as this stage is deciding for the persistence of *P. aeruginosa*. Three cases of *P. aeruginosa* survival in the CF airways for more than two years revealed very different evolutionary trajectories for early stage adaptation. One persisting lineage displayed adaptation mainly by modification of the expression of a global regulatory network while in another a few adaptations with small impact on the global phenotype were observed. The lineages are consequently climbing different peaks of the fitness landscape and further reveal that a specific sequence of events is not necessary to achieve sufficient fitness towards the CF lung for persistence. What seems necessary however, is increased fitness towards the selective pressures of the CF lung as the third lineage, displaying no signs of adaptation, eventually was eradicated.

The sinuses were shown to be the likely habitat of *P. aeruginosa* during the specific pattern of recurrent intermittent infection observed in some CF patients. Interestingly, before the onset of chronic lung infection, the sinus environment selected for the mucoid phenotype, SCVs and antibiotic resistance i.e. variants pre-adapted to the CF lung environment. Consequently *P. aeruginosa* sinusitis can provide increased opportunity for persistence in the host. Targeting sinusitis by focusing on improved antibiotic treatment of the sinuses during intermittent colonization could potentially lower the risk for progression to a chronic lung infection. The direction of transfer between upper and lower airways in intermittent colonization appears to be mainly downward however the dynamics of transfer could potentially change during infection. After establishment of chronic infection the lung population would be expected to exceed the sinus population in quantity and should therefore present a higher potential for acquiring adaptive mutations. Consequently there could be a change in the direction of transfer over time.
The adaptation of *P. aeruginosa* towards the CF lung can be viewed as the initial stage of host-restriction adaptation. The observation of substantial reductive evolution in the DK2 clone supports this notion. In the shift from one environment to another many functions are rendered redundant and are therefore prone to be lost by genetic drift. The overall sign of negative selection during evolution of the DK2 lineage (Yang et al., 2011b) confirms that genetic drift plays a fundamental part in re-shaping the genome and many of the observed deletions could therefore be attributed to this. Since the *P. aeruginosa* lung population is likely to encounter bottlenecks during e.g. antibiotic treatment evolution by genetic drift should not be unexpected. However some deletions could also occur through positive selection, e.g. the deletions containing *hpc* genes. The deletion of large contiguous genomic regions carries with it the likelihood of simultaneous removal of beneficial genes. Consequently the *P. aeruginosa* lung population could be subject to the effects of Muller’s ratchet by accumulating such slightly deleterious mutations over time. Unlike previous reports, the mechanism of deletion in the early stage of host adaptation was not IS element mediated but occurred through homologous and illegitimate recombination. The involvement of IS elements in genome reduction will depend on their genomic density and therefore ultimately their proliferation rate within the genome. Thus, their involvement could vary substantially between host adapting strains.

Only limited uptake of novel DNA, if any at all, was observed within the DK2 lineage, possibly as a consequence of a lack of external beneficial DNA or decreased capability for DNA uptake. Typical phage receptors such as pili, flagella and LPS are reduced or altered in the DK2 lineage suggesting increased phage resistance. Possibly these phenotypes are adaptive towards evasion of both the immune system and phage predation. The limited DNA uptake emphasizes a general aspect of *P. aeruginosa* adaptation to the CF lung, i.e. a preponderance towards reduction of function compared to gain of function.
7 References


Campbell, J.I., Ciofu, O., and Hoiby, N. (1997) Pseudomonas aeruginosa isolates from patients with cystic fibrosis have different beta-lactamase expression phenotypes but are homogeneous in the ampC-ampR genetic region. _Antimicrob Agents Chemother_ **41**: 1380-1384.


Scanlan, P.D., and Buckling, A. (2011) Co-evolution with lytic phage selects for the mucoid phenotype of Pseudomonas fluorescens SBW25. ISME J.


Early adaptive developments of \textit{Pseudomonas aeruginosa} after the transition from life in the environment to persistent colonization in the airways of human cystic fibrosis hosts

Martin Holm Rau,† Susse Kirkelund Hansen,† Helle Krogh Johansen, Line Elnif Thomsen, Christopher T. Workman, Kristian Fog Nielsen, Lars Jelsbak, Niels Hoiby, Lei Yang and Søren Molin†

1Department of Systems Biology, Technical University of Denmark, Building 301, 2800 Lyngby, Denmark.
2Copenhagen CF Center, Department of Clinical Microbiology, Rigshospitalet, Juliane Mariæsvej 22, 2200 Copenhagen, Denmark.
3Department of Veterinary Disease Biology, University of Copenhagen, Stigboejlen 4, 1870 Frederiksberg C, Denmark.

Summary

\textit{Pseudomonas aeruginosa} is an opportunistic pathogen ubiquitous to the natural environment but with the capability of moving to the host environment. Long-term infection of the airways of cystic fibrosis patients is associated with extensive genetic adaptation of \textit{P. aeruginosa}, and we have studied cases of the initial stages of infection in order to characterize the early adaptive processes in the colonizing bacteria. A combination of global gene expression analysis and phenotypic characterization of longitudinal isolates from cystic fibrosis patients revealed well-known characteristics such as conversion to a mucoid phenotype by \textit{mucA} mutation and increased antibiotic resistance by \textit{nfxB} mutation. Additionally, upregulation of the \textit{atu} operon leading to enhanced growth on leucine provides a possible example of metabolic optimization. A detailed investigation of the mucoid phenotype uncovered profound pleiotropic effects on gene expression including reduction of virulence factors and the Rhl quorum sensing system. Accordingly, mucoid isolates displayed a general reduction of virulence in the \textit{Caenorhabditis elegans} infection model, altogether suggesting that the adaptive success of the mucoid variant extends beyond the benefits of alginate overproduction. In the overall perspective the global phenotype of the adapted variants appears to place them on paths in direction of fully adapted strains residing in long-term chronically infected patients.

Introduction

The bacterium \textit{Pseudomonas aeruginosa} is a microorganism that is usually found both in aquatic and terrestrial environments. Like other \textit{Pseudomonads} it is a versatile species with an extended metabolic repertoire allowing it to occupy many different niches. Its genome, one of the largest in the bacterial world with close to 6000 genes, displays the versatility of the organism with its multiple metabolic pathway genes, and the extraordinary large number of regulatory genes (more than 500) suggests that \textit{P. aeruginosa} has a great competitive capacity to move between different types of environments (Stover et al., 2000). The diversity of gene regulatory activities in the organism ensures adaptation to new environmental conditions. Another reason why \textit{P. aeruginosa} has attracted considerable interest is connected to its capacity to cause infections in plants, animals and humans (Spencer, 1996; Wolfgang et al., 2003). These infections are associated with the genomic content of a large number of virulence genes, which are important elements in most cases of infections.

One of the best studied human infections caused by \textit{P. aeruginosa} is airway infections in patients suffering from the genetic disorder cystic fibrosis (CF). Because of a severely reduced mechanical clearing of the airway mucus in these patients they acquire multiple infections of various bacteria and fungi, and a major cause of morbidity and mortality is chronic infection of \textit{P. aeruginosa} in the lungs (Koch and Hoiby, 1993; Govan and Deretic, 1996). Most patients acquire \textit{P. aeruginosa} early in their lives and initial colonization is typically caused by environmental strains displaying all the wild-type phenotypes associated with this species (Johansen and Hoiby, 1992; Burns et al., 2001; Koch, 2002). Because the majority of children in a given clinical setting carry unique genotypes of
P. aeruginosa it also seems clear that colonization in most cases is a separate event (Burns et al., 2001; Jelsbak et al., 2007). In many cases the original colonization proceeds into a chronic infection (or alternatively other clones may later establish chronic infections), without any possibility of eradication (Koch, 2002; Jelsbak et al., 2007). During infection the bacteria have to adapt to the new environment comprising the host immune response, antibiotics and a different substrate composition. It is commonly observed that the bacteria gradually loose the function of many genes normally associated with bacterial pathogenicity (Smith et al., 2006; Jelsbak et al., 2007); however, in a chronic infection scenario this does not per se imply a reduction in virulence but likely rather an alteration of virulence (Bragonzi et al., 2009).

Genetic and phenotypic investigations of P. aeruginosa isolates from chronically infected patients have shown that among the many specific traits developing in the bacteria during infection, the mucoid phenotype caused by excessive production of the extracellular polysaccharide alginate is occurring with high frequency (Pedersen, 1992; Martin et al., 1993; Govan and Deretic, 1996). Conversion to a mucoid phenotype is caused by mutations in regulatory genes, mainly in various positions of the mucA gene encoding the anti-sigma factor MucA (Boucher et al., 1997). Binding between MucA and AlgU, the corresponding sigma protein (homologous to the sigma E factor of Escherichia coli) prevents AlgU from acting as a necessary initiation factor required for transcription of the alg operon (encoding the enzymes required for alginate synthesis) (Govan and Deretic, 1996; Boucher et al., 1997; Rowen and Deretic, 2000). Mutations disabling the activity of MucA lead to activation of AlgU, and the resulting elevated transcription of the alg operon leads to alginate overproduction (Rowen and Deretic, 2000). In the CF airways overproduction of alginate has been associated with increased tolerance to the host immune defense (Cabral et al., 1987; Govan and Deretic, 1996; Mathee et al., 1999) and additionally described as a protective extracellular matrix in connection with biofilm development in the infected airways (Hoiby et al., 2001; Yang et al., 2008). It is thus assumed that the major impact of the mucA mutations on the infection success of the bacteria is connected to the alginate production.

Pseudomonas aeruginosa in its wild-type configuration is naturally resistant to many antibiotics, and if confronted with antimicrobial agents it will in many cases be able to develop tolerance to increased levels of the compounds (Hancock, 1997; Hancock and Speert, 2000). This capacity to resist antimicrobial treatment is most likely an important reason why P. aeruginosa has become the dominant airway infection problem for CF patients as antibiotic therapy was introduced to these patients. Before antibiotic treatment patients died very young from infections of indigenous pathogens such as Staphylococcus aureus and Haemophilus influenzae. The medical success leading to extended life spans of CF patients is coupled to antibiotic treatments with compounds towards which P. aeruginosa is not inherently resistant (Heiby et al., 2005; Hansen et al., 2008). Because antibiotics are used to combat the infections as soon as P. aeruginosa has been diagnosed, it means that persistent colonization and subsequent chronic infection will eventually be associated with development of resistance.

Some of the most common phenotypic traits reported for P. aeruginosa isolates from chronically infected CF patients (‘the chronic infection phenotype’) include slow growth, auxotrophy, alginate overproduction (mucoidy), antibiotic resistance and loss of virulence factors and motility (Mahenthiralingam et al., 1994; Barth and Pitt, 1996; Bagge et al., 2002; Lee et al., 2005; Smith et al., 2006; Jelsbak et al., 2007; Ciofu et al., 2008). Because these phenotypic characteristics have been observed in many different clinical settings it is highly likely that such a phenotypic profile is an end-point result of parallel evolution of the bacteria in many CF patients colonized by many different bacterial genotypes. The actual evolutionary trajectories towards this common phenotypic profile, however, could be quite different in different clinics, patients and colonizing genotypes. In three cases described here the colonizing bacteria are all at the beginning of their colonization history, they all constitute persistent colonization (present for thousands of generations), and none of them have at the present time reached the typical ‘chronic infection phenotype’. The question addressed here is: in each of these three cases – do the bacteria show indications of being on a direct path towards the ‘chronic infection phenotype’, and if so what is the mechanism behind this?

Results

From the environment to colonization of human airways

We have collected isolates from the majority of CF children infected with P. aeruginosa associated with the Copenhagen CF clinic during a 4.5 year period (2005–2009). In total 45 CF children were followed from first isolate of P. aeruginosa (or very early in their colonization history). In this communication we distinguish between cases of cured and persistent colonization of the patient airways (cf. Experimental procedures for definitions), and at the conclusion of the study a large group of patients (20 patients) had been cured (one or several times) of P. aeruginosa after a period of treatment with antibiotics, and for 13 patients status was still unknown because of recent
colonization within the last 12 months. These cases are not examined further here. Another large group of 13 patients comprises cases of recurrent colonization where the bacteria are not present in airway samples for a period of time (often several months) and then suddenly reappear with the same genotype (to be published). Based mainly on duration of bacterial persistence and frequency of isolation it is from this group of persistent clones we have chosen three cases for further characterization of the bacterial genotypes and phenotypes and for evolutionary changes over time.

The experimental approach has been to combine transcriptional profiling using standardized Affymetrix DNA arrays with phenotypic characterization. The DNA array experiments have been performed under standardized conditions in laboratory media (cf. Experimental procedures), and the results of these provide information about impacts of genetic changes on expression of specific genes, when comparing data sets from isolates harvested from the patients at different time points. Because the Affymetrix gene chips mainly cover the genes present in reference strains of *P. aeruginosa* such as PAO1, PA14 and PAK, the obtained data only cover expression of genes in the core genome. Parallel phenotypic investigations have been employed to verify the transcription analysis, and to bring about information that may indicate which mutational events may be responsible for the observed changes of gene expression and corresponding phenotype. Gene sequencing was employed in some cases to confirm such mutational events.

Characterization of a transiently colonizing strain of *P. aeruginosa*

Patient B6 was initially colonized by a strain of *P. aeruginosa* that persisted for 2 years after which it was no longer isolated from patient samples. The sample time line is shown in Fig. 1. B6-0 is the first isolate of *P. aeruginosa* in this patient and most likely represents the original environmental strain initiating colonization, while B6-4 is the latest isolate of this particular genotype in the patient. It is not known if this first colonizing clone was removed by the medical treatment, by competition from a new genotype that emerged in the patient airways (Fig. 1), or both. Colony morphology appearance of the two isolates was identical and like that normally observed for wild-type isolates, indicating an environmental origin. The cells are fully motile in agreement with a general wild-type phenotype. Transcriptional profiling was performed through isolation of RNA from cultures growing exponentially in Luria–Broth (LB) medium and harvested at an OD_{600} of 0.5 followed by analysis using Affymetrix Gene Chips for *P. aeruginosa*. Global gene expression comparison of isolates B6-0 and B6-4 shows that only six genes had *P*-values below 0.05, all in the range of a twofold downregulation. Four of these are PA4582-PA4585 of which PA4585 (*rtcA*) is involved in RNA processing. These results suggest that only a small number of genetic changes occurred during the 2 years of colonization of this particular *P. aeruginosa* clone.

![Fig. 1. Time line of *P. aeruginosa* infection in patients B6, B12 and B38. Black boxes indicate *P. aeruginosa*-positive samples while stars, for patient B6, indicate samples containing a new genotype only. Unlike patient B6, only one genotype was detected in patients B12 and B38. Arrows indicate sample isolates chosen for DNA microarray analysis. Sample numbers are not completely consecutive as not all *P. aeruginosa*-positive samples were available for analysis.](image-url)
Characterization of a persisting non-mucoid strain of P. aeruginosa

In patient B12 a specific clone of *P. aeruginosa* has persisted for more than 4 years and is still isolated from the patient. The isolate B12-0 does not represent the first acquisition for this patient, as *P. aeruginosa* was identified in a sample from the patient 4 months prior to sample B12-0. Although we have no stored isolate from this 4 month period it is quite likely that the *P. aeruginosa* clone was the same as the one described in the subsequent collected samples. DNA microarray analysis was performed on isolates from three samples: B12-0, B12-4 (isolated after 2 years and 10 months) and B12-7 (isolated after 3 and 8 months) (cf. Fig. 1). Gene expression changes are shown in Table 1 and in contrast to the lack of changes observed for the B6 isolates, the analysis of the B12 isolates showed that several genes and operons display changed expression levels in the later isolates compared with the first one. These changed expression levels include upregulation of the MexCD-OprJ multidrug efflux system involved in resistance towards, e.g. fluoroquinolones (Jalal *et al*., 2000; Jeannot *et al*., 2008). In agreement with this it was found that the MIC of ciprofloxacin, with which the patient was treated, increased 15-fold from a wild-type level of 0.19 to 3 μg ml⁻¹. The underlying genetic basis is a 13 bp (∆329–341) frameshift deletion in the MexCD-OprJ operon repressor, *nfxB*, present in isolate B12-7. The *atu* operon, also upregulated, is involved in the catabolism of the acyclic monoterpenic citronellol, but also partly involved in catabolism of leucine (Aguilar *et al*., 2006), which may be an important nutrient in the airway mucus. Growth on leucine as carbon source was accordingly significantly improved for isolate B12-7 (doubling time of 4.1 h, SD ± 0.4) compared with isolate B12-4 (doubling time of 9.1 h, SD ± 0.4). A frame-shift deletion (∆A326) in *atuR* in B12-7 could explain this. Several genes showed moderately changed expression, among which were genes encoding type III secretion proteins (downregulation). Because isolate B12-0 presumably had a 4 month period of colonization and adaptation prior to the sampling time, some other changes in expression relative to the original environmental colonizer could have occurred during this period. In fact, gene expression in B12-0 in comparison with the reference strain PA01 revealed up to 10-fold downregulation of *wbp* genes involved in O-antigen synthesis as well as up to 17-fold downregulation of *pil* and *fim* genes involved in motility, in agreement with the finding that B12 isolates are twitching motility deficient.

Characterization of a persisting mucoid strain of *P. aeruginosa*

*Pseudomonas aeruginosa* colonization in patient B38 occurred after a 10 year period of time without *P. aeruginosa*. The time line of this new infection in B38 is displayed in Fig. 1. The initial colonizing strain was wild-type non-mucoid (same phenotype as for isolates from B6 and B12) but rapidly converted to an alginate overproducing mucoid phenotype. In sample B38-2 five different clonal mucoid lineages could be identified by differences in the respective *mucA* sequences (Table 2). The five mucoid lineages displayed some phenotypic diversity and could in part also be distinguished based on the plate colony morphotypes of *in vitro* generated non-mucoid revertants. For the sequenced isolates each colony morphology type matched a specific lineage.

Sequencing the *mucA* gene in isolates from later samples indicated that only the type A lineage carrying the T349C *mucA* allele was continuously present in the patient airways. Investigating colony morphologies of 70 isolates from sample B38-6 revealed that all displayed the morphotype associated with isolates carrying the T349C *mucA* allele. Thus it is strongly indicated that the type A mucoid lineage has dominated the *P. aeruginosa* population from sample B38-6 and onwards and most likely is the only persisting mucoid lineage. Consequently, isolates of the type A mucoid lineage were chosen for transcriptome analysis that included strains B38-1NM, B38-6A-M, B38-2A-M and B38-2A-NM*; an isogenic non-mucoid strain having a PA14 wild-type *mucA* allele. The latter strain was generated *in vitro* by allelic replacement, cf. Experimental procedures. Inter-strain gene expression comparison enables not only the determination of temporal changes in gene expression between isolates from samples 1, 2 and 6, but also documents that changes between isolates from sample 1 and 2 are related or unrelated to the *mucA* mutation.

Gene expression in B38-1-NM was considered to be a wild-type reference data set. In B38-2A-M 761 genes had significantly changed expression relative to B38-1-NM with a *P*-value cut-off of 0.05 and fold change above 2. Most of the gene expression changes were also observed when comparing the isogenic strains B38-2A-M and B38-2A-NM* (668 with a fold change above 2) showing that the genetic basis for most gene expression changes found in B38-2A-M is the *mucA* mutation. A comparison between strains B38-1NM and B38-2A-NM* disclosed fewer and smaller gene expression changes as only 75 genes showed changed expression (*P*-value below 0.05 and fold change above 2), while only one gene was significantly changed above twofold in expression in strain B38-6A-M relative to strain B38-2A-M. Selected differentially expressed genes from isolates from samples 1, 2 and 6 are shown in Table 3.

It is evident that most changes in gene expression shown in Table 3 are caused by the *mucA* mutation, and as expected expression of both *mucA* and *algU* were upregulated in the mucoid strain as was expression of the...
Table 1. All gene expression changes between isolates B12-0 and B12-7.*

<table>
<thead>
<tr>
<th>Locus ID and gene name</th>
<th>Description</th>
<th>Fold change</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibiotic resistance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA4597 oprJ</td>
<td>Multidrug efflux outer membrane protein OprJ precursor</td>
<td>25</td>
</tr>
<tr>
<td>PA4598 mexD</td>
<td>Multidrug efflux transporter MexD</td>
<td>53</td>
</tr>
<tr>
<td>PA4599 mexC</td>
<td>Multidrug efflux membrane fusion protein MexC precursor</td>
<td>86</td>
</tr>
<tr>
<td>PA4600 nfxB</td>
<td>Transcriptional regulator NfxB</td>
<td>12</td>
</tr>
<tr>
<td><strong>Type III secretion</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA1694 pscQ</td>
<td>Translocation protein in type III secretion</td>
<td>–2</td>
</tr>
<tr>
<td>PA1699</td>
<td>Conserved hypothetical protein in type III secretion</td>
<td>–2</td>
</tr>
<tr>
<td>PA1700</td>
<td>Regulator protein PcrH</td>
<td>–2</td>
</tr>
<tr>
<td>PA1708 popB</td>
<td>Translocator protein PopB</td>
<td>–3</td>
</tr>
<tr>
<td>PA1709 popD</td>
<td>Translocator outer membrane protein PopD precursor</td>
<td>–2</td>
</tr>
<tr>
<td>PA1714 exsD</td>
<td>ExsD</td>
<td>–2</td>
</tr>
<tr>
<td>PA1715 pscB</td>
<td>Type III export apparatus protein</td>
<td>–3</td>
</tr>
<tr>
<td>PA1716 pscE</td>
<td>Type III export protein PscE</td>
<td>–3</td>
</tr>
<tr>
<td>PA3841 esoS</td>
<td>Exoenzyme S</td>
<td>–2</td>
</tr>
<tr>
<td><strong>Citronellol/leucine catabolism</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA2885 atuR</td>
<td>Putative repressor of atu genes</td>
<td>3</td>
</tr>
<tr>
<td>PA2886 atuA</td>
<td>Protein with apparent function in citronellol catabolism</td>
<td>15</td>
</tr>
<tr>
<td>PA2887 atuB</td>
<td>Putative dehydrogenase involved in catabolism of citronellol</td>
<td>12</td>
</tr>
<tr>
<td>PA2888 atuC</td>
<td>Geranyl-CoA carboxylase, β-subunit</td>
<td>23</td>
</tr>
<tr>
<td>PA2889 atuD</td>
<td>Putative citronellyl-CoA dehydrogenase, citronellol catabolism</td>
<td>18</td>
</tr>
<tr>
<td>PA2890 atuE</td>
<td>Putative isoheptenylglutaconyl-CoA hydratase</td>
<td>10</td>
</tr>
<tr>
<td>PA2891 atuF</td>
<td>Geranyl-CoA carboxylase, α-subunit (biotin-containing)</td>
<td>3</td>
</tr>
<tr>
<td>PA2892 atuG</td>
<td>GCase, α-subunit (biotin-containing)</td>
<td>8</td>
</tr>
<tr>
<td>PA2893 atuH</td>
<td>Putative very long-chain acyl-CoA synthetase</td>
<td>2</td>
</tr>
<tr>
<td><strong>Exopolysaccharide production</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA2232 pslB</td>
<td>Probable phosphomannose isomerase</td>
<td>–4</td>
</tr>
<tr>
<td>PA2233 pslC</td>
<td>Probable glycosyl transferase</td>
<td>–3</td>
</tr>
<tr>
<td>PA2234 pslD</td>
<td>PslD</td>
<td>–2</td>
</tr>
<tr>
<td>PA2235 pslE</td>
<td>Hypothetical protein</td>
<td>–3</td>
</tr>
<tr>
<td>PA2236 pslF</td>
<td>Hypothetical protein</td>
<td>–3</td>
</tr>
<tr>
<td><strong>LPS modification</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA3552 arnB</td>
<td>ArnB</td>
<td>–5</td>
</tr>
<tr>
<td>PA3553 arnC</td>
<td>ArnC</td>
<td>–3</td>
</tr>
<tr>
<td>PA3554 arnA</td>
<td>ArnA</td>
<td>–3</td>
</tr>
<tr>
<td>PA3556 arnT</td>
<td>ArnT</td>
<td>–2</td>
</tr>
<tr>
<td>PA3558 arnF</td>
<td>ArnF</td>
<td>–3</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA4085</td>
<td>Conserved hypothetical protein</td>
<td>2</td>
</tr>
<tr>
<td>PA4086</td>
<td>Repressor, PthB</td>
<td>–2</td>
</tr>
<tr>
<td>PA4087</td>
<td>Conserved hypothetical protein</td>
<td>–2</td>
</tr>
<tr>
<td>PA4088</td>
<td>Hypothetical protein</td>
<td>–3</td>
</tr>
<tr>
<td>PA4089</td>
<td>Hypothetical protein</td>
<td>2</td>
</tr>
<tr>
<td>PA4105</td>
<td>Probable acyl-CoA dehydrogenase</td>
<td>4</td>
</tr>
<tr>
<td>PA4106</td>
<td>Hypothetical protein</td>
<td>–4</td>
</tr>
<tr>
<td>PA4156</td>
<td>Hypothetical protein</td>
<td>–5</td>
</tr>
<tr>
<td>PA4193 hcnA</td>
<td>Hydrogen cyanide synthase HcnA</td>
<td>–2</td>
</tr>
<tr>
<td>PA4198</td>
<td>Probable dehydrogenase</td>
<td>–3</td>
</tr>
<tr>
<td>PA4206</td>
<td>Probable dehydrogenase</td>
<td>–10</td>
</tr>
<tr>
<td>PA4230 nucD</td>
<td>NADH dehydrogenase</td>
<td>–2</td>
</tr>
<tr>
<td>PA4355</td>
<td>Hypothetical protein</td>
<td>2</td>
</tr>
<tr>
<td>PA4356</td>
<td>Motility protein FimV</td>
<td>2</td>
</tr>
<tr>
<td>PA4340</td>
<td>d-Erythro-7,8-dihydroneopterin triphosphate epimerase</td>
<td>–5</td>
</tr>
<tr>
<td>PA4346</td>
<td>Hypothetical protein</td>
<td>–2</td>
</tr>
<tr>
<td>PA4347</td>
<td>Hypothetical protein</td>
<td>2</td>
</tr>
<tr>
<td>PA4348</td>
<td>Probable chaperone</td>
<td>–2</td>
</tr>
<tr>
<td>PA4349</td>
<td>Probable chaperone</td>
<td>–2</td>
</tr>
<tr>
<td>PA4350</td>
<td>Conserved hypothetical protein</td>
<td>–3</td>
</tr>
<tr>
<td>PA4351</td>
<td>Chemotactic transducer PctC</td>
<td>–2</td>
</tr>
<tr>
<td>PA4352</td>
<td>Probable transcriptional regulator</td>
<td>2</td>
</tr>
<tr>
<td>PA4353</td>
<td>Motility regulator</td>
<td>2</td>
</tr>
<tr>
<td>PA4354</td>
<td>Hypothetical protein</td>
<td>–3</td>
</tr>
<tr>
<td>PA4355</td>
<td>Two-component regulator system response regulator</td>
<td>–2</td>
</tr>
<tr>
<td>PA4356</td>
<td>Xanthine phosphoribosyltransferase</td>
<td>–3</td>
</tr>
<tr>
<td>PA4357</td>
<td>Probable two-component system</td>
<td>–4</td>
</tr>
<tr>
<td>PA4358</td>
<td>RNA, histidine</td>
<td>–2</td>
</tr>
<tr>
<td>PA4359</td>
<td>RNA, lysine</td>
<td>–2</td>
</tr>
<tr>
<td>PA4360</td>
<td>RNA, methionine</td>
<td>–3</td>
</tr>
<tr>
<td>PA4361</td>
<td>Intergenic region</td>
<td>3</td>
</tr>
</tbody>
</table>

a. Gene expression changes are displayed as discrete fold changes, converted from log (2) fold changes. Negative values correspond to downwards fold change.

b. Gene expression change present between isolates B12-4 and B12-7, i.e. likely an effect of mutations occurring between these samples.

© 2010 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*, 12, 1643–1658
AlgU controlled regulators *algR* and *algB*. Among the observed changes in gene expression in the *mucA* isolate it was striking that expression of several virulence factor genes was downregulated in the mucoid strain suggesting reduced production of phenazines involved in pyocyanin synthesis, hydrogen cyanide, rhamnolipids, type III and type VI secretion proteins, motility factors, chemotaxis proteins and Rhl quorum sensing proteins. These down-regulations may not be a direct regulatory effect of AlgU, but rather indirect, as several well-known quorum sensing and virulence factor regulatory genes were also differentially expressed. Other pleiotropic effects of the *mucA* mutation include the expression upregulation of the genes *osmC* and *osmE* induced by osmotic stress in *E. coli* (Gutierrez and Devedjian, 1991; Gutierrez et al., 1995), and a number of genes shown to be part of the osmotic stress response of PA14 encoding hypothetical proteins or proteins putatively involved in the synthesis of the osmo-protectant N-acetyl glutaminyl glutamine amide (NAGGN) (Aspedon et al., 2006).

Because the most pronounced phenotypic effect of *mucA* mutations is the constitutive high-level production of extracellular alginate, it was important to clarify the direct and indirect consequences of the synthesis of this polymer. Therefore a B38-A2-M *algD* mutant strain was constructed, cf. Experimental procedures. The resulting strain produces no alginate, is still *mucA* deficient and expresses wild-type AlgU. As shown in the transcriptome analysis (Table 3) most of the gene expression changes that could be specifically associated with alginate overproduction are related to metabolism, e.g. downregulation of lactate and glycerol metabolism genes. Otherwise the gene expression profiles of B38-2A-M and the *algD* mutant strain are quite similar (correlation coefficient factor of 0.95), especially concerning the virulence factors.

Some of the differentially regulated genes unrelated to the *mucA* mutation include the operon (PA3186-PA3190) involved in transport of glucose, fructose, glycerol and mannitol (Wylie and Worobec, 1995) in addition to an operon involved in glutamate and aspartate transport and metabolism, both of which are upregulated. The only differentially expressed gene with a fold change above 2 when comparing isolates B38-2A-M and B38-6A-M is the *mexX* gene (2.6-fold), while a fold change of 1.7 is present for *mexY*. These genes encode part of the MexXY OpmR efflux pump system involved in antibiotic resistance towards aminoglycosides, e.g. tobramycin (Aires et al., 1999; Islam et al., 2009). However the fold changes were low and assessing the MIC of tobramycin revealed only a minor increase in MIC from 0.75 for B38-2A-M to 1.5 for B38-6A-M.

### Phenotypic consequences of *mucA* mutations

The gene expression data presented here suggest that the mucoid phenotype is associated with reduced virulence caused by the regulatory consequences of the *mucA* mutation. Because reduced virulence is one of the characteristics of *P. aeruginosa* isolates from chronically infected patients it was important to determine if the changed gene expression pattern related to the virulence genes was indeed correlated with a reduced virulence phenotype of *mucA* strains. Additionally, the generality of a possible reduction of virulence because of *mucA* mutations should be assessed in different strain pairs of *P. aeruginosa*, as clinical mucoid isolates have different genomes and carry different specific *mucA* mutations that could express different associated phenotypes. Thus the phenotypic analysis described in the following is both a control of the transcription analysis described in the previous section and an expanded investigation of the pleiotropy of *mucA* mutations.

Comparative analysis of mucoid and non-mucoid strains of *P. aeruginosa* can only be carried out reliably from isogenic pairs of strains. For the B38 isolates allelic replacements of the mutated *mucA* gene in B38-2A-M with a wild-type allele resulted in non-mucoid B38-2A-NM*, and a replacement of the wild-type *algD* gene with a mutant *algD* allele in B38-2A-M resulted in non-mucoid B38-2A-M *algD*. Similar allelic replacements were constructed for four other different mucoid strains isolated during early infection of CF children, and for the non-mucoid reference strain PAO1 an isogenic mucoid *mucA* strain, PDO300, was employed. The *mucA* genes of the mucoid variants of these pairs harbour different mutations. The resulting isogenic mucoid/non-mucoid pairs of strains were analysed phenotypically as described below.

The suggested reduced virulence of *mucA* mutant strains was confirmed by phenotypic characterization of the strains. Overall virulence activities of the relevant strains were assessed using the *Caenorhabditis elegans* worm killing assay. It was previously shown that *P. aeruginosa* isolates from chronically infected patients do not kill the worms, whereas isolates from the environment or from early stages of infection in CF children efficiently killed the worms. The comparison between non-mucoid, *mucA* mucoid and non-mucoid *mucA*, *algD* strains showed that

<table>
<thead>
<tr>
<th>Strain</th>
<th><em>mucA</em> mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>B38-2A-M</td>
<td>T349C</td>
</tr>
<tr>
<td>B38-2B-M</td>
<td>T382C</td>
</tr>
<tr>
<td>B38-2C-M</td>
<td>C insert 251</td>
</tr>
<tr>
<td>B38-2D-M</td>
<td>ΔGA478–479</td>
</tr>
<tr>
<td>B38-2E-M</td>
<td>ΔGA430</td>
</tr>
</tbody>
</table>

© 2010 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*, 12, 1643–1658
Table 3. Selected gene expression changes between B38 non-mucoid and mucoid strains.\textsuperscript{a,b}

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alginate regulation and biosynthesis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA0762 algU</td>
<td>Sigma factor</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>PA0763 mucA</td>
<td>Anti-sigma factor</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>PA5261 algR</td>
<td>Alginate biosynthesis regulatory protein</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>PA5483 algB</td>
<td>Two-component response regulator</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>PA3540 algD</td>
<td>GDP-mannose 6-dehydrogenase</td>
<td>181</td>
<td>180</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Virulence factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA2386 pvdA</td>
<td>L-ornithine N5-oxygenase</td>
<td>-2</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>PA2396 pvdF</td>
<td>Pseudo-enzyme synthase P</td>
<td>&gt;</td>
<td>4</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>PA2426 pvdS</td>
<td>Sigma factor</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>PA1905 phzG2</td>
<td>Probable pyridoxamine 5'-phosphate oxidase</td>
<td>&gt;</td>
<td>-3</td>
<td>-6</td>
<td>-6</td>
</tr>
<tr>
<td>PA2421 phzB1</td>
<td>Probable phenazine biosynthesis protein</td>
<td>&gt;</td>
<td>-4</td>
<td>-6</td>
<td>-6</td>
</tr>
<tr>
<td>PA4217 phzS</td>
<td>Flavon-containing monoxygenase</td>
<td>&gt;</td>
<td>-5</td>
<td>-7</td>
<td>-7</td>
</tr>
<tr>
<td>PA2193 hcnA</td>
<td>Hydrogen cyanide synthase</td>
<td>-20</td>
<td>-24</td>
<td>-27</td>
<td></td>
</tr>
<tr>
<td>PA2194 hcnB</td>
<td>Hydrogen cyanide synthase</td>
<td>&gt;</td>
<td>-5</td>
<td>-6</td>
<td>-5</td>
</tr>
<tr>
<td>PA2195 hcnC</td>
<td>Hydrogen cyanide synthase</td>
<td>&gt;</td>
<td>-4</td>
<td>-6</td>
<td>-5</td>
</tr>
<tr>
<td>PA3479 rhlA</td>
<td>Rhamnosyltransferase chain A</td>
<td>&gt;</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>PA4229 pchD</td>
<td>Pyochelin biosynthesis protein</td>
<td>-7</td>
<td>-7</td>
<td>-5</td>
<td>-5</td>
</tr>
<tr>
<td>PA4230 pchB</td>
<td>Salicylic biosynthesis protein</td>
<td>-17</td>
<td>-10</td>
<td>-9</td>
<td></td>
</tr>
<tr>
<td>PA4229 pchC</td>
<td>Pyochelin biosynthetic protein</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td></td>
</tr>
<tr>
<td><strong>Secretion systems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA1713 exsA</td>
<td>Transcriptional regulator type III secretion</td>
<td>&gt;</td>
<td>-4</td>
<td>-7</td>
<td>-5</td>
</tr>
<tr>
<td>PA1718 pscE</td>
<td>Type III export protein</td>
<td>&gt;</td>
<td>-12</td>
<td>-18</td>
<td>-27</td>
</tr>
<tr>
<td>PA1719 pscF</td>
<td>Type III export protein</td>
<td>&gt;</td>
<td>-4</td>
<td>-6</td>
<td>-7</td>
</tr>
<tr>
<td>PA0083</td>
<td>Type VI secretion</td>
<td>-2</td>
<td>-27</td>
<td>-13</td>
<td>-5</td>
</tr>
<tr>
<td>PA0084</td>
<td>Type VI secretion</td>
<td>-2</td>
<td>-18</td>
<td>-11</td>
<td>-3</td>
</tr>
<tr>
<td>PA0085</td>
<td>HcpL, secreted protein</td>
<td>&lt;</td>
<td>-28</td>
<td>-17</td>
<td>-6</td>
</tr>
<tr>
<td><strong>Regulation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA0652 vfr</td>
<td>Quorum sensing regulator</td>
<td>&lt;</td>
<td>-2</td>
<td>-2</td>
<td>-2</td>
</tr>
<tr>
<td>PA0905 rsmA</td>
<td>Regulator of secondary metabolites and/or quorum sensing</td>
<td>2</td>
<td>&gt;</td>
<td>&lt;</td>
<td></td>
</tr>
<tr>
<td>PA1003 mvfR</td>
<td>Quorum sensing regulator</td>
<td>&lt;</td>
<td>-2</td>
<td>-2</td>
<td>-2</td>
</tr>
<tr>
<td>PA3385 amrZ</td>
<td>Alginate and motility regulator Z</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>PA4856 retS</td>
<td>Regulator of virulence factors and type III and VI secretion</td>
<td>-2</td>
<td>&lt;</td>
<td>&lt;</td>
<td></td>
</tr>
<tr>
<td><strong>Quorum sensing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA3476 rhl</td>
<td>Auto-inducer synthesis protein Rhl</td>
<td>-4</td>
<td>-4</td>
<td>-3</td>
<td></td>
</tr>
<tr>
<td>PA3477 rhlR</td>
<td>Rhl quorum sensing regulator</td>
<td>-2</td>
<td>-3</td>
<td>-2</td>
<td></td>
</tr>
<tr>
<td><strong>Motility</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA1078 flgC</td>
<td>Flagellar basal-body rod protein</td>
<td>-17</td>
<td>-17</td>
<td>-21</td>
<td></td>
</tr>
<tr>
<td>PA1092 flIC</td>
<td>Flagellin type B</td>
<td>-48</td>
<td>-44</td>
<td>-59</td>
<td></td>
</tr>
<tr>
<td>PA1094 flID</td>
<td>Flagellar capping protein</td>
<td>-23</td>
<td>-21</td>
<td>-23</td>
<td></td>
</tr>
<tr>
<td>PA5042 pilO</td>
<td>Type 4 fimbrial biogenesis protein</td>
<td>-5</td>
<td>-5</td>
<td>-4</td>
<td>-4</td>
</tr>
<tr>
<td>PA5043 pilN</td>
<td>Type 4 fimbrial biogenesis protein</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td></td>
</tr>
<tr>
<td>PA5044 pilM</td>
<td>Type 4 fimbrial biogenesis protein</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td></td>
</tr>
<tr>
<td><strong>Osmotic stress</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA0059 osmC</td>
<td>Osmotically inducible protein</td>
<td>&gt;</td>
<td>11</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>PA4876 osmE</td>
<td>Osmotically inducible lipoprotein</td>
<td>9</td>
<td>7</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>PA1323</td>
<td>Putative osmotic stress response</td>
<td>2</td>
<td>11</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>PA1324</td>
<td>Putative osmotic stress response</td>
<td>&gt;</td>
<td>7</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>PA2146</td>
<td>Putative osmotic stress response</td>
<td>8</td>
<td>8</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>PA3459</td>
<td>Putative osmotic stress response</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>PA3460</td>
<td>Putative osmotic stress response</td>
<td>&gt;</td>
<td>10</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>PA3461</td>
<td>Putative osmotic stress response</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>PA4860</td>
<td>Putative osmotic stress response</td>
<td>2</td>
<td>12</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td><strong>Metabolism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA3584 gldD</td>
<td>Glycerol-3-phosphate dehydrogenase</td>
<td>-16</td>
<td>-20</td>
<td>-12</td>
<td></td>
</tr>
<tr>
<td>PA4770 lipo</td>
<td>L-lactate permease</td>
<td>2</td>
<td>-11</td>
<td>-23</td>
<td></td>
</tr>
<tr>
<td><strong>Carbohydrate transport</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA3186 oppB</td>
<td>Glucose/carbohydrate outer membrane porin</td>
<td>3</td>
<td>3</td>
<td>&gt;</td>
<td>&lt;</td>
</tr>
<tr>
<td>PA3187</td>
<td>Probable ATP binding component of ABC transporter</td>
<td>7</td>
<td>8</td>
<td>&lt;</td>
<td></td>
</tr>
<tr>
<td>PA3188</td>
<td>Probable permease of ABC sugar transporter</td>
<td>8</td>
<td>12</td>
<td>&gt;</td>
<td>&lt;</td>
</tr>
<tr>
<td>PA3189</td>
<td>Probable permease of ABC sugar transporter</td>
<td>5</td>
<td>9</td>
<td>&gt;</td>
<td>&lt;</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Gene expression changes are displayed as discrete fold changes, converted from log (2) fold changes. Negative values correspond to downwards fold change.

\textsuperscript{b} \( \text{A > B or A < B} \) denotes a significant upwards or downwards, below twofold, gene expression change respectively \((P < 0.05)\). The presence of no fold change number signifies an insignificant change in gene expression \((P > 0.05)\).
mucA strains independent of the actual production of alginate were less virulent than wild-type strains (Fig. 2) in agreement with the gene expression profile showing reduced virulence factor expression in B38-2A-M and the isogenic algD mutant strain (Table 3). The less pronounced difference observed for the B38 strains is due to rapid reversion to fast growing non-mucoid cells (> 95% for B38-2A-M and about 40% for B38-2A-M algD). However, even for the B38 strains the mucA mutation resulted in reduced killing of the worms.

Among the many virulence factors suggested to be expressed with reduced rates in the mucA strains (based on the transcriptomic data set) we chose to investigate further by phenotypic analysis the following activities: motility, rhamnolipid production and quorum sensing signal (the C4 Rhl signal) production. Table 4 shows that compared with the isogenic non-mucoid strain the mucoid variant B38-2A-M displayed a reduction in all types of motility, rhamnolipid and Rhl quorum sensing signal synthesis. Although B38-2A-M grows substantially slower than its isogenic non-mucoid strain (doubling time 60 vs. 35 min) it is not the cause of reduced virulence as B38-2A-M algD is almost equally slow growing. There was an overall trend of reduction in motility for all investigated mucoid isolates with minor deviations as, e.g. two pair of strains almost retain full swimming motility, and also rhamnolipid production differs between strains. This could indicate that the effect of the mucA mutation is strain background-dependent (e.g. isolate B7 seems more virulent than most others). Nevertheless, the specific virulence-associated phenotypes assayed here support the overall virulence activity measured in the C. elegans assay showing a reduced activity for the mucA strains, which also confirms the transcriptomic data.

Discussion

Do the bacteria show indications of being on a path towards a ‘chronic infection phenotype’, and if so which underlying routes and mechanisms are involved?

The first simple answer to the question is that for two of the patients (B12 and B38) genetic changes have led to development of new bacterial phenotypes during the course of colonization, and in both cases the derived prominent phenotypes – antibiotic resistance and mucoidy – are associated with the ‘chronic infection phenotype’. In fact, colonization in these two patients has progressed to a chronic infection. In the former case (B12) the mechanism behind the increased antibiotic resistance of the bacteria is a mutation in nfxB leading to upregulation of expression of the MexCD-OprJ efflux pump (Jalal et al., 2000; Jeannot et al., 2008). In the latter case (B38) the mechanism behind the bacterial mucoidy is upregulation of the alginate biosynthesis enzymes encoded by the alg operon, caused by a mutation in the mucA anti-sigma factor. For the last patient (B6) no obvious phenotypic changes occurring during the 2 year persistence of P. aeruginosa seem related to the ‘chronic infection phenotype’, and persistence can therefore be explained by either introduction of unidentified adaptive mutations or simply the general adaptive repertoire of this organism. Our data do not permit us to suggest any specific mechanisms in this case.

The high resolution of our phenotypic analysis (DNA arrays for transcriptional profiling) allows us, however, to provide more detailed answers to the question. Thus, in addition to the increased ciprofloxacin resistance mediated by the mexCD-OprJ upregulation, the later B12 isolates also show phenotypic changes reducing expression of virulence factors (TTSS). Furthermore, all B12 isolates...
are twitching motility deficient and compared with PAO1 expression of genes (wbp) involved in O-antigen synthesis were significantly downregulated, likely causing the LPS modifications frequently found in CF infection (Goldberg and Pier, 1996). Increased expression of the atu operon conferring improved growth on leucine has not previously been reported but probably reflects another less studied but potentially important means of adaptation by metabolic optimization. Sputum of CF patients is known to be rich in amino acids (Barth and Pitt, 1996), and a 12-patient study showed that leucine is the fourth most abundant amino acid in CF sputum present in almost 1 mM concentrations on average (Palmer et al., 2007). Pleiotropic effects of the upregulation of MexCD-OprJ often include reduced virulence and growth rate (Sanchez et al., 2002; Linares et al., 2005; Jeannot et al., 2008), but in our investigations no growth rate differences were observed for B12-4 and B12-7 relative to that of B12-0. Nonetheless, it is clear that the mutations responsible for the changed phenotypes of the B12 isolates – whether pleiotropic or not – shift the colonizing bacteria in direction of a ‘chronic infection phenotype’.

For the B38 isolates, the shift towards the ‘chronic infection phenotype’ is even more apparent. With a single mutation (mucA) the ‘chronic infection phenotype’ is approached covering most of the characteristics, e.g. alginate overproduction and reduction in virulence, motility and growth rate. The transcription data of the B38 isolates show that the mucA mutation has a substantial pleiotropic effect on gene expression. Here we will focus most of our discussion on gene expression changes that have been confirmed by phenotypic analysis (cf. Supporting information for complete data sets). Interestingly, it turned out that several well-known virulence genes in P. aeruginosa showed clear reduction in expression in the mucoid mucA strain compared with an isogenic non-mucoid strain. Previous studies have investigated the pleiotropic effects of mucA mutations or investigated the role of the relevant regulators, e.g. AlgU and AlgR. In a transcriptome analysis of PAO381 – a derivative of PAO1 – and its isogenic mucA- strain very different results compared with this study were obtained (Firoyed and Deretic, 2003). Of the 27 differentially expressed genes published by Firoyed and Deretic only 3 genes were regulated in the same direction, 9 in the opposite direction and the rest unchanged in the B38 transcriptome analyses. Although Firoyed and Deretic (2003) applied a strict cut-off value and only included genes with fold changes above 4 it cannot explain the observed differences. Strain differences and/or the specific conditions of the experiments may explain the observed differences.

Several other publications seem to corroborate the findings of the mucoid phenotype observed in this study. Thus it was found that a mucA mutation resulted in reduced expression of TTSS-related genes, AlgU shown to repress flagellar synthesis, and AlgR found to repress the rhl QS system in biofilms (Wu et al., 2004; Tart et al., 2006; Morici et al., 2007). However, differences have also been observed as mucoid strains produced more hydrogen cyanide than non-mucoid strains (Carterson et al., 2004) using a different experimental setup than ours, which could suggest that hcn regulation through AlgU is dependent on growth conditions. Yet, in non-mucoid strains AlgR was shown to be a hydrogen cyanide repressor (Lizewski et al., 2004).

The reduced expression of virulence factors in the mucA strains suggests that these variants have lost part of their acute infection potential and gained potential for chronic infections. There are several plant and animal models for acute infections with P. aeruginosa, one of which is the C. elegans killing assay (Tan et al., 1999; Thomsen et al., 2006), which was chosen here for assessments of the virulence activity of the mucoid/non-mucoid strain pairs. Previous investigations have shown that P. aeruginosa isolates from the environment or from early stages of infection in CF children in most cases are fully virulent resulting in rapid killing of the nematodes, whereas isolates from older chronically infected patients most often displayed reduced or no virulence (Jelsbak et al., 2007). The results presented here show that mucA strains in general display reduced acute virulence, as measured in the C. elegans assay, thus confirming the in vitro data.

The gene expression consequences of the mucA mutations affecting the AlgU regulon could be further selectively advantageous in the stressful environment of the CF airways. In fact, a subset of genes putatively involved in the osmotic stress response was shown here to be upregulated in the mucA strain (expected to be beneficial in the CF airway mucus). It was furthermore found that a large fraction of cell wall stress-induced genes belong to the AlgU regulon (Wood and Ohman, 2009). In vitro conversion to a mucoid phenotype is known to occur during both oxidative and osmotic stress conditions (Terry et al., 1991; Mathee et al., 1999). In other Gram-negative bacteria the role of AlgU homologues (σE) is also stress-related, e.g. by ensuring homeostasis of lipopolysaccharide and outer membrane porins (Rhodius et al., 2006). It was further proposed (Rhodius et al., 2006) that the σE regulon in general may constitute an adaptation system to facilitate survival in vivo; this seems to be true for P. aeruginosa. The linkage between alginate production, increased oxidative and osmotic stress response and reduction of virulence provides a powerful explanation to why this regulon is repeatedly targeted for activation by MucA mutation in the CF lung environment. Importantly, the effective alteration of the MucA-AlgT regulatory circuitry constitutes...
the basis for the pleiotropic phenotypic switch by a single mutational event.

At least five different mucoid lineages arose almost simultaneously in patient B38 as evidenced by mucA sequences and while it seems clear that only one lineage persisted, the reason for this is less clear. The presence of additional adaptation specific for this lineage is one explanation whereas another could be the specific pleiotropic effects of the mucA mutation in this lineage. Despite their clonal relationship the five lineages displayed small phenotypic differences in the virulence properties tested here (data not shown). If not because of other genetic differences among the lineages the latter could indicate that different mucA mutations may cause different degrees of pleiotropic effects based on varying interactions with the AlgT or MucB proteins.

The appearance of mucA variants in samples from CF patients is often very dominating for a period after the occurrence of the mutation(s) in the bacterial airways population, but subsequently non-mucoid clones appear and constitute substantial sub-populations in the patients. In most cases of chronically infected patients it has been shown that by far most of these non-mucoid clones still harbour mucA mutations (nearly always the same as the mucoid counterparts) (Ciofu et al., 2001; Jelsbak et al., 2007). In laboratory conditions most of such second-site revertants carry mutations in the algU gene, but this is not the normal cause of non-mucoid phenotypes in the CF patients (Ciofu et al., 2008). Although it is still not clear which mutations are responsible for the phenotypic reversion of mucoidy in these cases, the findings suggest that there is selection for maintenance of the mucA phenotype combined with a wild-type allele of the algU gene, and it is tempting to speculate that one reason for this is the apparent beneficial pleiotropic effects described in this communication. Moreover it could indicate that the most important pleiotropic effect of AlgU activation might not be alginate overproduction, but rather one or a collection of the other aforementioned pleiotropic effects.

The three cases of early colonization of human hosts have shown that P. aeruginosa normally residing in natural environments can move successfully to highly different and indeed stressful environmental settings including the airways of humans and persist for hundreds or even thousands of generations without a significant number of genetic adaptive changes. The normal adaptive repertoire of wild-type strains has some limits, however, and without fitness-increasing genetic alterations the host defence or competition from other strains will result in eventual eradication of the cells. In the event that fitness-improving mutations do occur in the bacterial genome, persistence may be extended. In the present cases, of the three main characteristics of the CF lung environment, host immune system, antibiotics and substrate composition, the B12 isolates seem to possess increased fitness towards all of these while fitness increase in the B38 isolates seems to relate to the immune system and, less pronounced, antibiotics. Meanwhile, the route towards ‘the chronic phenotype’ and fitness increase seems to consist either of the additive effects of several mutations or, for the mucoid variant, by a mutation with profound pleiotropic effects. Identification of pleiotropic adaptive mutations in regulatory genes is in fact a recurrent finding in a number of experimental evolution experiments with microbial populations in the laboratory (Kolter, 1999; Zinser and Kolter, 1999; Elena and Lenski, 2003; Bantinaki et al., 2007), documenting net fitness increases despite highly complex resulting phenotypes. Pseudomonas aeruginosa infections in CF airways may help to further understand how organisms can move between highly different natural environments helped by a low number of mutations with high phenotypic impact.

Experimental procedures

Bacterial strains and growth conditions

Bacterial strains and clinical isolates used in this study are listed in Table 5. The general isolation procedure and identification of P. aeruginosa from sputum was performed as previously described (Høiby and Frederiksen, 2000). Sputum samples were obtained by expectoration or endotracheal suction, followed by Gram staining. Pseudomonas aeruginosa was isolated on selective plates and identified by conventional biochemical tests. Up to 84 isolates from each sample were selected and stored in microtitre plates at –80°C including all colonies with a distinct morphotype. The genotype of P. aeruginosa isolates was identified by single-nucleotide polymorphism (SNP) typing using AT biochips (Clondiag Chip Technologies, Germany) (Wiehlmann et al., 2007). Unless otherwise noted bacteria were cultured in LB at 37°C. Antimicrobial agents were used where appropriate at the following concentrations: ampicillin at 100 μg ml⁻¹, gentamycin sulfate at 30 μg ml⁻¹ (P. aeruginosa) and 15 μg ml⁻¹ (E. coli), carbenicillin at 200 μg ml⁻¹. Estimation of growth rate on leucine as carbon source was performed essentially as previously described (Martin et al., 1973; Aguilar et al., 2006) using M9 medium containing 0.6% wt/vol L-leucine and supplemented with 0.005% L-isoleucine and L-valine. Growth was performed shaking at 37°C in 250 ml flasks containing 50 ml of medium.

CF patients

As part of a larger study we investigated P. aeruginosa isolates from the majority of CF children attending the Copenhagen CF clinic from 2005 and onwards having
their first acquisition of *P. aeruginosa* during this period, or for a few patients just before 2004. As part of the general management structure at the CF Center, all patients were monitored on a monthly basis by evaluation of their clinical status, pulmonary function and microbiology of lower-airway secretions. Detection of *P. aeruginosa* in sputum resulted in treatment with antimicrobial eradication therapy consisting of 3 weeks of oral ciprofloxacin plus 3 months of inhaled colistin. Instead of applying the clinical definition of intermittent colonization and chronic infection (Heiby et al., 2005), we here distinguish between cases of cured and persistent colonization of the patient airways. Colonization with a specific genotype of *P. aeruginosa*, which is absent from all consecutive samples collected over a period of at least 1 year, is considered cured, whereas continued colonization for more than a year with the same genotype is considered a persistent airway colonization. Analysis was performed on bacterial isolates from three selected patients (B6, B12 and B38) having persistent *P. aeruginosa* colonization. Two of the patients (B12 and B38) had chronic *P. aeruginosa* infections according to the clinical definition based on antibody titre level increases: for B12 the antibody titre level passed 2 after 51 months while reaching 7 after 56 months, and for B38 a level of 2 was passed after 7 months while reaching 21 after 19 months. In contrast the antibody titre level was 1 after 48 months for B6.

**mucA allelic replacement**

Isogenic non-mucoid strains were generated from mucoid clinical isolates by allelic replacement of the mutated mucA allele with a wild-type *P. aeruginosa* PA14 mucA allele (the B38-1 mucA sequence is identical to the PA14 mucA sequence). Initially, a 1.8 kb fragment containing the PA14 mucA allele was generated by PCR using primers CCATGTTGCAAGAAAGGCCGAGTCTAT and GAATTCAGACTGGGTAAGCTGAAAC containing an Ncol and EcoRI site respectively. The PCR fragment was cloned into the EcoRI and Ncol site of plasmid pEX18ApGW. A PCR fragment containing a gentamycin resistance cassette was amplified from pPS856 using primers CCATGCGAATTAGCTTCAAAAGCGCTCTGA and GGATCCCGAATTGGGGATCTTGAAGTTCCT and

### Table 5. Strains and plasmid used in this study.

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Description</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA01</td>
<td>Wild-type</td>
<td>Holloway and Morgan (1986)</td>
</tr>
<tr>
<td>PA14</td>
<td>Wild-type</td>
<td>Rahme and colleagues (1995)</td>
</tr>
<tr>
<td>PDO300</td>
<td>PA01 mucA22 (∆G430)</td>
<td>Mathee and colleagues (1999)</td>
</tr>
<tr>
<td>PDO300 algD</td>
<td>algD, non-mucoid strain derived from PDO300</td>
<td>This study</td>
</tr>
<tr>
<td>E. coli HB101</td>
<td>supE44 hsdR20 (rK12) recA14 ara-14 proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1 leuB thi-1</td>
<td>Kessler and colleagues (1992)</td>
</tr>
<tr>
<td>B3-M</td>
<td>Mucoid isolate from sample 3 from patient B3 having mucA mutation ∆A368</td>
<td>This study</td>
</tr>
<tr>
<td>B3-NM*</td>
<td>Non-mucoid strain derived from B3-M. mucA allele replaced</td>
<td>This study</td>
</tr>
<tr>
<td>B6-0</td>
<td>Non-mucoid isolate from sample 0 from patient B6</td>
<td>This study</td>
</tr>
<tr>
<td>B6-4</td>
<td>Non-mucoid isolate from sample 4 from patient B6</td>
<td>This study</td>
</tr>
<tr>
<td>B6-M</td>
<td>Mucoid isolate from sample 6 from patient B6 having mucA mutation C352T</td>
<td>This study</td>
</tr>
<tr>
<td>B6-NM*</td>
<td>Non-mucoid strain derived from B6-M. mucA allele replaced</td>
<td>This study</td>
</tr>
<tr>
<td>B7-M</td>
<td>Mucoid isolate from sample 3 from patient B7 having mucA mutation C424T</td>
<td>This study</td>
</tr>
<tr>
<td>B7-NM*</td>
<td>Non-mucoid strain derived from B7-M. mucA allele replaced</td>
<td>This study</td>
</tr>
<tr>
<td>B7-M algD</td>
<td>algD, non-mucoid strain derived from B7-M</td>
<td>This study</td>
</tr>
<tr>
<td>B11-M</td>
<td>Mucoid isolate from sample 4 from patient B11 having mucA mutation T382C</td>
<td>This study</td>
</tr>
<tr>
<td>B11-NM*</td>
<td>Non-mucoid strain derived from B11-M. mucA allele replaced</td>
<td>This study</td>
</tr>
<tr>
<td>B12-0</td>
<td>Non-mucoid isolate from sample 0 from patient B12</td>
<td>This study</td>
</tr>
<tr>
<td>B12-4</td>
<td>Non-mucoid isolate from sample 4 from patient B12</td>
<td>This study</td>
</tr>
<tr>
<td>B12-7</td>
<td>Non-mucoid isolate from sample 7 from patient B12</td>
<td>This study</td>
</tr>
<tr>
<td>B38-1-NM</td>
<td>Non-mucoid isolate from sample 1 from patient B38</td>
<td>This study</td>
</tr>
<tr>
<td>B38-2A-M</td>
<td>Mucoid isolate from sample 2 from patient B38 mucA mutation T349C</td>
<td>This study</td>
</tr>
<tr>
<td>B38-2A-NM*</td>
<td>Non-mucoid strain derived from B38-2A-M. mucA allele replaced</td>
<td>This study</td>
</tr>
<tr>
<td>B38-2A-M algD</td>
<td>algD, non-mucoid strain derived from B38-2A-M</td>
<td>This study</td>
</tr>
<tr>
<td>B38-6A-M</td>
<td>Mucoid isolate from sample 6 from patient B38 mucA mutation T349C</td>
<td>This study</td>
</tr>
</tbody>
</table>

### Plasmid

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Description</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRK600</td>
<td>Cm&lt;sup&gt;+&lt;/sup&gt;, oriColE1, RK2-Mob&lt;sup&gt;+&lt;/sup&gt;, RK2-Tra&lt;sup&gt;+&lt;/sup&gt;; helper plasmid for triparental matings</td>
<td>de Lorenzo and Timmis (1994)</td>
</tr>
<tr>
<td>pDONR221</td>
<td>Km&lt;sup&gt;+&lt;/sup&gt;; Gateway donor vector</td>
<td>Rohwer and Edwards (2002)</td>
</tr>
<tr>
<td>pPS856</td>
<td>Ap&lt;sup&gt;+&lt;/sup&gt;, Gm&lt;sup&gt;+&lt;/sup&gt;; Gm cassette flanked with FRT sequences</td>
<td>Hoang and colleagues (1998)</td>
</tr>
<tr>
<td>pFLP2</td>
<td>Ap&lt;sup&gt;+&lt;/sup&gt;; source for Flp recombinase</td>
<td>Hoang and colleagues (1998)</td>
</tr>
<tr>
<td>pMHR1</td>
<td>Ap&lt;sup&gt;+&lt;/sup&gt;, Gm&lt;sup&gt;+&lt;/sup&gt;; Gm cassette from pPS856 and 1.8 kb algT-mucA-mucB fragment from PA14</td>
<td>This study</td>
</tr>
<tr>
<td>pMHR2</td>
<td>Ap&lt;sup&gt;+&lt;/sup&gt;, Gm&lt;sup&gt;+&lt;/sup&gt;; from pEX18ApGW, Gm cassette and FRT sequences from pPS856 flanked with algD gene fragment from B38-2A-M</td>
<td>This study</td>
</tr>
</tbody>
</table>

© 2010 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*, 12, 1643–1658
cloned in the NcoI and BamHI sites of pEX18ApGW containing the PA14 muaA PCR fragment, creating pMHR1. Plasmid pMHR1 was introduced into P. aeruginosa clinical isolates by triparental conjugation employing helper strain HB101/pRK600. Single recombinants were selected by screening for gentamicin resistance and verifying carbencillin resistance, while sacB-mediated sucrose counterselection was employed for plasmid excision leading to allele replacement. To avoid selecting non-mucoid spontaneous revertants it was always verified that sucrose counterselection generated both mucoid plasmid excised and non-mucoid plasmid excised strains. Allelic replacement was verified by muaA sequencing while potential AlgU mutations were excluded by AlgU sequencing.

**Generation of unmarked P. aeruginosa algD deletion mutants**

algD deletion mutants were generated as described by Choi and Schweizer (2005). Briefly, a gentamycin cassette flanked by B38-2A-M algD sequences were generated using primers up fw TTAACGGAAGCCATCAAG, up rev CCCAAACCAAAGATGCTGAT, down fw TCGAC CTGGTGAAACAGACC and down rev ATCAGCAGG CTGAGGACAC. The construct was inserted into pEX18ApGW creating pMHR1. Allelic exchange was indicated by a Gm’ sucrose’ Cb’ phenotype was verified by PCR using the primers, up fw and down rev. Generation of unmarked deletion mutants was achieved by transforming electro-competent P. aeruginosa with pFLP2 and deletion of the Gm’ marker was verified by colony PCR and sequencing.

**DNA microarray sample processing**

Transcriptomic profiles of clinical isolates were obtained using the Affymetrix P. aeruginosa gene chip. Triplicate experiments were performed for each strain. Bacteria were grown at 37°C in 50 ml of LB medium in a baffled 250 ml Erlenmeyer flask shaking at 240 r.p.m. Cell density starting conditions were 0.01 at OD\textsubscript{600} inoculated from an overnight culture grown in LB. Bacteria were harvested in late exponential phase at an OD\textsubscript{600} of 0.5 and immediately mixed with RNAprotect Bacteria Reagent (Qiagen) and stored at −80°C. For mucoid strains preservation of mucoidy was confirmed by plating the cultures and scoring the colonies for mucoidy. RNA was extracted with RNeasy mini kit (Qiagen) and transcribed to cDNA using SuperScript III Reverse Transcriptase (Invitrogen). Subsequent cDNA purification, fragmentation and labeling were performed based on prokaryotic sample and array processing protocol from Affymetrix (Santa Clara, CA). The labelled cDNA were then hybridized on Affymetrix P. aeruginosa gene chips and stained on the GeneChip® Fluidics Station. The probe arrays were scanned with the GeneChip® Scanner 3000.

**Microarray data analysis**

The raw data (Tables S1 and S2) were obtained using the Affymetrix GeneChip® Operating System 1.4 (GCOS). Microarray data analysis was performed using the BioConductor package for the R software environment (http://www.bioconductor.org). Normalization and expression index calculation was done with the RMA function while linear modelling was performed with the LIMMA package and multiple strain-pair contrasts. A P-value < 0.05 and absolute fold change > 2 was applied as cut-off values for each contrast (strain comparison). Transcription factor genes were selected using a P-value threshold < 0.05 only. The fold change was calculated using the difference of average (log2) expression levels of all strain replicate group while the annotations and functional classes were assigned according to the Pseudomonas Genome Database V2 (http://www.pseudomonas.com).

**Gene sequencing**

Sequencing of genes was performed by PCR using High Fidelity Polymerase (Fermentas) and primers muaA fw CTCTGCAGCCTTTTGTTCGAAGAG, muaA rev CTGC CAAGCAAAAGACGGAGGGG, AlgU fw CCTGCAGC CGATGCAATCCATTTTCG, and functional classes were assigned according to the Pseudomonas Genome Database V2 (http://www.pseudomonas.com).

**Acyl-homoserine lactone quantification**

Acyl-homoserinelactone production was quantified by high-pressure liquid chromatography and mass spectrometry. ABT medium containing 0.2% glucose and 1.0% cas-aminoacids was used as growth medium while growth conditions were 37°C in 50 ml of medium in a baffled 250 ml Erlenmeyer flask shaking at 240 r.p.m. Cell density starting conditions were 0.01 at OD\textsubscript{600} inoculated from an overnight culture grown in ABTGC. For mucoid strains preservation of mucoidy was confirmed by plating culture. Bacteria were harvested in late exponential phase at an OD\textsubscript{600} of about 1.0 and after centrifuging culture at 12 000 g for 5 min, 1.0 ml of supernatant transferred to an auto-sampler vial along with 50 μl of 3 YY μM D\textsubscript{3}C-4 internal standard (hydrogen’s in the acyl chain substituted by deuterium) in acetoneitrile (ACN) and frozen at −80°C.
HPLC-MS/MS was performed on an Agilent (Torrance, CA) 1100 HPLC system controlled by MassLynx V4.1. Samples were separated on a Gemini C18-phenyl 3 μm, 2 mm ID, 50 mm column (Phenomenex, Torrance, CA), using a flow rate of 0.300 ml min⁻¹ at 25°C. A linear water-ACN gradient was used, starting at 2% ACN, going to 40% ACN in 30 s, then 100% ACN in 4.5 min, the 1 min with a flow of 0.500 ml min⁻¹, before reverting to the start conditions for 1 min and holding this for 5 min. Both solvents contained 20 mM formic acid. The HPLC was a coupled Quattro Ultima triple mass spectrometer (Waters, Manchester, United Kingdom) with a Z-spray electrospray ionization source using a flow rate of 700 l h⁻¹ nitrogen at 350°C; hexapole 1 was held at 30 V, and the cone was held at 25 V. Nitrogen was used as collision gas, and the mass spectrometer operated in positive multiple-reaction monitoring mode (dwell time 100 ms). Multiple-reaction monitoring mode were: 0–4 min,(i) C-4 AHL retention time (RT) 2.98 min m/z 172→102 @ 15eV (quantifier ion) and 172→71 @ 15eV (qualifier), (ii) open lactone-C-4 AHL RT 2.68 min 190→120 @ 25eV (quantifier) and 190→71 @ 25eV (qualifier) and (iii) D5-C-4 (internal standard) 177→102 @ 15eV (same RT as C-4); 4–7 min, (iv) Oxo-C12 AHL RT 5.08 298→102 @ 25V (quantifier) and 298→197 @ 25eV (qualifier) and (v) open lactone Oxo-C12 AHL RT 4.67 min 298→102 @ 25eV (quantifier) and 298→197 @ 25eV (qualifier). Calibration was done using D5-C-4 as internal standard (isotope dilution) against standards in ACN. Data were processed in Quanlynx 4.1 (Waters) with s/n ratio of 10 (both transitions per compound), giving detection limits in the 10–30 nM range.

Motility assay

Twitching and swimming motilities were assayed on agar plates containing AB mimimal medium (Hansen et al., 2007) supplemented with 7.4 μM thiamine, 0.5% glucose and 0.5% casamino acids; 0.3% and 1.5% agar were used for swimming and twitching respectively. Plates were inoculated with nearly equal amounts of biomass as starting conditions based on OD₆₀₀ measurement of ON culture. The cells were inoculated onto the bottom of the dish for twitching plates and inside the agar for swimming plates. Twitching plates were incubated for 48 h at 37°C while swimming plates were incubated for 24 h at 30°C and subsequently maximum diameter was measured and values background corrected by subtracting values of a motility deficient strains (pilA, fliF). Preservation of mucoidy was tested by streaking on LB plates from motility deficient strains (Schnabel and Schnabel, 1990) were transferred from ON culture diluted to an OD₆₀₀ of 0.5 of which 100 μl was inoculated and spread with a Drigalski spatula on a predried LB plate. Subsequently E-test strips were carefully placed on the LB plate and MIC values were read after 24 h of incubation.

Rhamnolipid assay

The concentration of rhamnolipids in culture supernatants was determined by the orcinol method as previously described (Koch et al., 1991; Pamp and Tolker-Nielsen, 2007), with modifications. Briefly, P. aeruginosa strains were grown at 37°C in LB for 24 h, and mucoid strains were tested for preservation of mucoidy by streaking on LB plates. A 0.5 ml of aliquot of culture supernatant was extracted twice with 2 vols of diethyl ether. The ether fractions were pooled, evaporated to dryness, and subsequently 1 ml of orcinol reagent was added and the sample heated at 80°C for 30 min. Orcinol reagent was prepared immediately prior to use and consisted of 7.5 vols of 60% (vol/vol) sulfuric acid and 1 vol. of 1.6% (wt/vol) orcinol in distilled water. After heating, the samples were allowed to cool at room temperature for 15 min, and absorbance (A421) was measured and compared with rhamnose standards. Values were adjusted for differences in final cell density (OD₆₀₀).

Antibiotic minimum inhibitory concentration

Minimum inhibitory concentrations were estimated by E-test according to the manufacturer’s guidelines (AB Biodisk, Solna, Sweden) with minor modifications. Inoculums were prepared from an ON culture diluted to an OD₆₀₀ of 0.5 which 100 μl was inoculated and spread with a Drigalski spatula on a predried LB plate. Subsequently E-test strips were carefully placed on the LB plate and MIC values were read after 24 h of incubation.

C. elegans virulence assay

Virulence was assessed in C. elegans as described (Thomsen et al., 2006). A total of 20 μl of overnight culture of each strain was spread onto Nematode Growth Medium (NGM) plates and incubated at 37°C over night. For each strain, about 100 L4 hermaphrodites of the pha-1 (e2123ts) mutant (Schnabel and Schnabel, 1990) were transferred from NGM plates seeded with E. coli OP50 to the plates seeded with P. aeruginosa strains and incubated at 25°C. The plates were scored for live and dead worms every 24 h. Three independent trials were performed for each strain. The E. coli and C. elegans strains used in this work were provided by the Caenorhabditis Genetics Center (University of Minnesota, Minneapolis).

Acknowledgements

The authors wish to thank Dr Anders Folkesson, Department of Systems Biology, DTU, for his advice and assistance in relation to some of the genetic experiments. Also thanks to Dr Hanne Jarmer, Department of Systems Biology, DTU, for initial supervision of DNA array data analysis and Dr Sünje Johanna Pamp, Stanford University School of Medicine, for advice on the rhamnolipid assay. The Lundbeck Foundation and the Danish Council for Independent Research (FNU) supported the work by grants to S.M.
References


**Supporting information**

Additional Supporting Information may be found in the online version of this article:

**Table S1.** Gene expression (all) – fold changes.

**Table S2.** Gene expression (all) – *P*-values.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.
Evolution and diversification of *Pseudomonas aeruginosa* in the paranasal sinuses of cystic fibrosis children have implications for chronic lung infection

Susse Kirkelund Hansen\(^1,5\), Martin Holm Rau\(^1,5\), Helle Krogh Johansen\(^2\), Oana Ciofu\(^3\), Lars Jelsbak\(^1\), Lei Yang\(^1\), Anders Folkesson\(^1\), Hanne Østergaard Jarmer\(^1\), Kasper Aanæs\(^4\), Christian von Buchwald\(^4\), Niels Høiby\(^2,3\) and Søren Molin\(^1\)

\(^1\)Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark; \(^2\)Department of Clinical Microbiology, University Hospital, Rigshospitalet, Copenhagen, Denmark; \(^3\)Institute for International Health, Immunology and Microbiology, University of Copenhagen, Copenhagen, Denmark and \(^4\)Department of Otorhinolaryngology, Head and Neck Surgery, Rigshospitalet and Faculty of Health Sciences Copenhagen, Copenhagen, Denmark

The opportunistic pathogen *Pseudomonas aeruginosa* is a frequent colonizer of the airways of patients suffering from cystic fibrosis (CF). Depending on early treatment regimens, the colonization will, with high probability, develop into chronic infections sooner or later, and it is important to establish under which conditions the switch to chronic infection takes place. In association with a recently established sinus surgery treatment program for CF patients at the Copenhagen CF Center, colonization of the paranasal sinuses with *P. aeruginosa* has been investigated, paralleled by sampling of sputum from the same patients. On the basis of genotyping and phenotypic characterization including transcription profiling, the diversity of the *P. aeruginosa* populations in the sinuses and the lower airways was investigated and compared. The observations made from several children show that the paranasal sinuses constitute an important niche for the colonizing bacteria in many patients. The paranasal sinuses often harbor distinct bacterial subpopulations, and in the early colonization phases there seems to be a migration from the sinuses to the lower airways, suggesting that independent adaptation and evolution take place in the sinuses. Importantly, before the onset of chronic lung infection, lineages with mutations conferring a large fitness benefit in CF airways such as *mucA* and *lasR* as well as small colony variants and antibiotic-resistant clones are part of the sinus populations. Thus, the paranasal sinuses potentially constitute a protected niche of adapted clones of *P. aeruginosa*, which can intermittently seed the lungs and pave the way for subsequent chronic lung infections.

*The ISME Journal* (2012) 6, 31–45; doi:10.1038/ismej.2011.83; published online 30 June 2011

**Subject Category:** microbe–microbe and microbe–host interactions

**Keywords:** adaptive evolution; chronic infection; cystic fibrosis; protected environment

**Introduction**

*Pseudomonas aeruginosa* is an opportunistic pathogen capable of causing chronic infections in the airways of cystic fibrosis (CF) patients, where it is the predominant pathogen associated with morbidity and mortality (Koch, 2002). Patients often acquire intermittent colonization of the lungs from early childhood, which eventually proceeds into chronic lung infection where the same bacterial lineage can persist for decades (Høiby et al., 2005; Jelsbak et al., 2007). In this process, *P. aeruginosa* is subjected to strong selection in the host environment leading to substantial genetic change and diversification. Many characteristics of chronically infecting strains are consistently selected in different CF patients, suggesting that adaptation occurs with conserved patterns of evolution (Nguyen and Singh, 2006; Smith et al., 2006; Yang et al., 2011). These characteristics include enhanced antibiotic resistance, changes in nutrient utilization, overproduction of the mucoid exopolysaccharide alginate, reduction in growth rate and loss of O-antigen, motility, type III secretion and quorum sensing (Lam et al., 1980; Hancock et al., 1983; Mahenthiralingam...
et al., 1994; Dacheux et al., 2001; D’Argenio et al., 2007; Yang et al., 2008; Hoffman et al., 2010). In addition, the initial infecting strains of \( \text{P. aeruginosa} \) often diversify into subpopulations with distinct colony morphologies (Wahba and Darrell, 1965; Thomassen et al., 1979). Major selective factors in the CF airways are believed to be the host immune system, nutrient composition and antibiotic treatments, and consequential adaptive changes of the bacterial phenotypes can seriously affect disease outcome and progression (Deretic et al., 1995; Haussler et al., 1999; Parad et al., 1999; Hoffman et al., 2009).

Initial colonizing strains primarily originate from the environment, each event usually with a unique genotype, displaying wild-type characteristics with a non-mucoid and antibiotic-susceptible phenotype (Burns et al., 2001). The early stages of lung disease have therefore been recognized as windows of opportunity to eradicate \( \text{P. aeruginosa} \) (Burns et al., 2001), and clinical trials have shown that early aggressive treatment is beneficial for the patient and delays transition into a chronic infection (Høiby et al., 2003; Taccetti et al., 2005). Following eradication, a new acquisition is often with a different genotype, but some studies have shown that in approximately 25% of the cases re-colonization occurs with the same genotype of \( \text{P. aeruginosa} \) (Munck et al., 2001; Gibson et al., 2003; Doring et al., 2006). Re-colonization could be either from a persistent environmental source or from an undetected reservoir in the patient upper airways such as the paranasal sinuses (Jelsbak et al., 2007). Only few reports have addressed the role of the upper airways, and these studies suggest that the paranasal sinuses may constitute an important niche for bacterial adaptation from which subsequent persistent infections in the lungs of the patients may be established.

**Materials and methods**

**Bacterial strains, plasmids and culture conditions**

Strains and plasmids used in this study are listed in Table 1. \( \text{P. aeruginosa} \) and \( \text{Escherichia coli} \) were routinely grown in Luria-Bertani (LB) medium or LB-agar at 37°C, unless noted. Antibiotics were used at the following concentrations: 30 \( \mu \text{g ml}^{-1} \) gentamycin, 200 \( \mu \text{g ml}^{-1} \) carbenicillin and 100 \( \mu \text{g ml}^{-1} \) ampicillin for \( \text{P. aeruginosa} \); and 15 \( \mu \text{g ml}^{-1} \) gentamycin and 100 \( \mu \text{g ml}^{-1} \) ampicillin for \( \text{E. coli} \).

**CF patients and clinical isolates**

A total of 46 patients from the Copenhagen CF Center, Rigshospitalet, were included in a longitudinal study where first and subsequent lung colonizing isolates were stored from the beginning of 2005 until conclusion in July 2009 (Supplementary Table S1). This group included the majority of young CF patients who acquired their first \( \text{P. aeruginosa} \) within the study period. Patients were monitored on a monthly basis in the outpatient clinic by evaluation of their clinical status, pulmonary function and microbiology of lower airway secretions that were obtained by expectoration, endolaryngeal suction or bronchoalveolar lavage. Identification of \( \text{P. aeruginosa} \) in lower airway samples was carried out at the Department of Clinical Microbiology, Rigshospitalet, as described previously (Høiby and Frederiksen, 2000). The genotype of \( \text{P. aeruginosa} \) was determined by

**Table 1** Non-clinical strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant characteristics</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{P. aeruginosa} ) PAO1</td>
<td>Wild-type strain</td>
<td>Holloway and Morgan, (1986)</td>
</tr>
</tbody>
</table>

**Plasmids**

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevant characteristics</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMHRA</td>
<td>( \text{thIA::gfp(ASV)} ) monitor cassette, Gm'</td>
<td>Yang et al. (2009)</td>
</tr>
<tr>
<td>PMHLAS</td>
<td>( \text{lasI::gfp(ASV)}—\text{Plac::lasR} ) monitor cassette, Gm'</td>
<td>Hentzer et al. (2002)</td>
</tr>
</tbody>
</table>

Gm' and Ap' indicate gentamicin and ampicillin resistance, respectively.
single-nucleotide polymorphism (SNP) typing using the AT-Biochip (Clontag, Jena, Germany) as described by the manufacturer. Three months of oral ciprofloxacin and inhaled colistin were used to treat patients at first acquisition of P. aeruginosa as well as patients with intermittent colonization. Intravenous treatment was started if recurrent isolate during treatment, mucoid phenotype or increased antibodies occurred. Chronic P. aeruginosa lung infection was defined as the persistent presence of P. aeruginosa in the sputum for 6 consecutive months or less when persistence was combined with the presence of two or more precipitating antibodies against P. aeruginosa (Høiby, 1974, 1977).

Sample materials from the paranasal sinuses were obtained by functional endoscopic sinus surgery in the ENT Department at Rigshospitalet, Copenhagen. The location and side (left or right) of each sinus sample was noted. Identification and genotyping of P. aeruginosa from the sinus samples was performed as described for the lower airway samples.

Morphotype diversity
The morphotype diversity of P. aeruginosa in the sinus samples was assessed using Pseudomonas isolation agar (PIA; Difco, Lawrence, KS, USA) plates containing ampicillin (100 μg ml⁻¹). Sample material was spread and plates incubated overnight (ON) at 37°C. Colonies were screened using a Zeiss axioplan microscope equipped with a ×2.5 plan objective. The number of colonies screened from each sinus surgery ranged from several hundred to several thousand depending on the colony-forming units of the samples. Several colonies of each different morphotype were picked, clonally purified on PIA amp plates and frozen at −80°C after growth ON in LB. The isolates were then re-cultivated on PIA amp plates and the different colony morphotypes were classified by comparing the shape of single colonies grown ON simultaneously on the same batch of plates (as the colony morphologies vary with dryness of plate, crowdedness and other factors). Morphotypes were distinguished based on features such as colony size, color (dark or light), translucent or opaque, surface roughness, surface shape (mountain or flat), line pattern on top, sharp edge or fuzzy edge and form/shape of colony (round, oval and other). As only negligible intra-morphotype differences were present in the tested phenotypes, a representative of each colony morphotype was chosen for further investigations. Morphotype diversity of P. aeruginosa in the longitudinal lower airway samples were assessed as described for the sinus samples with a small modification: after spreading the sample on PIA amp plates, 96 colonies were picked (if possible) and frozen in a microtiter plate at −80°C. Isolates were re-cultivated to assess morphotype diversity and a representative of each colony morphotype was chosen for further investigations. All unique morphotypes from one patient was assigned a capital letter, which does not correlate between patients, except morphotype M (mucoid) and SCV (small colony variant). All colony morphotypes isolated from the sinuses and lower airways of patient B11, B13, B34, B22, B28 and B42 are listed in Supplementary Table S2.

Sequencing of mucA and lasR genes. A 687-bp fragment covering the mucA region and an 1300-bp fragment covering the lasR region plus 495 bp upstream were amplified by standard polymerase chain reaction. Primers used for polymerase chain reaction amplification and sequencing of mucA were mucA1fwd (CTCTGCAGCGCTTTGTGCGAGAAG), mucA1rev (CTGCAAGAAAGGAGGAGGAG), and for lasR were lasRfwd3 (CTGGAAAAAGTGCCGTAGTCG) and lasRrev3 (TGCCCTTCCCTATATATCTGC).

Motility assays
All motility assays were performed using (soft) agar plates with ABT minimal medium (Hentzer et al., 2002) containing 0.5% glucose and 0.5% casamino acids. Plates were inoculated from single colonies using a sterile toothpick and incubated at 37°C. Swimming and swarming motility was assayed on 0.3% and 0.6% agar plates (wt vol⁻¹), respectively, incubated for 24 h and twitching motility was assayed on 1.3% agar plates (wt vol⁻¹) incubated for 48 h. Maximum diameter of the motility zone was recorded for minimum three replicates per strain.

Quorum sensing
Clinical strains were examined for a metallic iridescent screen indicative of the presence of lasR mutations as described previously (D’Argenio et al., 2007). Confirmation of a lasR mutation was carried out by polymerase chain reaction and sequencing. Isolates were tested for the presence of C4-HSL signal molecules using pMHRA containing an RhlR-regulated rhlA:gfp(ASV) translational fusion (Yang et al., 2009). 3-Oxo-C12-HSL signal molecules were assayed using pMHLAS containing the lasB:gfp(ASV)—Plac::lasR monitor cassette (Hentzer et al., 2002). A lasI, rhlI double mutant of P. aeruginosa (JP2) containing C4 or C12 monitor plasmid was cross-streaked on LB-agar against a clinical isolate in a T-shaped pattern as described previously (Andersen et al., 2001). The plate was incubated for 24 h before it was examined for green fluorescent protein fluorescence using an Axiosplan Epifluorescence Microscope (Carl Zeiss, Copenhagen, Denmark).

Protease assay
Secreted protease production was assayed using the method described in Brown and Foster (1970). Single colonies of each strain were patched onto LB-agar containing 10% skimmed milk and
incubated at 37 °C for 24 h and 48 h. The diameter of the clearing zones surrounding bacterial growth was measured in triplicate experiments.

Biofilm formation
Biofilm formation was examined in 96-well U-bottom polystyrene plates using crystal violet staining (Pratt and Kolter, 1998). Briefly, ON cultures of \( P. \) aeruginosa were diluted 20 times in ABT minimal medium (Hentzer et al., 2002) containing 0.5% glucose and 0.5% casamino acids. A measure of 150 μl of diluted bacterial culture was incubated in the microtiter plates for 24 h at 37 °C. Staining was performed using a 0.02% crystal violet solution for 20 min and optical density measured at 595 nm. Results are representative of two separate experiments with a minimum of seven replicates in total for each strain.

DNA microarray sample processing
Transcriptomic profiles of clinical isolates were obtained using the Affymetrix \( P. \) aeruginosa gene chip. Cells were grown in beef broth (State Serum Institute, Copenhagen, Denmark) at 37 °C with shaking at 170 r.p.m. Triplicate experiments were performed for each strain. A 50 ml volume of the medium was inoculated with cells from an ON culture to yield a start optical density of approx. 0.1 at 600 nm. A measure of 4 ml cells were harvested at an optical density of approx. 1 at 600 nm. RNA isolation and purification was performed by RNA Protect Bacteria Reagent and RNeasy Mini Kit (Qiagen, Hilden, Germany). RQ1 RNAase-free DNase (Promega, Madison, WI, USA) was added to remove the contaminating DNA. Processing of the \( P. \) aeruginosa GeneChip was performed at the Department of Clinical Biochemistry, Microarray Core Unit, Rigshospitalet. In all, 10 mg of purified RNA was used to synthesize single-stranded cDNA with SuperScript Choice system (Invitrogen, Taastrup, Denmark) with a random primer. DNAse I (Amersham Biosciences, Uppsala, Sweden) was used for fragmentation of the DNA followed by biotin labeling (GeneChip DNA Labeling Reagent; Affymetrix Inc., Santa Clara, CA, USA) The labeled cDNA was then hybridized on Affymetrix \( P. \) aeruginosa gene chips and stained on the GeneChip Fluidics Station. The probe arrays were scanned with the GeneChip Scanner 3000. The Fluidics Station and GeneArray Scanner were operated and managed with the GeneChip Operation Software v.1.4.

Microarray data analysis
Microarray data analysis was performed using bioconductor for the R software (http://www.bioconductor.org). Normalization and expression index calculation was carried out with the rma function. A \( P \)-value < 0.05 and absolute fold change ≥ 2 was applied as cutoff values. The fold changes of expression between two strains were calculated as the ratio of the average expression levels (of the three replicates). The annotations and functional classes were assigned according to the Pseudomonas Genome Database v.2 (http://www.pseudomonas.com).

Antibiotic resistance
Minimum inhibitory concentrations (MICs) were estimated by E-test according to the manufacturer’s guidelines (AB Biodisk, Solna, Sweden) with minor modifications. ON cultures in LB were diluted with LB to an optical density of 0.5 at 600 nm. A measure of 100 μl was spread with a Drigalski spatula on a predried LB plate. Subsequently E-test strips were carefully placed on the LB plate and MIC values were read after 24 h of incubation at 37 °C. A minimum of two replicates were performed for each strain (if results were not consistent more replicates were performed).

Results
In children suffering from CF, it was previously found that early airway infections with \( P. \) aeruginosa can be effectively removed from the lungs by antibiotic treatment, but often the same clones reappeared after months of clearance and it was suggested that subpopulations of the bacteria were hiding in the sinuses (Jelsbak et al., 2007). The primary hypothesis of this study therefore has been that the paranasal sinuses of CF children constitute a protected environment in which \( P. \) aeruginosa may hide to avoid the antibiotics and the immune system of the host, and from this niche subsequent seeding of the bacteria to the lower airways can take place. The hypothesis predicts that children with recurrent intermittent colonization of the lower airways with the same genotype of \( P. \) aeruginosa would also be expected to have their sinuses colonized with the same genotype. This hypothesis was tested in several stages: First, by genotyping isolates from the different airway compartments to establish clonal relationships, second to assess the population structure of \( P. \) aeruginosa in the airways, third to characterize the developed phenotypes in the different airway populations and finally to find evidence for the direction of migration of the evolved lineages between the airway compartments.

Airway colonization and \( P. \) aeruginosa genotyping in intermittently colonized CF patients
In a large clinical study investigating the effect of sinus surgery on disease progression, a number of CF patients from the Copenhagen CF Center underwent sinus surgery. The procedures were performed as a standard computer-assisted functional endoscopic sinus surgery, where the maxillary sinuses and the ethmoidal sinuses as a standard were
targeted and opened widely. Multiple samples were collected for cultivation including nasal secretions, pus, mucosal tissue, polyps and bone. Colonization and genotype profile patterns of \textit{P. aeruginosa} in the airways of 45 children enrolled in the Copenhagen CF Center in the period from the beginning of 2005 until conclusion of the study in July 2009 have been investigated. Results are listed for each patient containing \textit{P. aeruginosa}-positive sputum samples within the study period labeled with sample date and genotype (Supplementary Table S1). Genotypes of \textit{P. aeruginosa} were determined using the AT-Biochip (Clondiag) method based on SNPs (Wiehlmann \textit{et al.}, 2007). This identification method was further supported by performing pulsed field gel electrophoresis of most clinical isolates, and in all cases identical conclusions concerning clonal relationships were obtained (not shown). The genotypic analysis allowed us to investigate if (1) each patient was colonized by a unique genotype of \textit{P. aeruginosa}, (2) if the patient carried more than one genotype and (3) if the same genotype colonized both upper and lower airways in the patient.

In concordance with previous reports (Burns \textit{et al.}, 2001; Munck \textit{et al.}, 2001; Jelsbak \textit{et al.}, 2007), the majority of the children had acquired unique genotypes suggesting initial lung colonization from environmental sources rather than transmission between patients in our CF Center, the latter being restricted owing to cohort isolation (Høiby and Pedersen, 1989). However, eight specific genotypes were isolated from more than one patient (indicating transmission or common sources of colonization), and 9 of 45 children had more than one genotype isolated from their sputum. On the basis of the colonization pattern and genotype profiles, the CF patients could be divided into three groups: (a) Patients with single or multiple events of short colonization periods (<6 months) followed by eradication. Each colonization event was with a unique genotype. This group includes patients with new colonization events within the last 6 months before conclusion of study. (b) Intermittently colonized patients with multiple recurrent colonization events (usually separated by several months) with the same genotype of \textit{P. aeruginosa} and a low number of precipitating antibodies (<2) indicating the absence of chronic lung infection (Høiby, 1974, 1977). (c) Patients with a rapid development of chronic lung infection with increasing numbers of precipitating antibodies (≥2). A total of 31 patients (69%) belonged to group A with successful eradication after aggressive antibiotic treatments, and only three patients (7%) progressed directly to chronic infection (group C). In all, 11 patients (24%) belonged to group B and it is a primary hypothesis here, as suggested by several reports (Munck \textit{et al.}, 2001, 2003; Jelsbak \textit{et al.}, 2007; Mainz \textit{et al.}, 2009) that the bacteria were hiding in the paranasal sinuses. At conclusion of the study, 6 of the 11 group B patients had developed chronic infection after a period of intermittent colonization (Supplementary Table S1).

Cultivation and identification of bacteria were performed on sample materials from the six group B patients (B11, B12, B22, B28, B34 and B42 (had surgery twice)). We additionally included a young CF patient (B13, had surgery twice) who had been chronically infected for approximately 6 years. Thus, a total of nine events of sinus surgery performed on seven patients are referred to in this study (Supplementary Table 2).

The \textit{P. aeruginosa} genotype (determined as SNP genotype from AT-Biochip assays supported by pulsed field gel electrophoresis) was unique for each patient, and the same specific genotype could be found consistently in all \textit{P. aeruginosa}-positive sputum samples in each patient as well as in the patient’s sinus samples (Table 2). Patient B13 is the only exception as three genotypes were present in the sputum samples, of which only two were present in the sinus samples; however, the third genotype was only isolated once from sputum. For six of the seven patients (incl. B13) \textit{P. aeruginosa} was identified in the paranasal sinus samples. In patient B12, only the left side of the nose was cultured and no \textit{P. aeruginosa} was found in the cultured sinuses. In the other eight events, both left and right side was cultured, an average of four sinus cavities were cultured, each with several samples and \textit{P. aeruginosa} was found in all cultured sinus samples.

Consequently, genotype identity exists between the lower airways and the sinuses supporting the hypothesis that the paranasal sinuses constitute a colonized niche in the CF airways. The results further suggest that the sinuses represent a protected environment in which bacteria may survive even after their eradication from the lungs (Low \textit{et al.}, 2001). Despite the genotypic identity of the isolates derived from each patient, it was noted, however, that upon plating of individual clones, a clear diversity was apparent among the obtained bacterial colonies. It was therefore decided to investigate further this apparent diversity of the bacterial populations.

\textit{The paranasal sinuses harbor diverse bacterial subpopulations}

Colony morphology (morphotype) diversity of the sinus isolates was assessed for each patient by microscopy as described in Materials and methods. Each morphotype could be stably maintained after repeated growth, and the SNP genotype was determined for all variants. All available sputum samples from the lower airways were similarly examined for different morphotypes, and sinus and sputum populations were carefully compared to determine the number of identical morphotypes isolated until the point of surgery (Table 2). A strain list of all isolates (including morphotypes) is displayed in
Table 2 Genotype and morphotype diversity of P. aeruginosa isolates from the sinuses and lower airways of group B patients

<table>
<thead>
<tr>
<th>Patient no.</th>
<th>Year of birth</th>
<th>Clinical status of P. aeruginosa at sinus surgery</th>
<th>Period of intermittent/chronic infection before sinus surgery (years)*</th>
<th>P. aeruginosa isolated from the sinuses</th>
<th>SNP genotype found in both lower airways and sinuses</th>
<th>No. of lower airway morphotypes</th>
<th>No. of sinus morphotypes</th>
<th>No. of identical morphotypes in both</th>
<th>Different morphotypes in the right and left side of the sinuses</th>
</tr>
</thead>
<tbody>
<tr>
<td>B11</td>
<td>1993</td>
<td>Intermittent</td>
<td>2.6</td>
<td>Yes</td>
<td>6FA6</td>
<td>6</td>
<td>11</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>B22</td>
<td>1996</td>
<td>Intermittent</td>
<td>1.4</td>
<td>Yes</td>
<td>3C2A</td>
<td>7</td>
<td>6</td>
<td>3</td>
<td>Yes</td>
</tr>
<tr>
<td>B28</td>
<td>1994</td>
<td>Intermittent</td>
<td>1.9</td>
<td>Yes</td>
<td>2C1E</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>B34</td>
<td>1994</td>
<td>Intermittent/chronic infection</td>
<td>2.3</td>
<td>Yes</td>
<td>F679</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>Yes</td>
</tr>
<tr>
<td>B42</td>
<td>1996</td>
<td>Intermittent</td>
<td>2.1</td>
<td>Yes (both times)</td>
<td>F469</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>Yes*</td>
</tr>
<tr>
<td>B12</td>
<td>1998</td>
<td>Intermittent</td>
<td>3.25</td>
<td>No</td>
<td>—</td>
<td>ND</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>B13</td>
<td>1996</td>
<td>Chronic infection</td>
<td>9.25*</td>
<td>Yes (both times)</td>
<td>239A and EC5A</td>
<td>6^</td>
<td>9</td>
<td>5</td>
<td>—</td>
</tr>
</tbody>
</table>

Abbreviation: ND, not determined.

*The duration of intermittent colonization/chronic infection from first colonization event until first sinus surgery or second sinus surgery in the case of B42 and B13, which had first surgery 9 and 8 months before second, respectively.

The SNP genotype is given as hexadecimal code (Wiehlmann et al., 2007).

The total number of colony morphotypes was assessed for the lower airway samples collected until first sinus surgery or until second surgery in the case of B42 and B13.

Colony morphotypes found in the samples from sinus surgery includes both first and second surgery for patient B42 and B13.

Patient B34 was on the transition point from intermittent colonization to chronic infection (by definition) due to six consecutive samples just 1 month before surgery.

For patient B42, P. aeruginosa was only recovered from the left side of the sinuses at first surgery, but similar types were recovered from both sides at second surgery, possibly a result of cross-contamination by surgical instruments.

Note, only approx. one isolate per year from 2002 was available for B13 as this patient was included retrospectively.

Supplementary Table S2, and the morphotype diversity found in two of the patients (B22 and B34) is shown in Figure 1.

Interestingly, substantial diversification of P. aeruginosa was observed in the sinus cavities. In two patients (B11 and B13), 10 or more morphotypes were found and in average 6.7 ± 2.9 (mean ± s.d.) morphotypes per patient were isolated from the sinuses. For the lower airway samples, the observed diversity was in average 5.3 ± 1.2 (mean ± s.d.) morphotypes per patient. These numbers most likely represent minimum estimates owing to the partial sample material and limitations in screening method. The overlap of morphotypes between the sinuses and the lower airways was high; on an average 3.2 ± 1.2 (mean ± s.d.) or 59% of the morphotypes found in the lower airways were also cultivated from the paranasal sinuses. Patients B22 and B34 both had three morphotypes shared between the sinuses and the lower airways (Figure 1). Many of the lower airway morphotypes that were not found in the sinuses belonged to the very early isolates from the sputum. In addition, the population in the left and right sinus cavities on each side of the nose in most patients had evolved into distinct morphotype populations (Table 2). In one patient (B28), a few morphotypes were found in both sides, but it is not known if this was a result of cross-contamination during surgery, as in this case the same instruments were used in the right and left sinuses. Patient B13 who was chronically infected at the time of the sinus surgery carried different genotypes in each side of the sinuses, a result that was confirmed at the second surgery 8 months later.

The genotype found in the left sinuses was the only type recovered from the sputum samples until 2006, after which only the genotype from the right sinuses was identified in sputum (Supplementary Table S2). The apparent diversity of the P. aeruginosa populations in the sinus samples from the investigated group of patients strongly suggest that evolutionary changes occur, and that some of the evolved morphotypes may be similar to those documented from lung samples. These observations made us hypothesize that the bacteria colonizing the sinuses evolve to become distinct lineages in the CF upper airways, and that these lineages migrate to the lungs, where they may establish and become persistent. In the following sections, we will present data supporting this hypothesis.

Relationships between morphotype, phenotype and genotype

The finding that the sinuses and the lower airways are often colonized by identical P. aeruginosa morphotypes suggests that the two compartments harbor highly similar or identical lineages. If the isolates indeed are identical, it would be evidence of transfer between the sinuses and the lower airways. Therefore, phenotypic and genetic profiles of all isolates from patients B22 and B34 were compared. Specifically, phenotypic known to change during chronic infections in CF patients (Mahenthiralingam et al., 1994; Demko et al., 1995; Smith et al., 2006) such as motility, quorum sensing, secreted virulence factors and biofilm formation were analyzed, and the lasR and mucA genes were sequenced.
for certain isolates. In general, it was found that each morphotype had a distinct phenotypic profile displaying variable reduction or loss of the analyzed traits (Table 3). In patient B22, all isolates from the left sinuses had the same insertion mutation in the *lasR* regulator gene, and all were quorum sensing negative in contrast to the isolates from the right sinuses. In patient B34, the isolates from the right sinuses were mucoid owing to mutations in *mucA* causing alginate overproduction, whereas the left side isolates were non-mucoid and *mucA*\(^+\*\).

In patient B22, morphotypes E, F and G were found both in the sinuses and lower airways. The two lower airway isolates, B22-7_F and B22-8_G, were collected within the last 3 months before sinus surgery, and they displayed phenotypic profiles identical to the sinus isolates B22-sin_F and B22-sin_G, respectively (motility and biofilm formation values were within the standard deviation for each morphotype) (Table 3). The IS insertion causing a *lasR* mutation found in the entire left sinus population was also identified in the lower airway isolate B22-7_F, further suggesting a common origin. The E morphotypes had identical profiles as well, except for an 1.5-fold difference in biofilm formation capacity.

In patient B34, the shared morphotypes were B, M (mucoid) and SCV. All M isolates showed very similar phenotypic profiles with minor differences, which could in some part be due to instability of this phenotype, as it was observed to revert during experiments (DeVries and Ohman, 1994). However, the last mucoid morphotype isolated from the lower airways before surgery (B34-5_M) was not different...
Table 3 Phenotypic profile of isolates from patients B22 and B34

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Sample date</th>
<th>Twitching motility*</th>
<th>Swimming motility*</th>
<th>Swarming motility*</th>
<th>QS (C4)*</th>
<th>QS (C12)*</th>
<th>Protease†</th>
<th>Biofilm formation*</th>
<th>lasR‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>B22-1_A</td>
<td>13.03.06</td>
<td>3.0 (± 0.10)</td>
<td>2.1 (± 0.23)</td>
<td>0.4 (± 0.06)</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>1.70 (± 0.19)</td>
<td>WT</td>
</tr>
<tr>
<td>B22-2_C</td>
<td>20.03.06</td>
<td>3.0 (± 0.25)</td>
<td>1.9 (± 0.26)</td>
<td>0.5 (± 0.12)</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>2.05 (± 0.26)</td>
<td>WT</td>
</tr>
<tr>
<td>B22-3_J</td>
<td>22.08.06</td>
<td>2.6 (± 0.10)</td>
<td>3.4 (± 0.31)</td>
<td>0.6 (± 0.15)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>0.47 (± 0.08)</td>
<td>1.4 kb IS insertion</td>
</tr>
<tr>
<td>B22-4_J</td>
<td>22.08.06</td>
<td>1.6 (± 0.25)</td>
<td>3.2 (± 0.21)</td>
<td>1.0 (± 0.06)</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>2.18 (± 0.25)</td>
<td>WT</td>
</tr>
<tr>
<td>B22-5_C</td>
<td>10.01.07</td>
<td>2.3 (± 0.10)</td>
<td>3.2 (± 0.17)</td>
<td>0.6 (± 0.10)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>0.35 (± 0.10)</td>
<td>1.4 kb IS insertion</td>
</tr>
<tr>
<td>B22-6_E</td>
<td>24.04.07</td>
<td>2.2 (± 0.12)</td>
<td>3.4 (± 0.10)</td>
<td>0.5 (± 0.06)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>0.35 (± 0.12)</td>
<td>1.4 kb IS insertion</td>
</tr>
<tr>
<td>B22-7_F</td>
<td>26.06.07</td>
<td>0.5 (± 0.10)</td>
<td>2.0 (± 0.15)</td>
<td>5.5 (± 0.61)</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>1.33 (± 0.21)</td>
<td>WT</td>
</tr>
<tr>
<td>B22-8_G</td>
<td>01.08.07</td>
<td>0.5 (± 0.15)</td>
<td>3.5 (± 0.31)</td>
<td>1.1 (± 0.12)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>0.32 (± 0.12)</td>
<td>1.4 kb IS insertion</td>
</tr>
<tr>
<td>B22-2_A</td>
<td>29.10.07</td>
<td>0.7 (± 0.16)</td>
<td>2.4 (± 0.19)</td>
<td>5.7 (± 1.10)</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>0.84 (± 0.13)</td>
<td>WT</td>
</tr>
<tr>
<td>B22-sin_F</td>
<td>08.11.07</td>
<td>0.5 (± 0.21)</td>
<td>3.7 (± 0.15)</td>
<td>1.1 (± 0.12)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>0.28 (± 0.13)</td>
<td>1.4 kb IS insertion</td>
</tr>
<tr>
<td>B22-sin_I</td>
<td>08.11.07</td>
<td>2.5 (± 0.15)</td>
<td>3.1 (± 0.56)</td>
<td>0.8 (± 0.15)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>0.73 (± 0.14)</td>
<td>1.4 kb IS insertion</td>
</tr>
<tr>
<td>B22-sin_J</td>
<td>08.11.07</td>
<td>1.7 (± 0.14)</td>
<td>3.4 (± 0.35)</td>
<td>1.9 (± 0.12)</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>0.55 (± 0.09)</td>
<td>1.4 kb IS insertion</td>
</tr>
<tr>
<td>B22-sin_K</td>
<td>08.11.07</td>
<td>0.5 (± 0.20)</td>
<td>2.9 (± 0.10)</td>
<td>0.7 (± 0.06)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0.32 (± 0.09)</td>
<td>1.4 kb IS insertion</td>
</tr>
<tr>
<td>B22-sin_E</td>
<td>08.11.07</td>
<td>0.6 (± 0.25)</td>
<td>2.2 (± 0.25)</td>
<td>4.9 (± 0.46)</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>0.89 (± 0.15)</td>
<td>WT</td>
</tr>
<tr>
<td>B22-sin_G</td>
<td>08.11.07</td>
<td>0.6 (± 0.25)</td>
<td>2.6 (± 0.08)</td>
<td>5.5 (± 2.12)</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>0.80 (± 0.13)</td>
<td>WT</td>
</tr>
<tr>
<td>B34-1_A</td>
<td>08.05.06</td>
<td>3.8 (± 0.31)</td>
<td>4.9 (± 0.61)</td>
<td>1.9 (± 0.32)</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>0.58 (± 0.09)</td>
<td>WT</td>
</tr>
<tr>
<td>B34-2_M</td>
<td>29.08.06</td>
<td>0.3 (± 0.06)</td>
<td>1.1 (± 0.35)</td>
<td>0.3 (± 0.10)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.41 (± 0.27)</td>
<td>WT</td>
</tr>
<tr>
<td>B34-3_B</td>
<td>18.12.06</td>
<td>0.5 (± 0.15)</td>
<td>2.0 (± 0.15)</td>
<td>0.9 (± 0.06)</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>0.52 (± 0.10)</td>
<td>WT</td>
</tr>
<tr>
<td>B34-4_B</td>
<td>11.06.07</td>
<td>0.5 (± 0.25)</td>
<td>2.0 (± 0.06)</td>
<td>0.8 (± 0.31)</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>0.53 (± 0.13)</td>
<td>WT</td>
</tr>
<tr>
<td>B34-5_M</td>
<td>05.03.08</td>
<td>0.5 (± 0.15)</td>
<td>0.7 (± 0.20)</td>
<td>0.4 (± 0.20)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.10 (± 0.08)</td>
<td>T261C, 262C insert</td>
</tr>
<tr>
<td>B34-6_SCV</td>
<td>12.08.08</td>
<td>0.5 (± 0.10)</td>
<td>—</td>
<td>0.4 (± 0.06)</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>2.99 (± 0.17)</td>
<td>WT</td>
</tr>
<tr>
<td>B34-7_M</td>
<td>05.01.09</td>
<td>0.5 (± 0.10)</td>
<td>0.8 (± 0.00)</td>
<td>0.7 (± 0.38)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.26 (± 0.08)</td>
<td>T261C, 262C insert</td>
</tr>
<tr>
<td>B34-8_C</td>
<td>07.05.09</td>
<td>0.5 (± 0.20)</td>
<td>—</td>
<td>0.2 (± 0.06)</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>0.26 (± 0.10)</td>
<td>ND</td>
</tr>
<tr>
<td>B34-sin_B</td>
<td>18.09.08</td>
<td>0.5 (± 0.10)</td>
<td>2.1 (± 0.15)</td>
<td>0.6 (± 0.06)</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>0.64 (± 0.14)</td>
<td>WT</td>
</tr>
<tr>
<td>B34-sin_SCV</td>
<td>18.09.08</td>
<td>0.4 (± 0.10)</td>
<td>—</td>
<td>0.4 (± 0.06)</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>2.92 (± 0.18)</td>
<td>WT</td>
</tr>
<tr>
<td>B34-sin_M</td>
<td>18.09.08</td>
<td>0.6 (± 0.10)</td>
<td>0.8 (± 0.10)</td>
<td>0.4 (± 0.20)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.09 (± 0.07)</td>
<td>T261C, 262C insert</td>
</tr>
</tbody>
</table>

Abbreviations: ND, not determined; WT, wild type.
*Motility zone diameters are presented as mean ± s.d. (standard deviation of three replicates). –, motility diameter is zero.
†Presence (+) or absence (–) of C4-HSL or 3-oxo-C12-HSL signal molecules.
‡Extracellular protease level. Increasing number of +’s denotes increasing protease level, while – denotes no detectable protease level.
§Small colony variants did produce detectable levels of proteases after prolonged incubation.
*Biofilm formation (OD595) presented as mean ± s.d. of minimum seven replicates.
The lasR sequences in B22 isolates compared with PAO1 (WT). The IS element was inserted 4 bp upstream of the coding region in lasR.
The mucA sequence in B34 isolates compared with PAO1 (WT). All isolates from B34 had six additional nucleotides (GGGGGAC) inserted in position 360, which did not result in mucoid conversion.

from the mucoid sinus isolate B34-sin_M with respect to the tested phenotypes (within the standard deviation), and only six genes were differentially expressed at the transcriptomic level, \( P<0.05 \) and \( >2\)-fold change (Supplementary Table S3). The sinus and lower airway mucoid isolates all had the same two mutations in the mucA gene (T261C, 262 C insert), except the first mucoid isolate, B34-2_M, which had a one base pair deletion in mucA (585 ΔA), suggesting that this lineage became outcompeted and possibly went extinct. Both the B and SCV morphotype isolates from the
Sinuses and lower airways (B34-4_B vs B34-sin_B and B34-6_SCV vs B34-sin_SCV) displayed identical phenotypic profiles (within the standard deviation). Thus, the morphotypes found in both the sinuses and lower airways of patients B22 and B34 were highly similar or identical at a phenotypic and genetic level documenting transfer between the two compartments. Intra-morphotype diversity seemed negligible as phenotypic profiling of several isolates of the same morphotypes from B22 showed that they were in general very similar, although small differences could occur (data not shown). The strong correlation between morphotypes and phenotypes shows that colony morphologies can be used successfully to screen population diversity, a method that has also been applied in several other studies (Rainey and Travisano, 1998; Boles et al., 2005; Hansen et al., 2007). The data also suggest that sinus and lower airway isolates from other patients having identical colony morphologies may reflect identical or at least closely related variants, suggesting transfer between the two compartments for these patients as well. Most importantly, however, the phenotypic and genotypic characterization of several isolates indicate that diversification of the sinus bacterial population leads to the development of lineages with the potential of establishing chronic infections also in the lower airways of the patients.

**Sinus isolates with increased antibiotic resistance from patient B34**

The presence of antibiotic-resistant strains in the sinuses could implicate the sinuses as reservoirs for colonization of the lower airways with pre-adapted strains. Therefore, the level and mechanism of antibiotic susceptibility was investigated for isolates from patient B34 by determining MICs and global gene expression analysis. In these isolates, susceptibility to ciprofloxacin decreased during the infection (Figure 2). Isolate B34-2_M, the subsequent mucoid (M) isolates and the SCV isolates were moderately less susceptible than the initial B34-1_A isolate, while B34-sin_MB type isolates were found to be resistant. Notably, the highest levels of resistance for any isolate type were always associated with the sinus isolate of the respective type. Isolates chosen for gene expression analysis included the first isolate B34-1_A, the last B morphotype before surgery (B34-4_B), the last mucoid isolate before surgery (B34-5_M) and the mucoid sinus isolate B34-sin_M (see Figure 1b for morphotypes and Supplementary Table S3 for gene expression changes). All isolates showed increased expression of the mexCD-oprJ operon, 1.9–2.6-fold for B34-4_B, 1.4–1.9-fold for B34-5_M and 5–10-fold for B34-sin_M relative to the first isolate B34-1_A (Supplementary Table S3). Overproduction of the MexCD-OprJ efflux pump is known to reduce susceptibility to antibiotics such as ciprofloxacin (Poole, 2004). Both MIC and gene expression data indicate that at least two or more separate events of increased ciprofloxacin resistance occurred in the B34 lineage. It is not certain in which part of the airways the increase in resistance evolved, yet interestingly the mucoid lower airway isolate B34-7_M shows no further increase in ciprofloxacin resistance in contrast to what was found for B34-sin_M isolated at an earlier time point. It is possible that the increased level of mexCD-oprJ expression observed in B34-sin_M evolved in the sinuses, and therefore that the increased resistance was selected for in the sinuses. It should also be noted from the transcriptomic data presented in Supplementary Table S3 that genes known to display reduced expression in isolates of *P. aeruginosa* from chronically infected CF patients (Yang et al., 2011), such as genes connected to motility, type III secretion and other virulence factors, also showed reduced expression in the B34 sinus isolates, supporting the hypothesis that the sinus populations evolve towards phenotypes associated with chronic infections.

**Parallel adaptive mutations in sinuses and lower airways**

Screening of sinus isolates from all patients revealed the occurrence of mutations and phenotypes, which are frequently observed in CF lung isolates. An intriguing finding was mutations in *mucA* and *lasR* and the isolation of SCVs, as these genetic changes are known to promote disease progression in relation to chronic lung infection (Deretic et al., 1995; Haussler et al., 1999; Parad et al., 1999; Hoffman et al., 2009). To determine the frequency of these mutants in the sinuses, we screened the phenotypes of all sinus morphotypes isolated in the additional patients (Table 4). Colony morphology
studies documented the presence of SCVs in the sinus populations, but CF lung SCV isolates have some additional characteristics besides the small colony form (Haussler et al., 1999; Starkey et al., 2009). Sinus SCV isolates were therefore further analyzed for these specific phenotypes: clumping in liquid culture, decreased motility, enhanced biofilm formation and antibiotic resistance (Figure 3). Susceptibility towards three antibiotics routinely given to the patients, ciprofloxacin, tobramycin and colistin, were tested. Two of the SCVs were less susceptible to tobramycin when compared with the initial isolate, whereas one strain in addition to (B28-sin_SCV) displayed decreased susceptibility to ciprofloxacin. No SCVs displayed increased susceptibility to colistin. As the potential SCVs also displayed clumping, decreased motility and enhanced biofilm formation, it could be concluded that four out of six patients possessed SCVs in the sinus population. Morphotypes that did not produce

Table 4 Sinus isolates from several patients display mutations and phenotypes associated with poor lung disease prognosis

<table>
<thead>
<tr>
<th>SCV</th>
<th>Loss of Las QS signals</th>
<th>Mucoid (mucA−)</th>
<th>Reduced or loss of motility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Abbreviations: SCV, small colony variants.

A morphotype was classified as a SCV when it displayed characteristics as previously described for SCVs: small colony morphology, clumping in liquid culture and greatly enhanced biofilm formation (Haussler et al., 1999, 2004).

Only one of the two patients had lost production of 3-oxo-C12-HSL molecules owing to mutation in lasR.

Minimum one sinus morphotype had lost or was significantly reduced in swimming and twitching motility.

No. of patients out of the six sinus colonized patients.

| Figure 3 | SCVs from the paranasal sinuses display similar characteristics as SCVs from chronic lung infection. (a) Colony morphology of first isolate and SCV from patients B34, B28, B42 and B11. SCVs were routinely grown on LB and not PIA amp plates owing to a higher instability on the latter. On PIA amp plates, the morphology of SCVs resembled isolate B42-sin_SCV. (b) Motility was generally reduced or lost in SCVs when compared with the first isolate. Swimming motility was also reduced in all SCVs; however, some activity still remained (except for B34-sin_SCV). Motility zone diameters are presented as mean ± s.d. for at least three replicates. White bars, twitching motility; gray bars, swimming motility; black bars, swarming motility. (c) Biofilm formation abilities in microtiter plates were significantly increased for all SCVs. Data are presented as mean ± s.d. for seven replicates. (d) The antibiotic resistance profiles of the SCVs were strain specific. Resistances to ciprofloxacin (black bars) and tobramycin (white bars) were determined by E-test. At least two replicate experiments were performed for each strain. |
the Las-dependent signaling molecules were only observed in two patients (B22 and B34). In addition to patient B34, a mucoid variant was also present in the sinuses of patient B11, whereas loss or severe reduction in twitching and swimming motility was observed for some sinus isolates from all six patients (Table 3 and Supplementary Table S2). The finding of multiple phenotypes associated with chronic CF lung isolates among the sinus populations suggests that these phenotypes are adaptive both in the lower airways and in the sinus environment.

Transfer of P. aeruginosa between sinuses and lower airways
The results presented above strongly suggest that P. aeruginosa migrates between sinuses and lower airways, although without addressing the directionality. The fact that most patients have distinct morphotype populations in each side of the sinuses indicate little or no mixing of the two sinus populations. If the sinuses were occasionally seeded and invaded with different morphotypes that had evolved in undetectable populations residing in the lower airways (or oropharynx), one would expect to find mixed and more similar populations in both sides of the sinuses. Sinus surgery also seems to affect the lower airway population as one patient (B22) had no P. aeruginosa-positive sputum samples for almost 1 year and 6 months after sinus surgery, and the following colonization occurred with a new (different) genotype, suggesting a complete eradication of the previous genotype. Patient B42 also did not have recurrent colonizations after the second surgery until conclusion of the study in July 2009, and small reductions in lower airway colonization frequencies were seen for a couple of the other patients (Supplementary Table S1). These findings suggest that the sinus populations are intermittently seeding the lower airways, and that removal of these may offer a temporary relief of sinus symptoms and possibly lung colonization events (HK Johansen, personal communication; Jones et al., 1993; Nishioka et al., 1995).

On the basis of colony morphology and other phenotypic characteristics, we have constructed a model for a tentative development of the sinus populations of P. aeruginosa in patient B22 (Figure 4). According to the phenotypic data, it seems that two different phenotypic lineages were found among the longitudinal lower airway samples, each lineage correlating with a specific subpopulation from the right or left side of the sinuses. The most likely scenario for the evolutionary trajectory of the B22 isolates (Figure 4) shows that in one of the apparent lineages several phenotypes were lost or reduced in a step-wise manner over time, whereas the other lineage is almost unchanged for the tested phenotypes. Probably the two lineages branched even before the first isolate as...
none of the early evolutionary events are shared. Indeed, the point of divergence of the lineages could be the time of segregation into left and right sinus cavities and hence the divergent phenotypes between each side could reflect evolutionary events actually occurring in the sinuses. The lower airway isolates could consequently represent ‘snapshots’ of the sinus populations at a given time in accordance with the possibility that migration of *P. aeruginosa* in the early colonization stages mainly occurs in a downward direction from the sinuses towards the lungs.

**Discussion**

The initial hypothesis of this study was that the paranasal sinuses of CF patients constitute an alternative colonization site with reduced chances of antibiotic- and immune response-mediated clearance. This was supported by the findings that most CF children carry *P. aeruginosa* populations in their sinuses, and in all cases studied here the genotypes of these were the same as those colonizing the lungs. Surprisingly, it turned out that the sinus bacterial populations diversify into mixed populations of clonally related variants that migrate to the lungs and establish colonization with pre-adapted lineages, which eventually may result in chronic infection. It is practically and ethically impossible to prove a definite causal relationship between sinus colonization and chronic lung infection, and that the infection originates from the patients’ sinuses. However, our study provides arguments, which support the proposition that the sinuses constitute a focus site for early *P. aeruginosa* adaptive evolution directing the colonization towards chronic lung infection.

Some intermittent colonization in CF patients end after aggressive antibiotic therapy as *P. aeruginosa* cannot be cultured from the sputum upon subsequent microbiological examination. The interpretation of this is that the bacteria are eradicated from the combined airways of the treated patients, and any later colonization events thus derive from new environmental strains. Depending on the success of the treatment regimen, this situation can continue for many years, suggesting that the *P. aeruginosa* population size and time of residence in the airways do not allow adaptive mutations important for the establishment of chronic infection to evolve. Other intermittent colonization events, however, display a different pattern with multiple recurrent colonization events (usually separated by several months) with the same genotype. In these patients, sinus infections with *P. aeruginosa* may provide a source for the occasional detection and low number of *P. aeruginosa* in the sputum samples. Our results show that the second-site colonization events in the sinuses, often being more persistent than those in the lungs, may provide extended opportunities for evolution of the bacteria towards phenotypes, with greatly increased potential of creating chronic lung infections as well. Although it has long been known that the microbiotas of the upper and lower airways are similar in chronically infected CF patients, only a small number of studies have investigated the possibility of cross-infections between the paranasal sinuses and the lungs (Dosanjh et al., 2000; Muhlebach et al., 2006; Mainz et al., 2009). In a large cohort study, almost half of the patients with a history of chronic *P. aeruginosa* lung infection also had an infection with the same genotype in the upper airways (Mainz et al., 2009). Identical genotypes in the sinus and bronchoalveolar lavage (lung) samples have also been recovered from one study of CF children with chronic sinusitis, and the frequency of identical genotypes increased from 9% of children up to 30% of children more than 8 years old (Muhlebach et al., 2006).

In our study, the sinuses were found to display substantial diversity in the *P. aeruginosa* population, and in most cases the population had segregated into two distinct populations on each side of the nose. Although we cannot totally exclude that specific subpopulations of bacteria from the lungs are transmitted to individual sides of the sinuses from time to time, it seems unlikely that a specific population in the lower airways would always be transferred to a specific sinus side, in particular during the period of intermittent lung colonization, during which antibiotic treatment effectively removes the lung populations of *P. aeruginosa*. Therefore, these results suggest that the direction of migration is mainly downwards at the early stages of infection. If it is assumed that recurrent intermittent lung colonization with the same genotype is associated with persistent sinus colonization, ours and other studies show that around 25% of the CF children may begin with a sinus colonization (Munck et al., 2001; Gibson et al., 2003; Doring et al., 2006), and more than half of all patients will eventually be affected (Mainz et al., 2009). The study further shows that phenotypic and genotypic changes observed in sinus isolates of *P. aeruginosa* are similar or identical to those often reported for isolates from chronic lung infections in CF patients. Several of the mutations and phenotypes that we have observed among sinus isolates, evolved in parallel in the different patients, including those conferring loss of motility and quorum sensing signals (virulence), alginate overproduction, antibiotic resistance and increased biofilm formation (SCVs) (Cabral et al., 1987; Mahenthiralingam et al., 1994; Haussler et al., 1999; Smith et al., 2006; D’Argenio et al., 2007; Starkey et al., 2009; Hoffman et al., 2010). Most of them are generally thought to be beneficial also in the lung environment through the development of resistance to host defenses and antibiotics or to optimized utilization of available nutrients. The effective removal of lung-associated *P. aeruginosa* during intermittent colonization of young CF patients and the consequential directional dominance of migration from sinuses to lungs.
strongly support our proposition that the observed diversity of the sinus-associated bacterial populations is predominantly caused by localized adaptive evolution in the sinuses.

The suggested role of the paranasal sinuses as persistent reservoirs for evolving bacterial populations may be associated with differences in the sinus and lung environments. Although the physiological properties and environment of the paranasal sinuses and conductive bronchi (where the lung infection is mainly residing) are likely to be comparable, as they have similar mucous lining and the same defect in CF transmembrane regulator, some differences in the immune response seem to exist. Lung infection with *P. aeruginosa* usually elicits stimulation of the innate immune system (inflammation) and an increase in precipitating antibodies against *P. aeruginosa* (Heiby, 1977), but the CF children from this study had colonization of the sinuses for long periods without an elevated systemic (immunoglobulin G) antibody response. Instead, it was recently found that immunoglobulin antibodies were prevalent and inflammation low, as also indicated by low numbers of neutrophils associated with the bacteria in the sinus samples (HK Johansen, personal communication).

Another feature of the sinuses is the detention of thickened mucus due to mucosal edema and closing of the sinus ostia (connection to the nose). It is very likely that the nutritional conditions in a partially or fully concealed sinus cavity are different from the lung environment where the sputum is produced and replaced continuously. Airflow into the sinus cavities can be reduced and lower oxygen tension in the sinuses and anoxic conditions on the sinus mucosa are seen in CF patients. These findings can vary from one side to the other of the noses (Carenfelt and Lundberg, 1977; Aanaes et al., 2011). Hence, bacteria may be locally faced with nutrient limitations, starvation, or nutrient limitation in the paranasal sinuses.

Nutrient limitation including low oxygen tension may be part of the conditions that facilitated the generation of mucoid variants (Terry et al., 1991, 1992) in two of six sinus-colonized patients, possibly in combination with other advantages such as biofilm formation and protection from phagocytosis facilitated by the alginate exopolysaccharide (Cabral et al., 1987) and evasion of host defense via downregulation of motility and virulence factors (Rau et al., 2010). Nutrient availability may also explain, in combination with antibiotic treatment, the selection of the lasR mutant population found in the left side of the sinuses of patient B22. Recent findings have shown that lasR mutants exhibit a dramatic metabolic shift with decreased oxygen and increased nitrate utilization and in addition they display increased resistance to antibiotics such as ciprofloxacin and tobramycin (Hoffman et al., 2010). Therefore, the lasR mutation is also likely to confer increased fitness in the sinuses.

Partially obstructed sinus cavities lead to reduced access of administered antibiotics, further resulting in lower antibiotic load compared with the lower airways. Sublethal antibiotic concentrations could provide opportunities both for survival as well as sufficient time for evolving increased antibiotic resistance. Altogether, the environment of the sinuses seems to differ from that of the lower airways in several important aspects: The immune response is apparently much weaker in the sinuses, the antibiotic bioavailability is low, antibiotics may be less effective owing to the physiological state of the bacteria in the sinuses and the environmental conditions in the sinuses may stimulate the occurrence of antibiotic-resistant mutants. In essence, what is important for the clearance of intermittently colonizing *P. aeruginosa* in the CF lungs is only partially functional in the sinuses, providing opportunities for the bacteria to adapt through evolution of resistance mechanisms with severe impacts on subsequent treatment possibilities in relation to bacteria migrating from the sinuses to the lungs.

Our study suggests that the paranasal sinuses can be an evolutionary 'nest' in early colonizations, where the bacteria are diversifying, evolving antibiotic resistance and other phenotypes associated with adaptation to the CF airways in general. From there the population is intermittently colonizing the lungs and may ultimately cause a chronic lung infection. On the basis of our findings, we suggest that the paranasal sinuses could have an important role for some CF patients in the development of chronic lung infection, and that chronicity in such cases is in fact established before it is usually diagnosed in the clinic. A precise localization and diagnosis of the CF airway colonizations and infections are therefore crucial to target treatments. Early diagnosis and successful treatment of *P. aeruginosa* colonizations in the paranasal sinuses could be an important therapeutic approach to prevent or delay transition to chronic lung infection and ultimately prolong the life of the patients.

Acknowledgements

We gratefully acknowledge Matthew Parsek and Liang Yang for experimental assistance and we also thank the nurses and doctors at the Copenhagen CF Center for their help in recruiting and preparing the patients for surgery. This work was supported by grants from the Danish Research Agency and the Lundbeck Foundation to SM.

References

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)
Deletion and acquisition of genomic content during early stage adaptation of *Pseudomonas aeruginosa* to a human host environment

Martin Holm Rau¹,²,*, Rasmus Marvig¹,*, Garth Ehrlich, Lars Jelsbak¹, Søren Molin¹,²

¹ Department of Systems Biology, Technical University of Denmark, 2800 Lyngby, Denmark.
² Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark

Corresponding authors: Søren Molin and Lars Jelsbak, Department of Systems Biology, Technical University of Denmark, Building 301, 2800 Lyngby, Denmark. E-mail: sm@bio.dtu.dk/lj@bio.dtu.dk

* These authors contributed equally to the work.
Abstract

Adaptation of bacterial pathogens to a permanently host-associated lifestyle by means of deletion or acquisition of genetic material is usually examined through comparison of present-day isolates to a distant theoretical ancestor. This limits the resolution of the adaptation process. We conducted a retrospective study of the dissemination of the P. aeruginosa DK2 clone type among patients suffering from cystic fibrosis, sequencing the genomes of 45 isolates collected from 16 individuals over 35 years. Analysis of the genomes provides a high-resolution examination of the dynamics and mechanisms of the change in genetic content during the early stage of host adaptation by this P. aeruginosa strain as it adapts to the cystic fibrosis (CF) lung of several patients. Considerable genome reduction is detected predominantly through the deletion of large genomic regions, and up to 8% of the genome is deleted in one isolate. Compared to in vitro estimates the resulting average deletion rates are 12 to 36 fold higher. Deletions occur through both illegitimate and homologous recombination, but they are not IS element mediated as previously reported for early stage host adaptation. Uptake of novel DNA sequences during infection is limited as only one prophage region was putatively inserted in one isolate, confirming that the main mechanism of early host adaptation is loss of function instead of gain of function. Finally, we also describe the complete genome of this highly adapted pathogenic strain of P. aeruginosa to strengthen the genetic basis, which serves to help our understanding of microbial evolution in a natural environment.

Introduction

Entry of a bacterium into a new environment provides an opportunity for expanding the possible niches it can inhabit. The inherent fitness of the bacterium towards the new environment as well as improvement of fitness by genetic adaptation will decide if entry is also followed by successful persistence. Modification of the genetic makeup of the organism may be the result of multiple mechanisms. For example, already present DNA can be changed by single nucleotide polymorphisms, small insertions and deletions and genome rearrangements. Also, larger genetic elements can be entirely lost or novel DNA may be gained by the processes of transformation, conjugation and transduction. Depending on the environment certain mechanisms might be more prevalent than others. Acquisition of novel DNA can e.g. be important for adaptation towards environments
containing toxic compounds as is antibiotic resistance adaptation and acquisition of virulence enhancing elements for host infection (Hacker and Kaper, 2000; Aguilar-Barajas et al., 2010; Campos-García, 2010). On the other hand adaptation towards permanent association with a host may be characterized by reductive evolution through extensive gene loss affecting functions that are no longer essential.

Studies of genome reduction are normally carried out either on present isolates compared to a theoretical ancestor or using in vitro evolutionary experiments. The former method is limited by a low resolution of the succession of events as the most recent common ancestor is usually millions of years older than the present day species that are investigated. In vitro evolution experiments allow for high resolution of the genomic change but the necessary time-scales are difficult to obtain and extrapolating directly from in vitro to in vivo should be done with caution.

Based on these two types of studies late stage reductive evolution is known to occur by pseudogene formation and subsequent gradual loss of small genomic regions, whereas the process of reduction in the early stages is less clear with reports emphasizing either gene-by-gene non-functionalization (Silva et al., 2001; Dagan et al., 2006) or the loss of large contiguous regions (Moran and Mira, 2001; Nilsson et al., 2005).

For the lack of useful model organisms for in vivo monitoring of adaptation to a distinctly different environment, Pseudomonas aeruginosa infection of cystic fibrosis patients provides a useful alternative as well as an opportunity for studying the initial stages of host adaptation and the associated mechanisms of evolution. P. aeruginosa is a highly versatile bacterium present in water and soil habitats but also capable of acute infection in a range of hosts from amoeba and plants to humans (Rahme et al., 2000; Pukatzki et al., 2002). Its vast array of functionalities also enables colonization of the lungs of cystic fibrosis (CF) patients, a colonization which is rapidly followed by genetic adaptation to be able to successfully persist in the new CF lung environment (Smith et al., 2006; Rau et al., 2010).

A specific lineage of P. aeruginosa, DK2, with a history of infecting a large number of CF patients over at least 35 years was recently described by (Yang et al., 2011b), providing an opportunity for studying the early stages of adaptation to a host environment. The initial study correlated the phenotypic change with a SNP-based phylogeny based on genome sequences of 12 P. aeruginosa DK2 isolates collected from six individuals over 35 years. Here, we describe the complete genome of this highly adapted pathogenic strain of P. aeruginosa. In addition,
by sequencing the genomes of additional 33 DK2 clones sampled from 16 CF patients we have investigated the dynamics in the overall genomic content of the DK2 lineage, and focused on the role of genome reduction, acquisition of new DNA, and movement of mobile genetic elements in long-term adaptation to the new environment offered by its human hosts.

Materials and Methods

Bacterial strains and genome sequencing
Forty-five isolates of the P. aeruginosa DK2 clone type were sampled over 35 years from 16 CF patients attending the Copenhagen Cystic Fibrosis Clinic at the University Hospital. Isolation and identification of P. aeruginosa from sputum was done as previously described (Hoiby and Frederiksen, 2000). Sequencing of all isolates, except CF510-2006, was performed as earlier described (Yang et al., 2011b) on Illumina’s GAIIx or Hiseq2000 platforms generating 75 base single-end reads using a multiplexed protocol to an average coverage depth of 11-99 fold (except CF333-1991 which was 454-sequenced only cf. Yang PNAS).

Assembly and annotation of the P. aeruginosa DK2 reference genome
The initial study of the P. aeruginosa DK2 clone type was based on a draft genome sequence of isolate CF333-2007 that was represented in 87 contigs assembled from 524,464 reads with an average length of 232 nucleotides generated from 454 pyrosequencing (Yang et al., 2011b). Isolate CF333-2007 was also sequenced on an Illumina GAII platform which generated 5,791,129 75-bp reads to complement the 454 generated reads and to increase the overall read coverage. By combining the two data sets, and by using knowledge of gene synteny in the genomes of other strains of P. aeruginosa genomes, we were able to complete the genome of CF333-2007 over several rounds of improvement: First, the Illumina reads were de novo assembled with the assistance of genome sequences of P. aeruginosa PAO1, PA14, LESB58 and CF510-2006, respectively, using the Columbus module of Velvet version 1.0.16 (Zerbino and Birney, 2008). Each reference assisted de novo assembly was used to extend and close gaps between the 87 original contigs using Minimus2 (Sommer et al., 2007), resulting in 49 new contigs.

We then manually inspected the 49 contigs and found all gaps to be caused by repetitive regions due to the following previously described reasons (Nagarajan et al., 2010; Cerdeira et al., 2011); (1) copies of repeats collapsed onto one another into a single contig, (2) misassembly caused by collapse of tandem repeats, and (3) incorrectly assembled chimeric contigs (false positive re-arrangements). For example, reads from the four nearly identical ribosomal RNA operons co-assembled into a single contig. Although this resulted in gaps in the de novo draft genome, we closed these gaps by first ordering the de novo contigs based on other available P. aeruginosa genome sequences followed by local re-assembly of the reads flanking the gaps. In the same way, gaps caused by tandem units of nearly identical repeats could be resolved by local re-assembly where only exact match reads were used, and chimeric contigs leaving orphan contig ends were broken up and re-assembled. A single gap consisting
of 10 identical repeat units could not be resolved in silico, but was instead confirmed by PCR. In addition, the in silico gap closure were validated by PCR spanning 13 other gaps, including the gaps flanking the ribosomal RNA operons.

Finally, the Illumina reads was iteratively mapped back against the closed genome to correct small sequence errors which would be identified as mutations. Pileups of the reads mapped by Novoalign (Novocraft Technologies) (Krawitz et al., 2010) were produced by SAMtools release 0.1.7 (Li et al., 2009), and single nucleotide polymorphisms (SNPs) called by the varFilter algorithm in SAMtools (samtools.pl varFilter -d 3 -D 10000 –N 10) and SNP quality scores ≥50) were corrected in the genome sequence. In total 531 nucleotides were corrected. All putative indels were manually inspected, and 37 assigned to be true were corrected.

The final genome sequence was annotated using the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (http://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html) and deposited at GenBank (accession no. CP003149).

The accessory genome of P. aeruginosa DK2 was defined as genomic regions not present in at least two other fully sequenced P. aeruginosa genomes (strains PAO1, PA14, LESB, PA7). Alignment through MUMmer3 was used to locate accessory regions. Remaining regions were designated as core regions.

Genome sequencing and phylogenetic analysis of isolate CF510-2006

Clone CF510-2006 was isolated from a CF patient in 2006 with an acute P. aeruginosa infection. Initial molecular genotyping using ArrayTube microarrays showed that the isolate was genetically related to the DK2 lineage. However, the phenotype characteristics of the isolate were incompatible with contemporary DK2 strains present in chronically infected patients which strongly suggest that the CF510-2006 isolate has little or no history of growth in the CF environment prior to its isolation. The sequencing of the CF510-2006 genome was done by AGOWA (Berlin, Germany). First, 454 pyrosequencing reads were generated to an approximate genomic read-coveraged of 50 fold that was assembled into 1,205 contigs. Second, Sanger sequence reads from end-sequencing and primer walking of a fosmid genome library were used to close the majority of gaps to obtain a final genome containing 6,683,163 bp assembled into 71 contigs, which we annotated using the NCBI Prokaryotic Genomes Automatic Annotation Pipeline and deposited at GenBank (accession no. AJHI00000000).

The genome of CF510-2006 was analysed with respect to the set of 363 SNPs that was previously identified to reconstruct the phylogenetic relationship of the P. aeruginosa DK2 lineage (Yang et al., 2011b). Of these 363 informative positions we found 358 in the CF510-2006 genome to be identical to the previously predicted most recent common ancestor (MRCA) of the DK2 lineage. The MRCA was predicted based on the assumption that its sequence at the informative SNP sites would be identical to the homolog sequences found in P. aeruginosa reference strains PAO1, PA14, LESB58, and PA7 (Winsor et al., 2011).

Analysis of genome reduction in multiple clones of the DK2 lineage

To enable the identification of deletions larger than 1 kbp the genome sequence of a DK2 clone we used the CF510-2006 sequence. Illumina reads from all isolates were mapped onto the genome of CF510-2006 using Novoalign (Novocraft Technologies) (Krawitz et al., 2010)
and further processed by SAMtools release 0.1.7. A custom script was employed to identify candidate regions larger than 1 kbp that was not present (i.e. lost) in the sequenced isolate as compared to CF510-2006. Subsequently the candidate regions were manually inspected and exact deletion coordinates were verified by retrieving reads bridging the two regions adjacent to each deletion end. The genome sequence of isolate CF333-1991, which was assembled from 454 pyrosequencing reads, was aligned to the CF510-2006 genome using MUMmer3 (Kurtz et al., 2004) and deleted regions extracted from the alignment. Some deletions spanned several contigs in CF510-2006 with unassembled sequence in between. For such cases, the corresponding contiguous sequence was identified in CF333-2007 and used instead based on the assumption that synteny would be preserved between the two isolates.

**Mobile genetic elements**

Genomic islands in *P. aeruginosa* DK2 were detected by comparative genomics and prediction approaches. The first approach involved alignment to reference *P. aeruginosa* strains PAO1, PA14 and LEB genomes using MUMmer3 (Kurtz et al., 2004). Regions in PADK2 above 8 kbp and not present in two or more reference strains were designated as a genomic island. Additional genomic islands above 8 kbp were predicted using Islandviewer (Langille and Brinkman, 2009). Prophage regions in PADK2 were predicted using Prophinder (Lima-Mendez et al., 2008) using a normalized score of 0.5 as a cutoff and Prophage Finder (Bose and Barber, 2006) using number of hits and predicted phage components as selection criteria. IS elements were located in the DK2 genome using ISsaga (Varani et al., 2011) ([http://www-is.biotoul.fr](http://www-is.biotoul.fr)). The expansion or loss of these known IS elements during infection for Illumina sequenced isolates was tracked by detecting all flanking regions to each IS element in reads for each isolate. A change in location of an IS element or its duplication was then detected as a change or increase in flanking regions while the loss of an IS element was detected as a loss of a flanking region. The presence of CRISPRs in the DK2 genome was detected using CRISPRfinder (Grissa et al., 2007).

**Deletion rates**

Deletion rates per year were assessed for the most recently sampled isolate for each patient. The size in bp of deleted regions occurring since isolate CF30-1979 was divided by number of years from 1979 up until isolation year. The year 1979 and isolate CF30-1979 was chosen as reference time-point since CF30-1979 was isolated this year and based on SNP analysis (Yang et al. 2011b) it is very closely related to the ancestor of all isolates from 1979 and onwards. Deletion rates per generation were calculated using a previously described estimation of in vivo doubling times of the DK2 lineage (Yang et al., 2008; Yang et al., 2011b). These doubling times were used to estimate the number of generations from 1979 to the isolation year of the most recent isolate from each patient, and deletion rates per generation. was then calculated using the number of generations from 1979 and the quantity of genome reduction that occurred after CF30-1979.

**Uptake of novel DNA**


The genomes of Illumina sequenced isolates were *de novo* assembled using Velvet (version 1.0.17) (Zerbino and Birney, 2008) with the Columbus module providing reference assisted assembly. First, reads were mapped onto contigs of CF510-2006 using Novoalign (Novocraft Technologies) (Krawitz et al., 2010) and the mapping data was used during Velvet read assembly with CF510-2006 as reference. Coverage cutoff of 5 and minimum contig length of 100 bp were used as parameters for velvetg while k-mer length of 35 was used for velveth. *De novo* assembled genomes were aligned to the genome of CF510-2006 using MUMmer3 (Kurtz et al., 2004) to locate novel sequence in each isolate. Any novel sequence compared to CF510-2006 was subsequently aligned to CF114-1973, to verify if the sequence was present in the least evolved isolate.

**Results**

Isolates of the DK2 clone type have been extensively collected at the Copenhagen CF center at Rigshospitalet, Denmark, over a long time period ranging from 1973 and onwards. By sequencing the genomes of 45 DK2 isolates sampled from 16 individuals over a period of 35 years we have conducted a retrospective study of the dissemination of the *P. aeruginosa* DK2 clone type among CF patients attending the Copenhagen CF center (Figure 1).

![Figure 1](image-url)

*Figure 1: Overview the patient origin and sampling time of the P. aeruginosa DK2 isolates used in this study. Bacterial isolates are represented by grey circles and their location in the phylogentic tree in Figure 3 is indicated by capital letters. If multiple isolates are sampled the same year from a patient, they are represented by stacked circles.*
A previous study of DK2 strains by Yang et al. (2011b) dealt with the evolutionary relationship of isolates from six patients as determined by genome-wide SNP analysis based on a draft genome scaffold. To enable a more detailed investigation of the genetic content of this highly successful invader of the airways of CF patients, we completed and annotated the genomic features of clone CF333-2007, which has been sampled in 2007 from patient CF333 chronically infected with the DK2 strain for more than 16 years.

_P. aeruginosa DK2 genome features_

The entire genome sequence of _P. aeruginosa_ DK2 clone CF333-2007 was completed using a combination of Illumina and 454 reads together with a previously proposed hybrid _de novo_ assembly strategy (Nagarajan et al., 2010; Cerdeira et al., 2011). The CF333-2007 genome of _P. aeruginosa_ DK2 (which we label PADK2) has a size of 6,402,658 bp comprising 5884 genes, which corresponds to a coding density of 89.3 % (Figure 2). These numbers are similar to those of other assembled _P. aeruginosa_ genomes as is the GC content of 66.3 %. A large fraction of the genome is homologues to the other completed _P. aeruginosa_ genomes of strains PAO1, LESB58, PA14 and PA7 (92.5 %, 93.4 %, 93.2 %, 86.6 %, respectively) (Winsor et al., 2011), whereas 216 kbp containing 195 genes were found to be unique to the DK2 strain relative to the other four strains.
Figure 2: Genome map of *P. aeruginosa* DK2. First and second outer circles denotes CDS (blue) and rDNA (red); The third circle shows the location of genomic islands; the subsequent three rings display blast comparisons of the DK2 genome to *P. aeruginosa* PAO1 (red), PA14 (green), LEB (blue); the next ring depicts the GC content of the PADK2 genome.

The majority of the unique DNA (55%) was contained in three genomic islands related to heavy metal resistance (GI4, GI7 and GI12, respectively) indicating a former habitat of environmental origin with the presence of heavy metals. In total there are 12 genomic islands (GIs) in the PADK2 genome constituting roughly 489 kbp (Table 1).

<table>
<thead>
<tr>
<th>GI</th>
<th>Start</th>
<th>End</th>
<th>Size</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>GI1</td>
<td>667531</td>
<td>680249</td>
<td>12719</td>
<td>Prophage</td>
</tr>
<tr>
<td>GI2</td>
<td>1836969</td>
<td>1862655</td>
<td>25687</td>
<td>LPS/O-antigen biosynthesis</td>
</tr>
<tr>
<td>GI3</td>
<td>2296150</td>
<td>2341644</td>
<td>45494</td>
<td></td>
</tr>
<tr>
<td>GI4</td>
<td>2495033</td>
<td>2601031</td>
<td>105998</td>
<td>Copper resistance</td>
</tr>
<tr>
<td>GI5</td>
<td>2794504</td>
<td>2804831</td>
<td>10328</td>
<td>CRISPR region</td>
</tr>
<tr>
<td>GI6</td>
<td>2828700</td>
<td>2863808</td>
<td>35108</td>
<td>Pyoverdin biosynthesis</td>
</tr>
<tr>
<td>GI7</td>
<td>3056897</td>
<td>3144981</td>
<td>88084</td>
<td>Prophage, chromate/copper/silver resistance</td>
</tr>
<tr>
<td>GI8</td>
<td>3275228</td>
<td>3286591</td>
<td>11363</td>
<td></td>
</tr>
<tr>
<td>GI9</td>
<td>3708420</td>
<td>3723852</td>
<td>15432</td>
<td>Type III secretion</td>
</tr>
<tr>
<td>GI10</td>
<td>4214902</td>
<td>4228275</td>
<td>13373</td>
<td>Non-ribosomal peptide synthetase modules</td>
</tr>
<tr>
<td>GI11</td>
<td>5097794</td>
<td>5188335</td>
<td>90541</td>
<td>Type IVb pilus</td>
</tr>
<tr>
<td>GI12</td>
<td>6054401</td>
<td>6091531</td>
<td>37131</td>
<td>Arsenic resistance, nucleoside metabolism</td>
</tr>
</tbody>
</table>

Table 1: Genomic islands present in the *P. aeruginosa* DK2 genome.
The genome also contains two relatively small prophage regions, approximately 8 kbp and 16 kbp each as well as eight different insertion sequence (IS) element types belonging to five different IS families. One IS element, IS222, is previously characterized while the other seven (ISPa36 – ISPa42) are new and added to the IS Finder Database (Siguier et al., 2006). Three IS elements belong to the TN3 family and all carry passenger genes (3, 6 and 14), while the other IS elements carry no passenger genes. Two CRISPR loci are present in the genome. One consists of two CRISPR regions flanking CRISPR associated (cas) genes, while the other locus is devoid of cas genes. Apart from the spacers within the CRISPR regions, the former locus is identical to a CRISPR-cas system in *P. aeruginosa* strain PA14, where it has a different functionality than typical systems as it does not confer phage resistance but rather is mediating phage regulation of biofilm formation (Zegans et al., 2009; Cady et al., 2011).

**Reductive evolution is prevalent in the DK2 lineage**

Bacterial long-term adaptation to a host environment often includes reductive evolution in which non-essential functions are lost over time. We located all contiguous genome reductions of ≥1 kbp among 44 DK2 isolates sampled from chronically infected CF patients, and in a maximum-parsimonious phylogenetic manner we reconstructed the sequence of reductive events that was most likely to explain the observed genome reductions (Figure 3).
Figure 3: Maximum-parsimonious phylogenetic tree reconstructing the sequence of large deletion events within the DK2 lineage. The tree is based on the genome sequences of 45 DK2 isolates which according to their genetic content locates in the nodes of the tree designated by capital letters. Refer to Table 2 or Figure 1 for the location of each isolate. The size of each node represents the total size (kbp) of genomic regions that have been deleted since the previous node. If more than one deletion has occurred, the circle is divided, and each subdivision corresponds to the size of each deletion. The total quantity deleted regions in a isolate represented by node $J$ can for instance be found by summing all the deletion events along the path from the root to node $J$ (52+1.8+80=133.8 kbp). The root, represents the most recent common ancestor (MRCA) of all isolates.

The phylogenetic reconstruction displayed a high consistency (0.93) allowing for an accurate depiction of the evolutionary relationship among the clones. It predicted the clones to share a common ancestor located at the node between the branch point of the four earliest isolates sampled in 1973 (CF114-1973, CF105-1973, CF43-1973 and CF66-1973). The ancestor had a predicted approximate genome size of 6,535,548 bp. In total, 968 kbp (15 %) of the genome were found to be deleted in at least one of the descendant DK2 clones sampled from chronically infected CF patients. Consequently, 968 kbp (~ 968 genes) of the ancestral *P. aeruginosa* DK2 genome is non-essential for survival in the human hosts.
Table 2: Location of isolates on the nodes in the phylogenetic tree in Figure 2.

The extent of reductive evolution varies significantly between isolates, but every isolate displays reduced genomic content compared to the predicted ancestor. The level of genomic reduction is most extensive in CF240-2002 with a total reduction of at least 525 kbp since infection initiation corresponding to a loss of 8% of the entire genome. The majority of the reduction observed in CF240-2002 is specific to this isolate lineage only (472 kbp), the largest event being the loss of 146 kbp spanning mainly core genomic regions as well as minor accessory regions. The large deletions in CF240-2002 are caused by six separate events while most other reductions in individual isolates are the cause of only one event. The high number of large deletions in CF240-2002 could be stochastic, however, the genome of CF240-2002 contains missense mutations in two exonuclease genes, sbcB and sbcC, involved in recombination. In *E. coli*, mutation of sbcB enhances illegitimate recombination and increases the rate of deletion (Allgood and Silhavy, 1991), hence, the increased number of deletion events and the large quantity of lost regions in CF240-2002 could have a deterministic origin. In total 27 reductive events above 1 kbp are observed during the infection period with an average loss of 44.5 kbp per event (median 26.6 kbp). The lost genomic regions constitute both core regions and accessory regions of various functions (Table 3). Overall, 56% of the lost regions belong to the core genome while 44% belong to the accessory genome, a significant over-representation considering the overall percentage of the accessory genome is 8.5%. 

|--------------|-------------|-------------|--------------|-------------|-------------|-------------|-------------|--------------|-------------|-------------|-------------|-------------|-------------|-------------|

|--------------|-------------|--------------|-------------|

Column defines whether deleted sequence belongs to core or accessory genome. If both types are present within the sequence both are noted with the major type listed first.

Table 3: Description of deleted regions.

Deletion rates are high compared to in vitro estimates

Based on the knowledge of lost sequences and time points of strain isolation it is possible to obtain an estimation of the deletion rate during the reductive evolution. All DK2 clones sampled from chronic infections after 1973 share the same ancestor as recognized by proposed key-adaptive mutations in three global regulators (mucA, lasR and rpoN) and SNP-based phylogenetic analyses (Yang et al., 2011b). The latter suggested clone CF30-1979, which is the earliest clone to possess all three key-adaptive mutations, to closely resemble this common ancestor of all isolates sampled from chronic infections from 1979 and onwards (Yang et al., 2011b). In agreement with this, our analysis of genome reduction predicted CF30-1979 to represent an ancestor of all subsequently sampled clones (Figure 3 and Table 2). Therefore all deletions present in clones isolated after CF30-1979 most likely occurred
after the year 1979 which therefore can be used as a reference point. The lineages share evolutionary trajectory over several time periods and the deletions of each lineage are therefore not independent measures however the results will nonetheless provide a good indication of the deletion rate. Averaging the reduction rate per year for only the latest isolate for each patient (12 isolates) gives an overall deletion rate of 3,723 bp per year. Estimating the deletion rate for the CF240-2002 sub-lineage gives a deletion rate of 22,637 bp per year.

Using previous estimates of in vivo doubling times, it is also possible to estimate the number of generations from 1979 to each endpoint isolate (Yang et al., 2008; Yang et al., 2011b). A deletion rate of 3,723 bp per year then corresponds to 0.62 bp per generation which is 12 to 36 fold higher than previously reported in vitro deletion rates of 0.05 and 0.017 bp per generation for Salmonella enterica grown in rich medium (Nilsson et al., 2005). Although the extensive reduction observed in CF240-2002 significantly contributes to the high deletion rate, exclusion of this isolate from the calculations still result in a high rate of deletion (0.41 bp per generation). For bacteria usually 20 bp of homologous sequence is required for homologous recombination (Ehrlich et al., 1993). Out of the 27 large deletions found here, seven of these contain homologous regions above 20 bp in length in the deletion end-point regions. The homology of three of these are long and spans more than 100 bp. The remaining 20 deletions caused by illegitimate recombination have an average and median length of deletion endpoint identity of 5.5 bp with 10 of the deletions having homologous regions below 5 bp.

Parallel reductive evolution

Two genomic regions have been lost in parallel for more than two sub-lineages during the infection period indicating a selective pressure against encoded functions in the region. In addition to this there are seven examples of overlap between two lost regions. One of the two parallel reductions is the fivefold parallel loss of genes encoding enzymes for homoprotocatechuate (hpc) catabolism, a sub-pathway in tyrosine catabolism. Although some of the five observed reductions also include the loss of large genomic regions surrounding hpc catabolism genes (up to 79kbp) all five reductions include genes for this pathway making it the likely selective target of deletion.

The second example of parallel reductive evolution is the threefold loss of a region highly similar to the pathogenicity island PAGI-5, although the region in the DK2 isolates do not contain the main region of virulence genes (NR-II) identified in PAGI-5. This GI spans 90
kbp and the three-fold independent loss therefore constitutes a large fraction of the accessory genome reductions (50%). Two genomic islands closely related to PAGI-5 are pKLC102 and PAPI-1 both of which are mobilizable at respective frequencies of 10 % and 0.16 % (Qiu et al., 2006; Klockgether et al., 2007). Since PAGI-5 shares similar integration and transfer gene modules as PAPI-1 and pKLC102 it could possess similar mobilization frequencies. Accordingly the threefold independent loss of PAGI-5 homologue need not signify the GI has no fitness contribution, contrary the preservation of PAGI-5 in other lineages could indicate otherwise. We also note that deletion of different parts of the psl gene cluster occurred on three occasions (Table 3) which suggest selective pressure for loss of Psl exopolysaccharide production.

Acquisition of novel DNA

For rapid adaptation, acquisition of novel DNA can play a crucial role. However, during evolution of the DK2 lineage in the CF airways, acquisition of novel genetic elements, e.g. viruses or plasmids, seem to be less important in changing the genomic content. De novo assemblies of all genomes were performed to detect genetic elements not already present in the original infecting clone. Only 54 kbp of novel DNA sequence was found in CF114-1973 with half of it being homologous to phage D3 sequence indicating the insertion of a related prophage. However, since CF114-1973 is the first of the early 1973 clones to have diverged, the sequence could have been present also in the ancestor and merely lost in a lineage leading to all other isolates. No novel sequence could be found in other isolates. The above is the only putative example of prophage insertion in the genomes of any DK2 strains during infection indicating a low level of temperate phage attack or high level of phage resistance in the DK2 population. Contrary, a GI containing two prophage regions was lost in the branch leading to CF66-1973 and all strains isolated subsequent to 1973. The spacers in the two CRISPR loci present in the DK2 lineage also remain unchanged for the duration of infection indicating further the absence of temperate phage attack or simply that the systems are not functional. The identical system described by Cady et al. for strain PA14 was functional although 105 SNPs (within 8140 bp) differentiate it from the strain DK2 system.

Limited IS element expansion during infection
The *P. aeruginosa* DK2 genome contains eight different IS element types of which all exist in the 1973 isolates and have consequently not been acquired during infection. However one type, IS222, existed in 7 copies in the genome of CF333-2007 while only in 5 copies in CF114-1973 which based on SNP analysis is the least evolved isolate. The increase in copy number of this IS element occurred on branch C and it is therefore present in CF66-1973 and all subsequent isolates. Apart from this only two other multiplication events of IS222 occurs in CF43-1973 and in an ancestor to the CF66 isolates, leading to the presence of an extra copy in CF66-1992, CF66-2002 and CF66-2008.

**Discussion**

The early stages of *P. aeruginosa* adaptation to the cystic fibrosis airways is characterized by the loss of large genomic regions. Whether the deletions are caused by drift or selection is unknown, but large deletions of core genome could facilitate the loss of beneficial genes making the bacteria susceptible to Muller’s ratchet. Although the largest proportion of reductions reside in the core genome there is a significant overrepresentation of reductive evolution in the accessory genome. This might also be expected as accessory functions often ensures proliferation in certain niches and here we e.g. see the loss of some of the genes encoding heavy metal resistance in certain isolates. Probably many of the accessory functions in the genome are neutral at best and therefore more prone to loss by drift or even selection against their functions.

The time scale in other studies of naturally evolved human pathogens compared to a known or theoretical ancestor are usually of a different magnitude spanning millions of years. *Burkholderia mallei* provides such an example with 1.4 Mbp lost over 3.5 million years (Song et al., 2010) and *Mycobacterium leprae* another with 408 out of 2977 genes deleted over 66 million years (Gomez-Valero et al., 2007). In these cases there is no information on the deletion dynamics. Consequently, due to a lack of intermediary strains rough estimates of *in vivo* deletion rates and dynamics have not been reported until now.

*In vitro* deletion rates have been reported for *Salmonella enterica* grown in rich medium rendering several biosynthesis functions superfluous (Nilsson et al., 2005). The observed deletion rates for the DK2 lineage is however 12 to 36 fold higher than those reported for the *S. enterica* evolutionary experiment. The *S. enterica* genome is around 4.8 Mbp while the *P.*
aeruginosa DK2 genome is 6.4 Mbp, which could suggest that *P. aeruginosa* has a larger potential for genome reduction. In addition, *P. aeruginosa* is a versatile organism with a diverse metabolic repertoire, for which there is no need in the more confined human airways environment.

The CF lung contains a multitude of microbial species suggesting that this environment could constitute a rich source of new genomic material. However, during the infection period of the DK2 strain lineage only little or no novel DNA was acquired in the lineage. Only part of a prophage region was possibly acquired in CF114-1973. The prophage region was not present in other isolates suggesting either that it entered in the branch leading to CF114-73 or was present in the shared ancestor but subsequently lost before the isolation of other 1973 strains. The overall lack of genetic expansion could e.g. be due to the lack of beneficial novel DNA in the lung or reduced potential for uptake and acceptance of novel DNA in the isolates. Several receptors for DNA uptake, e.g. pili, flagella and LPS are either lost, reduced or altered during infection (Yang et al., 2011b; Yang et al., 2011a) likely diminishing the rate of DNA uptake. Since the DK2 lineage emerges as a highly successful colonizer of the CF lung it stands to reason that lateral acquisition of DNA is in no way essential for *P. aeruginosa* adaptation to the CF lung and probably plays a limited role in this. This is in contrast to the gut microbiota where lateral transfer of virulence and resistance determinants has been proposed to provide at mechanisms for opportunistic pathogens to emerge.

Transposable elements can impact evolution of an organism by inserting in new locations altering gene functionality or catalyzing deletions by providing opportunity for homologous recombination. In the infecting lineage only one transposable element, IS222, seems capable of movement and the limited increase in copy number from 5 to 7 for most isolates, only affecting two genes, demonstrates its low significance in generating direct mutations compared to the high number of SNPs occurring in the genome over time (Yang et al., 2011b). This is in contrast to recent *in vitro* experiments revealing an important contribution of IS elements to mutation generation (Barrick et al., 2009; Gaffe et al., 2011). In the early stages of reductive evolution it has been suggested that extensive proliferation of IS elements occurs based on the high densities of these in fairly recent host-restricted species such as *B. mallei*, *Bordetella pertussis*, *Shigella flexneri* and *Salmonella enterica* Typhi (Parkhill et al.,
In *B. mallei* and *B. pertussis* there seems to be a link between the location of IS elements and large deletions and inversions, suggesting that IS elements function as catalysts for reductive evolution, and this mechanism was therefore proposed to be a key component in the early stage. However, at the very earliest point in adaptation of *P. aeruginosa* to the CF airways the proliferation of IS222 is still too limited for catalysis of reductions, which creates a potential for only very large (usually deleterious) reductions. Although the smallest distance between IS222 in later isolates is only 6 kbp the second-smallest is close to 340 kbp. Possibly, with a higher density of IS222 or the presence of another IS element with faster duplication the rate of reduction would be increased. Even though IS elements were estimated to be involved in nearly all large deletions in *B. mallei* we see here in the earliest of stages substantial genome reduction of up to 8 % of the genome without any involvement of IS elements.

Instead of IS mediated genome reduction we observe mechanisms of homologous recombination and more frequently, illegitimate recombination. Likely this proportion is due to the abundance of non-homologous regions over homologous regions. Nilsson et al. (2005) also found mainly non-homologous end regions of reductions during *S. enterica* in vitro reductive evolution (6 regions) while one end region contained 40 bp with high identity when investigating the early stages of in vitro reductive evolution. These provide two examples of the earliest stage of host adaptation and both present reductive evolution through mainly illegitimate recombination.

The adaptation of the DK2 lineage to the CF lung environment provides for the first time insight into the early stages of adaptation of a bacterial pathogen to its human host. This early stage is exemplified by substantial reductive evolution predominantly through deletions of large genomic regions. Genomic change through uptake of novel DNA is on the other hand very limited which is likely a reflection of gain of function being less important during transition to a permanent host association. The movement of mobile genetic elements is likewise scarce as IS element multiplication is low. Accordingly, IS elements are not involved in the process of genome reduction observed here. In spite of this high deletion rates are readily occurring through mechanisms of deletion mainly consisting of illegitimate recombination and to a lesser extent homologous recombination. Consequently, the involvement of IS elements is dispensable even for extensive genome reduction. Deletion
rates are most likely determined by several factors, one of them being the genome reduction potential. This could explain the high deletion rate of the DK2 lineage as a significant portion of the coding capacity of \textit{P. aeruginosa} should be geared towards its normal environmental habitats and therefore dispensable in the host environment.

References


