Long-term evolution of *Burkholderia multivorans* bacteria during chronic respiratory infections of cystic fibrosis patients

Filipa Duarte Pessoa

Thesis to obtain the Master of Science Degree in

Biotechnology

Supervisor: Doctor Leonilde de Fátima Morais Moreira

Co-Supervisor: Doctor Inês Nunes Silva

Examination Committee:

Chairperson: Doctor Miguel Nobre Parreira Cacho Teixeira

Supervisor: Doctor Leonilde de Fátima Morais Moreira

Members of the Committee: Doctor Arsénio do Carmo Sales Mendes Fialho

November 2017

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor Prof. Leonilde Moreira for the constant support and guidance during the development of this work, and my co-supervisor Dr. Inês Silva for all the help and valuable advice. I also thank every member of the Biological Sciences Research Group (BSRG) that helped me throughout my experimental work, in particular Marcelo Ramires, for all the teachings and for always showing great availability, and Filipe Bica, for making the hardest days at the lab a lot more fun.

Additionally, I would like to acknowledge Prof. David P. Speert from the University of British Columbia, Canada, for providing the clinical isolates I used in the experiments, and to Prof. Vaughn Cooper and Dr. Chris Marshall from the University of Pittsburgh, USA, for all their efforts in the *de novo* assembly step required for this work.

Financial support by Programa Operacional Regional de Lisboa 2020 (LISBOA-01-0145-FEDER-007317) is also thankfully acknowledged.

Finally, I would like to thank my family and friends, especially my parents, for their unconditional support and invaluable motivation. A special thanks to my dear friends Inês and Rebeca, who have been by my side since the beginning of this five-year journey, and Mariana, who has been by my side for as long as I can remember. To Laura, the one person that always understands how I'm feeling, even when I can't put it in words, and knows exactly what to say in every single moment. And to João, for the endless support, patience and for constantly encouraging me to be better.

RESUMO

A fibrose quística é uma doença genética autossómica recessiva que afeta principalmente as funções pulmonares dos doentes, sendo que infeções respiratórias causadas por agentes patogénicos oportunistas podem ter consequências letais. Infeções crónicas causadas por bactérias pertencentes ao complexo *Burkholderia cepacia* são associadas a uma pior condição clínica para os doentes. Constituindo um ambiente complexo, as vias respiratórias de doentes com fibrose quística apresentam diversas fontes de stress que proporcionam o surgimento de variantes bacterianas com uma capacidade superior de subsistir nesse ambiente. Neste trabalho foi estudada a evolução genómica e fenotípica de 37 isolados clínicos de *Burkholderia multivorans* recolhidos sequencialmente de dois doentes com fibrose quística ao longo de 18 anos. Com este fim, a estratégia compreendeu a sequenciação de todos os genomas e identificação de mutações entre os isolados de cada doente e a análise de fenótipos relacionados com a adaptação bacteriana às vias respiratórias de doentes com fibrose quística, tais como a produção do antigénio O do lipopolissacárido, a suscetibilidade a antibióticos, a formação de biofilmes, a motilidade e a virulência em *Galleria mellonella*. Observou-se que os últimos isolados recolhidos dos doentes apresentaram resistência a antibióticos e formação de biofilmes superiores, e motilidade, taxas de crescimento e virulência em *G. mellonella* inferiores. Complementando estes resultados, as mutações mais relevantes que diferenciaram os isolados recolhidos posteriormente dos iniciais ocorreram em genes relacionados com resistência a antibióticos, metabolismo de lípidos e regulação da formação de biofilmes, motilidade e virulência. Adicionalmente, o regulador global *fixL* foi identificado como um possível candidato para a regulação de propriedades que influenciam a adaptação de bactérias do complexo *Burkholderia cepacia* ao hospedeiro. A compreensão das estratégias inerentes à adaptação evolutiva de bactérias responsáveis pelo desenvolvimento de infeções crónicas em doentes com fibrose quística poderá elucidar novos alvos para a criação de terapias com maior eficácia.

PALAVRAS-CHAVE

Fibrose quística; *Burkholderia multivorans*; evolução bacteriana; infeções respiratórias crónicas; genómica comparativa

ABSTRACT

Cystic fibrosis is an autosomal recessive genetic disorder that affects mainly the pulmonary functions of patients, creating a particular environment for airway infections with opportunistic pathogens that might have fatal consequences. Chronic infections with *Burkholderia cepacia* complex bacteria have been associated with a worse clinical outcome for patients. As a complex and challenging environment, the cystic fibrosis airways display several stress sources that lead to the emergence of bacterial variants with improved capacity to subsist in that environment. The aim of this work was to assess the genomic and phenotypic evolution of 37 *Burkholderia multivorans* isolates sequentially retrieved from two infected cystic fibrosis patients over 18 years. The approach included genome sequencing of all isolates and the identification of mutations among the isolates of each patient, and the study of phenotypes associated with adaptation to the cystic fibrosis airways. Alterations in lipopolysaccharide O-antigen production, susceptibility to antibiotics, biofilm formation, motility, growth rate, and virulence in *Galleria mellonella* are reported here, with the latter isolates collected from both patients displaying increased resistance to antibiotics, increased biofilm formation, decreased motility, decreased growth rates and decreased virulence in *G. mellonella*. Furthermore, the most relevant mutations distinguishing the latter isolates from the first isolates occurred in genes that may relate to antibiotic resistance, lipid metabolism and regulation of biofilm formation, motility and virulence, with the identification of the global regulator *fixL* as a possible candidate for regulating traits that impact *Burkholderia cepacia* complex bacteria adaptation to the host. Insights into the strategies underlying the adaptive evolution of bacteria during the course of cystic fibrosis chronic infections may provide novel targets for the development of therapies.

KEYWORDS

Cystic fibrosis; *Burkholderia multivorans*; bacterial evolution; chronic respiratory infections; comparative genomics

LIST OF ABBREVIATIONS

ABC	ATP-binding cassette
AHL	*N*-acyl-homoserine lactone
ATP	Adenosine triphosphate
Bcc	*Burkholderia cepacia* complex
BDSF	*Burkholderia* diffusible signal factor
BLAST	Basic local alignment search tool
cAMP	Cyclic adenosine monophosphate
c-di-GMP	Cyclic diguanosine monophosphate
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane conductance regulator
CFU	Colony-forming unit
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EPS	Extracellular polymeric substance (exopolysaccharide)
FBS	Fetal bovine serum
HMW	High molecular weight
LPS	Lipopolysaccharide
MFS	Major facilitator superfamily
MOI	Multiplicity of infection
mRNA	Messenger RNA
MSD	Membrane spanning domain
NBD	Nucleotide binding domain
OD	Optical density
PBS	Phosphate-buffered saline
PMN	Polymorphonuclear leukocytes
PTS	Phosphotransferase system
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
SCFM	Synthetic cystic fibrosis medium
SDS	Sodium dodecyl sulfate
SNP	Single-nucleotide polymorphism
tRNA	Transfer RNA
YEM	Yeast extract mannitol medium

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Cystic fibrosis

1.1.1. Cystic fibrosis overview

Cystic fibrosis is the most common inherited disease among Caucasians (O'Sullivan & Freedman 2009). The genetic defect responsible for the CF disorder results from mutations in the gene that encodes a membrane chloride channel, known as the cystic fibrosis transmembrane conductance regulator (CFTR; Riordan 2005). Recurrent symptoms of CF are exocrine pancreatic insufficiency, chronic airway infections with opportunistic pathogens, elevated sweat electrolytes and male infertility (Gibson *et al.* 2003; Ratjen & Döring 2003). This condition affects mainly the pulmonary functions of the infected patients, as it impairs normal mucociliary clearance of bacterial pathogens and consequently originates bronchopulmonary infections that can be extremely severe (Ratjen & Döring 2003; Filkins & O'Toole 2015). These infections can ultimately lead to the destruction of lung tissue, as they cause a persistent inflammatory response (Chmiel & Davis 2003).

In 1974, the international median age at death of CF patients was 8 years, since the infections were not treated (Folkesson *et al.* 2012). More recently, the mean expected lifetime has increased drastically due to intensive antibiotic therapies and the median predicted survival age in 2015 was 41.6 years (Cystic Fibrosis Foundation US 2016). Patients with cystic fibrosis receive antibiotic treatment intermittently for the management of chronic infection and in a more aggressive manner upon hospitalization due to pulmonary exacerbation (Filkins & O'Toole 2015). Despite the remarkable therapeutic advances accomplished to present time, life expectancy and quality of life for individuals with CF are still limited.

1.1.2. Cystic fibrosis transmembrane conductance regulator (CFTR)

The cystic fibrosis transmembrane conductance regulator (CFTR) is a 1480 amino acid protein encoded by a 230-kb gene known as the CFTR gene (Ratjen & Döring 2003). This protein is a member of the ATP-binding cassette (ABC) family of membrane transport proteins that functions as a cyclic AMP-regulated chloride ion-selective channel (Riordan 2005; Farinha & Matos 2016). Although functioning mainly as a chloride channel, CFTR has additional regulatory functions, such as inhibition of sodium transport through the epithelial sodium channel and regulation of ATP channels, intracellular vesicle transport and acidification of intracellular organelles (O'Sullivan & Freedman 2009).

CFTR channels are present at the luminal surface of epithelia and are composed of five domains (figure 1): two transmembrane domains (MSD1 and MSD2) that form the channel through the membrane, two nucleotide binding domains (NBD1 and NBD2), responsible for gating the flow of chloride (Cl^-) through the channel, and a regulatory domain (R), which mediates the regulation of channel activity (Riordan 2005).

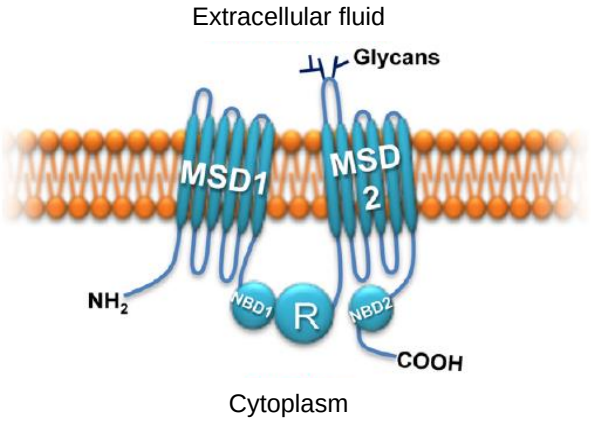

Figure 1 - Illustrative scheme of the CFTR protein at the cell membrane, highlighting the five constituting domains, the amino terminus (NH_2), the carboxy terminus (COOH) and the two N-linked oligosaccharides (Glycans). MSD1/2: membrane spanning domains; NBD1/2: nucleotide-binding domains; R: regulatory domain. Adapted from Farinha & Matos 2016.

Mutations in the CFTR gene lead to malfunction or loss-of-function of this chloride ion channel, causing cystic fibrosis (McKone *et al.* 2006). The most severe phenotypic manifestations of the absence of functional CFTR channels at the surface of several epithelia occur at the airways, where impaired CFTR function disturbs ionic homeostasis and airway hydration, leading to increased mucus viscosity (O'Sullivan & Freedman 2009; Farinha & Matos 2016). This leads to obstructed mucociliary clearance of inhaled microorganisms and early recruitment of inflammatory defense mediators, ultimately leading to respiratory failure (Gibson *et al.* 2003).

1.1.3. Mutations in the CFTR gene

There are more than 1500 possible mutations in the CFTR gene, the most frequent of which is the deletion of a phenylalanine residue at position 508, identified as F508del (O'Sullivan & Freedman 2009). The F508del mutation is present in at least one allele of 80-90% of patients with CF (Bell *et al.* 2015) and the most common CFTR genotype of CF patients is homozygous F508del/F508del (Farinha & Matos 2016). Individuals that are homozygous for this mutation generally present a more severe clinical phenotype than F508del heterozygotes and individuals with no F508del allele (McKone *et al.* 2006).

Mutations in the CFTR gene can affect the biogenesis, transport, activity and stability at the cell surface of the chloride channel (Riordan 2005; Bell *et al.* 2015), but all CF-causing mutations ultimately lead to hindered cAMP-regulated chloride secretion by epithelial cells. CFTR mutations have been grouped into six classes, according to their functional defect and the effect it has on the biosynthesis, transport or function of the CFTR channel (figure 2). Class I mutations impair protein production, often by nonsense mutations or splicing disruption; class II mutations influence CFTR processing and traffic, causing the misfold of the protein and directing it to degradation (this class includes most of F508del

mutations); class III mutations disrupt channel regulation through impaired gating, by affecting ATP binding or hydrolysis; class IV mutations reduce chloride conductance upon cAMP stimulation, preventing the correct flow of ions through the CFTR channel; class V mutations decrease CFTR protein levels, frequently by affecting splicing and generating a low amount of functional mRNA transcripts; and class VI mutations decrease the stability of the CFTR channel at the cell surface (reviewed in Farinha & Matos 2016).

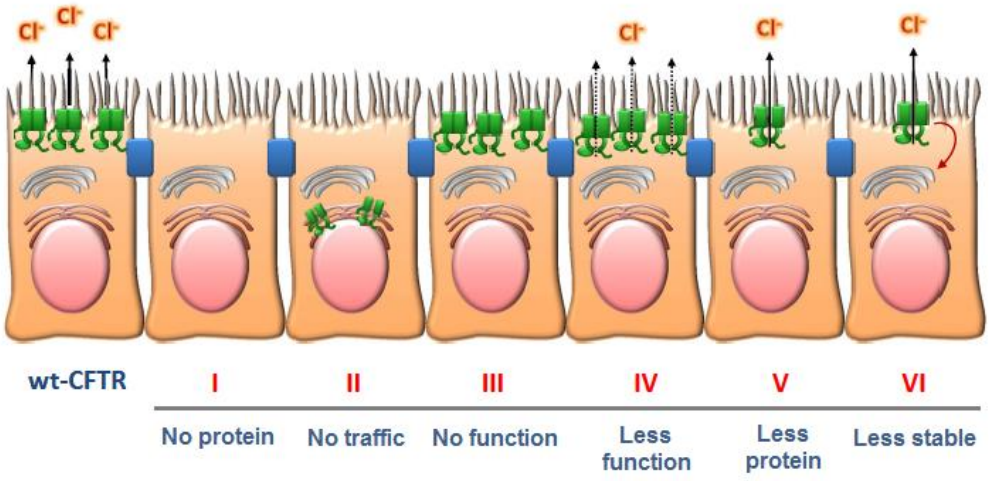

Figure 2 - Scheme of CFTR mutations grouped in classes I to VI based on the associated functional defect (adapted from Amaral & Farinha 2013). Class I: no protein is produced; class II: misfolded protein is retained at the endoplasmic reticulum and subsequently degraded in the proteasome; class III: disrupted channel regulation, impairing channel opening; class IV: reduced flow of ions through the channel; class V: decreased protein levels; class VI: reduced plasma membrane stability.

1.1.4. Cystic fibrosis lung environment

The airways of patients with CF constitute a warm, humid and nutrient-rich environment for bacterial growth (Oliver *et al.* 2000). However, when transitioning from the environment into the CF airways, bacteria are subjected to nutritional and physicochemical alterations leading to highly stressful conditions (Folkesson *et al.* 2012). Several of the new conditions bacteria become exposed to are the presence of antimicrobial agents, both antibiotics and host immune mediators, oxygen limitation and osmotic stress due to the high viscosity of the mucus produced in the airways (reviewed in Döring *et al.* 2011; Folkesson *et al.* 2012). Other important factors that need to be considered are competition from other resident microorganisms and the continuous changing of the CF airway environment as the infection advances, since the host immune response results in severe inflammation and extensive structural changes to the airways (Regamey *et al.* 2011).

The airways of humans constitute a vastly compartmentalized environment and each compartment has distinct characteristics. In the upper airways, the paranasal sinuses contain thick mucus that can provide a site for bacterial growth, and sinus cavities have less airflow and less exposure

to host immune mediators and to antibiotics than the lower airways (Hansen *et al.* 2008). In the lower part of the respiratory tract, the conductive zone refers to the tracheobronchial region and the respiratory zone includes the bronchioles and the alveoli (reviewed in Folkesson *et al.* 2012). The viscous layer of mucus produced in the bronchi is the major site for bacterial reproduction (figure 3). It has shifting levels of nutrients, oxygen and antibiotics, and is subject to osmotic and oxidative stresses. The presence of bacteria in the respiratory zone is uncommon and usually associated with destruction of the lung (Bjarnsholt *et al.* 2009; Ulrich *et al.* 2010). In individuals with CF, local differences in the inflammatory response among different areas of infection and due to competition between co-colonizing bacterial populations may lead to the development of microhabitats (reviewed in Folkesson *et al.* 2012).

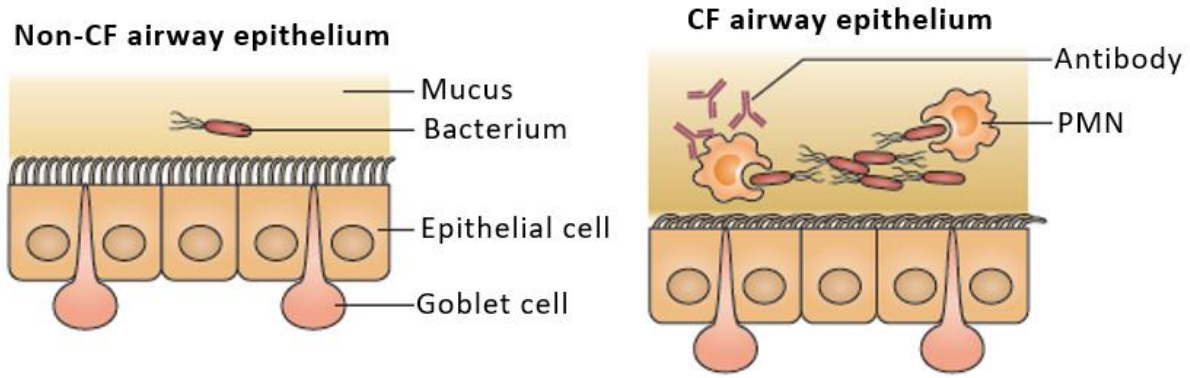

Figure 3 - Differences in mucociliary clearance between a non-CF airway epithelium and a CF airway epithelium. In the airways of individuals without CF, microbial cells that are trapped in the thin mucus layer produced by goblet cells are removed by the cilia of epithelial cells (mucociliary clearance). In the airways of CF patients, the cilia are unable to efficiently clear the viscous mucus layer, leading to colonization by bacteria. Immune responses mediated by antibodies and polymorphonuclear leukocytes (PMNs) lead to damages in lung tissue and impaired lung function. Adapted from Folkesson *et al.* 2012.

The thick mucus combined with the decreased clearance contribute to the formation of mucus plugs that lead to airway obstruction and may form a protected niche for the development of microbes (Boucher 2004). Within these mucus plugs, hypoxic and anoxic regions may form, as a steep oxygen gradient develops (Filkins & O'Toole 2015). Microbes traditionally thought of as aerobes may often be prepared to grow in environments with low oxygen concentrations or even with no oxygen. For example, *P. aeruginosa* is able to grow in anaerobic conditions, by using nitrate and nitrite for respiration (Alvarez-Ortega & Harwood 2007). The role of anaerobic organisms in the development of polymicrobial communities and regarding disease progression is still poorly understood (Klepac-Ceraj *et al.* 2010; Filkins & O'Toole 2015) and requires further study.

In conclusion, multiple niches may be present in CF airways, as this complex environment shows spatial and temporal heterogeneity. Since each niche comprises different selective pressures, infecting bacterial populations are likely to evolve into genetically distinct sublineages, as a consequence of adaptation to the niche they occupy (Marvig *et al.* 2014).

1.1.5. Pathogens and adaptation to the cystic fibrosis airways

The airways of each infected patient have their own individualized polymicrobial community, with complex inter-microbial and host-pathogen interactions that may affect the lung environment, response to treatment and clinical outcome (Filkins & O'Toole 2015). Although multiple bacterial species can coexist in the respiratory tract of CF patients and there is a great inter-individual variation (Klepac-Ceraj *et al.* 2010), there is a usual pattern of pathogen succession characteristic of most infections. In the initial years of patients' lives, the most abundant pathogens are *Haemophilus influenzae* and *Staphylococcus aureus*, followed by *Pseudomonas aeruginosa* and *Burkholderia cepacia* complex (Bcc) bacteria, which are important CF pathogens and can cause rapid decline in lung function (Ratjen & Döring 2003; Filkins & O'Toole 2015). In most patients, the establishment of *P. aeruginosa* infection is associated with a more rapid disease progression, yet there is significant variation in the following rate of pulmonary decline (Klepac-Ceraj *et al.* 2010). In addition to bacterial pathogens, diverse fungi (such as *Candida* spp., *Aspergillus* spp. and *Malassezia* spp.) and some viruses (such as influenza and syncytial virus) have been recognized to co-inhabit the lungs of patients with CF (Etherington *et al.* 2014; Gilligan 2014; Willger *et al.* 2014).

The primary sources of infection are the environment, patient-to-patient transmission and the patient's own microbiota (Filkins & O'Toole 2015). *P. aeruginosa* and members of Bcc are typical soil bacteria and patients with CF usually acquire them from environmental reservoirs (reviewed in Gilligan 2014).

During infection, microorganisms have to adapt to a new environment, facing the host immune system, antibiotics and a variable substrate composition. Throughout this adaptive process, it is frequently noticed that the function of several genes is progressively lost, mainly genes associated with bacterial virulence factors (Rau *et al.* 2010). Microbial communities in the CF lung come across host immune factors, frequent and prolonged antibiotic treatment and a challenging environment that includes the presence of hypoxic and anoxic regions. These characteristics promote the development of chronic communities that could occasionally have decreased microbial diversity, as organisms that become highly adapted to the environment and resilient to treatment are the only ones capable of survival (Filkins & O'Toole 2015). This short-term decreased microbial diversity, which occurs mostly after antibiotic treatment, is generally associated with decreased health (Zhao *et al.* 2012; Daniels *et al.* 2013). Some approaches to promote microbial diversity in the airways of CF patients have been proposed, such as decreasing the use of antibiotics or increasing the use of prebiotics and probiotics (Filkins & O'Toole 2015). It seems clear that the balance between intensive antibiotic treatment, which contributes to the extended life span of CF patients, and minimizing the decrease in microbial diversity needs to be explored.

Patients suffering from CF may be chronically infected for years and this long-term infection is associated with genomic evolution and adaptation of the bacterial population colonizing the lungs of CF patients. Throughout these infections, bacterial pathogens have been frequently reported to become more resistant to antibiotics (Smith *et al.* 2006; Silva *et al.* 2016). Little is known about within-host

evolutionary processes, but a molecular understanding of the strategies developed by bacterial pathogens to adapt inside their hosts should lead to the development of more efficient treatments for these infections.

To fully grasp the mechanisms underlying within-host evolution of pathogens, it is essential to understand the function of infecting microbes in the context of their community and not only as individuals. It is also critical to distinguish between microbes that do not directly contribute to the progression of disease and those that have pathogenic roles with high impact in the clinical outcome (Filkins & O'Toole 2015). Moreover, during the colonization of the airways of CF patients, bacteria can accumulate several mutations that allow them to adapt to the human host (Smith *et al.* 2006; Lieberman *et al.* 2011; Silva *et al.* 2016) and it is crucial to identify which of those mutations occur by chance and which arise as a mechanism of adaptation (Lieberman *et al.* 2014).

In the CF airway, the inflammatory response of the host, the presence of antibiotics, the oxygen limitation and the high osmotic pressure constitute stress sources that select for protective mutations (reviewed in Folkesson *et al.* 2012). The evolution of organisms that colonize changing environments may be directed to the development of specialists in a specific niche or generalists that can subsist in many different environments (Folkesson *et al.* 2012). The CF airway, as a spatially structured and complex environment, is thought to promote diversification and specialization of the microorganisms that colonize the different compartments. The observed phenotypic diversity of *P. aeruginosa* strains supports this hypothesis. Many colony morphology variants, the development of hypermutable strains and resistance to a wide range of antimicrobial agents have been documented (Oliver *et al.* 2000; Mena *et al.* 2008; Hoboth *et al.* 2009).

At the present time, it is possible to conduct a detailed and time-resolved characterization of the evolutionary paths of bacteria infecting CF patients, as regular sampling from sputum normally produced by patients is feasible and genome-sequencing, transcriptomics, metabolomics and metagenomics techniques allow thorough analyses of the clonal isolates sampled (reviewed in Folkesson *et al.* 2012). This has enabled the study of evolutionary traits of bacterial pathogens infecting CF patients by analysis of the genomes of the same strain during long-term infections. Studies of this nature have shown how these pathogens evolve over time and accumulate mutations in specific genes of a single clonal lineage, occasionally leading to diversifying lineages that coexist for many years (Lieberman *et al.* 2014). However, little is known about evolutionary paths concerning different strains of the same species and in what way the mechanisms of adaptation of genotypically distinct isolates can be related (Marvig *et al.* 2014). Efforts have been made to understand how lineages of bacterial pathogens with different genetic backgrounds can end up in a similar pathway leading to typical adaptive phenotypes, while chronically colonizing CF airways.

Genomic and phenotypic evolution strategies *of P. aeruginosa* and Bcc bacteria chronically infecting CF patients will be reviewed in this work, as infections with these pathogens are particularly challenging and have been associated with increased decline in lung function in CF patients.

1.2. *Pseudomonas aeruginosa* infections in cystic fibrosis

1.2.1. *P. aeruginosa* overview

Pseudomonas aeruginosa is a versatile Gram-negative bacterium capable of colonizing many different environments and causing severe infections in human hosts (Mathee *et al.* 2008). It has one of the largest genomes in the bacterial world, with nearly 6000 genes, with more than 500 constituting regulatory genes, suggesting that this microorganism has a great competitive ability to occupy diverse types of niches and adapt to different environmental conditions (Rau *et al.* 2010).

P. aeruginosa genome comprises one circular chromosome and a variable number of plasmids, with sizes ranging from 5.5 to 7 Mbp (Eberl & Tümmler 2004; Klockgether *et al.* 2011). It contains a conserved core component that shows a minor interclonal sequence diversity and a variable part with clone- or strain-specific genomic islands (Klockgether *et al.* 2011). Furthermore, the genome of *P. aeruginosa* includes a large number of virulence genes, such as *lasR* (quorum sensing regulator), *exsA* (type III secretion regulator), *retS* (motility, biofilm and type VI secretion system regulator) and *mexA* (multidrug efflux), which are essential for its capacity to cause infections in plants, animals and humans (Rau *et al.* 2010).

1.2.2. *P. aeruginosa* in the cystic fibrosis airways

P. aeruginosa is responsible for severe chronic infections in the lungs of CF patients and this is a major cause of morbidity and mortality among these patients (Folkesson *et al.* 2012). Initial colonization by this pathogen is usually caused by environmental strains and in many cases this original colonization progresses into a chronic infection that is virtually impossible to eradicate (Rau *et al.* 2010). It is well known that the genetic and phenotypic traits displayed by isolates from chronically infected CF patients are typically distinct from the characteristics exhibited by the clonal isolates that gave rise to the initial infection years before (Folkesson *et al.* 2012; Markussen *et al.* 2014), pointing towards an evolutionary adaptation of the bacteria that is the basis of the long-term persistence of this pathogen in the airways during chronic infections.

Regardless of the host inflammatory response and antibiotic treatment, infections by *P. aeruginosa* tend to persist in the airways of individuals with CF, becoming chronic and leading to lung decline and in some cases death (Hogardt & Heesemann 2010). Even though the specific ways by which patients acquire this bacterium are not known, studies show that infection by *P. aeruginosa* is present by the age of 20 in 60-70% of CF patients (Folkesson *et al.* 2012). Usually, the initial colonizing strains originate from unidentified environmental reservoirs, but transmission between patients may also occur (Burns *et al.* 2001).

During the period of intermittent colonization prior to the establishment of a chronic infection, which can last for several years, intensive antibiotic therapy can effectively postpone the onset of chronic

infection (Høiby *et al.* 2005). Chronic CF infections are characterized by constant growth of *P. aeruginosa* in airway secretions and are associated with a much higher degree of inflammation than that found in intermittent colonization, leading to increased lung tissue damage and decreased lung function (Folkesson *et al.* 2012). Even though antibiotic treatment seems to be effective in extending the lives of patients with CF, it is essential to identify alternative targets for the development of new therapeutic strategies, as antibiotic resistance and the lack of new drugs are important hurdles linked to this approach. Moreover, it has been reported that the presence of *P. aeruginosa* along with long-term antibiotic treatment decreases taxonomic richness within samples and phylogenetic diversity within each community (Flanagan *et al.* 2007; Klepac-Ceraj *et al.* 2010). This suggests that the decline in diversity caused by antibiotic treatment allows *P. aeruginosa* to become the dominant organism in infected CF patients (Flanagan *et al.* 2007).

1.2.3. Genomic and phenotypic evolution of *P. aeruginosa* isolates in cystic fibrosis

Upon colonization of the airways of CF patients, bacteria undergo changes in gene expression directing them to adaptation to the new environment, which presents several challenging stress sources, such as host immune defenses and antibiotic therapy. These selection conditions are prone to increase the occurrence of mutations (Oliver *et al.* 2000), leading to evolution of the infecting microorganisms to improve adaptive processes.

In *P. aeruginosa*, several of the changes that occur in the gene expression profile when bacteria face stressful conditions are specific for the particular stress. Nevertheless, there is a common regulator that is associated with the responses to many different antibiotics and to osmotic shock, among others, which is the *algU* regulator (Jones *et al.* 2010). The *algU* regulator is activated under stress conditions due to mutations in the *mucA* gene and causes downregulation of central metabolism, motility and production of virulence factors, while leading to upregulation of genes involved in membrane permeability and efflux (figure 4; Folkesson *et al.* 2012). This regulator is also associated with one of the specific traits changing in the bacteria during chronic infection that occurs frequently, which is the mucoid phenotype (Jones *et al.* 2010; Rau *et al.* 2010). The mucoid phenotype is caused by an excessive production of alginate, the extracellular polysaccharide commonly produced by *P. aeruginosa* strains, and it is known that the production of extracellular polysaccharides is part of the general envelope stress response of some bacterial species, forming a glycocalyx that conceals the surface of the bacteria and thus providing protection from environmental stresses (Folkesson *et al.* 2012). Mutations in regulatory genes are responsible for this conversion to a mucoid phenotype, including mutations in several positions of the *mucA* gene (Jones *et al.* 2010). This gene encodes the anti-sigma factor MucA that binds to the sigma factor AlgU and prevents the transcription of the *algD* operon, which encodes the enzymes involved in alginate synthesis (Rau *et al.* 2010; Marvig *et al.* 2014). Therefore, mutations that deactivate the function of the MucA protein trigger the activation of AlgU (figure 4) and lead to alginate overproduction, as they cause an excessive transcription of the *algD* operon. Overproduction of this exopolysaccharide in the airways of CF patients has been linked to increased

tolerance toward the host immune response and also to biofilm development (Rau *et al.* 2010; Folkesson *et al.* 2012).

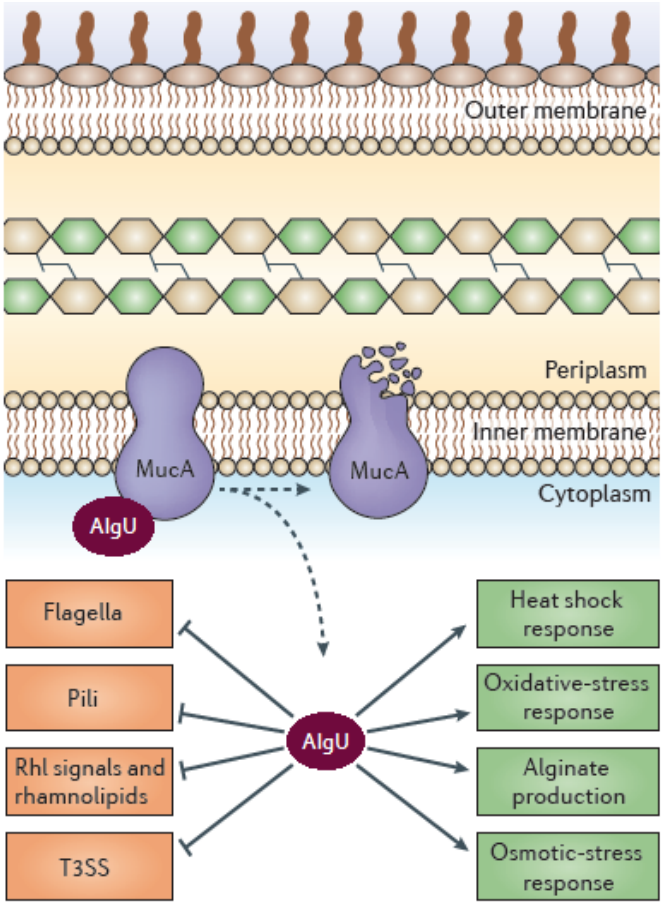

Figure 4 - Regulation network of MucA–AlgU. The function of the sigma factor AlgU is impeded through binding by the anti-sigma factor MucA. Mutations in *mucA* gene lead to release of AlgU, which activates the transcription of genes involved in alginate production and in the responses to heat shock, oxidative stress and osmotic stress, among others. AlgU also negatively regulates several *P. aeruginosa* virulence factors, such as flagella, pili, Rhl quorum sensing signals, Rhl-controlled rhamnolipids and the type III secretion system (T3SS). Adapted from Folkesson *et al.* 2012.

An approach relying on the genome comparison of two *P. aeruginosa* isolates retrieved from a CF patient over a period of infection of 90 months identified 68 mutational differences between the two subsequent isolates (Smith *et al.* 2006). The mutated genes in the isolate recovered later on in the course of infection were related to virulence, antibiotic resistance and iron acquisition, suggesting that these genes may be important in adaptation to the host environment. Moreover, 91 additional isolates longitudinally collected from 29 CF patients were analysed in the study, by sequencing 24 genes and regulatory regions that were mutated in the first patient and 10 genes and regulatory regions that were candidates for mutation as determined in other studies, for each of the 91 isolates. The most frequently mutated gene in this set of isolates was *mexZ*, which is a negative regulator of the *mexX* and *mexY* genes that encode the components of the MexXY-OprM multidrug-efflux pump. Loss-of-function

mutations in the *mexZ* gene lead to increased expression of *mexX* and *mexY*, which in turn is associated with resistance to antibiotics (Smith *et al.* 2006).

In a more recent work, Marvig and colleagues sequenced the genomes of 474 longitudinally collected isolates of *P. aeruginosa* from the airways of young individuals with CF, to study within-host evolution of the different strains sampled (Marvig *et al.* 2014). They identified and compared the genomes of 36 *P. aeruginosa* lineages to determine their evolutionary history, finding convergent molecular evolution in 52 genes, which were associated with remodeling of regulatory networks, acquisition of antibiotic resistance and loss of virulence factors and motility, thus suggesting an important role of the identified genes in adaptation to the host environment. Moreover, they observed that 25 out of 28 *algU* mutations were present in isolates that were also mutated in the *mucA* gene, suggesting that the mucoid phenotype caused by loss-of-function mutation in *mucA* is eliminated by a subsequent mutation in *algU*. They hypothesized that this evolutionary path occurs in CF because even though mucoidy could have immediate benefits, it is an energy-consuming process and thus its regulation could be modified by mutation of the *algU* gene. Another possibility suggested is that the primary selective advantage of mutations in *mucA* could be the activation of stress response mechanisms, since *algU* is a positive regulator of the stress response. According to this hypothesis, which has been considered in other works (Rau *et al.* 2010; Folkesson *et al.* 2012), mucoidy could be a pleiotropic effect of this mutation that is subsequently eliminated by mutation of *algU*. These constrains in the order of appearance of mutations constitute an example of historical contingency, as mutations in downstream transcriptional regulators are contingent upon mutations in upstream regulators. The secondary mutations may follow the initial mutation to rebalance the regulatory network towards the wild-type phenotype, or to facilitate the evolutionary adaptation process.

Along with overproduction of alginate, *P. aeruginosa* isolates sampled from chronically infected individuals usually display antibiotic resistance, lack of motility, slow growth and loss of virulence factors and quorum sensing (Oliver *et al.* 2000; Rau *et al.* 2010; Folkesson *et al.* 2012), which has been designated the ‘chronic infection phenotype’. This phenotypic profile is reported in isolates sampled from different individuals in distinct clinical settings, suggesting that it is an end-point result of parallel evolution of the bacteria in several different patients colonized by different genotypes of *P. aeruginosa* (Rau *et al.* 2010; Folkesson *et al.* 2012).

Frequent mutations found in *P. aeruginosa* isolates from chronic CF-associated infections occur in genes associated with antibiotic resistance (*gyrA, gyrB, mexA, mexB, mexR, mexS, mexZ, nalD, nfxB, oprD*) and in genes associated with regulation of biofilm formation (*lasR,* a quorum sensing regulator; *morA*, a regulator of flagellar development; *retS*, a sensor that regulates motility and transition between type III and VI secretion systems; *bifA*, *wspA* and *wspE,* involved in modulation of c-di-GMP levels) (Marvig *et al.* 2014). The *lasR* gene encodes a transcriptional regulator of quorum sensing and loss-of-function mutations in this gene are often found (Smith *et al.* 2006). Inactivation of LasR reduces the expression of several virulence genes (Folkesson *et al.* 2012). However, this is more likely associated with an alteration of virulence than with a reduction in virulence (Rau *et al.* 2010).

Furthermore, mutations in global regulatory genes are often found in *P. aeruginosa* isolates from patients with CF. The occurrence of mutations in global regulators, and possibly the sequence by which these mutations appear, may be responsible for determining the evolutionary paths undertaken by the bacteria (Folkesson *et al.* 2012). Some combinations of regulatory mutations might direct evolution towards increased adaptive behavior, still further investigations are required to support and complement this premise.

Regarding resistance to antibiotics, wild-type *P. aeruginosa* is naturally resistant to numerous antibiotics and it is generally able to develop tolerance to increased levels of antimicrobial compounds (Rau *et al.* 2010). Before the current antibiotic therapies were introduced, CF patients would succumb to *S. aureus* infections and it was only after the implementation of adequate antibiotic treatments that *P. aeruginosa* began to be considered a relevant pathogen in CF infections (Folkesson *et al.* 2012). Currently, antibiotic treatments are implemented in CF patients as soon as infection with *P. aeruginosa* has been diagnosed (Klepac-Ceraj *et al.* 2010) and bacteria are not intrinsically resistant to the compounds used in these treatments (Rau *et al.* 2010). Thus, persistent colonization and chronic infection are most likely related to a development of resistance to the antibiotics used in the treatment. *P. aeruginosa* isolates from CF patients are known to present resistance to all clinically relevant classes of antibiotics and, although chromosomally encoded mechanisms of resistance are usually involved, imported genes may also play an important role (Folkesson *et al.* 2012). As mentioned before, genes associated to antibiotic resistance are some of the most commonly mutated genes during *P. aeruginosa* adaptation to the CF environment, and continuous treatment with antibiotics plays a major part in shaping the adaptation process by originating antibiotic resistant lineages of *P. aeruginosa* evolving under this selective pressure (Folkesson *et al.* 2012).

With the purpose of studying the existence of a direct path towards the 'chronic infection phenotype', characterized by antibiotic resistance, lack of motility, slow growth and loss of virulence factors, Rau and colleagues collected *P. aeruginosa* isolates from three CF patients in the beginning of their colonization history (Rau *et al.* 2010). Combining transcriptional profiling with phenotypic characterization, they concluded that for two of the patients there had been genetic alterations that led to different phenotypes throughout the course of colonization. In one of those patients, the developed phenotype was increased antibiotic resistance and this was due to a mutation in the gene that encodes the transcriptional regulator NfxB (*nfxB* gene), which caused the upregulation of expression of the MexCD-OprJ multidrug efflux pump. Additionally, some isolates of this patient showed reduced expression of virulence factors and all of them were twitching motility deficient, thus it was clear that the mutations responsible for the phenotypic alterations were shifting the bacteria towards the typical 'chronic infection phenotype'. For another patient, the new phenotype acquired during colonization was bacterial mucoidy, which resulted from the upregulation of enzymes responsible for alginate synthesis, due to a mutation in the *mucA* gene. This mutation had a pleiotropic effect on gene expression and led to the acquisition of several characteristic traits of the 'chronic infection phenotype', such as alginate overproduction and reduced virulence, motility and growth rate.

There seems to be a significantly conserved pattern of evolution underlying this adaptation to the CF airway environment, as several phenotypic traits are recurrently observed and many mutations are consistently found in *P. aeruginosa* lineages evolving independently (Markussen *et al.* 2014). Besides this parallel evolution in different lineages, there is a high level of phenotypic diversity in isolates within each host (Oliver *et al.* 2000; Folkesson *et al.* 2012; Markussen *et al.* 2014). *P. aeruginosa* isolates belonging to the same clonal type often show distinct phenotypic profiles, such as different colony morphology, antibiotic resistance and quorum-sensing regulation, and genome-based studies have indicated that there is genomic diversity among these clonal isolates (Markussen *et al.* 2014). One of the suggested models to elucidate this intra-clonal diversity relies on the structural complexity of the CF airway environment, as it contains multiple compartments that exhibit different conditions (Markussen *et al.* 2014). As a result, population diversity may arise from the evolution of *P. aeruginosa* clones that occupy different niches in the CF airway. Still, other processes could also contribute to this diversity, such as the evolution of variants with high mutation rates and interactions that occur between variant subpopulations (Markussen *et al.* 2014).

Several studies suggest that the need to adapt to a complex and challenging environment would promote the spontaneous emergence of hypermutable bacterial strains (Taddei *et al.* 1997; Oliver *et al.* 2000). According to this premise, *P. aeruginosa* isolates sampled from chronically infected CF patients should have a high frequency of hypermutable strains, as high mutation rates could be beneficial for the appearance and fixation of advantageous mutations for a rapid adaptation to the CF airways environment. In *P. aeruginosa*, hypermutable strains result from mutations in the *mutS* and *mutL* genes, encoding proofreading proteins responsible for correcting errors that may occur during DNA replication (Mena *et al.* 2008). As a changing environment, the CF airways may ensure the fixation of the hypermutable genotype, since the selective pressure on bacteria is always present (Oliver *et al.* 2000). This creates a major therapeutic challenge, as the hypermutable strains are resistant to many antibiotics and, moreover, are more likely to become resistant to new antimicrobials.

In a study combining genome sequencing and phenotypic profiling, Markussen and colleagues demonstrated a rapid diversification of an original CF airway infecting *P. aeruginosa* strain from a chronically infected patient into three sublineages, during a 32-year period of infection (Markussen *et al.* 2014). Each of the subpopulations identified had a distinct functional and genomic profile, displaying differences in mucoidy, doubling time, metabolism and gene expression profiles. Additionally, they showed that one of the sublineages resided in the paranasal sinuses and the other two inhabited the lower airways, possibly occupying different niches in the lungs of the patient. This spatial distribution of the three subpopulations could explain the expression of lineage-specific sets of genes, as different selective pressures would be involved. Thus, this work demonstrated that diversity in *P. aeruginosa* populations from a single individual can have long-term stability, as the distinct sublineages coexisted for decades.

Genome analysis of sequentially sampled isolates from CF patients chronically infected with *P. aeruginosa* combined with phenotypic assessments has provided remarkable insights into the adaptive strategies that these pathogens use to evolve in the CF airways. Several genes have been identified

and correlated with pathogenic phenotypes, such as antibiotic resistance, overproduction of alginate and biofilm formation.

1.3. *Burkholderia cepacia* complex infections in cystic fibrosis

1.3.1. *B. cepacia* complex overview

The *Burkholderia cepacia* complex (Bcc) constitutes a group of at least 20 closely related Gram--negative bacterial species that emerged in the 1980s as opportunistic human pathogens, particularly in patients suffering from cystic fibrosis (Coenye *et al.* 2001; Depoorter *et al.* 2016). The genome of Bcc strains typically consists of three chromosomes and several strains also contain plasmids (Mahenthiralingam *et al.* 2005), resulting in a variable genome size ranging from 7.5 Mbp to 8.5 Mbp (Vandamme & Dawyndt 2011).

Bcc bacteria have the ability to occupy many ecological niches, such as soil, water and the rhizosphere of plants (Parke & Gurian-Sherman 2001; Mahenthiralingam *et al.* 2005), showing a notable capacity to adapt to hostile environments that is mainly attributed to their metabolic complexity and high genomic plasticity (Lessie *et al.* 1996). Although participating in ecologically beneficial interactions with plants (Parke & Gurian-Sherman 2001), Bcc bacteria can produce a wide range of virulence factors and are able to cause infection in plants and humans (Govan *et al.* 2007).

1.3.2. *B. cepacia* complex in the cystic fibrosis airways

Bcc bacteria are considered important opportunistic pathogens, capable of causing severe infections in immunocompromised individuals and in patients with CF (Jones *et al.* 2004; Mahenthiralingam *et al.* 2005). Infection with Bcc may cause a rapid and fatal necrotizing pneumonic illness with the occurrence of high temperatures and respiratory failure known as '*cepacia* syndrome', demonstrating an invasive and systemic infection not seen for other CF-associated pathogens (Isles *et al.* 1984). Although strains from every Bcc species are able to cause infection in CF patients, the two most frequently isolated species are *Burkholderia cenocepacia* and *Burkholderia multivorans* (Jones *et al.* 2004; LiPuma 2010). Infections caused by Bcc have been associated with a worse clinical outcome for these individuals than infections caused by *P. aeruginosa* alone, possibly depending on the species and even individual strains within species (Frangolias *et al.* 1999; Aris *et al.* 2001; Chaparro *et al.* 2001).

Infection with Bcc in CF patients is particularly feared due to the intrinsic resistance of these bacteria to many antimicrobial compounds, to the potential for patient-to-patient transmission and to the highly unpredictable clinical outcome (Govan *et al.* 2007). Moreover, during chronic colonization of CF airways, Bcc bacteria are exposed to changing selective pressures, as a consequence of the host immune defenses, antibiotic therapy and oxygen deprivation (Harrison 2007). In response to this

challenging environment, pathogens undergo genetic adaptations that lead to the emergence of phenotypic variants that have the capacity to establish long-term chronic infections in patients with CF.

After the acknowledgement of Bcc bacteria as highly transmissible among CF patients, infection control measures were adopted in several CF clinics, such as cohorting of infected patients and segregation both from each other and from other patients with CF (Mahenthiralingam *et al.* 2005; Zlosnik *et al.* 2015). Although the implementation of strict segregation policies prevents epidemic outbreaks of Bcc in CF clinics, it evidently has devastating consequences for the quality of life of CF patients, in physical, psychological and social ways (LiPuma 2002; Mahenthiralingam *et al.* 2005). Thus, new strategies to attempt the eradication of Bcc bacteria from the airways of CF patients need to be developed and the study of the mechanisms involved in the adaptive evolution of these pathogens to the CF environment is of great importance in this context.

1.3.3. Genomic and phenotypic evolution of *B. cepacia* complex isolates in cystic fibrosis

Infection of the respiratory tract of CF patients with *B. cepacia* complex bacteria can be transient, but for most patients it leads to a chronic infection that subsequently causes the deterioration of lung function (Govan *et al.* 2007; LiPuma 2010). Despite usually involving one single strain, long-term coinfection with different strains or species and the replacement of the strain that initiates the infection with another throughout the infection may also occur in Bcc chronic infections (Harrison 2007). During these long-term chronic infections, the genome and the phenotypic traits of Bcc bacteria evolve in response to the challenging selective pressures present in the CF airways.

One of the most frequently reported phenotypic alterations in Bcc bacteria is the variation in colony mucoidy associated with exopolysaccharide (EPS) production (Zlosnik *et al.* 2008; Silva *et al.* 2011). Cepacian is the most common exopolysaccharide produced by Bcc strains and is considered a virulence factor of these bacteria, as it has been shown to inhibit neutrophil chemotaxis, to neutralize reactive oxygen species *in vitro* (Bylund *et al.* 2006), to affect the phagocytosis of bacteria by human neutrophils and to facilitate persistent bacterial infection in mice (Conway *et al.* 2004). As high-molecular weight polymers, EPS create a hydrated anionic matrix that surrounds the bacterial cell, protecting it from environmental sources of stress (reviewed in Ferreira *et al.* 2011).

In a comprehensive study involving 560 Bcc isolates sampled from 100 chronically infected CF patients, Zlosnik and colleagues demonstrated that strains from all species of the Bcc can express the mucoid phenotype (Zlosnik *et al.* 2008). In this work, they showed that sequential clonal isolates from CF patients experienced mucoid phenotypic conversion mainly from mucoid to non-mucoid (figure 5), with *Burkholderia multivorans* mucoid to non-mucoid transitions occurring in 9 patients and *Burkholderia cenocepacia* mucoid to non-mucoid transitions occurring in 3 patients. Moreover, in a retrospective clinical review of CF patients attending clinics in Vancouver during a 26-year period, Zlosnik and colleagues observed that patients infected with non-mucoid Bcc showed a more dramatic decline in lung

function than those infected with mucoid bacteria, suggesting that the non-mucoid morphotype is associated with increased disease severity and the mucoid morphotype with persistence in the lungs (Zlosnik *et al.* 2011).

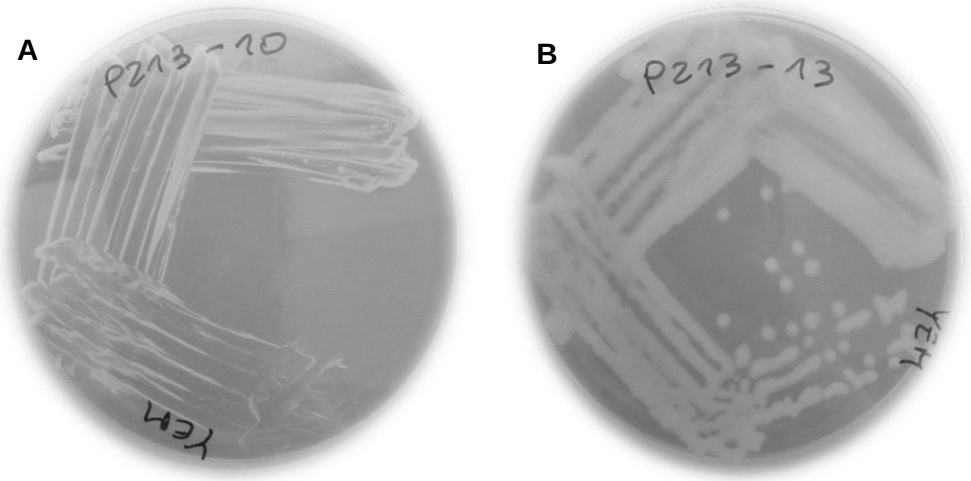

Figure 5 - *Burkholderia multivorans* mucoid phenotypes on yeast extract mannitol medium (YEM) plates: (A) non-mucoid and (B) mucoid.

The production of extracellular polymers by members of the Bcc has been associated with biofilm formation, which constitute communities within which bacteria adopt a sessile lifestyle, being protected from environmental stresses. It was demonstrated that, although not required for biofilm formation, EPS is involved in the development of thick and mature biofilms *in vitro* by Bcc bacteria (Cunha *et al.* 2004). Furthermore, bacteria within biofilms have been shown to be significantly more resistant to antibiotics (Caraher *et al.* 2006). Thus, by enhancing the ability of bacteria to resist host immune responses and antibiotic exposure in the form of biofilms, the production of EPS could improve the survival of Bcc bacteria in the CF lung environment (Cunha *et al.* 2004).

To gain further insights into the importance of the Bcc mucoid variation in the adaptive process to the CF airways, Silva and colleagues studied two *B. multivorans* clonal isolates sampled from a chronically infected CF patient, where a mucoid to non-mucoid transition had occurred (Silva *et al.* 2011). This approach combined transcriptional profiling and phenotypic characterization to analyse the mucoid and non-mucoid isolates and the results revealed that the non-mucoid isolate had a decreased expression of genes encoding proteins linked to virulence-associated characteristics and metabolism. In addition, the non-mucoid isolate displayed lower motility, no exopolysaccharide production, increased biofilm formation and higher survival in minimal medium. The mucoid variant was shown to be more virulent in the acute infection model *Galleria mellonella*, likely associated with the higher expression of genes related to virulence traits in comparison with the non-mucoid morphotype. The reduced expression of virulence traits in the non-mucoid variant may enhance the ability of bacteria to escape from the host immune system, developing an increased fitness to persist in the CF lung (Silva *et al.* 2011). In a subsequent study, the stability of the same two isolates after exposure to stress conditions

was investigated and the results revealed that mucoid to non-mucoid transitions in Bcc bacteria may occur under stress conditions such as subinhibitory concentrations of antibiotics and oxidative and osmotic stress (Silva *et al.* 2013). The role of EPS production in Bcc virulence is still poorly understood, with some evidences pointing to EPS as a virulence factor and others pointing to a role in persistence in the CF lung (Ferreira *et al.* 2011). Therefore, further efforts need to be put into the study of this phenotypic trait and how it relates to virulence and adaptation of Bcc strains to the CF environment. Besides EPS production, several other phenotypic traits are associated with Bcc persistence and pathogenicity in the CF airways, such as resistance to antibiotics, biofilm formation, and motility (Mahenthiralingam *et al.* 2005).

Another relevant phenotypic characteristic of Bcc is the expression of O-antigen repeats of lipopolysaccharide (LPS), as several isolates recovered from chronically infected CF patients lack the O-antigen (reviewed in Mahenthiralingam *et al.* 2005). LPS is considered a virulence factor produced by Bcc bacteria, as it may inhibit phagocytosis and reduce bacterial elimination (Saldías *et al.* 2009). Nevertheless, the O-antigen can interfere with adhesion to bronchial epithelial cells and stimulate the host immune response, which indicates that its loss could be beneficial to evade the host immune response and also to cells living in the form of biofilms, like in most CF airways (Saldías *et al.* 2009; Silva *et al.* 2016).

In a phenotypic assessment of 11 *B. cenocepacia* sequential clonal isolates sampled from a CF patient chronically infected for 3.5 years, Coutinho and colleagues observed that the phenotypic traits of the isolate believed to have initiated the infection were different from those of the isolates retrieved during the course of infection (Coutinho *et al.* 2011a), which is in agreement with the hypothesis that Bcc bacteria undergo several adaptive changes throughout the colonization of CF lungs. In comparison with the first *B. cenocepacia* isolate, the subsequent isolates showed increased antimicrobial resistance against different classes of antibiotics, decreased swarming motility and increased ability to grow under iron limiting conditions. Moreover, the authors described the reduction of the degree of fatty acid saturation as an adaptive response to growth under severe oxygen-limiting conditions. Complementing this study with a transcriptomic analysis, the genomic expression levels of the first *B. cenocepacia* isolate and of a subsequent isolate retrieved from the same CF patient were compared (Mira *et al.* 2011). Approximately 1000 genes were differently expressed by the two clonal isolates, indicating a significant alteration in gene expression. Among the genes that were upregulated in the isolate obtained further on during the course of infection were genes involved in efflux of drugs, iron uptake and assembly of flagella, the latter ones possibly playing a role in adhesion to epithelial cells.

In addition to *B. cenocepacia* and *B. multivorans*, which are the most commonly isolated species from the airways of individuals with CF (LiPuma 2010), studies on *Burkholderia dolosa* have also been conducted. Although rarely retrieved from CF patients, this bacterium has been linked to accelerated decline in lung function and decreased survival (Kalish *et al.* 2006). In a retrospective study of 112 *B. dolosa* isolates recovered from 14 CF patients over a 16-year period, Lieberman and colleagues sequenced the genomes of the isolated opportunistic pathogens and identified 17 genes that acquired mutations in isolates retrieved from several patients, revealing parallel adaptive evolution (Lieberman *et*

al. 2011). The mutated genes were associated with pathogenic phenotypes, such as antibiotic resistance and membrane composition, including LPS biosynthesis. In addition, one of the most mutated genes was a homolog of *fixL*, which is a global regulator involved in biofilm formation, motility and virulence (Schaefers *et al.* 2017), and homologs of two other genes related to this process (*fixJ* and *fnr*) were also mutated. In a more recent study, samples retrieved from five of the same CF patients were examined regarding intraspecies diversity and the authors found extensive intrasample diversity, leading to the hypothesis that mutations rarely become fixed in a population (Lieberman *et al.* 2014). Instead, several different lineages may coexist for many years in the CF airways of one patient, with genes associated with outer membrane composition, antibiotic resistance and iron scavenging shown to be under strong selective pressure (Lieberman *et al.* 2014). Furthermore, a phenotypic assessment of 14 sequential *B. dolosa* isolates retrieved from a chronically infected CF patient over 5.5 years showed that, in comparison with the first isolate, the subsequently retrieved isolates displayed increased resistance to antibiotics and decreased swarming motility (Moreira *et al.* 2014), indicating an adaptive evolution throughout the course of infection.

In a recent work that combined genome sequencing technologies and phenotypic profiling of 22 *B. multivorans* isolates retrieved from a CF patient during a 20-year infection, several genes and pathways that were likely under strong selection within the host were identified (Silva *et al.* 2016). The genomes of the 22 isolates were sequenced and mutations that were found to have evolved independently in multiple isolates affected genes encoding regulatory proteins, proteins associated with envelope biogenesis and enzymes involved in lipid and amino acid metabolism. It was observed that multiple lineages coexisted in this patient's airways for years, evolving at different rates in the course of infection. In the lineage that became dominant during the infection, multiple mutations were found in a *fixL* homolog, which had been previously found in the study of *B. dolosa* isolates by Lieberman and colleagues (Lieberman *et al.* 2011) and regulates biofilm formation, motility and virulence (Schaefers *et al.* 2017). Other genes found to be under strong positive selection were *rpfR* (encoding a diguanylate cyclase/phosphodiesterase with regulatory functions), *plsX* (involved in the biosynthesis of phospholipids) and an *ompR*-like gene (OmpR is an osmolarity response regulator). Moreover, the phenotypic analysis showed increased antibiotic resistance and biofilm formation, and decreased motility and growth rate throughout the evolutionary trajectory of the isolates, shifting the bacteria toward a 'chronic infection phenotype', which was associated with the period of accelerated lung function decline (Silva *et al.* 2016).

More recently, a study combining whole-genome sequencing and phenotypic assays on 215 *B. cenocepacia* isolates sampled from 16 CF patients over a span of 2 to 20 years of infection was conducted (Lee *et al.* 2017). The tested phenotypes were growth in liquid media, motility, biofilm formation, mucoidy and acute virulence in *Galleria mellonella*, and the results showed extensive phenotypic variation among isolates within patient bacterial series, with progressive changes in several series leading to decreased motility, growth rate and acute virulence, while for biofilm formation there were two patient series progressively increasing this trait and three decreasing it. Alongside the phenotypic assessments, the genomic analysis provided insights into genetic variations that could

underlie phenotypic changes, with the identification of four candidate genes that are associated with motility and biofilm formation (*dnaK*, *papC*, *gcvA* and *qseC*). It was also observed that strains isolated on the same date often belonged to distinct phylogenetic clades, supporting the coexistence of multiple strains in the infection. Moreover, genome reduction was observed in seven patient series, with two *B. cenocepacia* strains having lost chromosome 3 completely. Genome reduction by large deletions of genes encoding nonessential traits had been previously reported in studies on long-term bacterial evolution (Price *et al.* 2013; Rau *et al.* 2012; Silva *et al.* 2016).

Bacteria undergo genetic and phenotypic changes when in need to adapt to new and challenging environments, leading to improved survival ability and fitness. Genomic and phenotypic characterizations of longitudinally collected Bcc isolates from CF patients with long-term chronic infections have greatly contributed to the identification of evolutionary strategies underlying the adaptive processes that take place during bacterial colonization of the CF airways. Among the reported adaptation-related traits of Bcc bacteria in this complex environment are increased antibiotic resistance and biofilm formation and decreased motility and growth rate. This knowledge may point to potential therapeutic targets for the development of improved treatments against Bcc bacteria.

1.4. Objectives

The aim of this work is to analyse the genomic and phenotypic evolution of 37 *B. multivorans* isolates that were sequentially retrieved from two infected CF patients (patients P213 and P426) over 15 to 18 years of chronic infection. The approach comprises the identification of important mutations among the isolates sampled from each patient and the assessment of phenotypes associated with bacterial adaptation to the CF airways, such as EPS production, LPS O-antigen production, resistance to antibiotics, biofilm formation, motility, adhesion to epithelial cells, growth rate and acute virulence in *Galleria mellonella*. Insights into the adaptive strategies that drive within-host bacterial evolution during colonization of the CF airways are expected to lead to improved therapies against infections caused by Bcc bacteria.

2. MATERIALS AND METHODS

2.1. Bacterial strains and growth conditions

Bacterial isolates were sampled from two cystic fibrosis patients (patients P213 and P426) attending a clinic in Vancouver, Canada, and were provided by Prof. David P. Speert from the University of British Columbia. 16 bacterial isolates consist of a single clone of *Burkholderia multivorans* sampled from patient P213 at certain time points in the period between 1996 and 2010 (table 1) and 21 isolates consist of a single clone of *B. multivorans* sampled from patient P426 at certain time points in the period between 1997 and 2014 (table 2). Isolates were grown in lysogeny broth (LB medium) or in extracellular polymeric substance (EPS)-producing salt-mannitol medium (SM medium) at 37ºC with agitation at 250 rpm. SM medium is a minimal medium with a high carbon to nitrogen ratio to stimulate the production of EPS: 12.5 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 1 g/L K_2SO_4, 1 g/L NaCl, 0.2 g/L $MgSO_4.7H_2O$, 0.001 g/L $CaCl_2.2H_2O$, 0.001 g/L $FeSO_4.7H_2O$, 1 g/L casaminoacids, 1 g/L yeast extract, 20 g/L D-mannitol.

Table 1 - Date of isolation and ID of each *B. multivorans* isolate recovered from CF patient P213.

Isolate	ID	Date of isolation
P213-1	VC7495	13-02-1996
P213-2	VC7704	04-06-1996
P213-3	VC7804	13-08-1996
P213-4	VC8896	06-05-1998
P213-5	VC9452	20-01-1999
P213-6	VC10037	18-01-2000
P213-7	VC10458	07-11-2000
P213-8	VC11002	23-10-2001
P213-9	VC11446	23-07-2002
P213-10	VC12063	26-09-2003
P213-11	VC12556	30-09-2004
P213-12	VC13750	11-04-2007
P213-13	VC13748	11-04-2007
P213-14	VC13833	20-06-2007
P213-15	VC13834	20-06-2007
P213-16	VC15071	11-01-2010

Table 2 - Date of isolation and ID of each *B. multivorans* isolate recovered from CF patient P426.

Isolate	ID	Date of isolation
P426-1	VC8086	11-02-1997
P426-2	VC8136	17-03-1997
P426-3	VC8585	20-01-1998
P426-4	VC9177	25-09-1998
P426-5	VC9783	12-08-1999
P426-6	VC10411	12-10-2000
P426-7	VC11369	24-05-2002
P426-8	VC11982	17-07-2003

Table 2 - Date of isolation and ID of each *B. multivorans* isolate recovered from CF patient P426 (cont.).

Isolate	ID	Date of isolation
P426-9	VC12458	22-07-2004
P426-10	VC13030	22-09-2005
P426-11	VC13616	12-01-2007
P426-12	VC13617	12-01-2007
P426-13	VC13777	02-05-2007
P426-14	VC13778	02-05-2007
P426-15	VC14363	06-07-2008
P426-16	VC14863	11-08-2009
P426-17	VC15667	23-06-2011
P426-18	VC15867	17-11-2011
P426-19	VC16979	04-12-2013
P426-20	VC17389	22-09-2014
P426-21	VC17390	22-09-2014

2.2. Analysis of lipopolysaccharide

Cells from overnight cultures were first suspended in 1x PBS and then suspended in 100 µL of lysis buffer (2% SDS, 4% β-mercaptoethanol, 0.5M TrisCl pH 6.8), followed by proteinase K digestion for at least 1 hour at 60°C. LPS was purified by adding 150 µL of a 90% phenol solution and then 10 volumes of ethylic ether. LPS samples were resolved by electrophoresis in 16% polyacrylamide gels with a Tricine-SDS system (Lesse *et al.* 1990), followed by silver staining.

2.3. Exopolysaccharide production

2.3.1. Liquid medium

The amount of exopolysaccharide produced was measured based on the dry weight of the ethanol-precipitated polysaccharide recovered from 50 mL cultures grown in EPS-producing medium (SM medium) for 6 days at 37°C with agitation at 250 rpm (Silva *et al.* 2011). The 50 mL cell suspensions were centrifuged for 15 minutes at 9000 rpm at 20°C and 150 mL of 96% ethanol was added to the supernatant to precipitate the polysaccharide. Results are the means of data from three replicates of three independent experiments with duplicate flasks.

2.3.2. Solid medium

Mucoidy was assessed on yeast extract mannitol medium (YEM) containing 4 g/L mannitol, 0.5 g/L yeast extract and 15 g/L agar (Zlosnik *et al.* 2008). After inoculation, YEM plates were incubated for 48 h at 37°C and mucoidy was then observed by visual inspection.

2.4. Antimicrobial susceptibility

To evaluate the susceptibility of isolates against paper discs containing antibiotics, the agar disc diffusion method (Bauer *et al.* 1966) was used. Müller-Hinton agar (Sigma-Aldrich) plates were

inoculated with 100 µl of a suspension at an OD_{640} of 0.1 prepared from exponential-phase cells growing on LB medium at 37ºC. Paper discs (BD BBL Sensi-Disc) containing piperacillin (100 µg) plus tazobactam (10 µg), kanamycin (30 µg), ciprofloxacin (5 µg) or aztreonam (30 µg) were applied onto the surfaces of the inoculated plates. The diameter of the growth inhibition zone was measured after 24 h of incubation at 37°C. Results are the means of data from at least six replicates of three independent experiments, each with three discs per isolate.

2.5. Biofilm formation

Bacteria were grown in LB medium at 37°C to mid-exponential phase and diluted to an OD_{640} of 0.05. 200 µL samples of the cell suspensions were used to inoculate 96-well polystyrene microtiter plates. Plates were incubated at 37°C statically for 48 h, after which the wells were washed three times with 0.9% (wt/vol) NaCl. The biofilm was stained with 200 µL of a 1% (wt/vol) crystal violet solution for 20 minutes at room temperature (Ferreira *et al.* 2007), followed by washing three times with 200 µL of 0.9% NaCl. The dye was then solubilized with 200 µL of 96% ethanol and the solution's absorbance at 590 nm (A_{590}) was measured in a microplate reader (Spectrostar nano, BMG LabTech). Results are the means of data from at least eight replicates of three independent experiments.

2.6. Motility

2.6.1. Swimming motility

The swimming motility was assessed on swimming agar plates: 1% (wt/vol) tryptone, 0.5% (wt/vol) NaCl, 0.3% (wt/vol) noble agar (Difco) (Silva *et al.* 2016). Plates were spot inoculated with a 5 µl drop of a culture at an OD_{640} of 1.0 and were incubated for 24 h at 37°C. The diameter of the swimming zone was then measured. Results are the means of data from at least ten replicates of three independent experiments.

2.6.2. Swarming motility

The swarming motility was assessed on swarming agar plates: 0.04% (wt/vol) tryptone, 0.01% (wt/vol) yeast extract, 0.0067% (wt/vol) $CaCl_2$, 0.6% (wt/vol) bacto agar (Difco) (Silva *et al.* 2016). Plates were spot inoculated with a 5 µl drop of a culture at an OD_{640} of 1.0 and were incubated for 48 h at 37°C. The diameter of the swarming zone was then measured. Results are the means of data from at least ten replicates of three independent experiments.

2.7. Growth rate and doubling time determination

Isolates were grown overnight in LB medium at 37°C. A volume was then centrifuged (2 min, 8 rpm), and the pellet was washed with saline solution (0.9% NaCl) and used to inoculate a flask with 50 mL of liquid synthetic cystic fibrosis medium (SCFM) (Palmer *et al*. 2007), generating an initial optical density at 640 nm (OD_{640}) of 0.1. Flasks were incubated at 37°C with agitation at 250 rpm, and growth rates were measured by monitoring the OD_{640} for 24 h. Growth rates were calculated from the exponential phase of growth from two independent experiments.

2.8. Virulence determination in *Galleria mellonella*

Killing assays were performed as described previously (Seed & Dennis 2008). *Galleria mellonella* larvae were injected with cell suspensions containing a total CFU of approximately $1x10^6$ in 10 mM $MgSO_4$ with 1.2 mg mL^{-1} ampicillin, and incubated at 37°C. Survival rates were assessed during the following 3 days post-infection. As a negative control, 10 mM $MgSO_4$ with 1.2 mg mL^{-1} ampicillin was used. Triplicates of ten larvae were used in each experiment.

2.9. Adhesion to epithelial cells

The *B. multivorans* isolates under study were analysed for adhesion to the bronchial epithelial cell line CFBE41o$^-$, derived from a patient homozygous for the cystic fibrosis transmembrane conductance regulator F508del mutation (Goncz *et al.* 1999). Host cell attachment was performed as described before (Ferreira *et al.* 2015). Bacterial strains grown in SM medium for 4 hours were used to infect epithelial cells at a multiplicity of infection (MOI) of 10 (10 bacterial cells to 1 epithelial cell). Bacteria were applied to a 24-well plate previously seeded with CFBE41o$^-$ cells in minimal essential medium supplemented with 10% (vol/vol) FBS and 1% (vol/vol) L-glutamine, and the plates were centrifuged at 700 x g for 5 min. The plates were then incubated for 30 min at 37°C in an atmosphere of 5% CO_2. Afterwards, each well was washed three times with phosphate-buffered saline (PBS) to remove unbound bacteria and cells were lysed with lysis buffer (0.01 M PBS, 10 mM EDTA, 0.25% [vol/vol] Triton X-100; pH 7.4) for 20 min at 4°C. Serial dilutions were plated onto LB agar and adhesion was quantified by determining CFU counts after 48h of incubation at 37°C. Duplicates with each strain were performed per assay and the results presented were obtained from three independent experiments. Results are shown as the percentage of adhesion, which was calculated as the number of CFU recovered divided by the number of CFU applied to the epithelial cells multiplied by 100.

2.10. Detection of single-nucleotide polymorphism (SNP) and indel mutations

Genomic DNA from the *B. multivorans* clinical isolates was extracted and purified using the DNeasy blood and tissue kit (Qiagen). DNA samples were processed according to Illumina's instructions for generating paired-end libraries of the 37 *B. multivorans* isolates, which were sequenced at least to 38-fold coverage using an Illumina MiSeq system at the Instituto Gulbenkian de Ciência (Portugal). The genome of the first bacterial isolate of each patient was also sequenced using a PacBio system at Icahn School of Medicine at Mount Sinai, New York (USA) and polished using the reads generated by the Illumina MiSeq sequencing system, after *de novo* assembly. To obtain these corrected assemblies, the program MeDuSa v1.6 (Bosi *et al.* 2015) was used to align the assemblies to a reference (*Burkholderia multivorans* ATCC 17616) to reorder contigs; all the Illumina reads were trimmed using Trimmomatic (Bolger *et al.* 2014); the trimmed reads were aligned to MeDuSa-assemblies with BWA v0.7.15 (Li & Durbin 2009); the mapped reads were used to predict and correct misassemblies with Pilon v1.19 (Walker *et al.* 2014); and genes were predicted and annotated using Prokka v1.11 (Seemann 2017). This polished assembly resulted in 5 contigs for the genome of the first isolate retrieved from patient P213 and 7 contigs for the genome of the first isolate retrieved from patient P426, which were used as the reference genomes for the subsequent detection of SNP and indel mutations. To identify which

contig corresponded to each *B. multivorans* chromosome, Basic Local Alignment Search Tool (BLAST, Altschul *et al.* 1990) was used to align each contig to the genome of *B. multivorans* ATCC 17616. A standard nucleotide BLAST optimized for highly similar sequences (megablast) was used. To detect SNP and indel mutations, raw paired-end reads from the 37 *B. multivorans* isolates generated with Illumina MiSeq were filtered based on Phred quality scores, followed by trimming off of adapter contamination and ambiguous nucleotides using the fastq-mcf tool (Aronesty 2013). Only reads with Phred scores higher than 30 (99.9% accuracy) were considered for further analysis, while maintaining at least 70% of the initial coverage. The resulting filtered paired-end data sets were aligned against the reference genome sequences of the first *B. multivorans* isolates of each patient (16 data sets corresponding to the isolates of patient P213 were aligned to the polished genome sequence of the first isolate, P213-1; and 21 data sets corresponding to the isolates of patient P426 were aligned to the polished genome sequence of the first isolate, P426-1) using BWA-MEM (Li 2013) and NovoAlign v.3.02.13 (Novocraft; website last accessed January 2017) to determine a first list of putative mutations. Only SNPs and indels detected in alignments using both BWA-MEM and NovoAlign and occurring in at least three forward and three reverse reads were further analysed. Additionally, a visual inspection of the alignments using Geneious v.6.1.8 (Kearse *et al.* 2012) allowed the confirmation of each indel mutation and the identification of large deletions that were also included in the final lists of mutations.

2.11. Phylogenetic analyses

The evolutionary history of the *B. multivorans* isolates retrieved from each CF patient under study was inferred by using the Maximum Likelihood method based on the Tamura-Nei model (Tamura & Nei 1993). Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. There was a total of 229 positions in the final dataset for patient P213 and 249 positions for patient P426. Evolutionary analyses were conducted in Molecular Evolutionary Genetics Analysis (MEGA) software v.7.0.26 (Kumar *et al.* 2016).

2.12. Statistical analyses

The statistical significance of differences in the data was determined using the one-way analysis of variance (ANOVA) followed by Dunnett's multiple comparisons test, which were performed using GraphPad Prism software v.7.03. Kaplan-Meier survival curves were also performed with GraphPad Prism software v.7.03. Differences were considered statistically significant for *P*-values lower than 0.05.

3. RESULTS

3.1. Analysis of *B. multivorans* isolates retrieved from CF patient P213

3.1.1. *B. multivorans* isolates from CF patient P213

The 16 bacterial isolates consist of a single clone of *Burkholderia multivorans* sampled from cystic fibrosis patient P213 at certain time points in the period between 1996 and 2010 (figure 6). The first three isolates and isolates 6 and 7 were recovered in different time points of years 1996 and 2000, respectively. Isolate pairs 12/13 and 14/15 were recovered in two different time points of 2007, each pair in one time point.

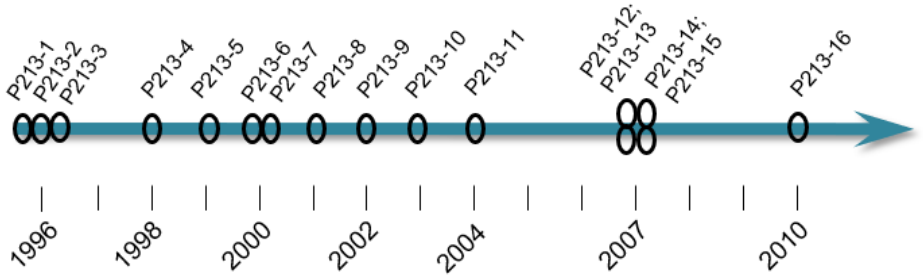

Figure 6 - *B. multivorans* isolates recovered from CF patient P213.

3.1.2. Analysis of single-nucleotide polymorphism and indel mutations

In order to identify genetic mutations among the clinical isolates, the genomes of the 16 *B. multivorans* isolates were sequenced and analysed. The genome of the first isolate was used as the reference genome for the subsequent steps of this analysis, after sequencing and *de novo* assembly. The assembly of the genome of the first isolate generated 5 contigs with lengths between 813 and 3298492 bp (table 3). To identify which contig corresponded to each *B. multivorans* chromosome, Basic Local Alignment Search Tool (BLAST, Altschul *et al.* 1990) was used to align each contig to the genome of *B. multivorans* ATCC 17616. Contigs 1, 2 and 3 correspond to chromosomes 1, 3 and 2, respectively, and contig 4 also aligned to chromosome 1. Contig 5 did not map onto the reference genome of *B. multivorans* ATCC 17616 and is likely a plasmid.

Single-nucleotide polymorphism (SNP) and indel mutations were identified by aligning sequence reads for each isolate against the genome sequence of the first isolate. Only mutations detected in alignments using two alignment tools and occurring in at least three forward and reverse reads were subsequently analysed. The resulting list of SNPs and indels showed mutations existing among the latter 15 *B. multivorans* isolates when compared to the genome of the first isolate retrieved from the patient. In addition, a visual inspection of the alignments led to the identification of large deletions that were also included in this study. In this analysis, a total of 291 mutations was identified, comprising 240 SNPs, 35 indels and 16 large deletions that were detected among the isolates recovered from patient P213 over 15 years of chronic colonization (supplementary table A1).

Table 3 - Lengths of the 5 contigs generated from the assembly of the genome of the first *B. multivorans* isolate recovered from patient P213 and lengths of the corresponding chromosomes of *B. multivorans* ATCC 17616.

P213-1 contig	P213-1 contig length (bp)	*B. multivorans* ATCC 17616 chromosome	*B. multivorans* ATCC 17616 chromosome length (bp)
1	3 298 492	1	3 448 421
2	706 259	3	919 805
3	2 464 683	2	2 473 162
4	813	1	-
5	28 489	-	-

3.1.2.1. Identification of polymorphic genes

In this analysis, ten genes were found to have two different mutations and two genes had three mutations, all within the coding sequence (table 4). The polymorphic genes with three mutations encode a Fis family transcriptional regulator conserved in the *Burkholderia* genus (gene PROKKA_02577) and a TetR family transcriptional regulator (gene PROKKA_02681). The TetR homolog in *B. vietnamiensis* (*amrR*) has been demonstrated to be involved in aminoglycoside resistance (Jassem *et al.* 2014). Genes *accA* and *accD*, encoding acetyl-coA carboxylase subunits alpha and beta, respectively, were also polymorphic, possibly affecting fatty acid biosynthesis. Moreover, two different mutations were identified in a gene encoding a methyl-accepting chemotaxis protein, which may be involved in motility and biofilm formation. Other polymorphic genes are *murF* and *dacB*, involved in peptidoglycan biosynthesis; gene *nuoF*, involved in oxidative phosphorylation; gene *fhuA*, involved in iron metabolism; and gene PROKKA_03885, which encodes a highly repetitive protein and has two different mutations in all isolates except the first three.

Table 4 - Polymorphic genes identified among the 16 *B. multivorans* isolates recovered from CF patient P213.

Contig	Position	Mutation	Locus tag	Gene name	Annotation	Homolog in *B. multivorans* ATCC 17616	Effect of mutation	Type of mutation	Isolates with mutation
contig 1	866409	GTC>ATC	PROKKA_00810	*murF*	UDP-N-acetylmuramoyl-tripeptide-D-alanyl-D- alanine ligase	Bmul_2840	nonsyn	SNP	2
contig 1	866580	CTC>ATC					nonsyn	SNP	8
contig 1	1764172	-12cgacg cgctcga	PROKKA_01653	*dacB*	D-alanyl-D-alanine carboxypeptidase DacB precursor	Bmul_0679	4 aa deletion	indel	12
contig 1	1764306	+1t					frameshift	indel	5
contig 1	2152189	GGC>GCC	PROKKA_02018	*nuoF*	NADH-quinone oxidoreductase subunit F	Bmul_1033	nonsyn	SNP	11
contig 1	2152190	GGC>GGA					syn	SNP	11
contig 1	2341503	GAT>GAA	PROKKA_02189	*accA*	acetyl-CoA carboxylase carboxyltransferase subunit alpha	Bmul_1205	nonsyn	SNP	4-16
contig 1	2341659	GAG>GAA					syn	SNP	8
contig 1	2542874	CCG>CCA	PROKKA_02387	-	major facilitator transporter	Bmul_1342	syn	SNP	4-16
contig 1	2543389	TGC>CGC					nonsyn	SNP	8
contig 1	2755251	GCT>GCG	PROKKA_02577	*fis*	Fis family transcriptional regulator	Bmul_1516	syn	SNP	14
contig 1	2755284	GCA>GCG					syn	SNP	8
contig 1	2755289	CGA>CCA					nonsyn	SNP	11

Table 4 - Polymorphic genes identified among the 16 *B. multivorans* isolates recovered from CF patient P213 (cont.).

Contig	Position	Mutation	Locus tag	Gene name	Annotation	Homolog in *B. multivorans* ATCC 17616	Effect of mutation	Type of muta-tion	Isolates with mutation
contig 1	2837644	TGA>CGA	PROKKA_02659	*fhuA*	TonB-dependent siderophore receptor	Bmul_1594	nonsyn	SNP	10;15
contig 1	2838056	CGG>CGA					syn	SNP	10;15
contig 1	2878065	GCT>GCA	PROKKA_02681	*amrR*	TetR family transcriptional regulator	Bmul_1617	syn	SNP	6;7; 9-11;13-16
contig 1	2878400	-2ac					frameshift	indel	12
contig 1	2878405	+6atttgt					2 aa insertion	indel	12
contig 2	98034	CAT>CGT	PROKKA_03146	-	Sensor histidine kinase	Bmul_5359	nonsyn	SNP	11
contig 2	98413	CTC>TTC					nonsyn	SNP	8
contig 3	128522	GAC>GAT	PROKKA_03779	*mcpC*	Methyl-accepting chemotaxis protein McpC	Bmul_4946	syn	SNP	5-7;9-16
contig 3	128525	CAC>CAT					syn	SNP	4-16
contig 3	246465	GCC>TCC	PROKKA_03885	*bag*	IgA FC receptor precursor	-	nonsyn	SNP	4-16
contig 3	246467	CGT>CTT					nonsyn	SNP	4-16
contig 3	2374729	ATT>ATC	PROKKA_05730	*accD*	Acetyl-CoA carboxylase subunit beta	Bmul_4619	syn	SNP	4
contig 3	2374762	CGA>CGC					syn	SNP	5-7;9-16

3.1.2.2. Mutation rate determination

A representation of the SNPs identified per isolate over time since the first isolate was recovered (figure 7) allowed the identification of two distinct patterns. First, isolates 1 to 7 and 9 to 11 seem to acquire mutations at a constant rate of about 8.2 SNPs/year. Second, the set of the latter 5 isolates recovered after 12 and 14 years since the first isolate showed significantly less SNPs than expected if following the linear tendency of the first set of isolates. Isolate 8 showed the highest number of mutations, with 75 SNPs, 16 indels and 7 large deletions, and also did not follow the linear tendency described for the first set of isolates. One important achievement of this analysis was the identification of a 48-bp deletion in the mismatch repair gene *mutL* for isolate 8, which could explain the increased number of SNPs, as mutations in this gene lead to a hypermutator phenotype.

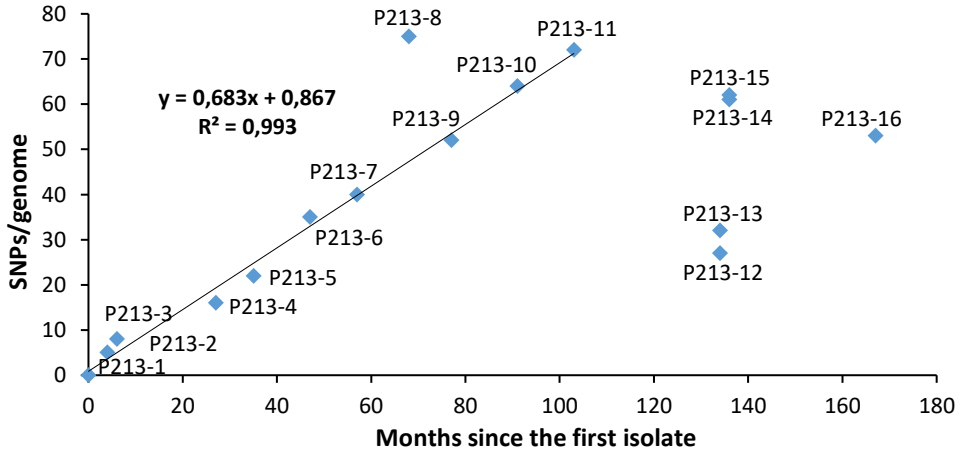

Figure 7 - Number of single-nucleotide polymorphisms (SNPs) distinguishing each *B. multivorans* isolate retrieved from patient P213 from the first isolate over time. A linear fit for isolates 1 to 7 and 9 to 11 is shown.

3.1.2.3. Analysis of large deletions

A visual inspection of genome alignments led to the identification of large deletions with sizes between 0.148 and 151.1 kb (figure 8; table 5). Similar studies with *B. multivorans*, *B. cenocepacia* and *P. aeruginosa* have also reported this loss of large genomic portions in clinical isolates recovered from CF patients (Lee *et al.* 2017; Rau *et al.* 2012; Silva *et al.* 2016).

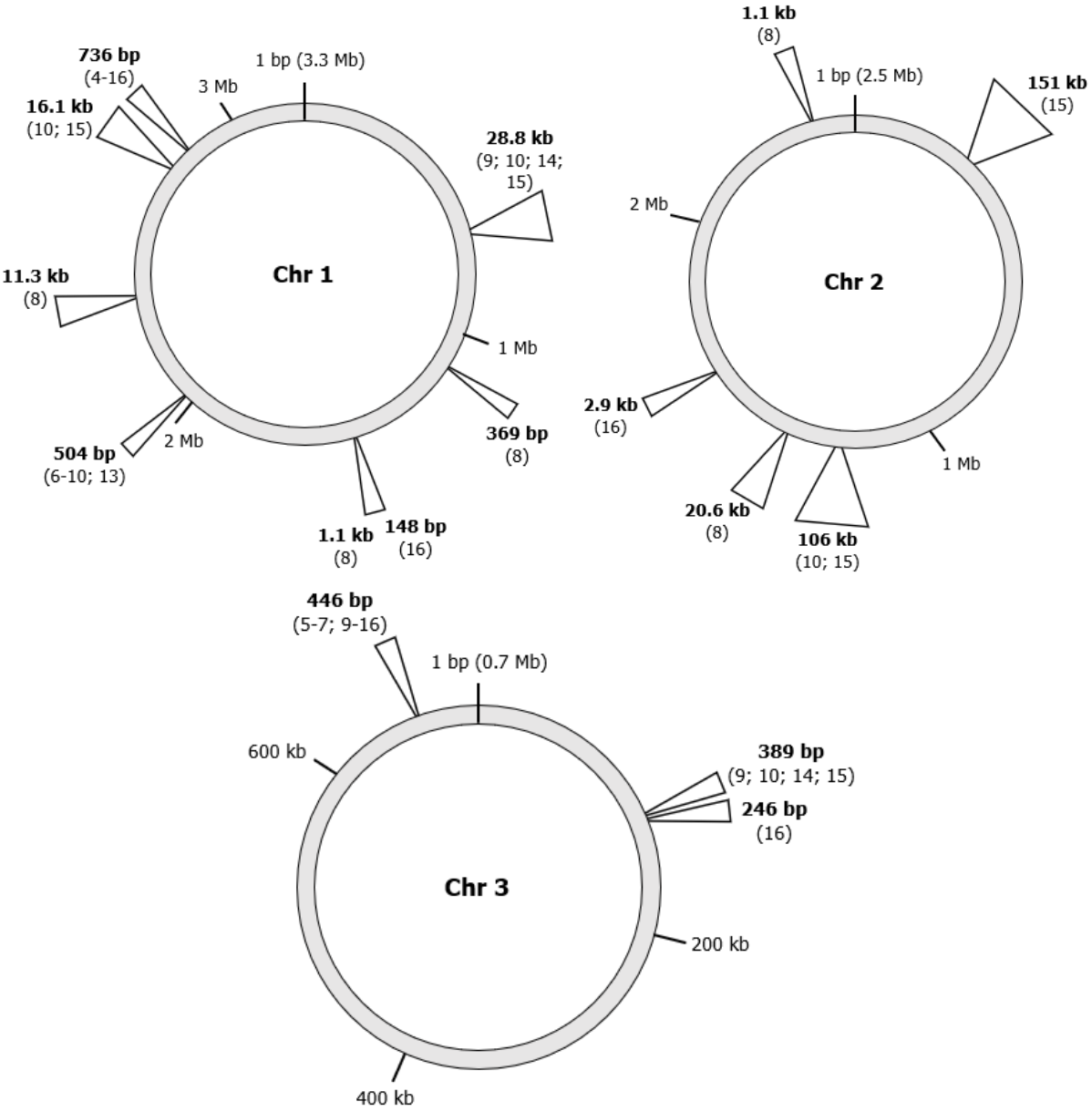

Figure 8 - Schematic representation of large deletions identified among the 16 *B. multivorans* isolates recovered from CF patient P213, divided by chromosomes 1 (Chr1), 2 (Chr2) and 3 (Chr3). Deletions are represented by triangles; isolates with each deletion and its size are also shown. Chromosome representation is not at scale.

An examination of the genes included in these deletions enabled the identification of several protein functions that might be affected in mutated isolates. In chromosome 1, a 504-bp deletion in a gene encoding the DNA translocase FtsK (PROKKA_01941) is present in isolates 6 to 10 and 13, possibly affecting cell division. Further, a 28.8-kb deletion in isolates 9, 10, 14 and 15 comprises genes

involved in type VI secretion, with possible effects in virulence, and a 16.1-kb deletion in isolates 10 and 15 likely influences iron uptake, as it includes genes involved in siderophore biosynthesis. In chromosome 2, two large deletions include genes encoding important regulators of quorum sensing. Deleted in isolate 15, genes PROKKA_03962 and PROKKA_03963 encode the diguanylate cyclase/phosphodiesterase RpfR, the receptor of the *Burkholderia* diffusible signal factor (BDSF), that regulates biofilm formation, motility and virulence (Deng *et al.* 2012; Ryan *et al.* 2009). Deleted in isolate 8, gene PROKKA_04943 encodes the CepR regulator of the CepRI quorum sensing system, which also regulates biofilm formation, motility and production of extracellular virulence factors (Huber *et al.* 2001; Lewenza *et al.* 1999). Moreover, a 106-kb deletion comprising genes involved in type III secretion is present in isolates 10 and 15, in which type VI secretion genes are also deleted from chromosome 1, likely determining changes in virulence in these isolates.

Table 5 - Large deletions identified among the 16 *B. multivorans* isolates recovered from CF patient P213.

Chromosome	Start position	Deletion size (kb)	Locus tag	Isolates with mutation	Annotation / affected protein functions
Chr 1	763710	28.8	PROKKA_00718 to PROKKA_00742	9; 10; 14; 15	Type VI secretion; chaperone ClpB
Chr 1	1144822	0.369	PROKKA_01072	8	LysR transcriptional regulator
Chr 1	1576769	1.1	PROKKA_01479; PROKKA_01480	8	Hypothetical proteins
Chr 1	1577258	0.148	PROKKA_01479	16	Hypothetical protein
Chr 1	2073572	0.504	PROKKA_01941	6-10; 13	DNA translocase FtsK
Chr 1	2444332	11.3	PROKKA_02289 to PROKKA_02299	8	Hypothetical proteins
Chr 1	2839612	16.1	PROKKA_02660 to PROKKA_02663	10; 15	Siderophore biosynthesis
Chr 1	2934558	0.736	PROKKA_02740	4-16	Cystine-binding periplasmic protein
Chr 2	275479	151.1	PROKKA_03911 to PROKKA_04047	15	Diguanylate cyclase/phosphodiesterase RpfR; chaperone DnaK
Chr 2	1264028	106	PROKKA_04773 to PROKKA_04857	10; 15	Type III secretion; flagellar biosynthesis; galactose metabolism
Chr 2	1447293	20.6	PROKKA_04923 to PROKKA_04945	8	CepR regulatory protein
Chr 2	1643082	2.9	PROKKA_05088	16	Adhesion protein
Chr 2	2418658	1.1	PROKKA_05769; PROKKA_05770	8	Transport proteins
Chr 3	120468	0.389	PROKKA_03164	9; 10; 14; 15	23S rRNA
Chr 3	122606	0.246	PROKKA_03165; PROKKA_03166	16	tRNA-Ala; tRNA-Ile
Chr 3	672318	0.446	PROKKA_03626; PROKKA_03627	5-7; 9-16	Trehalohydrolase; glucanotransferase

3.1.2.4. Mutations distinguishing first and subsequent groups of isolates

Analysing the complete list of mutations (supplementary table A1) and considering that genes with mutations that became fixed over time are the most likely targets of evolution during chronic *B. multivorans* colonization, a collection of 12 SNPs and one large deletion that distinguished the latter 13 isolates from the first 3 isolates was selected for further analysis. Table 6 shows the list of mutations found in isolates 4 to 16 when compared with the genome of the first isolate and absent in isolates 2 and 3. In addition to these mutations, there was one gene that was mutated in all the isolates apart from the first one – gene PROKKA_01466, which encodes a LacI family transcription regulator. This gene is located upstream several genes encoding an ABC transporter.

Table 6 - List of mutations that fixed in the infecting *B. multivorans* population, distinguishing the first three isolates and the subsequent isolates sampled from CF patient P213. CDS, coding sequence; nonsyn, nonsynonymous; syn, synonymous.

Gene locus	Gene name	Category	Annotation	Mutations found	Effect in protein	Effect of mutation	Homolog in *B. multivorans* ATCC 17616
PROKKA_01273	*pheA*	CDS	Chorismate mutase	T>G	L280R	nonsyn	Bmul_2261
PROKKA_01500	*hprK*	CDS	HPr kinase	G>A	R114C	nonsyn	Bmul_0521
PROKKA_01663	*acrB*	CDS	Multidrug efflux transporter permease subunit	T>C	V178A	nonsyn	Bmul_0689
PROKKA_02125	*folD*	CDS	Methenyltetrahydrofolate cyclohydrolase	C>T	A210T	nonsyn	Bmul_1130
PROKKA_02127	*fixL*	CDS	Sensor protein FixL	A>G	S707P	nonsyn	Bmul_1132
PROKKA_02189	*accA*	CDS	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha	T>A	D186E	nonsyn	Bmul_1205
PROKKA_02387	-	CDS	Major Facilitator Superfamily transporter	C>T	P244P	syn	Bmul_1342
PROKKA_02752	-	CDS	LysR family transcriptional regulator	G>A	P157L	nonsyn	Bmul_1695
PROKKA_03101	-	CDS	Endoribonuclease	G>A	A147V	nonsyn	Bmul_5405
PROKKA_03779	*mcpC*	CDS	Methyl-accepting chemotaxis protein	G>A	H311H	syn	Bmul_4946
PROKKA_03885	*bag*	CDS	IgA FC receptor precursor	C>A	A10S	nonsyn	-
PROKKA_03885	*bag*	CDS	IgA FC receptor precursor	C>A	R9L	nonsyn	-
PROKKA_02740	*fliY*	-	Cystine-binding periplasmic protein	736-bp deletion	-	-	Bmul_1679

Notably, most of the mutations that fixed in the population were nonsynonymous, suggesting positive evolutionary selection. Genes such as PROKKA_02387 and PROKKA_01663, encoding a major facilitator superfamily (MFS) transporter and a multidrug efflux transporter permease subunit, respectively, could play a role in antimicrobial resistance. Gene PROKKA_02189 encodes the acetyl-CoA carboxylase subunit alpha, which is involved in fatty acid biosynthesis, suggesting that changes in lipid metabolism may be under selection during adaptation of *B. multivorans* to the CF airways of this patient. Other mutated genes that might affect metabolism are *pheA*, encoding the enzyme chorismate mutase that is involved in the synthesis of phenylalanine and tyrosine; gene *hprK*, encoding a serine/threonine kinase that phosphorylates/dephosphorylates the phosphotransferase system (PTS) protein HPr, regulating the entrance of metabolites into the cell (catabolite repression); gene *folD*, encoding an enzyme involved in the folate-activated one-carbon pool, a process required for nucleotide biosynthesis and, consequently, DNA and RNA synthesis and DNA repair; and gene PROKKA_02740, encoding a subunit of an ABC transporter involved in the uptake of cystine, the oxidized dimer form of cysteine. Another mutated gene, PROKKA_03101, encodes a protein of the RidA family, which prevents accumulation of reactive intermediates generated by pyridoxal-5'-phosphate--dependent enzymes and thus cellular damage. Some mutations also targeted regulators, namely a LysR family transcriptional regulator (gene PROKKA_02752); a methyl-accepting chemotaxis protein (gene PROKKA_03779) homolog to the *P. aeruginosa* WspA receptor involved in biofilm formation (Hickman *et al.* 2005); and gene PROKKA_02127, a *fixL* homolog, which encodes a sensor histidine kinase of a two-component regulatory system. This *fixLJ* regulatory system is involved in biofilm formation, motility, virulence and intracellular invasion in *B. dolosa* (Schaefers *et al.* 2017).

In summary, the most relevant mutations distinguishing the two groups of sequentially retrieved isolates from patient P213 identified in this analysis occurred in genes that may relate to antibiotic resistance, metabolism of several macromolecules, motility and biofilm formation.

3.1.2.5. Phylogenetic analysis

To determine the evolutionary relationship among the 16 *B. multivorans* isolates, SNPs were used to construct a maximum likelihood phylogenetic tree (figure 9). The root of the tree was placed on the branch of isolate P213-1 and the analysis clusters the first three isolates together and separates them from the latter ones. Isolate 8 seems to be phylogenetically more distant from the other isolates, which is likely explained by its increased number of SNPs due to a deletion in the mismatch repair gene *mutL*. Interestingly, isolates that were collected on the same date do not seem to be closely related, as is the case for isolate pairs P213-12/P213-13 and P213-14/P213-15, showing coexistence of several lineages within the same CF patient.

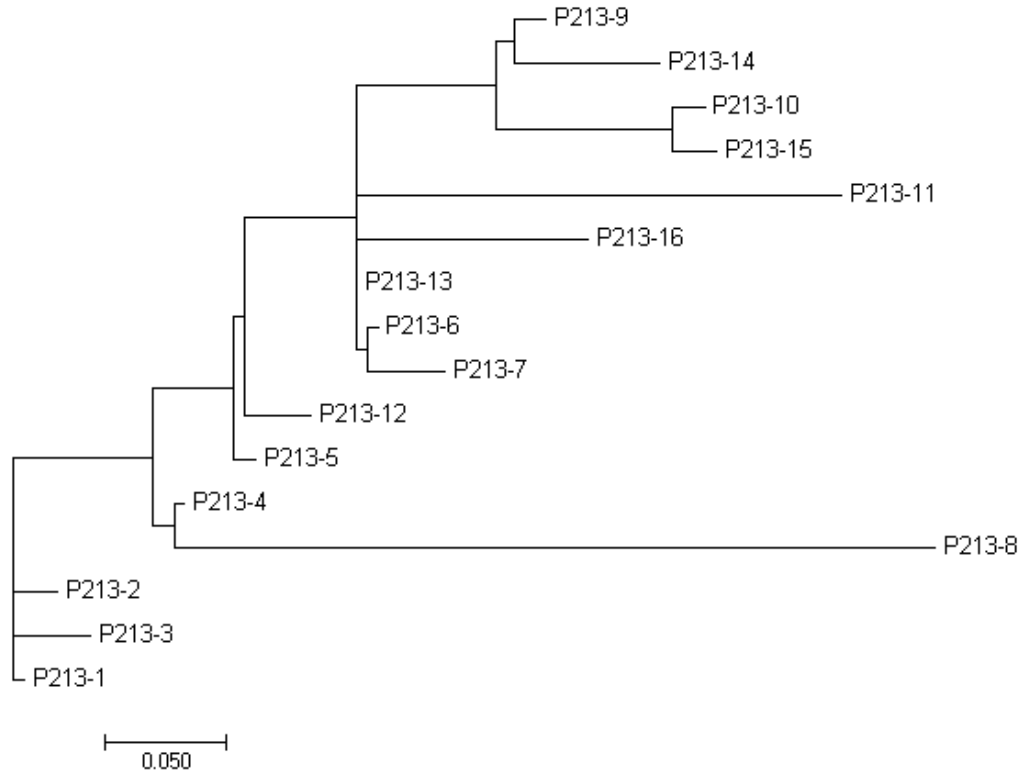

Figure 9 - Phylogenetic tree of the 16 *B. multivorans* isolates recovered from CF patient P213. The evolutionary history was inferred by using the maximum likelihood method based on the Tamura-Nei model (Tamura & Nei 1993). The tree with the highest log likelihood is shown and it is drawn to scale, with branch lengths measured in the number of substitutions per site (229 positions).

3.1.3. Analysis of lipopolysaccharide

LPS plays an important role in virulence for Bcc species, and O-antigen repeats have been shown to interfere with adhesion to abiotic surfaces and bronchial epithelial cells (Saldías *et al.* 2009). Nevertheless, the O-antigen is highly immunogenic, which means that bacteria that lack this portion of the LPS could have an advantage in evading the host immune system. To evaluate the production of the O-antigen repeats of LPS by the 16 *B. multivorans* isolates sampled from CF patient P213, LPS was extracted and purified, and resulting samples were resolved by electrophoresis. Figure 10 shows that the LPS of the first three isolates presented the O-antigen repeats, but the subsequently retrieved isolates lacked the O-antigen portion of LPS.

Examining the list of mutations generated by the comparative genomics analysis, mutations in genes involved in LPS biosynthesis were identified, but the observed phenotypes could not be explained for all the isolates based on this analysis. Nonsynonymous SNPs were found in genes encoding LPS export system protein LptA (in isolates 4 and 8), LPS export system permease LptG (in isolate 7) and LPS core heptosyltransferase (in isolates 10 and 15).

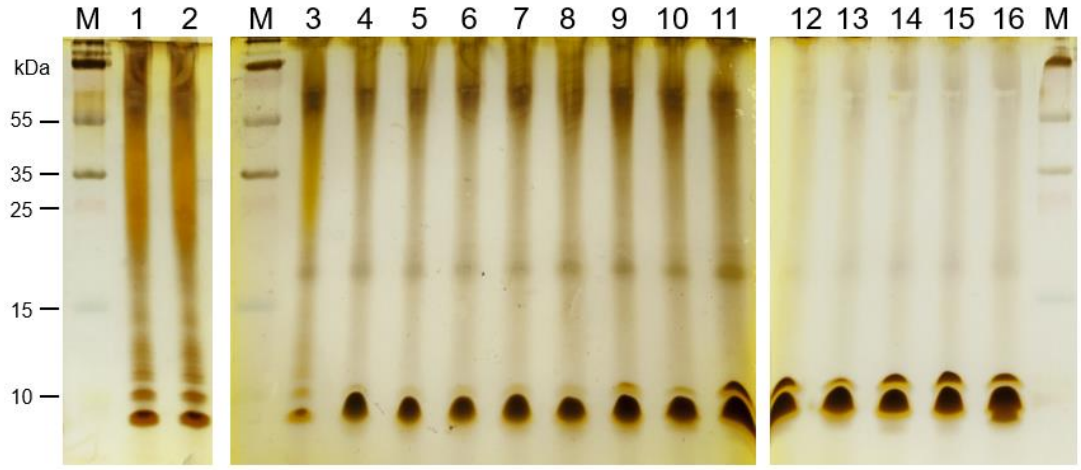

Figure 10 - Changes in the LPS pattern of the 16 *B. multivorans* isolates recovered from patient P213. Isolates 1, 2 and 3 are the only ones that exhibit the O-antigen. Lanes M, protein markers.

3.1.4. Exopolysaccharide production

A transition in mucoidy is commonly observed in *Burkholderia cepacia* complex (Bcc) bacteria during long-term chronic infections in CF patients. To analyse the mucoid phenotype, the amount of exopolysaccharide (EPS) produced by each isolate of *B. multivorans* sampled from CF patient P213 was assessed after 6 days of growth in liquid salt-mannitol (SM) medium. The recovered EPS concentration of the 16 isolates is presented in figure 11.

Isolates 5, 7, 11, 12, 14 and 15 were the only ones able to produce detectable amounts of high molecular weight (HMW) exopolysaccharide. However, isolates 1, 2, 4 and 13 produced minor amounts of a precipitate that resembled the extracellular polymeric substance but consisted of small particles that were difficult to detect and collect. The recovered HMW EPS dry-weight ranged from approximately 4 to 9 g/L.

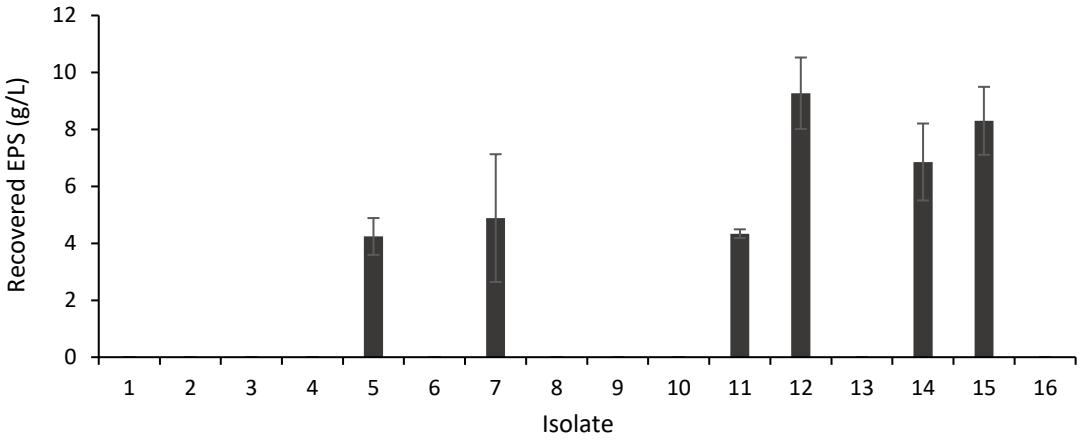

Figure 11 - Exopolysaccharide production of the 16 *B. multivorans* isolates sequentially retrieved from CF patient P213, based on the dry weight of ethanol-precipitated EPS after 6 days of growth at 37ºC (error bars represent the standard deviations of the mean values for three independent experiments). Only high molecular weight EPS was collected for this analysis.

In addition, the mucoid phenotype was also evaluated in solid medium (YEM, yeast extract mannitol medium), by visual inspection of the inoculated plates after 48 h of growth at 37°C. Isolates were classified as highly mucoid, mucoid and non-mucoid, and the results are presented in table 7. YEM plates of isolates 6, 9 and 10, corresponding to highly mucoid, mucoid and non-mucoid isolates, respectively, are shown in figure 12 (the plates of the remaining isolates are presented in supplementary figure A1). The first seven isolates displayed higher mucoidy than the remaining ones, and isolates 10, 12 and 15 were non-mucoid in YEM plates. The results for the mucoid phenotype analysis in solid and liquid media are not in agreement, which is likely due to the fact that different compositions were used for the solid and liquid media in this experiment.

Table 7 - Mucoid phenotype assessment in yeast extract mannitol medium (YEM) after 48 h of incubation at 37°C for the 16 *B. multivorans* isolates sequentially retrieved from CF patient P213. Isolates were classified as highly mucoid (++), mucoid (+) and non-mucoid (-) by visual inspection.

Isolate	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Mucoidy	++	++	++	++	++	++	++	+	+	-	+	-	+	+	-	+

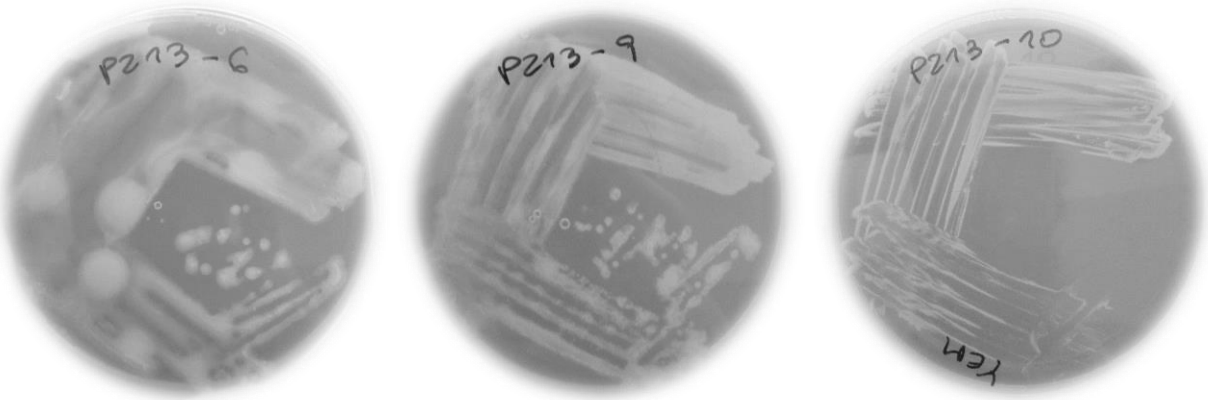

Figure 12 - Mucoid phenotype assessment in yeast extract mannitol medium (YEM) after 48 h of incubation at 37°C for *B. multivorans* isolates 6 (highly mucoid, ++), 9 (mucoid, +) and 10 (non-mucoid, -) of patient P213.

Although variations in EPS production were evident in the phenotypical assessments, no mutations in genes belonging to the EPS biosynthetic gene cluster (*bce* gene cluster; Moreira *et al.* 2003) were identified in the genome analysis.

3.1.5. Susceptibility against antimicrobials

Antibiotic resistance is a frequently reported characteristic of Bcc isolates chronically infecting CF patients. Therefore, the susceptibility of each isolate retrieved from patient P213 against the antibiotics piperacillin/tazobactam, ciprofloxacin, aztreonam and kanamycin was tested. Piperacillin is a β-lactam antibiotic that belongs to the class of penicillin and is widely used in combination with the β-lactamase inhibitor tazobactam, acting by inhibiting bacterial cell wall biosynthesis; ciprofloxacin is a

fluoroquinolone that works by inhibiting DNA replication; aztreonam is a β-lactam that inhibits cell wall biosynthesis; and kanamycin is an aminoglycoside that inhibits protein synthesis.

The agar disc diffusion method (Bauer *et al.* 1966) was used to assess the susceptibility of the isolates against these antimicrobials and the resulting average diameter of the growth inhibition zone is represented in figure 13.

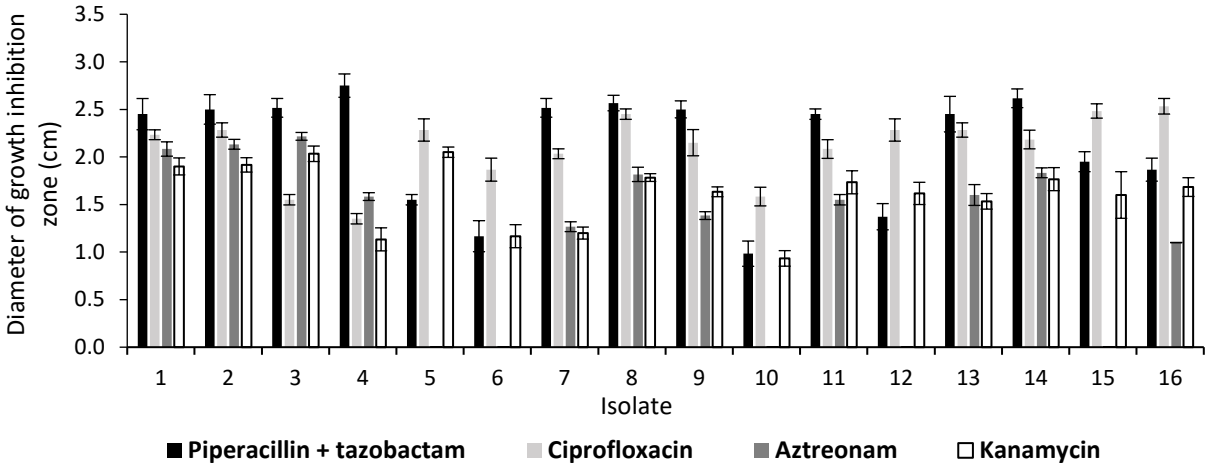

Figure 13 - Susceptibility of the 16 *B. multivorans* isolates sequentially retrieved from CF patient P213 to antibiotics (piperacillin/tazobactam, ciprofloxacin, aztreonam and kanamycin), measured as the diameter of cell growth inhibition, after growth for 24h at 37ºC (error bars represent the standard deviations of the mean values for three independent experiments).

Results show that the first three *B. multivorans* isolates were more susceptible to aztreonam and kanamycin, while the subsequently sampled isolates were more resistant to these two antibiotics. Statistical significance of differences between the first isolate and the subsequent ones was determined for these two antimicrobials using ANOVA followed by Dunnett's multiple comparisons test. For susceptibility to aztreonam, differences were extremely significant for all isolates (*P*-value < 0.0001; supplementary figure A2). For kanamycin, while most isolates presented significant differences in susceptibility to this antimicrobial (*P*-value < 0.01, at least), isolates 5, 8, 11 and 14 showed no statistically significant alterations (supplementary figure A3). Concerning susceptibility to piperacillin/tazobactam and ciprofloxacin, there seems to be two distinct populations, one that remains susceptible and one that becomes more resistant to these antimicrobials during the course of infection. Moreover, isolates 5, 6, 10, 12 and 15 were completely resistant to aztreonam and were simultaneously more resistant to piperacillin/tazobactam than the first three isolates.

In search of mutations that could underlie the observed increase in resistance to antibiotics, two SNPs were found in genes encoding a MFS transporter and a multidrug efflux transporter permease subunit (genes PROKKA_02387 and PROKKA_01663, respectively) in the latter 13 isolates retrieved from this CF patient.

3.1.6. Biofilm formation

During chronic CF infections, Bcc isolates have been reported to undergo alterations regarding biofilm formation. Biofilms consist in communities within which bacteria present a sessile lifestyle, being protected from environmental stresses. To evaluate the biofilm formation of the *B. multivorans* isolates retrieved from CF patient P213, the absorbance at 590 nm (A_{590}) of cell suspensions grown in microtiter plates for 48h was measured after adequate staining (results in figure 14).

Results show that the first three isolates produced less biofilm than the subsequent isolates, except for isolates 6 and 7, for which differences in this trait were found statistically non-significant (differences were considered significant for *P*-values < 0.05). Interestingly, isolate 4 produced a much higher amount of biofilm than all the other isolates.

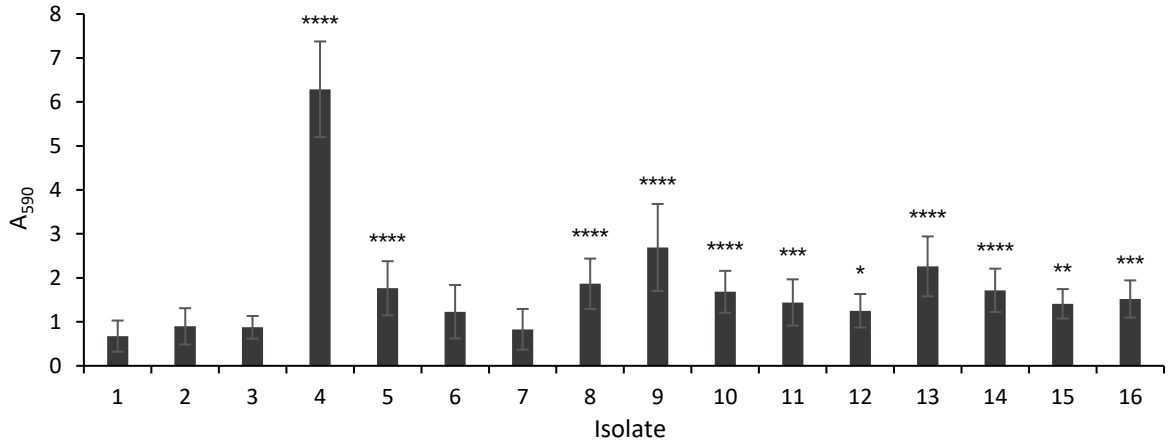

Figure 14 - Biofilm formation of the 16 *B. multivorans* isolates sequentially retrieved from CF patient P213, after growth in polystyrene microplates for 48h at 37ºC. Error bars represent the standard deviations of the mean values for at least three experiments (statistical significance of differences between the first isolate and the subsequent ones was determined: *, *P*-value < 0.05; **, *P*-value < 0.01; ***, *P*-value < 0.001; ****, *P*-value < 0.0001 by ANOVA followed by Dunnett's multiple comparisons test).

Possible mutations underlying differences in biofilm formation are SNPs in a *fixL* homolog and in a gene encoding a methyl-accepting chemotaxis protein, which were identified in the latter 13 isolates of this patient. In *B. dolosa*, *fixL* encodes a sensor histidine kinase of a two-component regulatory system (FixLJ) that regulates biofilm formation, motility and virulence (Schaefers *et al.* 2017). Chemotaxis proteins sense and respond to environmental stimuli and may also be involved in motility and biofilm formation.

3.1.7. Swimming and swarming motility

The following phenotypic trait associated with long-term CF chronic infections studied in this work was motility. It has been demonstrated that longitudinally retrieved Bcc isolates chronically infecting the airways of a CF patient experience a decrease in motility during the course of infection. To assess the swimming motility of the 16 *B. multivorans* isolates sampled from patient P213, swimming agar plates were incubated for 24 h and the diameter of the swimming zone was measured. Results are

shown in figure 15. The first three isolates clearly display increased swimming motility in comparison with the subsequently recovered isolates. Moreover, isolates 9, 10 and 16 showed no swimming motility.

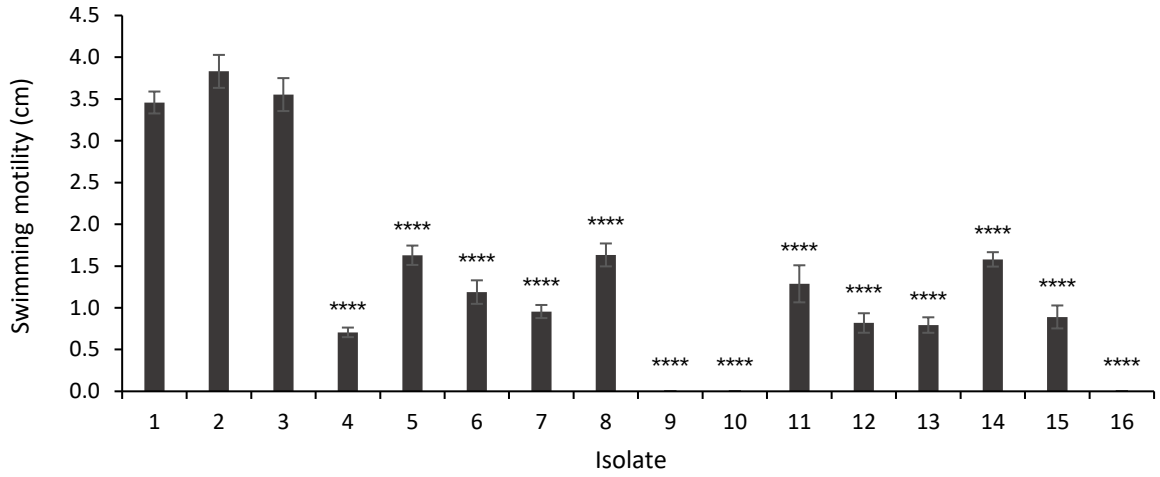

Figure 15 - Swimming motility of the 16 *B. multivorans* isolates sequentially retrieved from CF patient P213, after growth in swimming agar plates for 24 h at 37ºC. Error bars represent the standard deviations of the mean values for three independent experiments (statistical significance of differences between the first isolate and the subsequent ones was determined: ****, *P*-value < 0.0001 by ANOVA followed by Dunnett's multiple comparisons test).

To evaluate the swarming motility of the 16 *B. multivorans* isolates from patient P213, swarming agar plates were incubated for 48 h and the diameter of the motility zone was measured. Results are shown in figure 16. The latter 13 isolates show decreased swarming motility in comparison with the first three (except for isolate 5, that showed no statistically significant changes), with isolates 9 to 16 displaying no motility in swarming agar plates.

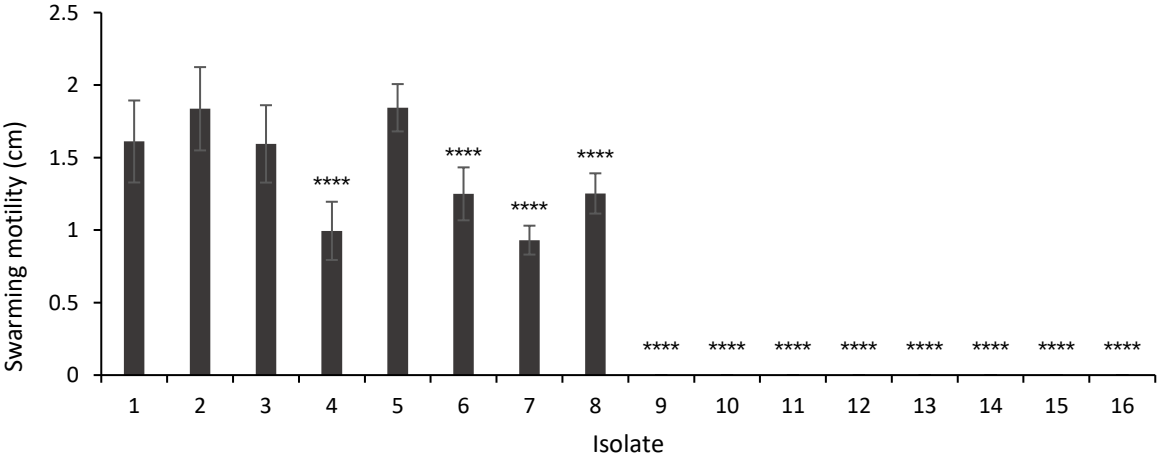

Figure 16 - Swarming motility of the 16 *B. multivorans* isolates sequentially retrieved from CF patient P213, after growth in swarming agar plates for 48 h at 37ºC. Error bars represent the standard deviations of the mean values for three independent experiments (statistical significance of differences between the first isolate and the subsequent ones was determined: ****, *P*-value < 0.0001 by ANOVA followed by Dunnett's multiple comparisons test).

Some mutations involved in motility were identified among the 16 *B. multivorans* isolates, namely in genes encoding flagellar proteins (FliL in isolate 10, FlgE in isolate 11 and FlgI in isolate 16). Nonetheless, the observed differences in motility between the two sets of isolates might be a result of mutations in the homolog of the regulatory gene *fixL* or in the gene encoding a chemotaxis protein, both associated with regulation of motility.

3.1.8. Growth rate and doubling time determination

To evaluate differences in growth among the isolates retrieved from patient P213, growth rates were measured by monitoring the OD_{640} of cultures grown in synthetic cystic fibrosis medium (SCFM) (Palmer *et al*. 2007). Table 8 presents the doubling times calculated for each isolate and the corresponding growth curves are shown in supplementary figure A4. Results show that the first seven isolates, except for isolate 4, show shorter doubling times when compared with the latter ones.

Table 8 - Doubling times (in hours) calculated for the 16 *B. multivorans* isolates sampled from CF patient P213, based on growth rates measured from cultures grown in synthetic cystic fibrosis medium (SCFM). Results are the means of data from two independent experiments.

Isolate	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Doubling time (h)	1.6	1.6	1.6	2.9	1.6	1.6	1.6	1.7	2.8	3.0	2.0	2.0	2.0	1.8	2.0	2.7

In the comparative genomics analysis, two mutations that could be involved in amino acid biosynthesis were identified in the latter 13 isolates of this patient, possibly influencing bacterial growth. These mutations are a 736-bp deletion in a gene encoding a cystine-binding periplasmic protein and a nonsynonymous SNP in a chorismate mutase. Moreover, nonsynonymous SNPs in genes encoding ribosomal proteins were found, with possible implications in protein synthesis and, consequently, cell growth. Isolates 9, 10, 14 and 15 have a SNP in a gene encoding the 50S ribosomal protein L25 and isolate 16 has a SNP in a gene that encodes the 50S ribosomal protein L4. However, not all observed growth rates can be explained by these mutations.

3.1.9. Adhesion to epithelial cells

Adhesion of the *B. multivorans* isolates recovered from CF patient P213 to the bronchial epithelial cell line CFBE41o$^-$, derived from a patient homozygous for the CFTR F508del mutation, was assessed. Alterations in this phenotypic trait have been associated with adaptation to the airways of CF patients (Silva *et al.* 2016). Epithelial cells were infected at a multiplicity of infection (MOI) of 10 (10 bacterial cells to 1 epithelial cell) and results are shown in figure 17, as percentage of adhesion.

Results show that the first three isolates retrieved from patient P213 exhibit decreased adhesion to bronchial epithelial cells in comparison to the latter isolates. This may be due to the presence of LPS O-antigen in these isolates, as O-antigen has been reported to inhibit adhesion to epithelial cells in *B. cenocepacia* (Saldías *et al.* 2009). A 2.9-kb deletion was identified in a gene encoding an adhesin (PROKKA_05088) in isolate 16, but ability to adhere to epithelial cells seems to not have been affected.

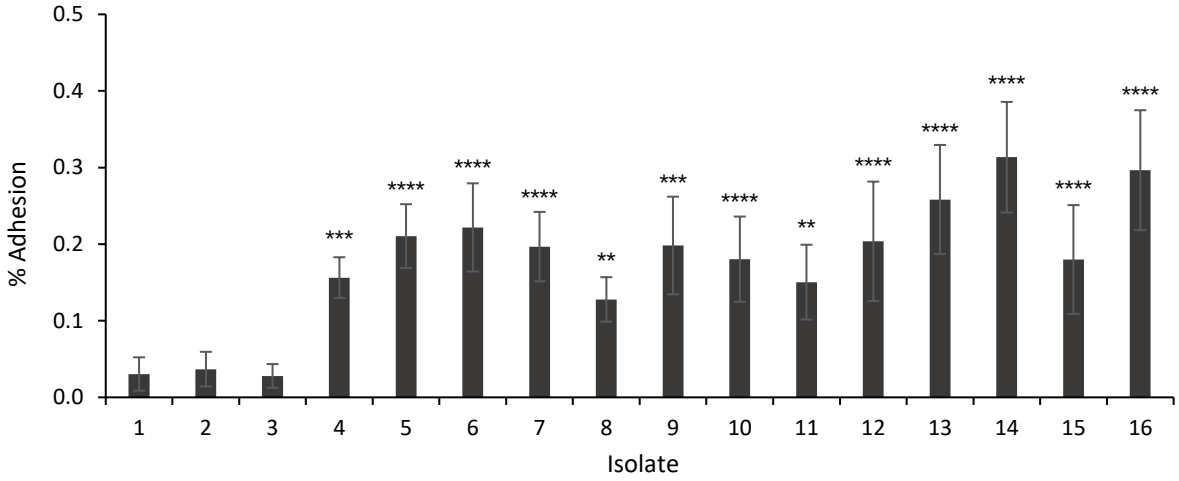

Figure 17 - Adhesion of the 16 *B. multivorans* isolates sequentially retrieved from CF patient P213 to the bronchial epithelial cell line CFBE41o⁻. Results are shown as the percentage of adhesion, which was calculated as the number of CFU recovered divided by the number of CFU applied to the epithelial cells multiplied by 100. Error bars represent the standard deviations of the mean values for three independent experiments (statistical significance of differences between the first isolate and the subsequent ones was determined: **, *P*-value < 0.01; ***, *P*-value < 0.001; ****, *P*-value < 0.0001 by ANOVA followed by Dunnett's multiple comparisons test).

3.1.10. Virulence in *Galleria mellonella*

Galleria mellonella was used as an infection model to study virulence of the *B. multivorans* isolates under study. This insect was chosen as an infection model for having an innate immune system with a high degree of structural and functional homology to the innate immune systems of mammals (Seed & Dennis 2008). Ten larvae were injected per isolate in each experiment and survival was followed for three days post-infection (results in figure 18). A clear difference in terms of virulence is observed when comparing the first three isolates with the latter ones, except for isolate 13, as these four *B. multivorans* strains cause 0% survival only 24 hours post-infection.

As *fixL* regulates virulence in *B. dolosa*, the mutation in its homolog that is present in the latter 13 isolates of patient P213 might underlie the decreased virulence of those isolates (except for isolate 13). Nevertheless, this change in acute virulence could be associated with the presence of LPS O-antigen in the first three isolates, as O-antigen has been shown to inhibit phagocytosis by macrophages, thus reducing bacterial elimination, in *B. cenocepacia* (Saldías *et al.* 2009). Furthermore, large deletions found in some isolates include genes involved in type III and type VI secretion, likely influencing virulence ability. A 28.8-kb deletion in isolates 9, 10, 14 and 15 (in chromosome 1) comprises genes involved in type VI secretion and isolates 10 and 15 also present a 106-kb deletion (in chromosome 2) including genes involved in type III secretion. While virulence of isolate 9 generates a Kaplan-Meier survival curve similar to other moderately virulent isolates, acute virulence in *G. mellonella* is notably attenuated for isolates 10, 14 and 15. This attenuated virulence is also observed for isolate 12, but no mutations that could be responsible for this phenotype were identified.

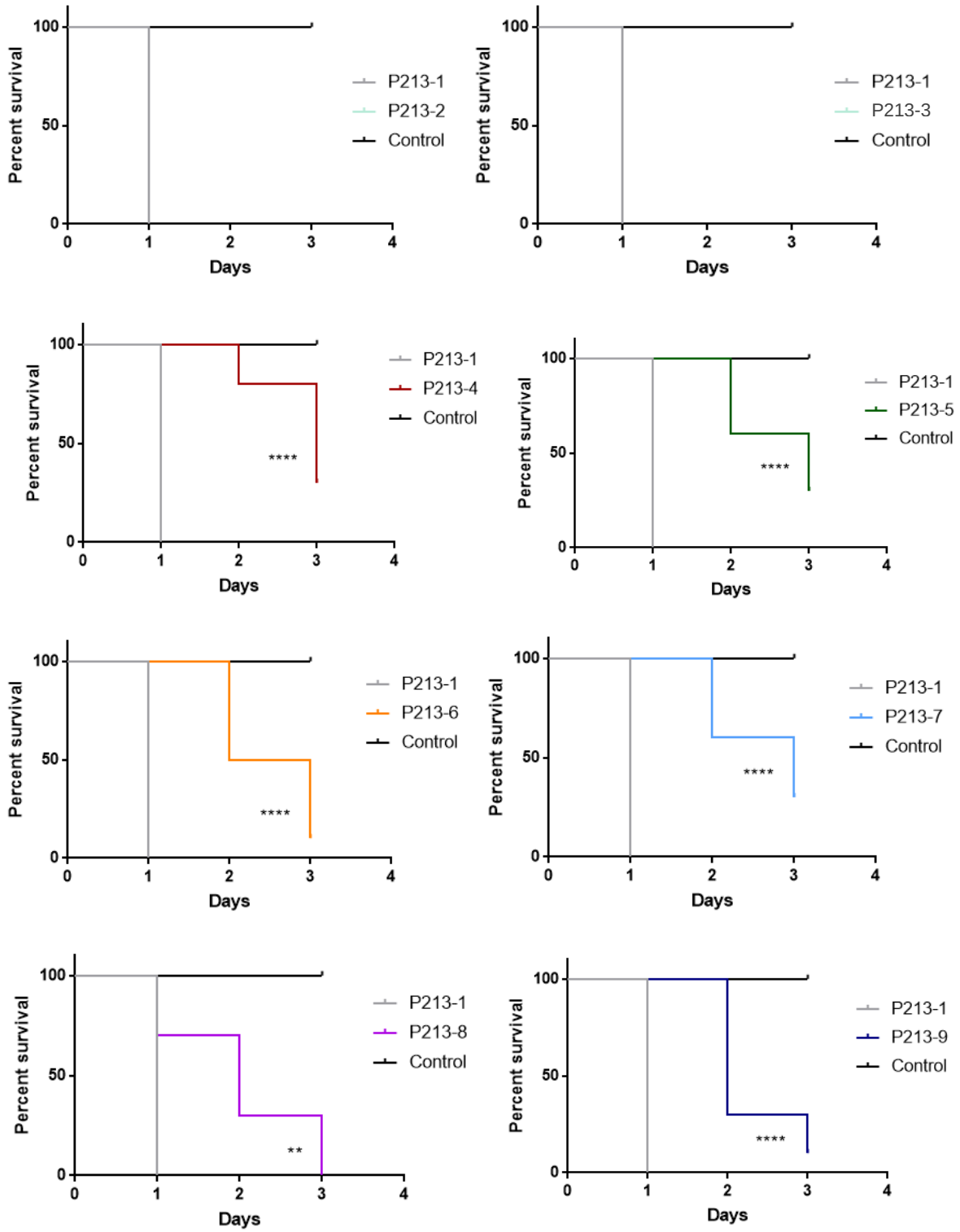

Figure 18 - Survival of *Galleria mellonella* larvae inoculated with *B. multivorans* isolates retrieved from CF patient P213. Triplicate groups of 10 larvae were inoculated with each isolate and survival was followed for three days post-infection. Larvae were injected with approximately $1x10^6$ bacterial cells. The control experiment without bacteria is also represented. Statistical significance of differences between the Kaplan-Meier curve of the first isolate and the subsequent ones was determined: **, *P*-value < 0.01; ***, *P*-value < 0.001; ****, *P*-value < 0.0001.

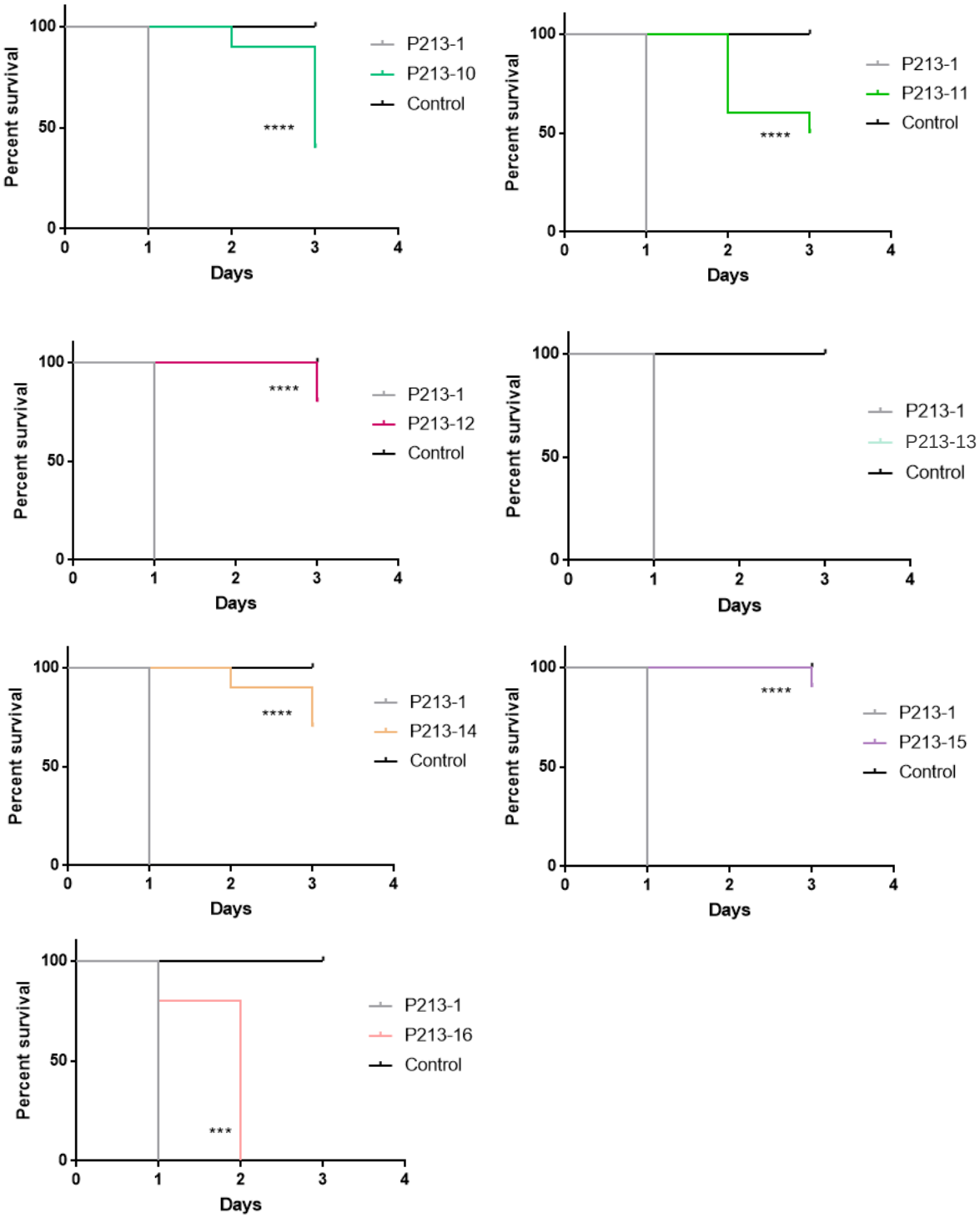

Figure 18 - Survival of *Galleria mellonella* larvae inoculated with *B. multivorans* isolates retrieved from CF patient P213 (cont.).

3.2. Analysis of *B. multivorans* isolates retrieved from CF patient P426

3.2.1. *B. multivorans* isolates from CF patient P426

The 21 bacterial isolates consist of a single clone of *Burkholderia multivorans* sampled from cystic fibrosis patient P426 at certain time points in the period between 1997 and 2014 (figure 19). Isolate pairs 1/2, 3/4 and 17/18 were recovered in different time points of years 1997, 1998 and 2011, respectively. Isolate pairs 11/12, 13/14 and 20/21 were each recovered in a single time point.

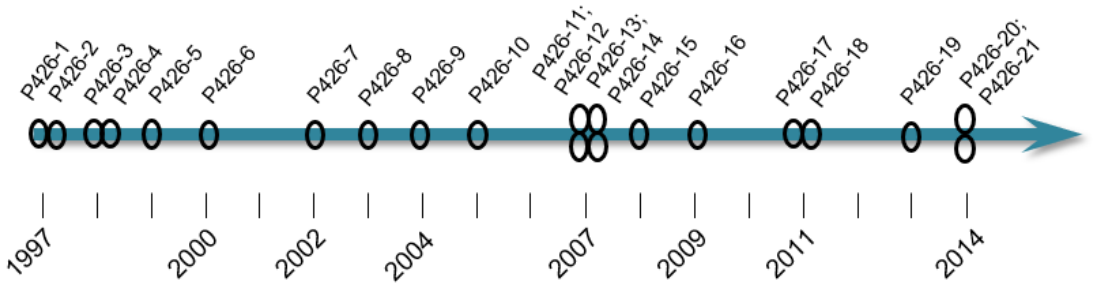

Figure 19 - *B. multivorans* isolates recovered from CF patient P426.

3.2.2. Analysis of single-nucleotide polymorphism and indel mutations

To detect genetic mutations that could originate phenotypic variation, the genomes of the 21 *B. multivorans* isolates sampled from patient P426 were sequenced and analysed. The genome of the first isolate was used as the reference genome for the identification of mutations, after sequencing and *de novo* assembly. The assembly of the genome of the first isolate yielded 7 contigs with lengths between 7045 and 3430912 bp (table 9). To identify which contig corresponded to each *B. multivorans* chromosome, Basic Local Alignment Search Tool (BLAST, Altschul *et al.* 1990) was used to align each contig to the genome of *B. multivorans* ATCC 17616. Contigs 1, 2 and 3 correspond to chromosomes 3, 2 and 1, respectively. Contigs 4 to 7 did not map onto the reference genome of *B. multivorans* ATCC 17616 and are likely plasmids. Given that contigs 4, 5 and 7 were simultaneously absent from five isolates, it is expected that these contigs comprise one plasmid that was lost in those isolates and that contig 6 corresponds to another plasmid that was maintained in all isolates.

Table 9 - Lengths of the 7 contigs generated from the assembly of the genome of the first *B. multivorans* isolate recovered from patient P426 and lengths of the corresponding chromosomes of *B. multivorans* ATCC 17616.

P426-1 contig	P426-1 contig length (bp)	*B. multivorans* ATCC 17616 chromosome	*B. multivorans* ATCC 17616 chromosome length (bp)
1	1 204 139	3	919 805
2	2 425 401	2	2 473 162
3	3 430 912	1	3 448 421
4	7 288	-	-
5	7 840	-	-
6	33 556	-	-
7	7 045	-	-

SNP and indel mutations were detected by aligning sequence reads for each isolate against the genome sequence of the first isolate. Only mutations identified in alignments using two alignment tools and occurring in at least three forward and reverse reads were subsequently analysed. The final list of SNPs and indels comprised mutations existing among the latter 20 *B. multivorans* isolates when compared to the genome of the first isolate. In addition, large deletions were identified by a visual inspection of the alignments and were also included in this analysis. A total of 337 mutations was identified, comprising 275 SNPs, 40 indels and 22 large deletions that were detected among the isolates recovered from patient P426 over 18 years of chronic colonization (supplementary table A2).

3.2.2.1. Identification of polymorphic genes

Polymorphic genes were investigated, with 24 genes presenting two different mutations and 5 genes presenting three mutations, all within the coding sequence (table 10). The polymorphic genes with three mutations encode a transport system permease protein (gene PROKKA_02456), an electron-transferring-flavoprotein dehydrogenase (PROKKA_04099), a hypothetical protein (PROKKA_01578), a LPS-assembly protein precursor (PROKKA_05461) and a lipid A deacylase precursor (PROKKA_05547). Interestingly, three genes related to LPS were polymorphic (PROKKA_05461, PROKKA_05547 and PROKKA_05090). One gene with two nonsynonymous mutations that seems relevant is a *fixL* homolog, which encodes a sensor histidine kinase of a two-component regulatory system that is involved in biofilm formation, motility, virulence and intracellular invasion in *B. dolosa* (Schaefers *et al.* 2017). A mutation in this gene was also found in 13 isolates of CF patient P213. Further, two distinct nonsynonymous mutations were found in gene PROKKA_05863, encoding a Fis family transcriptional regulator that was also identified as a polymorphic gene with three different mutations in the isolates of P213.

Table 10 - Polymorphic genes identified among the 21 *B. multivorans* isolates recovered from CF patient P426.

Contig	Position	Mutation	Locus tag	Gene name	Annotation	Homolog in *B. multivorans* ATCC 17616	Effect of mutation	Type of mutation	Isolates with mutation
contig 1	163483	CAG>CCG	PROKKA_00155	-	Hypothetical protein	Bmul_5602	nonsyn	SNP	7
contig 1	163487	ACG>CCG					nonsyn	SNP	7
contig 1	502063	GGT>AGT	PROKKA_00467	*ytnP*	Putative quorum-quenching lactonase	-	nonsyn	SNP	13
contig 1	502108	AGC>CGC					nonsyn	SNP	2-21
contig 2	24055	CCG>CCT	PROKKA_01149	-	Type VI secretion protein	Bmul_3902	syn	SNP	17
contig 2	24056	AGC>CGC					nonsyn	SNP	17
contig 2	144318	ACG>ACC	PROKKA_01249	-	LuxR family transcriptional regulator	Bmul_3738	syn	SNP	9-12; 14-16; 19; 21
contig 2	144529	CTG>CGG					nonsyn	SNP	13
contig 2	495662	GAA>GAC	PROKKA_01578	-	Hypothetical protein	-	nonsyn	SNP	20
contig 2	495668	GAC>GAG					nonsyn	SNP	20
contig 2	495701	TTT>TTG					nonsyn	SNP	20
contig 2	692133	AAC>GAC	PROKKA_01734	-	Outer membrane porin	Bmul_3342	nonsyn	SNP	13
contig 2	692531	TAC>TAG					nonsyn	SNP	13

Table 10 - Polymorphic genes identified among the 21 *B. multivorans* isolates recovered from CF patient P426 (cont.).

Contig	Position	Mutation	Locus tag	Gene name	Annotation	Homolog in *B. multivorans* ATCC 17616	Effect of mutation	Type of mutation	Isolates with mutation
contig 2	810557	CTC>ATC	PROKKA_01840	*aroG*	Phospho-2-dehydro-3-deoxyhepto-nate aldolase	Bmul_3232	nonsyn	SNP	20
contig 2	811057	GTC>GCC					nonsyn	SNP	11; 21
contig 2	1203055	-9ggtg cccga	PROKKA_02187	*bag*	IgA FC receptor precursor	-	3 aa deletion	indel	5
contig 2	1203065	GAG>GAA					syn	SNP	5
contig 2	1503211	CTC>CGC	PROKKA_02456	*potH*	Putrescine transport system permease protein PotH	Bmul_5430	nonsyn	SNP	3-7; 9-16; 18-21
contig 2	1503212	CTC>CTA					syn	SNP	3-7; 9-16; 18-21
contig 2	1503720	CTC>TTC					nonsyn	SNP	10
contig 3	271374	GGG>GGC	PROKKA_03495	-	Hypothetical protein	-	syn	SNP	8-12; 14-16; 19; 21
contig 3	271595	-1c					frameshift	indel	7; 18
contig 3	274762	TCG>TTG	PROKKA_03497	*pimB*	GDP-manno-se-dependent alpha-(1-6)-phosphatidyli nositol mono-mannoside mannosyl-transferase	-	nonsyn	SNP	18
contig 3	275439	AGC>AGT					syn	SNP	8
contig 3	420538	CTG>TTG	PROKKA_03649	-	Putative glutamine amidotrans-ferase	Bmul_2371	syn	SNP	11
contig 3	420883	-1a					frameshift	indel	20
contig 3	810899	GGC>AGC	PROKKA_04003	-	Acylaldehyde oxidase and xanthine de-hydrogenase molybdopterin binding protein	Bmul_1961	nonsyn	SNP	8; 17
contig 3	812372	TTC>CTC					nonsyn	SNP	11; 16; 21
contig 3	925543	GCG>GCC	PROKKA_04099	-	Electron transfer flavoprotein-ubiquinone oxidoreduc-tase	Bmul_1788	syn	SNP	14
contig 3	925885	GCG>GCC					syn	SNP	14
contig 3	925915	AAG>AAA					syn	SNP	14
contig 3	1113816	CAG>CGG	PROKKA_04305	-	Hypothetical protein	Bmul_1658	nonsyn	SNP	13
contig 3	1114149	ACC>AGC					nonsyn	SNP	4
contig 3	1155926	+3gag	PROKKA_04345	-	TetR family transcriptional regulator	Bmul_1617	1 aa deletion	indel	8; 17
contig 3	1155929	+4gatc					frameshift	indel	13
contig 3	1322758	-2gc	PROKKA_04491	-	Hypothetical protein	Bmul_1477	frameshift	indel	20
contig 3	1322782	CAG>CCG					nonsyn	SNP	11; 16; 21
contig 3	1611930	CCG>CAG	PROKKA_04760	*mepH*	Murein DD-endopepti-dase MepH precursor	Bmul_1218	nonsyn	SNP	3
contig 3	1612228	GGC>TGC					nonsyn	SNP	6; 13
contig 3	1745194	GGC>TGC	PROKKA_04896	*fixL*	Sensor protein FixL	Bmul_1132	nonsyn	SNP	8-12; 14-16; 19; 21
contig 3	1746932	ATG>ACG					nonsyn	SNP	8; 17

Table 10 - Polymorphic genes identified among the 21 *B. multivorans* isolates recovered from CF patient P426 (cont.).

Contig	Position	Mutation	Locus tag	Gene name	Annotation	Homolog in *B. multivorans* ATCC 17616	Effect of mutation	Type of mutation	Isolates with mutation
contig 3	1950433	TGA>CGA	PROKKA_05090	*rfaQ*	LPS core heptosyltransferase RfaQ	-	nonsyn	SNP	18
contig 3	1951341	GCC>GTC					nonsyn	SNP	6; 13
contig 3	2128370	CAG>CGG	PROKKA_05256	*livH*	Amino acid transport system permease	Bmul_0798	nonsyn	SNP	13
contig 3	2128371	CAG>AAG					nonsyn	SNP	13
contig 3	2206844	GAC>GTC	PROKKA_05327	*cmaA*	Cyclopropane-fatty-acyl-phospholipid synthase	Bmul_0727	nonsyn	SNP	13
contig 3	2207131	TAC>CAC					nonsyn	SNP	19
contig 3	2229130	CGC>GGC	PROKKA_05345	-	Beta-lactamase	Bmul_0708	nonsyn	SNP	20
contig 3	2229131	CCG>CCC					syn	SNP	20
contig 3	2356074	GTC>GAC	PROKKA_05461	*lptD*	LPS assembly protein LptD precursor	Bmul_0588	nonsyn	SNP	8-12; 14-16; 19; 21
contig 3	2356176	GGC>GCC					nonsyn	SNP	7
contig 3	2356177	GGC>AGC					nonsyn	SNP	7
contig 3	2429786	CGC>CTC	PROKKA_05526	*hprK*	HPr kinase/phosphorylase	Bmul_0521	nonsyn	SNP	8; 17
contig 3	2430037	GCC>ACC					nonsyn	SNP	13
contig 3	2448847	TAT>CAT	PROKKA_05547	*pagL*	Lipid A deacylase PagL precursor	Bmul_0487	nonsyn	SNP	13
contig 3	2448851	+1t					frameshift	indel	20
contig 3	2449067	-2ca					frameshift	indel	15
contig 3	2536493	-27tcacgcgcatcgcgtcgagctcctcgg	PROKKA_05629	*mlaE*	Putative phospholipid ABC transporter permease protein MlaE	Bmul_0403	9 aa deletion	indel	8
contig 3	2537284	CGC>CGT					syn	SNP	13
contig 3	2789456	GAA>GAC	PROKKA_05863	*fis*	Fis family transcriptional regulator	Bmul_3117	nonsyn	SNP	9-12; 14-16; 19; 21
contig 3	2789704	GAG>CAG					nonsyn	SNP	20
contig 3	3196331	+6tcggcc	PROKKA_06251	*pilQ*	Type IV pilus secretin PilQ	Bmul_0294	2 aa deletion	indel	9
contig 3	3197434	CTG>CTT					syn	SNP	3-16; 18-21

3.2.2.2. Mutation rate determination

Figure 20 shows a representation of the SNPs identified per isolate over time since the first isolate was recovered, leading to the identification of different patterns. Initially, isolates 1 to 11 seem to gain mutations at a constant rate of about 5.4 SNPs/year. The remaining isolates presented less SNPs than expected if following the linear tendency of the first set of isolates. Isolate 17 showed a significantly low number of mutations, with a total of 25 mutations comprising 21 SNPs, 3 indels and 1 large deletion. On the other hand, isolate 20 presented the highest number of mutations, with a total of 87 mutations, including 74 SNPs, 8 indels and 5 large deletions.

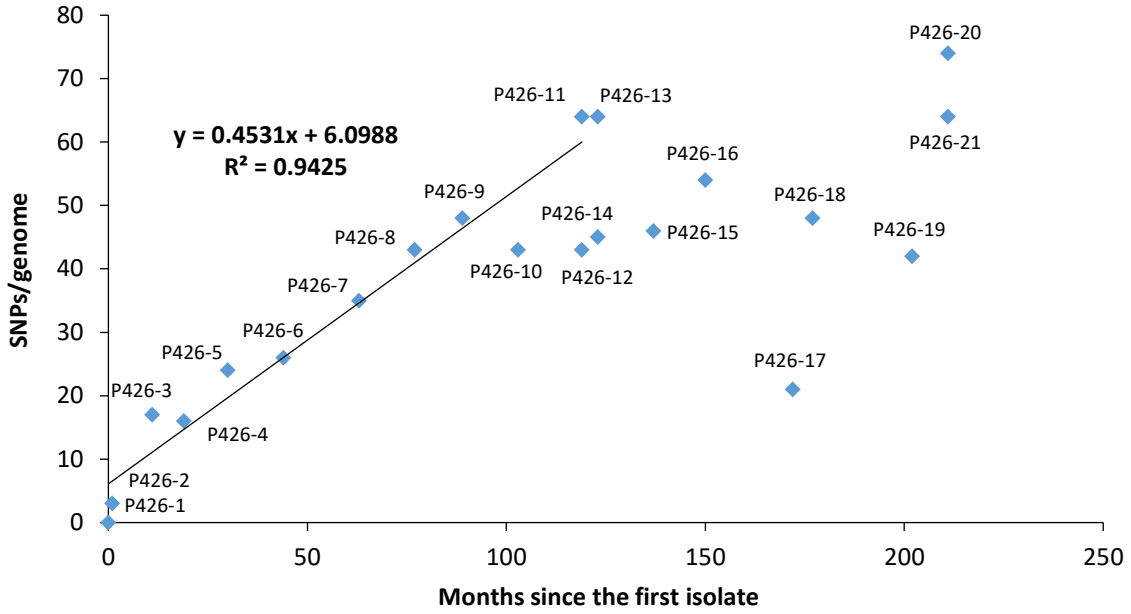

Figure 20 - Number of single-nucleotide polymorphisms (SNPs) distinguishing each *B. multivorans* isolate retrieved from patient P426 from the first isolate over time. A linear fit for isolates 1 to 11 is shown.

3.2.2.3. Analysis of large deletions

Several large deletions were identified among the isolates recovered from patient P426, with sizes between 0.508 and 217.7 kb (figure 21; table 11). It is important to note that 7 large deletions that were identified among the *B. multivorans* isolates were present in at least 5 isolates, the largest one including 209 genes (218-kb deletion present in 8 isolates). Further, isolates 5, 10, 11, 19 and 21 lacked contigs 4, 5 and 7 completely, likely corresponding to one of two plasmids that were identified for the isolates of P426. This loss of genomic portions shows a genome reduction that has been observed in similar studies with *B. multivorans*, *B. cenocepacia* and *P. aeruginosa* (Lee *et al.* 2017; Rau *et al.* 2012; Silva *et al.* 2016). Moreover, a large deletion of about 49 kb was found at the end of contig 1 (corresponding to chromosome 3) in isolates 7, 9, 11, 16 and 21, and in isolate 7 there was also a large deletion of about 156 kb at the start of this contig.

Main protein functions affected by each large deletion were analysed. In chromosome 2, isolate 21 has a 92.4-kb deletion that includes a gene encoding the diguanylate cyclase/phosphodiesterase RpfR (PROKKA_02109). RpfR is a cis-2-dodecenoic acid (BDSF) receptor that regulates quorum sensing, playing a role in regulation of biofilm formation, motility and virulence (Deng *et al.* 2012; Ryan *et al.* 2009). This gene was also deleted in isolate 15 of patient P213. In chromosome 3, large deletions comprised several genes involved in type IV secretion and pilus biosynthesis, associated with bacterial conjugation. Deletions including genes involved in type IV secretion are present in isolate 7 (156.2-kb and 109.1-kb deletions in chromosome 3), isolates 9 and 11 (109.1-kb and 217.7-kb deletions in chromosome 3), isolates 12, 14 to 16, 20 and 21 (217.7-kb deletion in chromosome 3) and isolate 13 (63.6-kb deletion in chromosome 2). Deletions including genes involved in pilus biosynthesis are present

in isolate 11 (89.4-kb and 109.1-kb deletions in chromosome 3), isolates 7 and 9 (109.1-kb deletion in chromosome 3) and isolates 5, 10, 19 and 21 (89.4-kb deletion in chromosome 3).

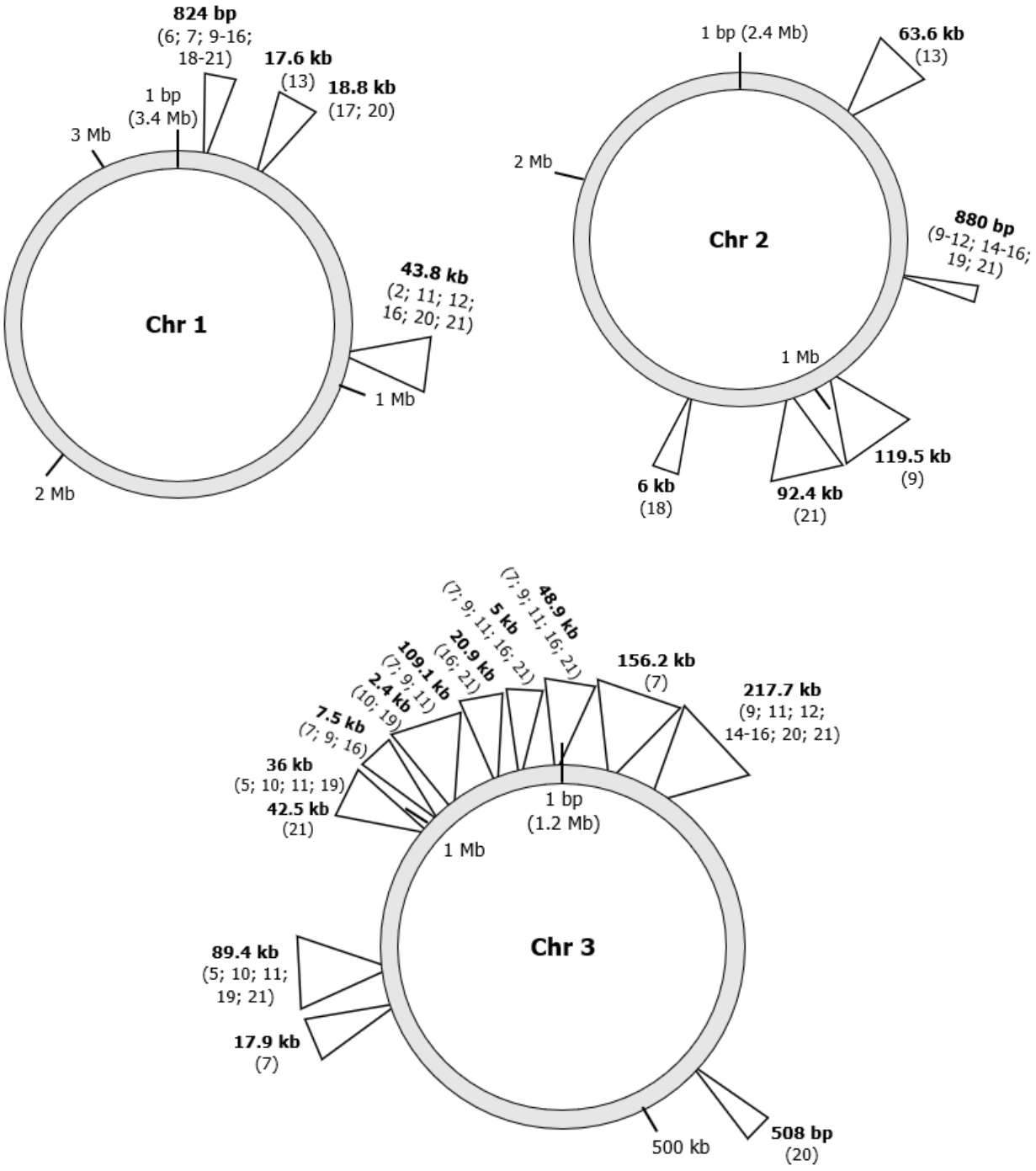

Figure 21 - Schematic representation of large deletions identified among the 21 *B. multivorans* isolates recovered from CF patient P426, divided by chromosomes 1 (Chr1), 2 (Chr2) and 3 (Chr3). Deletions are represented by triangles; isolates with each deletion and its size are also shown. Chromosome representation is not at scale.

Table 11 - Large deletions identified among the 21 *B. multivorans* isolates recovered from CF patient P426.

Chromosome	Start position	Deletion size (kb)	Locus tag	Isolates with mutation	Annotation / affected protein functions
Chr 1	68863	0.824	PROKKA_03313; PROKKA_03314	6; 7; 9-16; 18-21	Glutamine transport protein
Chr 1	269236	17.6	PROKKA_03494 to PROKKA_03505	13	Fructose and mannose metabolism
Chr 1	269236	18.8	PROKKA_03494 to PROKKA_03506	17; 20	Fructose and mannose metabolism
Chr 1	945339	43.8	PROKKA_04119 to PROKKA_04187	2; 11; 12; 16; 20; 21	Hypothetical proteins
Chr 2	262950	63.6	PROKKA_01353 to PROKKA_01422	13	Transport proteins; type IV secretion
Chr 2	691600	0.880	PROKKA_01734	9-12; 14-16; 19; 21	Outer membrane porin
Chr 2	979219	119.5	PROKKA_01981 to PROKKA_02095	9	Transcriptional regulators; sensor histidine kinases
Chr 2	1100890	92.4	PROKKA_02097 to PROKKA_02178	21	Diguanylate cyclase/phosphodiesterase RpfR; chaperone DnaK
Chr 2	1355298	6.1	PROKKA_02325 to PROKKA_02331	18	Transcriptional regulator
Chr 3	7203	156.2	PROKKA_00008 to PROKKA_00155	7	Type IV secretion; DNA replication
Chr 3	16531	217.7	PROKKA_00017 to PROKKA_00225	9; 11; 12; 14-16; 20; 21	Type IV secretion; DNA replication; LPS biosynthesis
Chr 3	450595	0.508	PROKKA_00422; PROKKA_00423	20	tRNA
Chr 3	855545	17.9	PROKKA_00785 to PROKKA_00800	7	Outer membrane porins
Chr 3	887178	89.4	PROKKA_00810 to PROKKA_00904	5; 10; 11; 19; 21	Pilus biosynthesis
Chr 3	995236	36.1	PROKKA_00918 to PROKKA_00953	5; 10; 11; 19	Hypothetical proteins
Chr 3	995236	42.5	PROKKA_00918 to PROKKA_00958	21	Hypothetical proteins
Chr 3	1030237	7.5	PROKKA_00954 to PROKKA_00958	7; 9; 16	Hypothetical proteins
Chr 3	1038909	2.4	PROKKA_00960; PROKKA_00961	10, 19	Hypothetical proteins
Chr 3	1038936	109.1	PROKKA_00960 to PROKKA_01069	7; 9; 11	Pilus biosynthesis; type IV secretion
Chr 3	1127241	20.9	PROKKA_01048 to PROKKA_01069	16; 21	Hypothetical proteins
Chr 3	1148967	5	PROKKA_01069 to PROKKA_01078	7; 9; 11; 16; 21	Hypothetical proteins
Chr 3	1155187	48.9	PROKKA_01080 to PROKKA_01126	7; 9; 11; 16; 21	Hypothetical proteins

3.2.2.4. Mutations distinguishing first and subsequent groups of isolates

As genes with mutations that became fixed over time are likely targets of evolution during chronic *B. multivorans* infection, a group of mutations that distinguished the latter isolates from the earlier isolates was selected. Table 12 shows a list of mutations that were found in at least 17 of the 21 isolates recovered from the patient. Genes coding for transporters possibly involved in resistance to antibiotics were found in the majority of the isolates (PROKKA_00501, PROKKA_04445 and PROKKA_02456). Two nonsynonymous mutations were found in genes associated with fatty acid biosynthesis, one in the acetyl-CoA carboxylase biotin subunit (PROKKA_05665) and the other in the enzyme 3-ketoacyl-ACP reductase (PROKKA_01182), indicating that alterations in lipid metabolism may be involved in adaptation of *B. multivorans* to the CF airways. A mutation in the gene encoding the acetyl-CoA carboxylase subunit alpha was found for patient P213, supporting this hypothesis. Gene PROKKA_01972, encoding a cyclic diguanosine monophosphate (c-di-GMP) phosphodiesterase had a nonsynonymous mutation in 18 isolates. This signaling molecule has been associated with regulation of biofilm formation, motility, virulence and adhesion (Boyd & O'Toole 2012; Hengge 2009; Traverse *et al.* 2013).

Table 12 - List of SNPs that exist in the majority of the isolates belonging to the infecting *B. multivorans* population sampled from CF patient P426. CDS, coding sequence; nonsyn, nonsynonymous; syn, synonymous.

Gene locus	Gene name	Category	Annotation	Mutations found	Effect in protein	Effect of mutation	Homolog in *B. multivorans* ATCC 17616	Isolates with mutation	Number of isolates with mutation
PROKKA_00467	*ytnP*	CDS	Putative quorum-quenching lactonase YtnP	T>G	S84R	nonsyn	-	2 - 21	20
PROKKA_00501	*pbuE*	CDS	Major facilitator transporter	G>T	L176M	nonsyn	Bmul_6097	2 - 21	20
PROKKA_01972	-	CDS	Cyclic-di-GMP phosphodiesterase	C>T	R145H	nonsyn	Bmul_5257	3 - 16; 18 - 21	18
PROKKA_02243	*gcvA*	CDS	LysR family transcriptional regulator	G>C	P114A	nonsyn	Bmul_4992	3 - 16; 18 - 21	18
PROKKA_04445	-	CDS	Acriflavin resistance protein	G>C	G452A	nonsyn	Bmul_1519	3 - 16; 18 - 21	18
PROKKA_04548	*hom*	CDS	Homoserine dehydrogenase	C>A	A121S	nonsyn	Bmul_1419	3 - 16; 18 - 21	18
PROKKA_05665	*accC*	CDS	Acetyl-CoA carboxylase biotin carboxylase subunit	G>A	D444N	nonsyn	Bmul_0367	3 - 16; 18 - 21	18
PROKKA_06251	*pilQ*	CDS	Type IV pilus secretin PilQ	G>T	L387L	syn	Bmul_0294	3 - 16; 18 - 21	18
PROKKA_01182	-	CDS	3-ketoacyl-ACP reductase	G>T	D219E	nonsyn	Bmul_3804	3 - 16; 18 - 21	18
PROKKA_02456	*potH*	CDS	Transport system permease protein	T>G	L225R	nonsyn	Bmul_5430	3 - 7; 9 - 16; 18 - 21	17
PROKKA_02456	*potH*	CDS	Transport system permease protein	C>A	L225L	syn	Bmul_5430	3 - 7; 9 - 16; 18 - 21	17

3.2.2.5. Phylogenetic analysis

To determine the evolutionary relationship among the 21 *B. multivorans* isolates, SNPs were used to construct a maximum likelihood phylogenetic tree (figure 22). The root of the tree was placed on the branch of isolate P426-1. The analysis groups the first two isolates together and shows that the first 5 isolates and isolate 17 are closely related. Isolate 17 is likely placed in this cluster due to its unexpected reduced number of SNPs. Isolate 20 seems to have a distant phylogenetic relation with the other isolates, which is likely a result of its increased number of SNPs, as this was the isolate with the highest number of mutations. It is important to note that isolates that were sampled from the patient on the same date do not seem to be closely related, as is observed for isolate pairs P426-11/P426-12, P426-13/P426-14 and P426-20/P426-21, indicating the coexistence of different lineages within this CF patient.

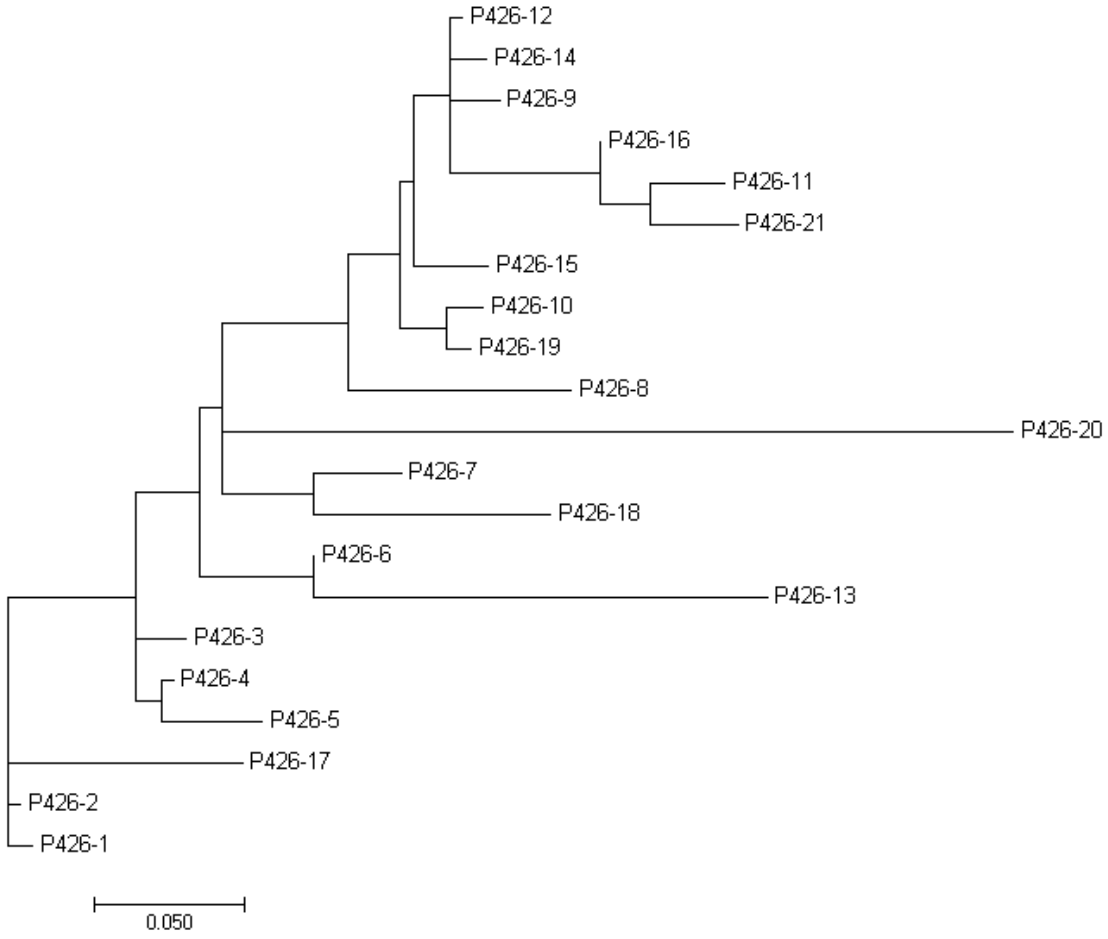

Figure 22 - Phylogenetic tree of the 21 *B. multivorans* isolates recovered from CF patient P426. The evolutionary history was inferred by using the maximum likelihood method based on the Tamura-Nei model (Tamura & Nei 1993). The tree with the highest log likelihood is shown and it is drawn to scale, with branch lengths measured in the number of substitutions per site (249 positions).

3.2.3. Analysis of lipopolysaccharide

To assess the production of the O-antigen portion of LPS by the 21 *B. multivorans* isolates recovered from patient P426, LPS was extracted and purified, and the resulting samples were resolved by electrophoresis. Results are presented in figure 23 and show that isolates 1 to 5, 7, 12, 14, 16, 18, 19 and 21 display the O-antigen repeats of LPS, while the remaining isolates lack this phenotypic trait.

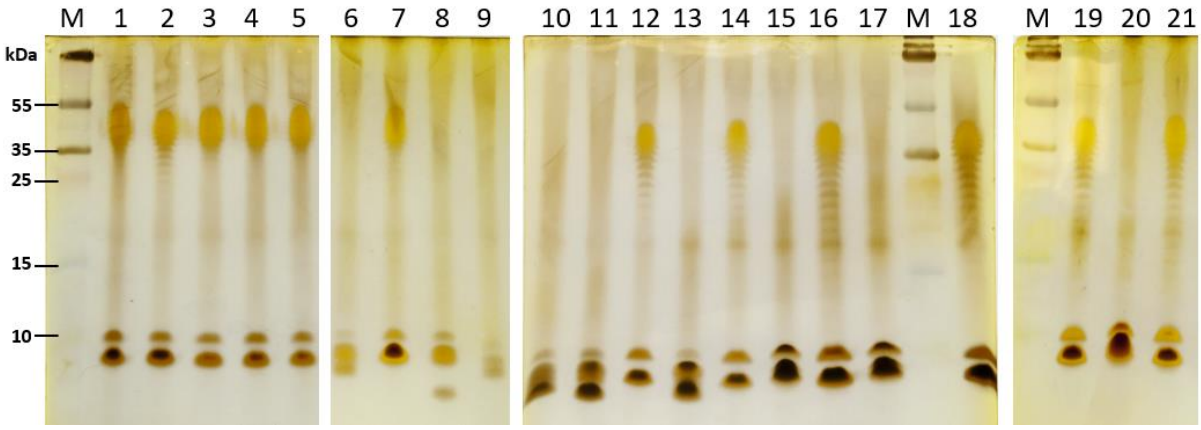

Figure 23 - Changes in the LPS pattern of the 21 *B. multivorans* isolates recovered from patient P426. Isolates 1 to 5, 7, 12, 14, 16, 18, 19 and 21 exhibit the O-antigen. Lanes M, protein markers.

Mutations in genes involved in LPS biosynthesis were identified in the comparative genomics analysis, but the absence of O-antigen could not be explained for all the isolates that lack this trait. Isolates 6 and 13 have a nonsynonymous SNP in the LPS core heptosyltransferase RfaQ; isolates 13, 17 and 20 have large deletions that include genes involved in the synthesis of GDP-D-rhamnose, the precursor of D-rhamnose, which is an O-antigen subunit commonly found in Bcc bacteria; and isolates 8 to 11 and 15 have a nonsynonymous SNP in a LPS assembly protein LptD precursor, but this mutation is also present in isolates that display the O-antigen.

3.2.4. Exopolysaccharide production

To evaluate the mucoid phenotype, the amount of exopolysaccharide produced by each isolate of *B. multivorans* sampled from CF patient P426 was assessed after 6 days of growth at 37°C. The recovered EPS concentration of the 21 isolates is presented in figure 24.

Isolates 10, 11, 12, 13 and 20 were the only ones able to produce detectable amounts of high molecular weight (HMW) EPS, with the recovered HMW EPS dry-weight ranging from approximately 3 to 6 g/L. Nevertheless, isolates 1, 2, 3, 4, 6, 9, 18 and 21 produced a precipitate that resembled the extracellular polymeric substance but consisted of small particles that were difficult to collect, which has already been described in the results for patient P213.

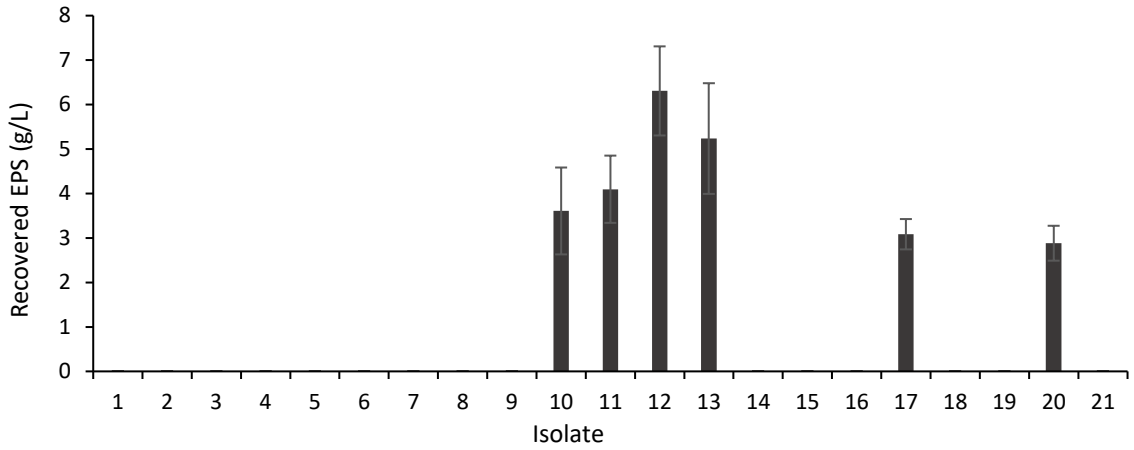

Figure 24 - Exopolysaccharide production of the 21 *B. multivorans* isolates sequentially retrieved from CF patient P426, based on the dry weight of ethanol-precipitated EPS after 6 days of growth at 37ºC (error bars represent the standard deviations of the mean values for three independent experiments). Only high molecular weight EPS was collected for this analysis.

The mucoid phenotype was also studied in solid medium (YEM, yeast extract medium), by visual inspection of the inoculated plates after 48 h of growth at 37°C. Isolates were classified as highly mucoid, mucoid and non-mucoid, and the results are presented in table 13. Figure 25 shows YEM plates of isolates 13, 17 and 19, corresponding to highly mucoid, mucoid and non-mucoid isolates, respectively (the plates of the remaining isolates are shown in supplementary figure A5).

Table 13 - Mucoid phenotype assessment in yeast extract mannitol medium (YEM) after 48 h of incubation at 37°C for the 21 *B. multivorans* isolates sequentially retrieved from CF patient P426. Isolates were classified as highly mucoid (++), mucoid (+) and non-mucoid (-) by visual inspection.

Isolate	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Mucoidy	++	++	++	++	++	++	++	++	-	++	++
Isolate	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	
Mucoidy	++	++	++	++	++	+	+	-	+	+	

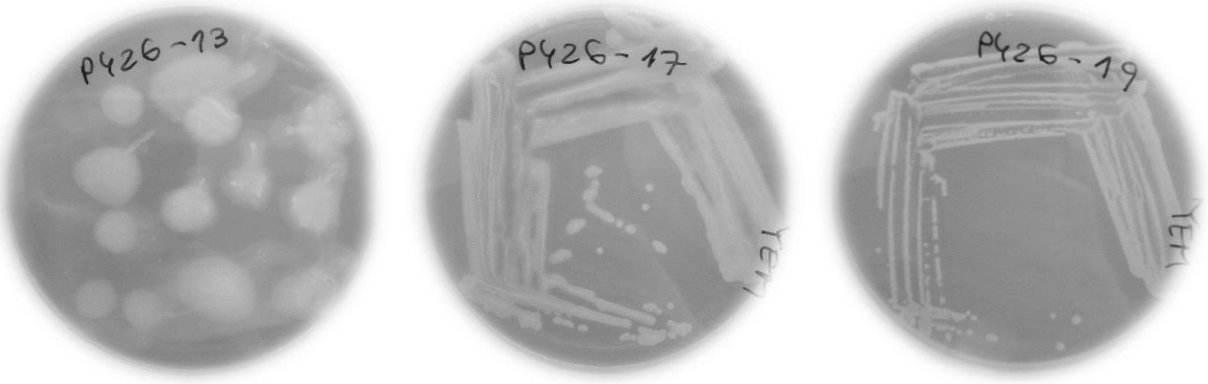

Figure 25 - Mucoid phenotype assessment in yeast extract mannitol medium (YEM) after 48 h of incubation at 37°C for *B. multivorans* isolates 13 (highly mucoid, ++), 17 (mucoid, +) and 19 (non-mucoid, -) of patient P426.

The first 16 isolates, apart from 9, displayed a higher mucoidy than the final 5 isolates. Isolate 9, along with isolate 19, was non-mucoid in YEM plates. Just as for patient P213, the results for the mucoid phenotype analysis in solid and liquid media are not in agreement, likely because different compositions were used for the solid and liquid media in this experiment.

Examining the list of mutations existing among the *B. multivorans* isolates, no mutations in genes belonging to the EPS biosynthetic gene cluster (*bce* gene cluster; Moreira *et al.* 2003) were identified.

3.2.5. Susceptibility against antimicrobials

The susceptibility of each *B. multivorans* isolate retrieved from patient P426 against antibiotics was tested, using piperacillin/tazobactam, ciprofloxacin, aztreonam and kanamycin to assess this phenotypic trait. The results regarding the susceptibility of the bacterial isolates to the antimicrobials are shown as the average measure of the diameter of the growth inhibition zone, represented in figure 26.

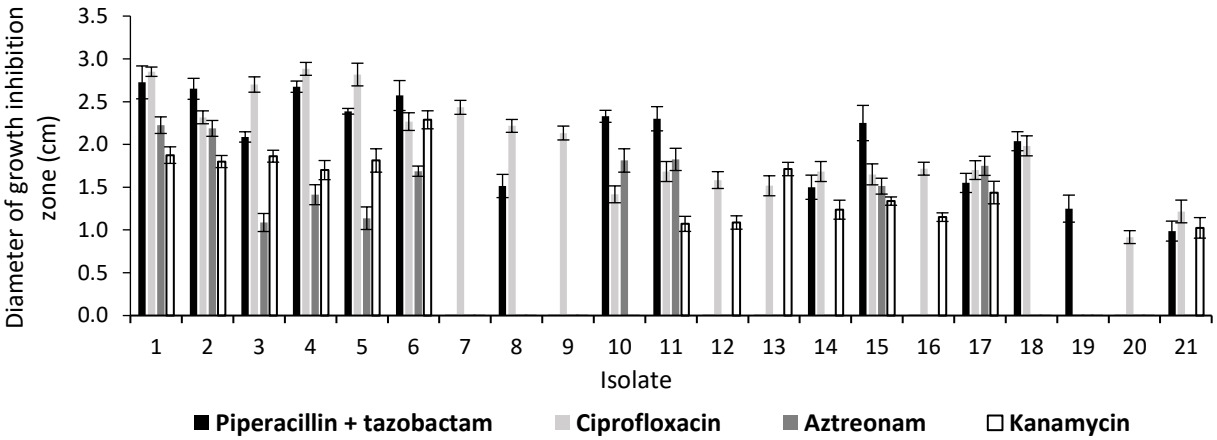

Figure 26 - Susceptibility of the 21 *B. multivorans* isolates sequentially retrieved from CF patient P426 to antibiotics (piperacillin/tazobactam, ciprofloxacin, aztreonam and kanamycin), measured as the diameter of cell growth inhibition, after growth for 24h at 37ºC (error bars represent the standard deviations of the mean values for three independent experiments).

Results show that the first two isolates are more susceptible to aztreonam than the subsequently retrieved isolates and that the first six isolates are generally more susceptible to the four antimicrobials used in this analysis than the latter isolates. Furthermore, isolates 7, 9, 19 and 20 are completely resistant to three of the four antibiotics and isolates 8, 12, 13, 16 and 18 showed complete resistance to two of the four tested antimicrobials. Statistical significance of differences between the first isolate and the subsequent ones was determined for susceptibility to aztreonam using ANOVA followed by Dunnett's multiple comparisons test, with extremely significant differences for all isolates (*P*-value < 0.0001; supplementary figure A6).

Mutations in genes encoding transport proteins that could be involved in resistance to antimicrobial compounds were identified in the genome analysis, such as genes encoding a major facilitator transporter (in isolates 2 to 21), an acriflavin resistance protein (in isolates 3 to 16 and 18 to

21), a transport system permease (in isolates 3 to 7, 9 to 16 and 18 to 21), a multidrug resistance protein (in isolates 8 to 12, 14 to 16, 19 and 21) and an efflux system outer membrane lipoprotein (in isolate 17).

3.2.6. Biofilm formation

Biofilm formation of the *B. multivorans* isolates sampled from CF patient P426 was assessed by measuring the absorbance at 590 nm (A_{590}) of cell suspensions grown in microtiter plates for 48h. Results are presented in figure 27 and show that the first two isolates produce a decreased amount of biofilm in comparison with the subsequent bacterial isolates, with the exception of isolate 18, for which no statistically significant differences in biofilm formation were found.

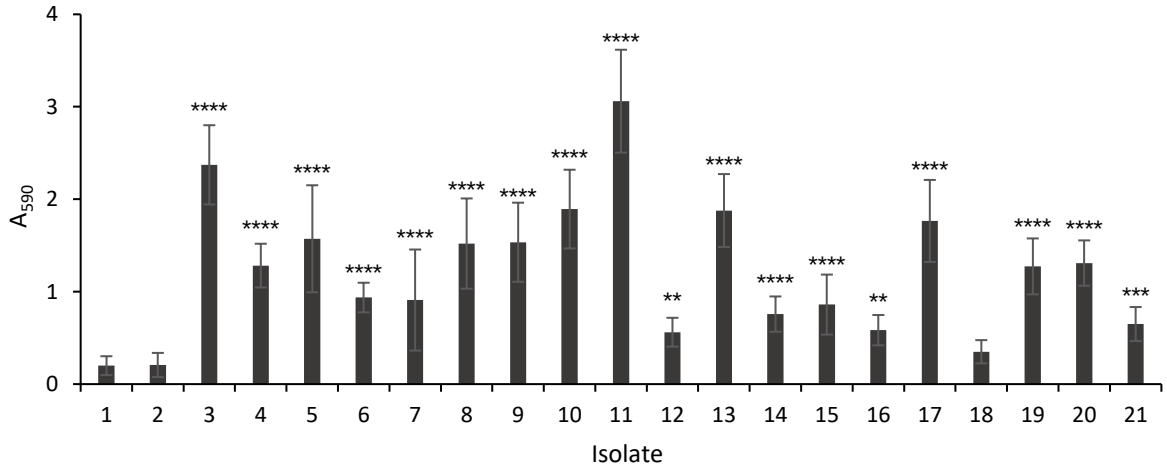

Figure 27 - Biofilm formation of the 21 *B. multivorans* isolates sequentially retrieved from CF patient P426, after growth in polystyrene microplates for 48h at 37ºC. Error bars represent the standard deviations of the mean values for at least three independent experiments (statistical significance of differences between the first isolate and the subsequent ones was determined: **, *P*-value < 0.01; ***, *P*-value < 0.001; ****, *P*-value < 0.0001 by ANOVA followed by Dunnett's multiple comparisons test).

Alterations in biofilm formation among the isolates could be due to the presence of a nonsynonymous mutation in a gene encoding a c-di-GMP phosphodiesterase in isolates 3 to 16 and 18 to 21. This signaling molecule is involved in regulation of biofilm formation, motility, virulence and adhesion (Boyd & O'Toole 2012; Hengge 2009; Traverse *et al.* 2013). Further, a SNP located 78 bp upstream the coding region of a fimbrial protein precursor was also present in isolates 3 to 16 and 18 to 21, possibly affecting biofilm formation ability. Given that these mutations were absent in isolate 17 and the biofilm formation displayed by this isolate was significantly higher than the one observed for the first two isolates, one mutation that could explain this phenotype is a nonsynonymous SNP in a *fixL* homolog that is also associated with regulation of biofilm formation, motility and virulence.

3.2.7. Swimming and swarming motility

To study the motility of the 21 *B. multivorans* isolates longitudinally sampled from patient P426, swimming and swarming agar plates were incubated for 24 h and 48 h, respectively, and the diameter of the motility zone was measured. Results are shown in figure 28 for swimming motility and no data are shown for swarming motility, since only isolate 1 was able to move in this medium (motility zone diameter = 1.64 cm; standard deviation = 0.33 cm; three independent experiments were performed). The first isolate clearly shows increased motility when compared with the remaining isolates, which display a much lower swimming motility and no swarming motility.

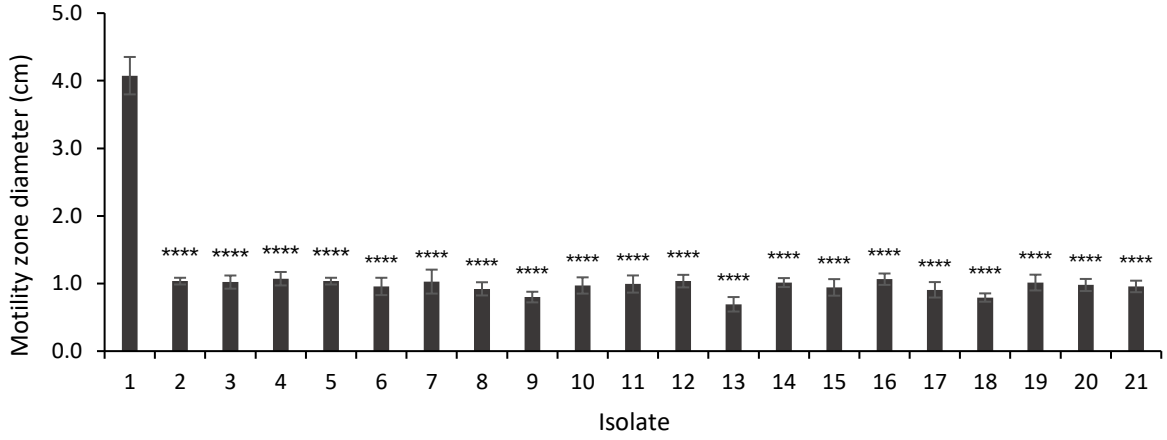

Figure 28 - Swimming motility of the 21 *B. multivorans* isolates sequentially retrieved from CF patient P426, after growth in swimming agar plates for 24 h at 37ºC. Error bars represent the standard deviations of the mean values for three independent experiments (statistical significance of differences between the first isolate and the subsequent ones was determined: ****, *P*-value < 0.0001 by ANOVA followed by Dunnett's multiple comparisons test).

Mutations involved in motility were identified in the comparative genomics analysis, namely in genes encoding flagellar proteins (FliJ in isolate 15 and FliI in isolate 17), the flagellar transcriptional regulator FlhC (isolate 6) and chemotaxis proteins (CheB in isolate 5 and CheA in isolate 20). Only two mutations were identified distinguishing the first isolate and the subsequent ones, in genes encoding a putative quorum-quenching lactonase and a MFS transporter. Lactonases degrade AHL molecules, thus inhibiting quorum sensing, which might influence bacterial motility. Additionally, the decrease in motility among the isolates could be explained by a mutation in a gene encoding a c-di-GMP phosphodiesterase in isolates 3 to 16 and 18 to 21, as c-di-GMP is involved in regulation of motility. Also, a mutation that could explain the observed motility of isolates 8 and 17 is a nonsynonymous SNP in a *fixL* homolog that could also regulate this phenotype.

3.2.8. Growth rate determination

Growth rates of the 21 isolates retrieved from patient P426 were measured by monitoring the OD_{640} of cultures grown in synthetic cystic fibrosis medium (SCFM) (Palmer *et al*. 2007). The doubling times calculated for each isolate are shown in table 14 and the corresponding graphs are in

supplementary figure A7. Results show that the first four isolates present lower doubling times than the remaining ones.

Table 14 - Doubling times (in hours) calculated for the 21 *B. multivorans* isolates sampled from CF patient P426, based on growth rates measured from cultures grown in synthetic cystic fibrosis medium (SCFM). Results are the means of data from two independent experiments.

Isolate	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Doubling time (h)	1.3	1.4	1.4	1.4	1.8	1.9	1.9	1.5	1.7	1.9	2.1
Isolate	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	
Doubling time (h)	2.1	1.9	2.0	2.1	2.1	2.1	2.1	2.0	2.1	2.2	

The genome analysis identified a mutation in a gene encoding a homoserine dehydrogenase, which could be involved in amino acid biosynthesis, in isolates 3 to 16 and 18 to 21 of this patient, possibly influencing bacterial growth. However, growth rates of isolates 3, 4 and 17 are not explained by this hypothesis. Two additional mutations with possible implications in growth were found in isolates 6 to 16 and 18 to 21 – a mutation in a gene encoding the cytochrome o ubiquinol oxidase subunit, involved in oxidative phosphorylation, and a mutation in a gene encoding the enzyme cobalt-precorrin--6A synthase, involved in cobalamin (vitamin B12) biosynthesis.

3.2.9. Adhesion to epithelial cells

To analyse adhesion of the *B. multivorans* isolates sampled from patient P426 to the bronchial epithelial cell line CFBE41o⁻, epithelial cells were infected at a multiplicity of infection (MOI) of 10 (10 bacterial cells to 1 epithelial cell) and results are shown in figure 29, as percentage of adhesion.

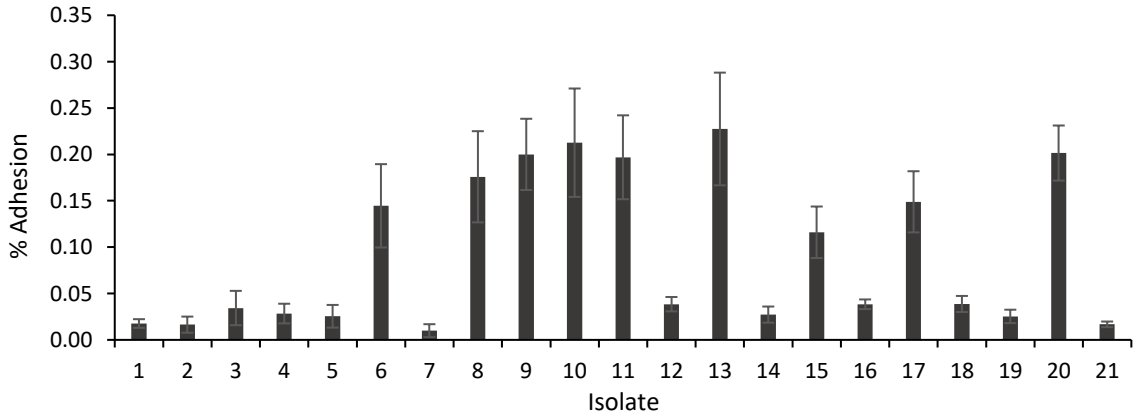

Figura 29 - Adhesion of the 21 *B. multivorans* isolates sequentially retrieved from CF patient P426 to the bronchial epithelial cell line CFBE41o⁻. Results are shown as the percentage of adhesion, which was calculated as the number of CFU recovered divided by the number of CFU applied to the epithelial cells multiplied by 100. Results are the mean values from two replicates from three independent experiments.

Results show a clear difference between isolates with O-antigen and isolates without this portion of the LPS. Isolates that display this phenotypic trait (1 to 5, 7, 12, 14, 16, 18, 19 and 21) show remarkably low values of adhesion to epithelial cells, whereas isolates that don't produce the LPS

O-antigen (6, 8 to 11, 13, 15, 17 and 20) present increased percent adhesion to the CFBE41o⁻ cell line. This distinction was also observed for CF patient P213 and O-antigen has been reported to inhibit adhesion to epithelial cells in *B. cenocepacia* (Saldías *et al.* 2009).

3.2.10. Virulence in *Galleria mellonella*

To assess virulence of the *B. multivorans* isolates recovered from CF patient P426, *Galleria mellonella* was used as an infection model. Triplicates of ten larvae were injected per isolate and survival was followed for three days post-infection. Results are shown in figure 30.

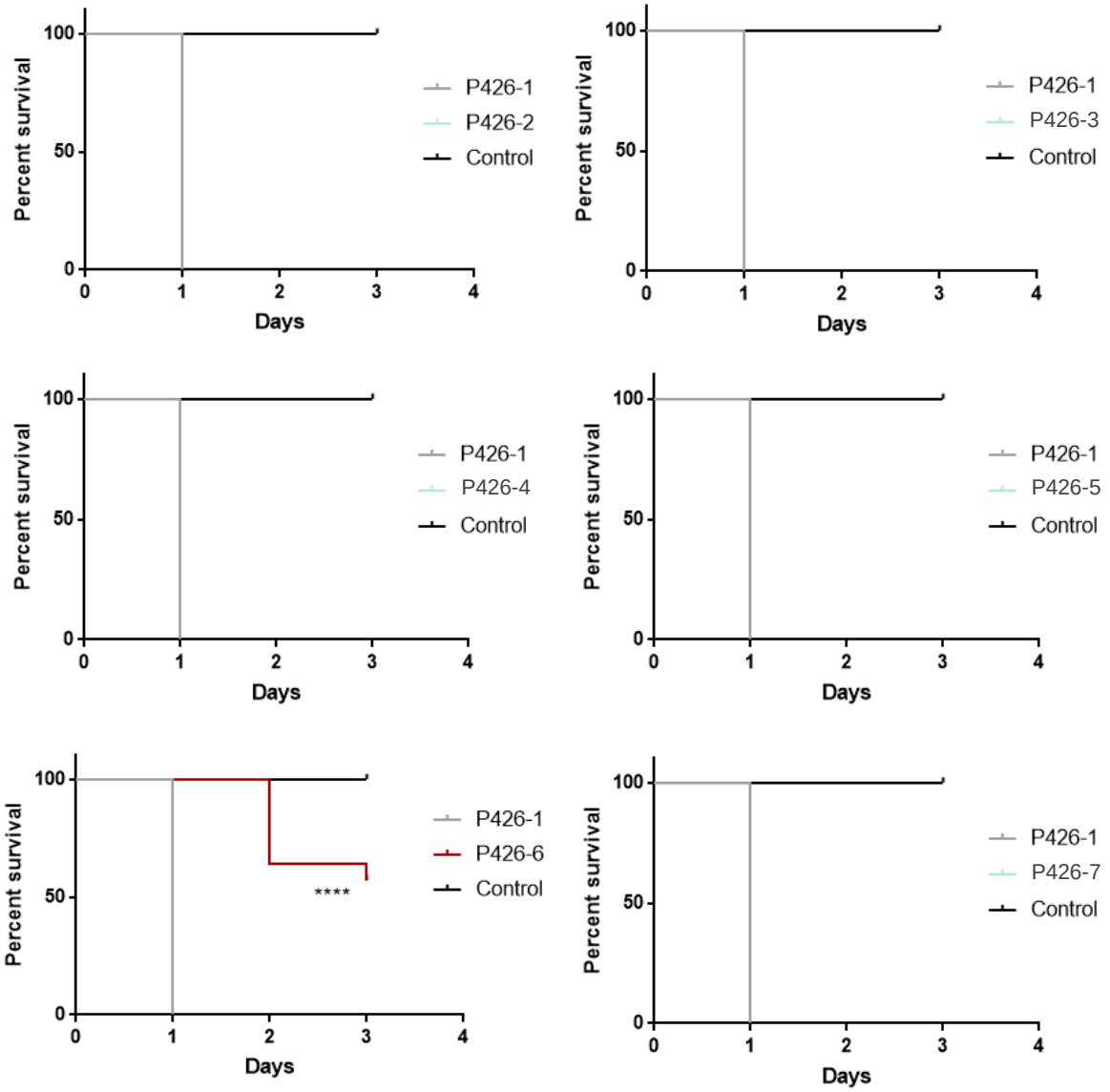

Figure 30 - Survival of *Galleria mellonella* larvae inoculated with *B. multivorans* isolates retrieved from CF patient P426. Triplicate groups of 10 larvae were inoculated with each isolate and survival was followed for three days post-infection. Larvae were injected with approximately $1x10^6$ bacterial cells. The control experiment without bacteria is also represented. Statistical significance of differences between the Kaplan-Meier curve of the first isolate and the subsequent ones was determined: ***, *P*-value < 0.001; ****, *P*-value < 0.0001.

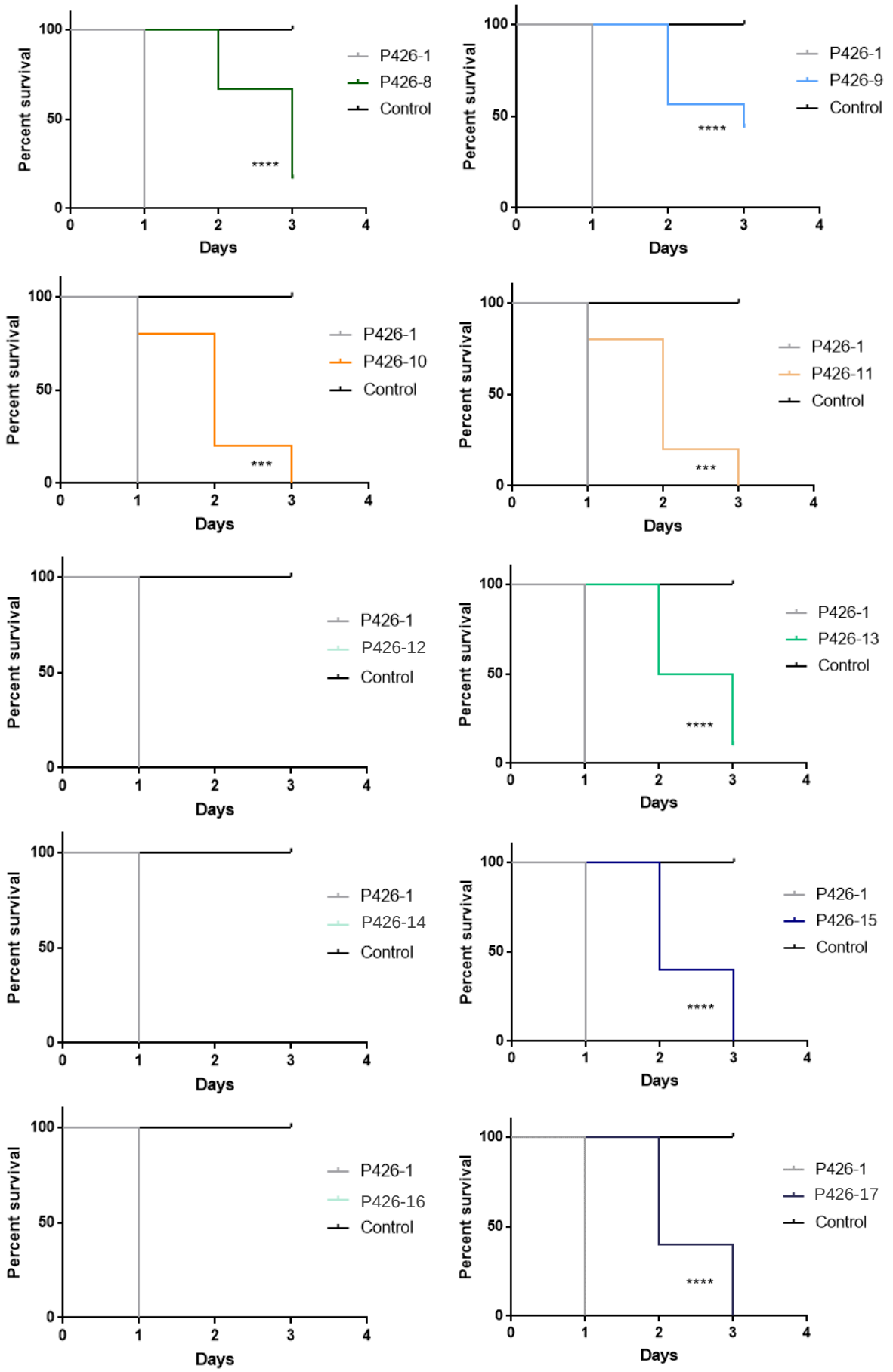

Figure 30 - Survival of *G. mellonella* larvae inoculated with *B. multivorans* isolates from CF patient P426 (cont.).

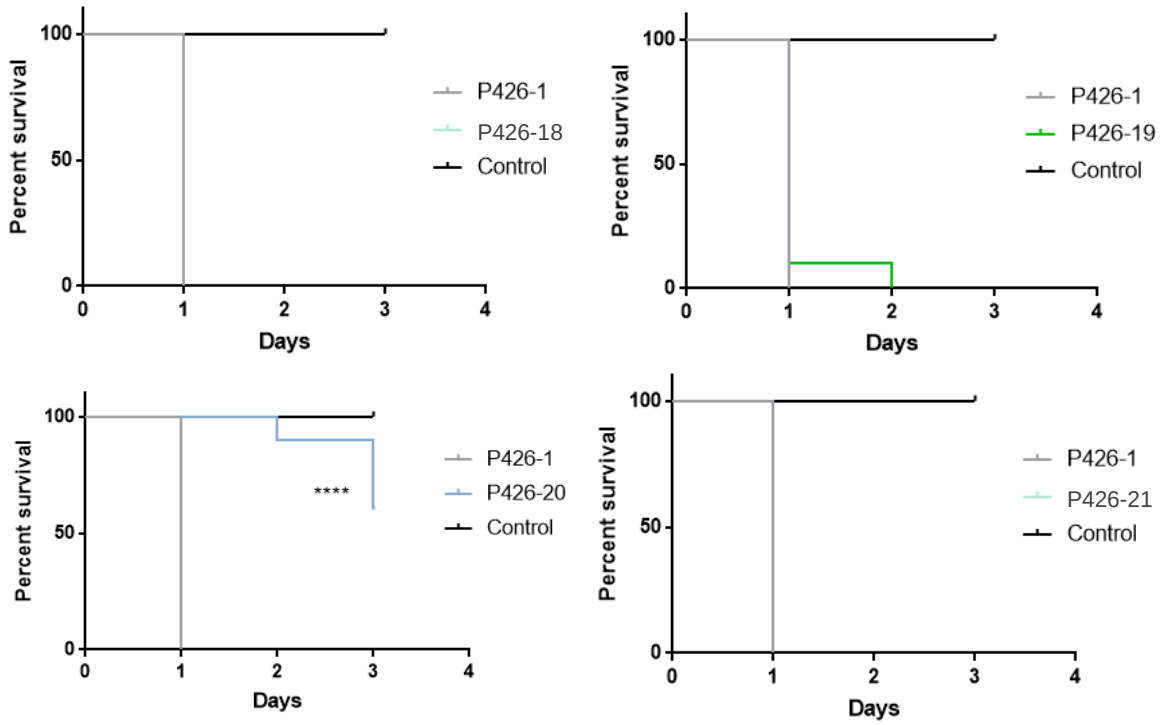

Figure 30 - Survival of *G. mellonella* larvae inoculated with *B. multivorans* isolates from CF patient P426 (cont.).

Results show that all the isolates that display the LPS O-antigen were more virulent than the ones without this portion of the LPS, being responsible for a 0% survival only 24 hours post-infection. This drastic decline in survival of larvae injected with bacteria that present the O-antigen was also observed in the experiments with isolates retrieved from CF patient P213. Although genes encoding regulators associated with virulence were found in isolates of patient P426 (*fixL* homolog and gene encoding c-di-GMP phosphodiesterase), acute virulence in *G. mellonella* seems to be determined by the presence of LPS O-antigen. Accordingly, this phenotypic trait has been shown to inhibit phagocytosis by macrophages, reducing bacterial elimination in *B. cenocepacia* (Saldías *et al.* 2009).

4. DISCUSSION

The aim of this work was to assess the genomic and phenotypic evolution of 37 *Burkholderia multivorans* isolates longitudinally sampled from two CF patients (P213 and P426) over 15 to 18 years of chronic infection. The strategy included a comparative genomics analysis, for the identification of important adaptive mutations among the isolates, and the analysis of phenotypes associated with adaptation of *B. cepacia* complex (Bcc) bacteria to the CF airways.

Detection of mutations that could influence adaptation of *B. multivorans* isolates to the CF airways was performed by comparing the genomes of the isolates with the first isolate for each patient, after sequencing and *de novo* assembly of the first isolates. To identify which contig generated by the assembly process corresponded to each *B. multivorans* chromosome, Basic Local Alignment Search Tool (BLAST, Altschul *et al.* 1990) was used to align each contig to the genome of *B. multivorans* ATCC 17616. This analysis also allowed the identification of one plasmid for the first isolate of patient P213 and two plasmids for the first isolate of patient P426.

Comparing the genomes of the 16 *B. multivorans* isolates recovered from patient P213 led to the identification of a total of 291 mutations, comprising 240 SNPs, 35 indels and 16 large deletions. Twelve polymorphic genes were found, all with mutations within the coding sequence. Two genes for which three different mutations were identified encode a TetR family transcriptional regulator and a Fis family transcriptional regulator conserved in the *Burkholderia* genus. The gene encoding a Fis family transcriptional regulator has unknown regulatory targets and it was recently found to have six mutations in *B. multivorans* isolates recovered from another chronically infected CF patient over 20 years (Silva *et al.* 2016). When evaluating the rate of accumulated SNPs per year, two distinct patterns were observed. While isolates 1 to 7 and 9 to 11 seem to have acquired mutations at a constant rate of about 8.2 SNPs per year, the set of the latter 5 isolates showed significantly less SNPs than expected if the linear tendency of the first set of isolates were followed. Assuming that bacterial isolates may have been collected from different regions of the airways of the CF patient, this difference in mutation acquisition could be due to the distinct selective pressures microorganisms encounter in those regions. The mutation rate of 8.2 SNPs per year for the first set of isolates is higher than the rate of about 2 SNPs per year observed for other long-term infections with *B. dolosa* and *B. multivorans* in CF patients (Lieberman *et al.* 2011; Silva *et al.* 2016). Isolate 8 was excluded from the mutation rate assessment due to the presence of a 48-bp deletion in the mismatch repair gene *mutL*, which could explain the increased number of mutations identified for this isolate (total of 98 mutations). Mutations in the *mutL* gene lead to a hypermutator phenotype, characterized by an increased mutation rate. Hypermutator phenotypes have been frequently reported for *P. aeruginosa* isolates retrieved from CF patients (Oliver *et al.* 2000; Smith *et al.* 2006; Hoboth *et al.* 2009) and also for Bcc isolates sampled from CF patients (Martina *et al.* 2014; Silva *et al.* 2016). Examining the genes comprised in the large deletions identified among the clinical isolates led to the identification of two important regulators of quorum sensing. Genes encoding the diguanylate cyclase/phosphodiesterase RpfR were deleted in isolate 15. This protein is the cis-2-dodecenoic acid (BDSF) receptor, being involved in c-di-GMP metabolism mainly as a c-di-GMP phosphodiesterase, and thus regulating biofilm formation, motility and virulence (Deng *et al.*

2012; Ryan *et al.* 2009). In a similar study of *B. multivorans* isolates chronically infecting a CF patient, a mutation in the gene encoding the RpfR protein fixed in the infecting population and was described as a strong candidate for governing global regulatory changes involved in adaptation to the CF host (Silva *et al.* 2016). The other quorum sensing system identified in this analysis was the CepRI system, as a gene encoding the CepR protein was deleted in isolate 8. CepR and CepI are the transcriptional regulator and the *N*-acyl-homoserine lactone (AHL) synthase, respectively, of the CepRI quorum sensing system, which regulates biofilm formation, motility and production of extracellular virulence factors (Huber *et al.* 2001; Lewenza *et al.* 1999). In *P. aeruginosa* clinical isolates sampled from CF patients, mutations in the *lasR* gene are often found (Marvig *et al.* 2014; Smith *et al.* 2006). LasR is a transcriptional regulator of quorum sensing and its inactivation reduces the expression of several virulence genes (Folkesson *et al.* 2012). Further, a selection of mutations that distinguished the latter 13 isolates of patient P213 from the first three was analysed and most of those mutations that became fixed in the population were nonsynonymous, indicating positive evolutionary selection. Relevant mutated genes may be involved in antibiotic resistance (genes encoding a MFS transporter and a multidrug efflux transporter), lipid metabolism (*accA* gene, encoding the acetyl-CoA carboxylase subunit alpha) and regulation of transcription targeting biofilm formation, motility and virulence (gene encoding a *fixL* homolog). In a similar study of *B. multivorans* isolates recovered from a chronically infected CF patient over 20 years, Silva and colleagues found four mutations in the gene encoding the acetyl-CoA carboxylase subunit alpha (Silva *et al.* 2016). A reduction of the degree of saturation of fatty acids produced by *B. cenocepacia* isolates infecting a CF patient had previously been associated with lung function deterioration (Coutinho *et al.* 2011). Furthermore, a homolog of *fixL* was found to be the most mutated gene in a study including 112 *B. dolosa* clinical isolates recovered from CF patients (Lieberman *et al.* 2011) and Silva and colleagues also found three mutations in the *B. multivorans fixL* homolog, which suggests that this gene could be an important global regulator in the adaptation of Bcc isolates to the CF environment.

The comparative genomics analysis of the 21 *B. multivorans* isolates recovered from patient P426 identified a total of 337 mutations, including 275 SNPs, 40 indels and 22 large deletions. Several polymorphic genes were found, with 24 genes presenting two different mutations and 5 genes presenting three mutations, all within the coding sequence. Among the polymorphic genes with three mutations were genes encoding a transport system permease protein, a LPS-assembly protein precursor and a lipid A deacylase precursor. Interestingly, two nonsynonymous mutations were present in a *fixL* homolog, which was identified as an important regulatory gene in the analysis of patient P213 and in similar studies (Lieberman *et al.* 2011; Silva *et al.* 2016). Moreover, two different nonsynonymous mutations were found in a gene encoding a Fis family transcriptional regulator that was also identified as a polymorphic gene with three mutations in the isolates of P213, suggesting a possible role for this regulator in the adaptation of *B. multivorans* isolates to the airways of CF patients. The mutation rate was assessed and isolates 1 to 11 seemed to gain mutations at a constant rate of about 5.4 SNPs per year, a value that is higher than the rate observed for other long-term infections with *B. dolosa* and *B. multivorans* in CF patients (Lieberman *et al.* 2011; Silva *et al.* 2016), as previously mentioned, but lower than the rate observed for the first set of isolates of patient P213. The remaining isolates presented less

SNPs than expected if following the linear tendency of the first group of isolates and this difference could be caused by the existence of distinct selective pressures in different areas of the airways. Isolate 17 showed a remarkably low number of mutations, with a total of 25 mutations. Further, it is important to note that several large deletions were found in *B. multivorans* isolates retrieved from this patient and seven of those were found in at least five isolates, the largest one including 209 genes (218-kb deletion in 8 isolates), showing a genome reduction that has been reported in other long-term bacterial evolution analyses with *B. multivorans*, *B. cenocepacia*, *B. pseudomallei* and *P. aeruginosa* (Lee *et al.* 2017; Price *et al.* 2013; Rau *et al.* 2012; Silva *et al.* 2016). Most of these large deletions were identified in contig 1, corresponding to chromosome 3 of *B. multivorans* ATCC 17616. Also, isolates 5, 10, 11, 19 and 21 completely lacked contigs 4, 5 and 7, likely corresponding to a plasmid that was lost over time in those isolates. Aiming to study the dynamics of genetic content alterations during host adaptation of *P. aeruginosa* to the lungs of CF patients, Rau and colleagues sequenced and analysed the genomes of 45 isolates sampled from 16 patients over 35 years, observing considerable genome reduction through deletions of large genomic regions (Rau *et al.* 2012). Genome reduction as an evolutionary strategy involves extensive gene loss affecting non-essential functions and has been related to adaptation towards permanent association with a host (Shigenobu *et al.* 2000). Interestingly, a gene encoding the diguanylate cyclase/phosphodiesterase RpfR, a regulator of biofilm formation, motility and virulence (Deng *et al.* 2012; Ryan *et al.* 2009), was deleted in isolate 21 and this gene was also deleted in isolate 15 of patient P213. Further analysis of the complete list of mutations reveals that several mutations were found in most of the isolates recovered from the patient, namely in genes coding for transporters possibly involved in resistance to antibiotics, in a gene encoding a c-di-GMP phosphodiesterase and in genes associated with fatty acid biosynthesis. One mutation was found in the acetyl-CoA carboxylase biotin subunit and the other in the enzyme 3-ketoacyl-ACP reductase, supporting the previously stated hypothesis that lipid metabolism may be involved in adaptation of *B. multivorans* to the airways of CF patients.

To determine the evolutionary relationship among the *B. multivorans* isolates of each patient, SNPs were used to construct maximum likelihood phylogenetic trees. For both patients, the analyses group the first two to five isolates together, separating them from the latter ones. Also, isolates that were collected on the same date fall into distinct areas of the trees, showing coexistence of different lineages within the same CF patient. Coexistence of distinct lineages in the same infection has also been observed in other long-term evolution studies with Bcc isolates from CF patients (Lee *et al.* 2017; Silva *et al.* 2016).

Regarding the phenotypical analyses performed in this work, one of the assessed phenotypes was the production of the LPS O-antigen, with the first three isolates sampled from CF patient P213 producing the O-antigen and the subsequently retrieved isolates lacking the ability to produce this portion of the LPS. LPS is considered a virulence factor produced by Bcc bacteria, as it seems to inhibit phagocytosis by macrophages and thus reduce bacterial elimination (Saldías *et al.* 2009). Nevertheless, the O-antigen can interfere with adhesion to abiotic surfaces and bronchial epithelial cells, and also stimulate the host immune response (Saldías *et al.* 2009), indicating that cells living in the form of

biofilms, like in most CF airways, could benefit from its loss. For the *B. multivorans* isolates retrieved from patient P426, there seem to be two different populations regarding this phenotypic trait, as isolates 1 to 5, 7, 12, 14, 16, 18, 19 and 21 display the O-antigen portion of LPS and the remaining isolates lack this trait. A few mutations in genes relating to LPS production were identified, namely in genes encoding LPS heptosyltransferases, LPS transport periplasmic protein LptA and LPS export system permease protein LptG. However, no mutations that could explain the absence of the LPS O-antigen in every isolate without this trait were found.

The mucoid phenotype of each *B. multivorans* isolate was analysed by the quantification of the dry weight ethanol-precipitated EPS in liquid medium and by visual inspection of mucoidy in solid medium. For patient P213, isolates 5, 7, 11, 12, 14 and 15 were the only ones able to produce detectable amounts of high molecular weight EPS in liquid medium. However, some of the isolates produced reduced amounts of a precipitate that consisted of small particles, which was not acknowledged as the typical EPS. One possible hypothesis is that the small particles could be repeating units that failed to be polymerized. Further analysis would need to be conducted in order to identify this precipitate. Regarding mucoidy in solid medium for this CF patient, the first seven isolates displayed a higher mucoidy than the remaining ones, and isolates 10, 12 and 15 were non-mucoid. For patient P426, isolates 10, 11, 12, 13 and 20 produced detectable amounts of high molecular weight EPS in liquid medium and isolates 1, 2, 3, 4, 6, 9, 18 and 21 produced the small particle precipitate that was described for patient P213. Mucoidy inspection in solid medium showed that the first 16 isolates of this patient, except for 9, displayed a higher mucoidy than the final 5 isolates. Isolates 9 and 19 were non-mucoid. For both patients, the results for the mucoid phenotype analysis in the two media are not in agreement, as different compositions were used for the solid and liquid media in this experiment. EPS is considered a virulence factor of Bcc bacteria, as it has been demonstrated to inhibit neutrophil chemotaxis, to neutralize reactive oxygen species *in vitro* (Bylund *et al.* 2006), to affect the phagocytosis of bacteria by human neutrophils and to facilitate persistent bacterial infection in mice (Conway *et al.* 2004). In *P. aeruginosa*, it is well established that longitudinally retrieved isolates from CF patients undergo a non-mucoid to mucoid conversion, which generally occurs due to mutations in the *mucA* gene that ultimately lead to an excessive production of alginate (Jones *et al.* 2010; Rau *et al.* 2010; Marvig *et al.* 2014). Contrarily, Bcc isolates sequentially sampled from CF patients have been shown to undergo a transition from a mucoid phenotype to a non-mucoid one (Zlosnik *et al.* 2008; Silva *et al.* 2011). The results for mucoidy analysis in solid medium for the two patients studied in this work are in agreement with a decrease in mucoidy over time. Nevertheless, it isn't clear whether EPS production associates with increased virulence or persistence in the lungs, as studies have suggested that the non-mucoid morphotype is associated with increased disease severity for patients and the mucoid morphotype with persistence in the lungs (Zlosnik *et al.* 2011), and others demonstrated that mucoid variants are more virulent in the acute infection model *Galleria mellonella* and non-mucoid variants are related to persistence in the CF airways (Silva *et al.* 2011). To study the virulence of the *B. multivorans* isolates under study in this work, virulence assays using the infection model *G. mellonella* were performed and will be discussed further on.

Concerning susceptibility to antibiotics, the first three isolates of patient P213 were highly susceptible to aztreonam and kanamycin, with the subsequent isolates retrieved from the patient displaying increased resistance to these antimicrobials. Regarding susceptibility to piperacillin/tazobactam and to ciprofloxacin, two populations seem to have evolved differently, with some isolates remaining susceptible and others gaining resistance. During long-term chronic CF infections, bacteria commonly become more resistant to antibiotics (Coutinho *et al.* 2011; Lieberman *et al.* 2011; Silva *et al.* 2016), as frequent and prolonged exposure to antimicrobials promotes the emergence of strains with improved ability to survive in the challenging CF environment. For the *B. multivorans* isolates recovered from patient P426, the first two isolates were more susceptible to aztreonam than the subsequently retrieved isolates and the first six isolates were generally more susceptible to the four antimicrobials used in this analysis than the latter isolates. Moreover, four isolates were completely resistant to three of the four antibiotics and five isolates showed complete resistance to two of the four tested antimicrobials. To draw further conclusions about antibiotic resistance, information regarding the administration of antibiotics to the patients under study should be acquired, such as the time points at which antimicrobials were administrated and which antimicrobials were included in the treatment. The comparative genomics analysis led to the identification of mutations distinguishing the earlier isolates of each patient from the latter ones in genes encoding major facilitator superfamily (MFS) and multidrug efflux transporters, which could underlie the increase in resistance to antibiotics observed for the two CF patients.

The production of biofilms was also assessed in this work, as alterations in this phenotypic trait are frequently reported for Bcc isolates chronically infecting CF patients. For patient P213, the first three isolates showed a lower biofilm formation than the subsequently retrieved isolates (except for isolates 6 and 7, for which no statistically significant differences in biofilm production were found). Living within biofilms, bacteria are protected from environmental stresses, such as the host immune response and the presence of antibiotics, which are common stress sources in the CF airways. It was demonstrated that bacteria within biofilms are more resistant to antibiotics (Caraher *et al.* 2006) and the observations that P213 isolates 4 to 16 produced more biofilm and showed increased resistance to two of the tested antibiotics is in agreement with these findings. A nonsynonymous SNP in a *fixL* homolog was identified in the latter 13 isolates of this patient. In *B. dolosa*, *fixL* encodes a sensor histidine kinase of a two-component regulatory system (fixLJ) that is activated by low oxygen levels and regulates biofilm formation, motility, virulence and intracellular invasion (Schaefers *et al.* 2017). Concerning biofilm formation of the isolates recovered from patient P426, the first two isolates produced a decreased amount of biofilm in comparison with the subsequent bacterial isolates (except for isolate 18, for which no statistically significant differences in biofilm production were found). This observation also agrees with bacteria living in biofilms showing increased resistance to antimicrobials. A nonsynonymous mutation in a gene encoding a c-di-GMP phosphodiesterase was found in 18 isolates of this patient (isolates 3 to 16 and 18 to 21). This signaling molecule is related to regulation of biofilm formation, motility, virulence and adhesion (Boyd & O'Toole 2012; Hengge 2009; Traverse *et al.* 2013). While this mutation was absent in isolate 17, biofilm formation displayed by this isolate was significantly higher

than the one observed for the first two isolates, which may be explained by the presence of a nonsynonymous SNP in a *fixL* homolog that was mentioned above for patient P213.

Differences in motility were evaluated in swimming and swarming agar plates. Regarding motility of the *B. multivorans* isolates sampled from patient P213, the first three isolates showed substantially increased swimming and swarming motility in comparison with the subsequently retrieved isolates. This decrease in motility may be a method of bacteria to adapt to the respiratory tract environment of patients with CF and has been frequently reported in members of the Bcc (Coutinho *et al.* 2011; Silva *et al.* 2016). These results also confirm the inverse correlation between biofilm formation and motility that is commonly observed. For patient P426, the first isolate displays increased motility when compared with the remaining isolates, which show a much lower swimming motility and no swarming motility. Relevant mutations possibly underlying this phenotypic transition were previously referred for biofilm formation, as sensor histidine kinase *fixL* and c-di-GMP phosphodiesterase have been associated with regulation of motility as well as biofilm production (Boyd & O'Toole 2012; Schaefers *et al.* 2017). Also, one mutation in a putative quorum-quenching lactonase, present in all isolates of patient P426 except the first one, could influence this phenotypic trait by affecting quorum sensing. Moreover, several mutations in genes associated with motility were identified in isolates of both patients, such as genes encoding flagellar proteins (*fliI*, *fliJ*, *fliL*, *flgE*, *flgI*), a flagellar transcriptional activator (*flhC*) and proteins related to chemotaxis (*cheA*, *cheB*), supporting the hypothesis that alterations in this phenotypic trait are involved in adaptation of Bcc bacteria to the airways of CF patients.

Growth rate of each bacterial strain was measured in synthetic cystic fibrosis medium (SCFM; Palmer *et al*. 2007). For patient P213, results show that the first seven isolates, except for isolate 4, show decreased doubling times when compared with the latter ones. Regarding patient P426, the first four isolates presented lower doubling times than the remaining ones. These findings suggest a possible role for decreasing growth rate in adaptation of *B. multivorans* isolates to the airways of CF patients during infection. This tendency to reduce growth rates over time during long-term chronic infections in CF patients has been reported both for *P. aeruginosa* and Bcc bacteria (Lee *et al.* 2017; Rau *et al.* 2010; Silva *et al.* 2016), and it might occur in response to the different nutrient composition of CF airways or as a consequence of a greater commitment to biofilm production. Mutations in genes involved in amino acid biosynthesis were identified in isolates of both patients, but no mutations that could explain variations in all the isolates were found.

Another phenotypic trait analysed in this work was adhesion of *B. multivorans* isolates to the bronchial epithelial cell line CFBE41o$^-$, derived from a CF patient homozygous for the CFTR F508del mutation. For both patients, results showed a clear difference between isolates with O-antigen and isolates without this portion of the LPS. Isolates that display this trait showed reduced values of adhesion to epithelial cells, while isolates that don't produce the LPS O-antigen presented increased percent adhesion to the CFBE41o$^-$ cell line. In *B. cenocepacia*, the O-antigen inhibits adhesion to epithelial cells (Saldías *et al.* 2009), possibly by concealing surface molecules that influence the adhesion process. Despite mutations possibly affecting c-di-GMP signaling, associated with regulation of adhesion, the presence of LPS O-antigen seems to be the cause of the observed changes in this phenotype.

To study virulence of the *B. multivorans* isolates, *Galleria mellonella* was used as an infection model, as it presents an innate immune system with a high degree of structural and functional homology to the innate immune systems of mammals (Seed & Dennis 2008). For both patients, the isolates that display the LPS O-antigen were more virulent than the ones without this portion of the LPS, being responsible for a 0% survival only 24 hours post-infection. Nonetheless, this drastic decline in survival of larvae was also observed for isolate 13 from patient P213, which lacks the LPS O-antigen. LPS has been shown to inhibit phagocytosis by macrophages, reducing bacterial elimination in *B. cenocepacia* (Saldías *et al.* 2009). Regardless of mutations in a gene encoding a c-di-GMP phosphodiesterase, which is associated with regulation of virulence (Boyd & O'Toole 2012), virulence among the isolates of P426 seems to be determined by the presence of the O-antigen. For patient P213, a mutation in *fixL* distinguishes the first three isolates from the latter ones, and this global regulator is also associated with virulence in Bcc (Schaefers *et al.* 2017). However, given that the first three isolates display the LPS O-antigen and the latter ones lack this trait, it isn't clear if the decreased virulence of these isolates is due to the *fixL* mutation or the absence of O-antigen. No mutation that could explain the increased virulence of isolate 13 from patient P213 was identified. In a similar analysis, increased virulence of *B. multivorans* isolates also correlated with the presence of LPS O-antigen (Ramires 2017). Further, similar studies have reported decreases in acute virulence during long-term CF infections with Bcc bacteria (Lee *et al.* 2017), possibly linked to the development of a persistent infection.

In conclusion, the latter sets of isolates longitudinally collected from the CF patients displayed increased resistance to antibiotics, increased biofilm formation, decreased motility and decreased growth rates, shifting the bacteria toward a 'chronic infection phenotype' that is well established for *P. aeruginosa* infections and seems to apply to infections caused by Bcc bacteria as well. Unlike *P. aeruginosa* isolates, which undergo a non-mucoid to mucoid transition, Bcc isolates often convert from a mucoid morphotype into a non-mucoid one, with results of mucoidy in solid medium reported here supporting a decrease in mucoidy over time. Virulence in *G. mellonella* and adhesion to epithelial cells were apparently determined by the presence of the LPS O-antigen, with isolates that display this trait showing increased virulence and decreased adhesion ability. Mutations in genes related to resistance to antimicrobials, lipid metabolism and regulation of transcription likely drive the transition to the 'chronic infection phenotype', and regulatory genes might play important roles in adaptation of Bcc bacteria to the CF environment. For *P. aeruginosa*, genes *mucA* and *lasR* have been identified as important global regulators, with *mucA* showing particularly extensive pleiotropic effects on gene expression and leading to changes in central metabolism, motility, production of virulence factors and membrane permeability (Rau *et al.* 2010; Smith *et al.* 2006). In the present work, the *fixL* gene seems to be a possible candidate for regulating traits that impact Bcc bacteria adaptation to the host. This gene has been identified as an important regulator in similar studies with *B. multivorans* and *B. dolosa* (Lieberman *et al.* 2011; Silva *et al.* 2016), supporting this hypothesis. Furthermore, a gene encoding a Fis family transcriptional regulator conserved in the *Burkholderia* genus was found to be polymorphic in isolates from both patients. This gene has unknown regulatory targets and it was found to have six mutations in *B. multivorans* isolates recovered from another chronically infected CF patient over 20 years (Silva *et al.* 2016), suggesting a possible role for this regulator in adaptation of *B. multivorans* to the CF airways. Additionally, deletions

including genes encoding quorum sensing proteins RpfR and CepR were identified in this analysis. The quorum sensing regulators RpfR and LasR have been suggested as good candidates for global regulators of adaptation to the CF host in studies with *B. multivorans* and *P. aeruginosa*, respectively.

The study of evolutionary paths of bacteria in specific environments is being increasingly used as an investigative strategy, as this type of experiments offers knowledge about adaptive changes that take place during evolution and allows the comparison of different evolutionary trajectories. It is expected that the combination of phenotypic assessments with information provided by comparative genomics analysis of Bcc isolates sampled from CF patients will help elucidate which mutations may underlie adaptation to the CF airways. The identification of adaptive strategies during bacterial colonization of the airways of individuals with CF may point to potential therapeutic targets for the development of improved treatments against Bcc bacteria.

5. REFERENCES

Altschul, S.F. *et al.*, 1990. Basic Local Alignment Search Tool. *Journal of Molecular Biology*, 215, pp.403–410.

Alvarez-Ortega, C. & Harwood, C.S., 2007. Responses of *Pseudomonas aeruginosa* to low oxygen indicate that growth in the cystic fibrosis lung is by aerobic respiration. *Molecular Microbiology*, 65, pp.153–165.

Amaral, M.D. & Farinha, C.M., 2013. Rescuing mutant CFTR : A multi-task approach to a better outcome in treating cystic fibrosis. *Current Pharmaceutical Design*, 19(19), pp.3497–3508.

Aris, R.M. *et al.*, 2001. Lung transplantation for cystic fibrosis patients with *Burkholderia cepacia* complex. *American Journal of Respiratory and Critical Care Medicine*, 164(11), pp.2102–2106.

Aronesty, E., 2013. Comparison of Sequencing Utility Programs. *The Open Bioinformatics Journal*, 7, pp.1–8.

Bauer, A.W. *et al.*, 1966. Antibiotic susceptibility testing by a standardized single disk method. *American Journal of Clinical Pathology*, 45(4), pp.493–496.

Bell, S.C., De Boeck, K. & Amaral, M.D., 2015. New pharmacological approaches for cystic fibrosis: Promises, progress, pitfalls. *Pharmacology and Therapeutics*, 145, pp.19–34.

Bjarnsholt, T. *et al.*, 2009. *Pseudomonas aeruginosa* biofilms in the respiratory tract of cystic fibrosis patients. *Pediatric Pulmonology*, 44(6), pp.547–558.

Bolger, A.M., Lohse, M. & Usadel, B., 2014. Genome analysis Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), pp.2114–2120.

Bosi, E. *et al.*, 2015. Genome analysis MeDuSa: a multi-draft based scaffolder. *Bioinformatics*, 31(September 2017), pp.2443–2451.

Boucher, R.C., 2004. New concepts of the pathogenesis of cystic fibrosis lung disease. *European Respiratory Journal*, 23, pp.146–158.

Boyd, C.D. & O'Toole, G.A., 2012. Second Messenger Regulation of Biofilm Formation: Breakthroughs in Understanding c-di-GMP Effector Systems. *Annual Review of Cell and Developmental Biology*, 28(1), pp.439–462.

Burns, J.L. *et al.*, 2001. Longitudinal assessment of *Pseudomonas aeruginosa* in young children with cystic fibrosis. *Journal of Infectious Diseases*, 183(3), pp.444–452.

Bylund, J. *et al.*, 2006. Exopolysaccharides from *Burkholderia cenocepacia* inhibit neutrophil chemotaxis and scavenge reactive oxygen species. *Journal of Biological Chemistry*, 281(5), pp.2526–2532.

Caraher, E. *et al.*, 2006. Comparison of antibiotic susceptibility of *Burkholderia cepacia* complex organisms when grown planktonically or as biofilm in vitro. *European Journal of Clinical Microbiology and Infectious Diseases*, 26(3), pp.213–216.

Chaparro, C. *et al.*, 2001. Infection with *Burkholderia cepacia* in cystic fibrosis outcome following lung transplantation. *American Journal of Respiratory and Critical Care Medicine*, 163, pp.43–48.

Chmiel, J.F. & Davis, P.B., 2003. State of the art: why do the lungs of patients with cystic fibrosis become infected and why can't they clear the infection? *Respiratory Research*, 4(1), p.8.

Coenye, T. *et al.*, 2001. Taxonomy and identification of the *Burkholderia cepacia* complex. *Journal of Clinical Microbiology*, 39(10), pp.3427–3436.

Conway, B.D. *et al.*, 2004. Production of exopolysaccharide by *Burkholderia cenocepacia* results in altered cell-surface interactions and altered bacterial clearance in mice. *Journal of Infectious Diseases*, 190(5), pp.957–966.

Coutinho, C.P. *et al.*, 2011. *Burkholderia cenocepacia* phenotypic clonal variation during a 3.5-year colonization in the lungs of a cystic fibrosis patient. *Infection and Immunity*, 79(7), pp.2950–2960.

Cunha, M. V. *et al.*, 2004. Studies on the involvement of the exopolysaccharide produced by cystic fibrosis-associated isolates of the *Burkholderia cepacia* complex in biofilm formation and in persistence of respiratory infections. *Journal of Clinical Microbiology*, 42(7), pp.3052–3058.

Cystic Fibrosis Foundation US, 2016. *Cystic Fibrosis Foundation Patient Registry 2015 Annual Data Report. Bethesda, Maryland*,

Daniels, T.W. V *et al.*, 2013. Impact of antibiotic treatment for pulmonary exacerbations on bacterial diversity in cystic fibrosis. *Journal of Cystic Fibrosis*, 12(1), pp.22–28.

Deng, Y. *et al.*, 2012. Cis-2-dodecenoic acid receptor RpfR links quorum-sensing signal perception with regulation of virulence through cyclic dimeric guanosine monophosphate turnover. *Proceedings of the National Academy of Sciences*, 109, pp.15479–15484.

Depoorter, E. *et al.*, 2016. *Burkholderia*: an update on taxonomy and biotechnological potential as antibiotic producers. *Applied Microbiology and Biotechnology*, 100(12), pp.5215–5229.

Döring, G., Parameswaran, I.G. & Murphy, T.F., 2011. Differential adaptation of microbial pathogens to airways of patients with cystic fibrosis and chronic obstructive pulmonary disease. *FEMS Microbiology Reviews*, 35(1), pp.124–146.

Eberl, L. & Tümmler, B., 2004. *Pseudomonas aeruginosa* and *Burkholderia cepacia* in cystic fibrosis: Genome evolution, interactions and adaptation. *International Journal of Medical Microbiology*, 294(2–3), pp.123–131.

Etherington, C. *et al.*, 2014. The role of respiratory viruses in adult patients with cystic fibrosis receiving intravenous antibiotics for a pulmonary exacerbation. *Journal of Cystic Fibrosis*, 13(1), pp.49–55.

Farinha, C.M. & Matos, P., 2016. Repairing the basic defect in cystic fibrosis - One approach is not enough. *FEBS Journal*, 283(2), pp.246–264.

Ferreira, A.S. *et al.*, 2011. Insights into the role of extracellular polysaccharides in *Burkholderia* adaptation to different environments. *Frontiers in Cellular and Infection Microbiology*, 1, pp.1–16.

Ferreira, A.S. *et al.*, 2015. The tyrosine kinase BceF and the phosphotyrosine phosphatase BceD of *Burkholderia contaminans* are required for efficient invasion and epithelial disruption of a cystic fibrosis lung epithelial cell line. *Infection and Immunity*, 83(2), pp.812–821.

Filkins, L.M. & O'Toole, G.A., 2015. Cystic fibrosis lung infections: Polymicrobial, complex, and hard to treat. *PLoS Pathogens*, 11(12), pp.1–8.

Flanagan, J.L. *et al.*, 2007. Loss of bacterial diversity during antibiotic treatment of intubated patients colonized with *Pseudomonas aeruginosa*. *Journal of Clinical Microbiology*, 45(6), pp.1954–1962.

Folkesson, A. *et al.*, 2012. Adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis airway: an evolutionary perspective. *Nature Reviews Microbiology*, 10(12), pp.841–51.

Frangolias, D.D. *et al.*, 1999. *Burkholderia cepacia* in cystic fibrosis: Variable disease course. *American Journal of Respiratory and Critical Care Medecine*, 160(11), pp.1572–1577.

Gibson, R.L., Burns, J.L. & Ramsey, B.W., 2003. Pathophysiology and management of pulmonary

infections in cystic fibrosis. *American Journal of Respiratory and Critical Care Medicine*, 168(8), pp.918–951.

Gilligan, P.H., 2014. Infections in patients with cystic fibrosis diagnostic microbiology update. *Clinics in Laboratory Medicine*, pp.1–21.

Goncz, K.K., Feeney, L. & Gruenert, D.C., 1999. Differential sensitivity of normal and cystic fibrosis airway epithelial cells to epinephrine. *British Journal of Pharmacology*, 128, pp.227–233.

Govan, J.R., Brown, A.R. & Jones, A.M., 2007. Evolving epidemiology of *Pseudomonas aeruginosa* and the *Burkholderia cepacia* complex in cystic fibrosis lung infection. *Future Microbiology*, 2(2), pp.153–164.

Hansen, C.R., Pressler, T. & Høiby, N., 2008. Early aggressive eradication therapy for intermittent *Pseudomonas aeruginosa* airway colonization in cystic fibrosis patients: 15 years experience. *Journal of Cystic Fibrosis*, 7(6), pp.523–530.

Harrison, F., 2007. Microbial ecology of the cystic fibrosis lung. *Microbiology*, 153(4), pp.917–923.

Hengge, R., 2009. Principles of c-di-GMP signalling in bacteria. *Nature Reviews Microbiology*, 7, pp.263–273.

Hickman, J.W., Tifrea, D.F. & Harwood, C.S., 2005. A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels. *Proceedings of the National Academy of Sciences*, 102(40), pp.14422–14427.

Hoboth, C. *et al.*, 2009. Dynamics of adaptive microevolution of hypermutable *Pseudomonas aeruginosa* during chronic pulmonary infection in patients with cystic fibrosis. *Journal of Infectious Diseases*, 200, pp.118–130.

Hogardt, M. & Heesemann, J., 2010. Adaptation of *Pseudomonas aeruginosa* during persistence in the cystic fibrosis lung. *International Journal of Medical Microbiology*, 300(8), pp.557–562.

Høiby, N., Frederiksen, B. & Pressler, T., 2005. Eradication of early *Pseudomonas aeruginosa* infection. *Journal of Cystic Fibrosis*, 4, pp.49–54.

Huber, B. *et al.*, 2001. The *cep* quorum-sensing system of *Burkholderia cepacia* H111 controls biofilm formation and swarming motility. *Microbiology*, (147), pp.2517–2528.

Isles, A. *et al.*, 1984. *Pseudomonas cepacia* infection in cystic fibrosis: an emerging problem. *Journal of Pediatrics*, 104(2), pp.206–210.

Jassem, A.N., Forbes, C.M. & Speert, D.P., 2014. Investigation of aminoglycoside resistance inducing conditions and a putative AmrAB-OprM efflux system in *Burkholderia vietnamiensis*. *Annals of Clinical Microbiology and Antimicrobials*, 13(2), pp.1–5.

Jones, A.K. *et al.*, 2010. Activation of the *Pseudomonas aeruginosa* AlgU regulon through *mucA* mutation inhibits cyclic AMP / Vfr Signaling. *Journal of Bacteriology*, 192(21), pp.5709–5717.

Jones, a M. *et al.*, 2004. *Burkholderia cenocepacia* and *Burkholderia multivorans*: influence on survival in cystic fibrosis. *Thorax*, 59, pp.948–951.

Kalish, L.A. *et al.*, 2006. Impact of *Burkholderia dolosa* on lung function and survival in cystic fibrosis. *American Journal of Respiratory and Critical Care Medicine*, 173(4), pp.421–425.

Kearse, M. *et al.*, 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12), pp.1647–1649.

Klepac-Ceraj, V. *et al.*, 2010. Relationship between cystic fibrosis respiratory tract bacterial communities

and age, genotype, antibiotics and *Pseudomonas aeruginosa*. *Environmental Microbiology*, 12(5), pp.1293–1303.

Klockgether, J. *et al.*, 2011. *Pseudomonas aeruginosa* genomic structure and diversity. *Frontiers in Microbiology*, 2, pp.1–18.

Kumar, S., Stecher, G. & Tamura, K., 2016. MEGA7 : Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33(7), pp.1870–1874.

Lee, A.H. *et al.*, 2017. Phenotypic diversity and genotypic flexibility of *Burkholderia cenocepacia* during long-term chronic infection of cystic fibrosis lungs. *Genome Research*, pp.1–13.

Lesse, A.J. *et al.*, 1990. Increased resolution of lipopolysaccharides and lipooligosaccharides utilizing tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Journal of Immunological Methods*, 126(1), pp.109–117.

Lessie, T.G. *et al.*, 1996. Genomic complexity and plasticity of *Burkholderia cepacia*. *FEMS Microbiology Letters*, 144(2–3), pp.117–128.

Lewenza, S., Conway, B. & Greenberg, E.P., 1999. Quorum Sensing in *Burkholderia cepacia*: Identification of the LuxRI Homologs CepRI. *Journal of Bacteriology*, 181(3), pp.748–756.

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997v2*.

Li, H. & Durbin, R., 2009. Fast and accurate short read alignment with Burrows – Wheeler transform. *Bioinformatics*, 25(14), pp.1754–1760.

Lieberman, T.D. *et al.*, 2014. Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures. *Nature Genetics*, 46(1), pp.82–87.

Lieberman, T.D. *et al.*, 2011. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nature Genetics*, 43(12), pp.1275–1280.

LiPuma, J.J., 2002. Preventing *Burkholderia cepacia* complex infection in cystic fibrosis: Is there a middle ground? *Journal of Pediatrics*, 141(4), pp.467–469.

LiPuma, J.J., 2010. The changing microbial epidemiology in cystic fibrosis. *Clinical Microbiology Reviews*, 23(2), pp.299–323.

Mahenthiralingam, E., Urban, T.A. & Goldberg, J.B., 2005. The multifarious, multireplicon *Burkholderia cepacia* complex. *Nature Reviews. Microbiology*, 3(2), pp.144–156.

Markussen, T. *et al.*, 2014. Environmental heterogeneity drives within-host diversification and evolution of *Pseudomonas aeruginosa*. *mBio*, 5(5), pp.1–10.

Martina, P. *et al.*, 2014. Hypermutation in *Burkholderia cepacia* complex is mediated by DNA mismatch repair inactivation and is highly prevalent in cystic fibrosis chronic respiratory infection. *International Journal of Medical Microbiology*, 304(8), pp.1182–1191.

Marvig, R.L. *et al.*, 2014. Convergent evolution and adaptation of *Pseudomonas aeruginosa* within patients with cystic fibrosis. *Nature Genetics*, 47, pp.57–64.

Mathee, K. *et al.*, 2008. Dynamics of *Pseudomonas aeruginosa* genome evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 105(8), pp.3100–3105.

McKone, E.F., Goss, C.H. & Aitken, M.L., 2006. CFTR genotype as a predictor of prognosis in cystic fibrosis. *Chest*, 130(5), pp.1441–1447.

Mena, A. *et al.*, 2008. Genetic adaptation of *Pseudomonas aeruginosa* to the airways of cystic fibrosis

patients is catalyzed by hypermutation. *Journal of Bacteriology*, 190(24), pp.7910–7917.

Mira, N.P. *et al.*, 2011. Genomic expression analysis reveals strategies of *Burkholderia cenocepacia* to adapt to cystic fibrosis patients' airways and antimicrobial therapy. *PLoS ONE*, 6(12).

Moreira, A.S. *et al.*, 2014. *Burkholderia dolosa* phenotypic variation during the decline in lung function of a cystic fibrosis patient during 5.5 years of chronic colonization. *Journal of Medical Microbiology*, 63, pp.594–601.

Moreira, L.M. *et al.*, 2003. Identification and physical organization of the gene cluster involved in the biosynthesis of *Burkholderia cepacia* complex exopolysaccharide. *Biochemical and Biophysical Research Communications*, 312, pp.323–333.

O'Sullivan, B.P. & Freedman, S.D., 2009. Cystic fibrosis. *The Lancet*, 373(9678), pp.1891–1904.

Oliver, A. *et al.*, 2000. High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science*, 288(May), pp.1251–1254.

Palmer, K.L., Aye, L.M. & Whiteley, M., 2007. Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum. *Journal of Bacteriology*, 189(22), pp.8079–8087.

Parke, J.L. & Gurian-Sherman, D., 2001. Diversity of the *Burkholderia cepacia* complex and implications for risk assessment of biological control strains. *Annual Reviews of Phytopathology*, 39(1), pp.225–258.

Price, E.P. *et al.*, 2013. Within-host evolution of *Burkholderia pseudomallei* over a twelve-year chronic carriage infection. *mBio*, 4, pp.e00388–13.

Ramires, M., 2017. Within-host evolution of *Burkholderia multivorans* during chronic infection of three cystic fibrosis patients (Master's thesis). Instituto Superior Técnico, Universidade de Lisboa.

Ratjen, F. & Döring, G., 2003. Cystic fibrosis. *The Lancet*, 361(9358), pp.681–689.

Rau, M.H. *et al.*, 2012. Deletion and acquisition of genomic content during early stage adaptation of *Pseudomonas aeruginosa* to a human host environment. *Environmental Microbiology*, 14, pp.2200–2211.

Rau, M.H. *et al.*, 2010. 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. *Environmental Microbiology*, 12(6), pp.1643–1658.

Regamey, N. *et al.*, 2011. Airway remodelling and its relationship to inflammation in cystic fibrosis. *Thorax*, 66(7), pp.624–629.

Riordan, J.R., 2005. Assembly of functional CFTR chloride channels. *Annual Review of Physiology*, 67(1), pp.701–718.

Ryan, R.P. *et al.*, 2009. Intraspecies signaling involving the diffusible signal factor BDSF (cis-2-dodecenoic acid) influences virulence in *Burkholderia cenocepacia*. *Journal of Bacteriology*, 191(15), pp.5013–5019.

Saldías, M.S., Ortega, X. & Valvano, M.A., 2009. *Burkholderia cenocepacia* O-antigen lipopolysaccharide prevents phagocytosis by macrophages and adhesion to epithelial cells. *Journal of Medical Microbiology*, 58(12), pp.1542–1548.

Schaefers, M.M. *et al.*, 2017. An oxygen-sensing two-component system in the *Burkholderia cepacia* complex regulates biofilm, intracellular invasion, and pathogenicity. *PLOS Pathogens*, 13(1), p.e1006116.

Seed, K.D. & Dennis, J.J., 2008. Development of *Galleria mellonella* as an alternative infection model for the *Burkholderia cepacia* complex. *Infection and Immunity*, 76(3), pp.1267–1275.

Seemann, T., 2017. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30(14), pp.2068–2069.

Shigenobu, S., Watanabe, H. & Hattori, M., 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature*, 407(6800), pp.81–86.

Silva, I.N. *et al.*, 2016. Long-term evolution of *Burkholderia multivorans* during a chronic cystic fibrosis infection reveals shifting forces of selection. *mSystems*, 1(3), pp.e00029-16.

Silva, I.N. *et al.*, 2011. Mucoid morphotype variation of *Burkholderia multivorans* during chronic cystic fibrosis lung infection is correlated with changes in metabolism, motility, biofilm formation and virulence. *Microbiology*, 157(11), pp.3124–3137.

Silva, I.N. *et al.*, 2013. Stress conditions triggering mucoid morphotype variation in *Burkholderia* species and effect on virulence in *Galleria mellonella* and biofilm formation in vitro. *PLoS ONE*, 8(12).

Smith, E.E. *et al.*, 2006. Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proceedings of the National Academy of Sciences of the United States of America*, 103(22), pp.8487–92.

Taddei, F. *et al.*, 1997. Role of mutator alleles in adaptive evolution. *Nature*, 387(6634), pp.700–702.

Tamura, K. & Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, 10, pp.512–526.

Traverse, C.C. *et al.*, 2013. Tangled bank of experimentally evolved *Burkholderia* biofilms reflects selection during chronic infections. *Proceedings of the National Academy of Sciences of the United States of America*, 110(3), pp.E250-9.

Ulrich, M. *et al.*, 2010. Alveolar inflammation in cystic fibrosis. *Journal of Cystic Fibrosis*, 9(3), pp.217–227.

Vandamme, P. & Dawyndt, P., 2011. Classification and identification of the *Burkholderia cepacia* complex: Past, present and future. *Systematic and Applied Microbiology*, 34(2), pp.87–95.

Walker, B.J. *et al.*, 2014. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE*, 9(11), p.e112963.

Willger, S.D. *et al.*, 2014. Characterization and quantification of the fungal microbiome in serial samples from individuals with cystic fibrosis. *Microbiome*, 2(1), pp.40–55.

Zhao, J. *et al.*, 2012. Decade-long bacterial community dynamics in cystic fibrosis airways. *Proceedings of the National Academy of Sciences of the United States of America*, 109(15), pp.5809–14.

Zlosnik, J.E.A. *et al.*, 2015. *Burkholderia* species infections in patients with cystic fibrosis in British Columbia, Canada: 30 years' experience. *Annals of the American Thoracic Society*, 12(1), pp.70–78.

Zlosnik, J.E.A. *et al.*, 2008. Differential mucoid exopolysaccharide production by members of the *Burkholderia cepacia* complex. *Journal of Clinical Microbiology*, 46(4), pp.1470–1473.

Zlosnik, J.E.A. *et al.*, 2011. Mucoid and nonmucoid *Burkholderia cepacia* complex bacteria in cystic fibrosis infections. *American Journal of Respiratory and Critical Care Medicine*, 183(1), pp.67–72.

6. APPENDIX

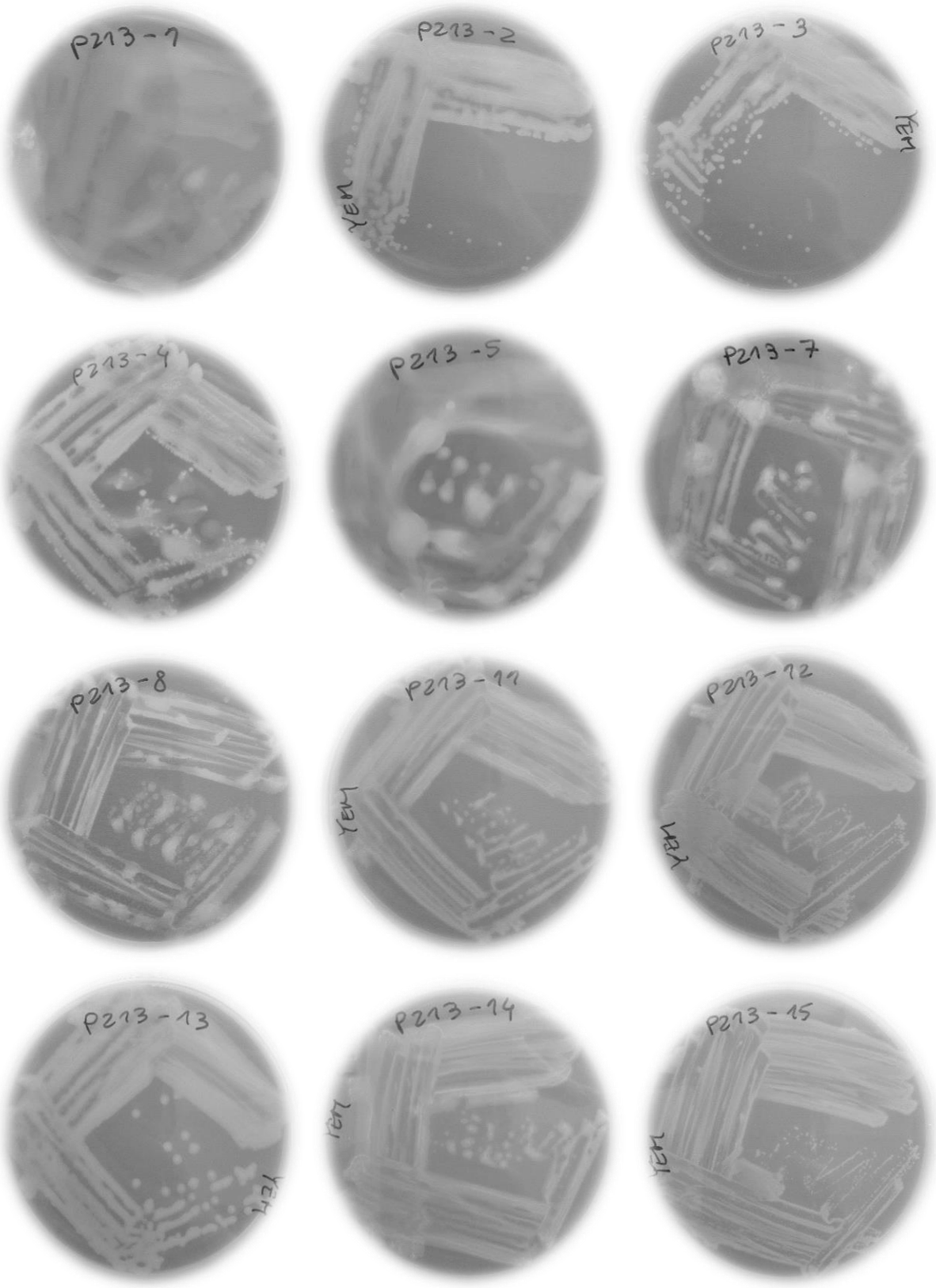

Figure A1 - Mucoid phenotype assessment in yeast extract medium (YEM) after 48 h of incubation at 37°C for *B. multivorans* isolates 1 to 5, 7, 8 and 11 to 16 of CF patient P213. Isolates were classified as highly mucoid (++), mucoid (+) and non-mucoid (-) based on visual inspection of the YEM plates.

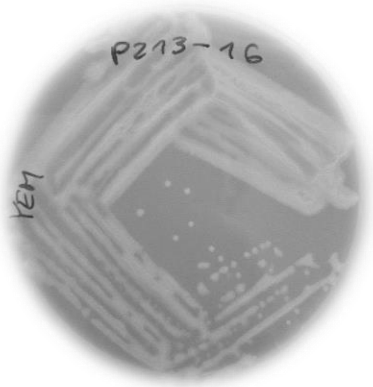

Figure A1 - Cont.

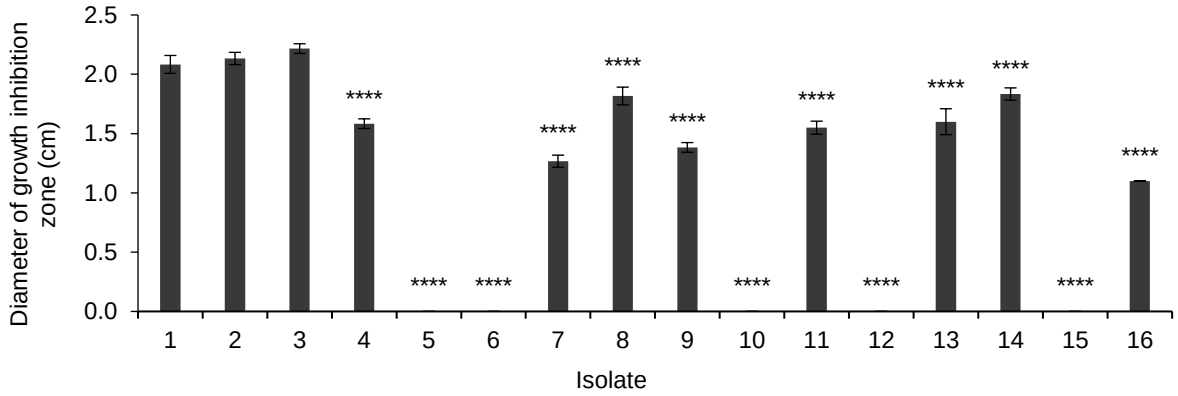

Figure A2 - Susceptibility of the 16 *B. multivorans* isolates sequentially retrieved from CF patient P213 to antibiotic aztreonam, measured as the diameter of cell growth inhibition, after growth for 24h at 37ºC. Error bars represent the standard deviations of the mean values for three independent experiments (statistical significance of differences between the first isolate and the subsequent ones was determined: ****, P-value < 0.0001 by ANOVA followed by Dunnett's multiple comparisons test).

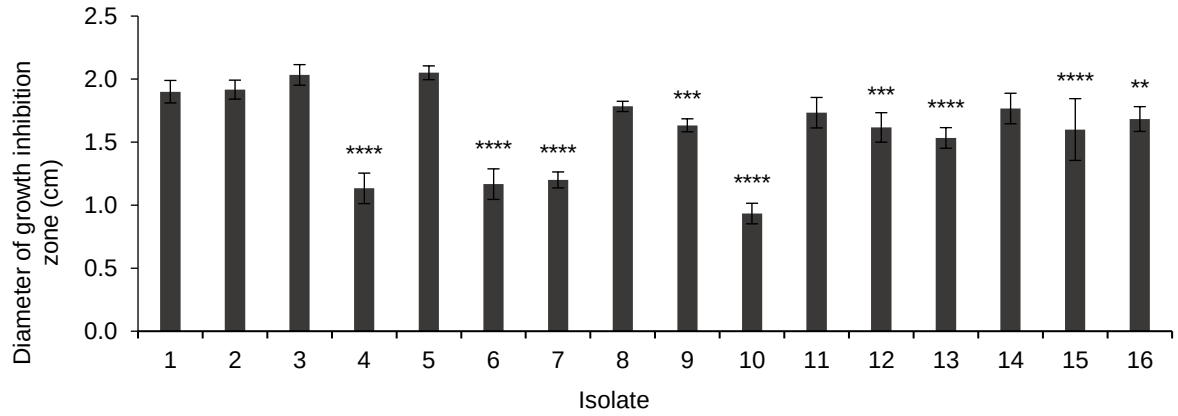

Figure A3 - Susceptibility of the 16 *B. multivorans* isolates sequentially retrieved from CF patient P213 to antibiotic kanamycin, measured as the diameter of cell growth inhibition, after growth for 24h at 37ºC. Error bars represent the standard deviations of the mean values for three independent experiments (statistical significance of differences between the first isolate and the subsequent ones was determined: **, P-value < 0.01; ***, P-value < 0.001; ****, P-value < 0.0001 by ANOVA followed by Dunnett's multiple comparisons test).

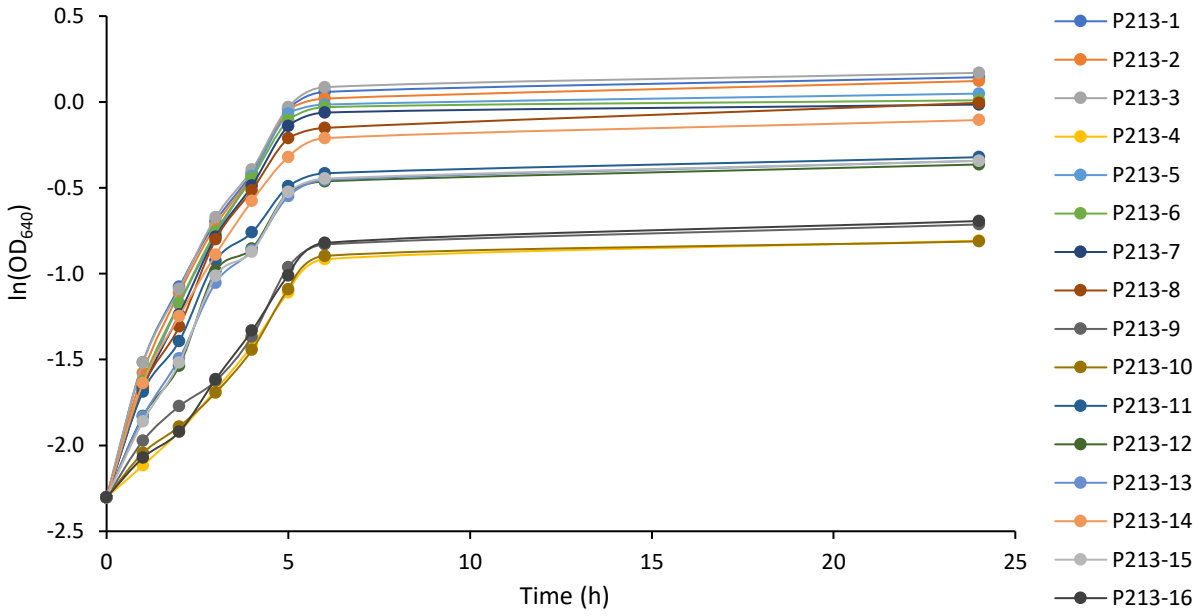

Figure A4 - Growth curves in synthetic cystic fibrosis medium (SCFM) measured for the 16 *B. multivorans* isolates sampled from CF patient P213. Results are the means of data from two independent experiments.

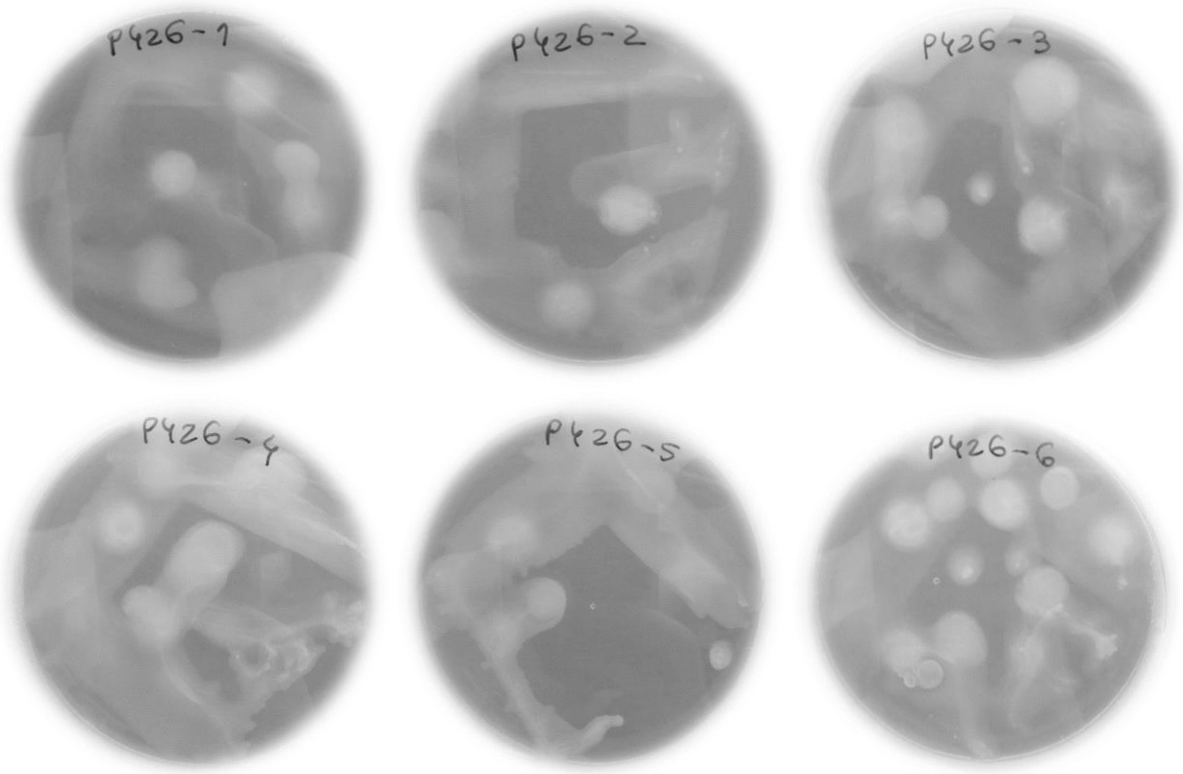

Figure A5 - Mucoid phenotype assessment in yeast extract medium (YEM) after 48 h of incubation at 37°C for *B. multivorans* isolates 1 to 12, 14 to 16, 18, 20 and 21 of CF patient P426. Isolates were classified as highly mucoid (++), mucoid (+) and non-mucoid (-) based on visual inspection of the YEM plates.

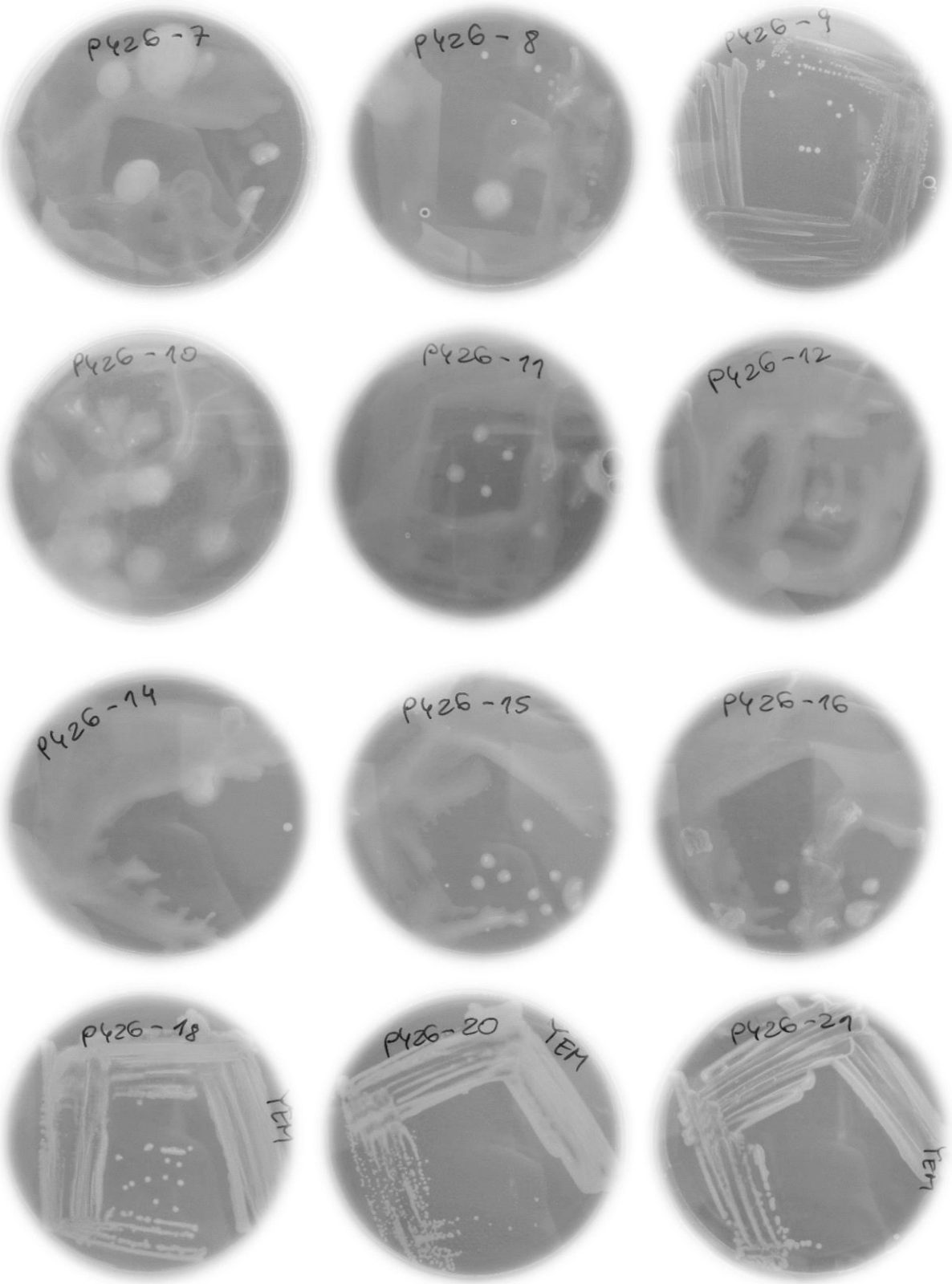

Figure A5 - Cont.

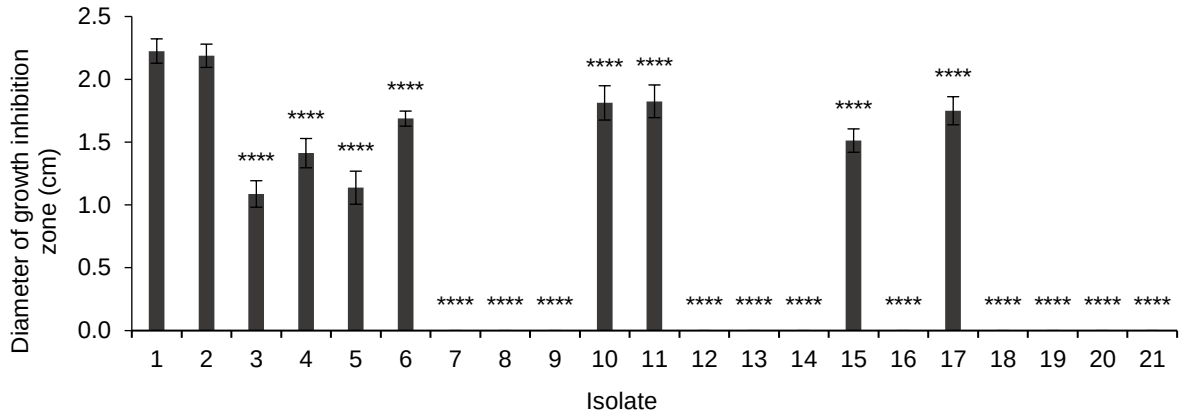

Figure A6 - Susceptibility of the 21 *B. multivorans* isolates sequentially retrieved from CF patient P426 to antibiotic aztreonam, measured as the diameter of cell growth inhibition, after growth for 24h at 37ºC. Error bars represent the standard deviations of the mean values for three independent experiments (statistical significance of differences between the first isolate and the subsequent ones was determined: ****, P-value < 0.0001 by ANOVA followed by Dunnett's multiple comparisons test).

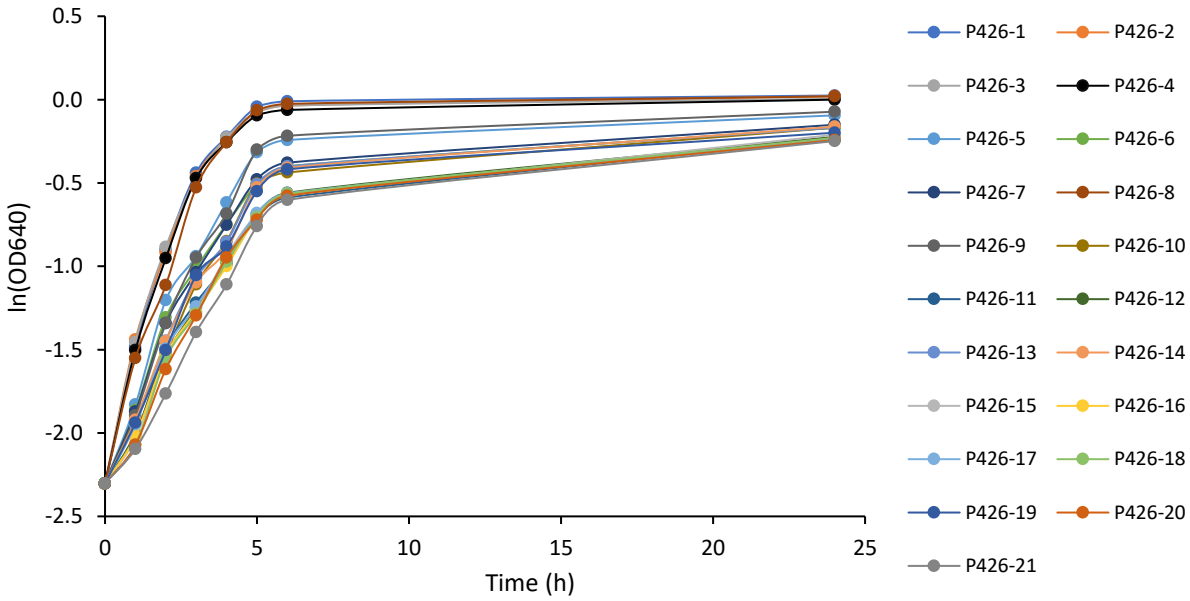

Figure A7 - Growth curves in synthetic cystic fibrosis medium (SCFM) measured for the 21 *B. multivorans* isolates sampled from CF patient P426. Results are the means of data from two independent experiments.

Table A1 - Mutations found among the 16 *B. multivorans* isolates of CF patient P213 (CDS = coding sequence, INT = intergenic, del = large deletion).

Contig	Position	Category	bp upstream or downstream CDS (if IG)	Mutation	Locus tag (bp upstream or downstream if IG)	Gene name	Annotation	Homolog in *B. multivorans* ATCC 17616	Mutations found (if CDS)	Type of mutation	Type of mutation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Nr of isolates with mutation
contig1	1562874	CDS		C>A	PROKKA_01466	*ccpA_1*	LacI family transcription regulator	Bmul_2073	S41I	nonsyn	SNP	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	15
contig1	1355287	CDS		T>G	PROKKA_01273	*pheA*	chorismate mutase	Bmul_2261	L281R	nonsyn	SNP	T	T	T	G	G	G	G	G	G	G	G	G	G	G	G	G	13
contig1	1596603	CDS		G>A	PROKKA_01500	*hprK*	HPr kinase/phosphorylase	Bmul_0521	R114C	nonsyn	SNP	G	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A	13
contig1	1774113	CDS		T>C	PROKKA_01663	*acrB_1*	Multidrug efflux pump subunit	Bmul_0689	V179A	nonsyn	SNP	T	T	T	C	C	C	C	C	C	C	C	C	C	C	C	C	13
contig1	2271249	CDS		C>T	PROKKA_02125	*folD*	Bifunctional protein FolD protein	Bmul_1130	A210T	nonsyn	SNP	C	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T	13
contig1	2273148	CDS		A>G	PROKKA_02127	*fixL_1*	Sensor protein FixL	Bmul_1132	S707P	nonsyn	SNP	A	A	A	G	G	G	G	G	G	G	G	G	G	G	G	G	13
contig1	2341503	CDS		T>A	PROKKA_02189	*accA*	acetyl-CoA carboxylase carboxyltransferase subunit alpha	Bmul_1205	D186E	nonsyn	SNP	T	T	T	A	A	A	A	A	A	A	A	A	A	A	A	A	13
contig1	2542874	CDS		C>T	PROKKA_02387	-	Major Facilitator Superfamily transporter	Bmul_1342	P245P	syn	SNP	C	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T	13
contig1	2947352	CDS		G>A	PROKKA_02752	*cysL_4*	LysR family transcriptional regulator	Bmul_1695	P158L	nonsyn	SNP	G	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A	13
contig2	50394	CDS		G>A	PROKKA_03101	*yabJ_2*	endoribonuclease	Bmul_5405	A148V	nonsyn	SNP	G	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A	13
contig3	128525	CDS		G>A	PROKKA_03779	*mcpC*	Methyl-accepting chemotaxis protein	Bmul_4946	H312H	syn	SNP	G	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A	13
contig3	246465	CDS		C>A	PROKKA_03885	*bag*	IgA FC receptor precursor	-	A10S	nonsyn	SNP	C	C	C	A	A	A	A	A	A	A	A	A	A	A	A	A	13
contig3	246467	CDS		C>A	PROKKA_03885	*bag*	IgA FC receptor precursor	-	R9L	nonsyn	SNP	C	C	C	A	A	A	A	A	A	A	A	A	A	A	A	A	13
contig1	2934558				PROKKA_02740	*fliY_2*	extracellular solute-binding protein	Bmul_1679		246 aa del	indel	G	G	G	736-bp del	736-bp del	736-bp del	736-bp del	736-bp del	736-bp del	736-bp del	736-bp del	736-bp del	736-bp del	736-bp del	736-bp del	736-bp del	13
contig1	1428203	INT	31 bp up	G>A	PROKKA_01351	*plsX*	putative glycerol-3-phosphate acyltransferase PlsX	Bmul_2184	-	-	SNP	G	G	G	G	A	A	A	G	A	A	A	A	A	A	A	A	11
contig1	3156946	CDS		G>A	PROKKA_02933	-	hypothetical protein	Bmul_1941	C220C	syn	SNP	G	G	G	G	A	A	A	G	A	A	A	A	A	A	A	A	11

contig2	124782	INT	285 bp up	C>T	PROKKA_03168	-	hypothetical protein	Bmul_5341	-	-	SNP	C	C	C	C	T	T	T	C	T	T	T	T	T	T	T	T	11
contig2	124788	INT	279 bp up	G>A	PROKKA_03168	-	hypothetical protein	Bmul_5341	-	-	SNP	G	G	G	G	A	A	A	G	A	A	A	A	A	A	A	A	11
contig3	43002	CDS		A>G	PROKKA_03700	*bphA*	Biphenyl dioxygenase subunit alpha	Bmul_4823	T105A	nonsyn	SNP	A	A	A	A	G	G	G	A	G	G	G	G	G	G	G	G	11
contig3	128522	CDS		G>A	PROKKA_03779	*mcpC*	Methyl-accepting chemotaxis protein McpC	Bmul_4946	D313D	syn	SNP	G	G	G	G	A	A	A	G	A	A	A	A	A	A	A	A	11
contig3	2374762	CDS		T>G	PROKKA_05730	*accD*	acetyl-CoA carboxylase subunit beta	Bmul_4619	R242R	syn	SNP	T	T	T	T	G	G	G	T	G	G	G	G	G	G	G	G	11
contig2	672318				PROKKA_03626; PROKKA_03627					149 aa del	indel	C	C	C	C	446-bp del	446-bp del	446-bp del	C	446-bp del	446-bp del	446-bp del	446-bp del	446-bp del	446-bp del	446-bp del	446-bp del	11
contig1	607806	INT	104 bp down	C>G	PROKKA_00561	*yofA_2*	LysR family transcriptional regulator	Bmul_0213	-	-	SNP	C	C	C	C	C	G	G	C	G	G	G	G	G	G	G	G	10
contig1	1825232	CDS		T>G	PROKKA_01707	*amiC_2*	N-acetylmuramoyl-L-alanine amidase	Bmul_0733	D505A	nonsyn	SNP	T	T	T	T	T	G	G	T	G	G	G	T	G	G	G	G	9
contig1	2878065	CDS		T>A	PROKKA_02681	*ttgR_3*	TetR family transcriptional regulator	Bmul_1617	A87A	syn	SNP	T	T	T	T	T	A	A	T	A	A	A	T	A	A	A	A	9
contig1	3024324	CDS		G>A	PROKKA_02813	*patA*	aspartate aminotransferase	Bmul_1760	A146V	nonsyn	SNP	G	G	G	G	G	A	A	G	A	A	A	G	A	A	A	A	9
contig2	509855	CDS		G>C	PROKKA_03499	-	Alpha/beta hydrolase family protein	-	R168G	nonsyn	SNP	G	G	G	G	G	C	C	G	C	C	C	G	C	C	C	C	9
contig1	45182	CDS		C>T	PROKKA_00044	*ptsI_2*	Phosphoenolpyruvate-protein phosphotransferase	Bmul_0443	V235M	nonsyn	SNP	C	C	C	C	C	T	T	C	T	T	T	C	T	T	T	T	8
contig1	1433121	CDS		A>G	PROKKA_01356	*fabF_1*	3-oxoacyl-(acyl carrier protein) synthase II	Bmul_2179	I149V	nonsyn	SNP	A	A	A	A	A	G	G	A	G	G	G	A	G	G	G	G	8
contig1	2571203	CDS		G>A	PROKKA_02407	*purL*	Phosphoribosylformyl glycinamidine synthase	Bmul_1360	V1065M	nonsyn	SNP	G	G	G	G	G	A	A	G	A	A	A	G	A	A	A	A	8
contig1	3044623	CDS		G>A	PROKKA_02834	*relA*	GTP pyrophosphokinase	Bmul_1780	R108C	nonsyn	SNP	G	G	G	G	G	A	A	G	A	A	A	G	A	A	A	A	8
contig2	102091	CDS		A>G	PROKKA_03149	*ykuV*	cytochrome c biogenesis protein transmembrane region	Bmul_5356	D492G	nonsyn	SNP	A	A	A	A	A	G	G	A	G	G	G	A	G	G	G	G	8
contig2	519694	CDS		G>A	PROKKA_03509	*bbsF_3*	Succinyl-CoA (R)-benzylsuccinate CoA-transferase subunit BbsF	-	G231S	nonsyn	SNP	G	G	G	G	G	A	A	G	A	A	A	G	A	A	A	A	8
contig3	368540	CDS		C>T	PROKKA_03992	*mlr_2*	4-methylamino butanoate oxidase (formaldehyde-forming)	Bmul_5151	R125Q	nonsyn	SNP	C	C	C	C	C	T	T	C	T	T	T	C	T	T	del	T	7

contig1	2073572				PROKKA_01941	*ftsK_2*	DNA translocase FtsK			168 aa del	indel	C	C	C	C	C	504-bp del	504-bp del	504-bp del	504-bp del	504-bp del	C	C	504-bp del	C	C	C	6
contig1	320176	INT	155 bp up	C>T	PROKKA_00298	*gatC*	Glutamyl-tRNA(Gln) amidotransferase subunit C	Bmul_3106	-	-	SNP	C	C	C	C	C	C	C	C	T	T	C	C	C	T	T	C	4
contig1	349906	INT	206 bp up	G>A	PROKKA_00326	-	ADP-heptose:LPS heptosyltransferase II	-	-	-	SNP	G	G	G	G	G	G	G	G	A	A	G	G	G	A	A	G	4
contig1	673279	CDS		G>C	PROKKA_00631	*rpoA*	DNA-directed RNA polymerase subunit alpha	Bmul_0275	D256H	nonsyn	SNP	G	G	G	G	G	G	G	G	C	C	G	G	G	C	C	G	4
contig1	1587107	CDS		C>T	PROKKA_01490	*rplY*	50S ribosomal protein L25	Bmul_0512	A164T	nonsyn	SNP	C	C	C	C	C	C	C	C	T	T	C	C	C	T	T	C	4
contig1	1833568	CDS		T>A	PROKKA_01715	-	putative nucleotide-binding protein	Bmul_0741	R82S	nonsyn	SNP	T	T	T	T	T	T	T	T	A	A	T	T	T	A	A	T	4
contig1	2375814	CDS		G>A	PROKKA_02226	-	Uracil DNA glycosylase superfamily protein	Bmul_1234	R221R	syn	SNP	G	G	G	G	G	G	G	G	A	A	G	G	G	A	A	G	4
contig1	2460009	CDS		T>G	PROKKA_02306	-	hypothetical protein	-	V65V	syn	SNP	T	T	T	T	T	T	T	T	G	G	T	T	T	G	G	T	4
contig1	2981579	CDS		A>T	PROKKA_02781	*hfq_2*	RNA chaperone Hfq	Bmul_1722	K12N	nonsyn	SNP	A	A	A	A	A	A	A	A	T	T	A	A	A	T	T	A	4
contig2	119366	INT	rRNA	G>A	PROKKA_03163	-	5S ribosomal RNA	-	-	-	SNP	G	G	G	G	G	G	G	G	A	A	G	G	G	A	A	G	4
contig2	122253	INT	rRNA	C>A	PROKKA_03164	-	23S ribosomal RNA	-	-	-	SNP	C	C	C	C	C	C	C	C	A	A	C	C	C	A	A	C	4
contig2	574129	INT	36 bp down	G>A	PROKKA_03552	*yjjL_1*	L-galactonate transporter	-	-	-	SNP	G	G	G	G	G	G	G	G	A	A	G	G	G	A	A	G	4
contig3	1228299	CDS		A>G	PROKKA_04746	*algE7*	Poly(beta-D-mannuronate) C5 epimerase 7	Bmul_3709	P239P	syn	SNP	A	A	A	A	A	A	A	A	G	G	A	A	A	G	G	A	4
contig1	763710				PROKKA_00718 to PROKKA_00742					9625 aa del	indel	C	C	C	C	C	C	C	C	28873-bp del	28873-bp del	C	C	C	28873-bp del	28873-bp del	C	4
contig2	120468				PROKKA_03164	*rRNA*				130 aa del	indel	A	A	A	A	A	A	A	A	389-bp del	389-bp del	A	A	A	389-bp del	389-bp del	A	4
contig1	1429157	CDS		G>C	PROKKA_01351	*plsX*	putative glycerol-3-phosphate acyltransferase PlsX	Bmul_2184	E308D	nonsyn	SNP	G	G	G	G	G	G	G	C	G	G	C	G	G	C	G	G	3
contig1	56082	CDS		A>G	PROKKA_00055	*dctD_1*	Fis family transcriptional regulator	Bmul_0432	V309A	nonsyn	SNP	A	A	A	A	A	A	A	A	A	G	A	A	A	A	G	A	2
contig1	840628	CDS		T>A	PROKKA_00789	*paaK_1*	phenylacetate-CoA ligase	Bmul_2855	E417D	nonsyn	SNP	T	T	T	T	T	T	T	T	T	A	T	T	T	T	A	T	2
contig1	871150	CDS		C>T	PROKKA_00813	*ftsW*	cell division protein FtsW	Bmul_2837	L240F	nonsyn	SNP	C	C	C	C	C	C	C	C	C	T	C	C	C	C	T	C	2

contig1	920893	INT	19 bp up	G>T	PROKKA_00861	*mpl*	UDP-N-acetyl muramate:L-alanyl-gamma-D-glutamyl-meso-diamino pimelate ligase	Bmul_2775	-	-	SNP	G	G	G	G	G	G	G	G	G	T	G	G	G	G	T	G	2
contig1	1025044	CDS		G>A	PROKKA_00958	*rlmL*	Ribosomal RNA large subunit methyltransferase L	Bmul_2677	R140R	syn	SNP	G	G	G	G	G	G	G	G	G	A	G	G	G	G	A	G	2
contig1	1045300	INT	1 bp up	C>T	PROKKA_00975	-	hypothetical protein	-	-	-	SNP	C	C	C	C	C	T	T	C	C	C	C	C	C	C	C	C	2
contig1	1169460	INT	19 bp up	A>C	PROKKA_01091	*groS3*	co-chaperonin GroES	Bmul_2529	-	-	SNP	A	A	A	A	A	A	A	A	A	C	A	A	A	A	C	A	2
contig1	1259446	CDS		C>G	PROKKA_01177	*aspC*	Aspartate aminotransferase	Bmul_2443	A225A	syn	SNP	C	C	C	C	C	C	C	C	C	G	C	C	C	C	G	C	2
contig1	1601125	CDS		C>T	PROKKA_01505	*lptA*	lipopolysaccharide transport periplasmic protein LptA	Bmul_0526	A90T	nonsyn	SNP	C	C	C	T	C	C	C	T	C	C	C	C	C	C	C	C	2
contig1	2086264	CDS		G>A	PROKKA_01950	*rfaQ_1*	LPS core heptosyl transferase RfaQ	-	G355R	nonsyn	SNP	G	G	G	G	G	G	G	G	G	A	G	G	G	G	A	G	2
contig1	2492395	INT	31 bp up	A>C	PROKKA_02339	-	NUDIX hydrolase	Bmul_1291	-	-	SNP	A	A	A	A	A	A	A	A	A	C	A	A	A	A	C	A	2
contig1	2807720	INT	220 bp up	C>T	PROKKA_02631	*cobL*	Precorrin-6Y C(5,15)-methyltransferase [decarboxylating]	Bmul_1567	-	-	SNP	C	C	C	C	C	C	C	C	T	C	C	C	C	T	C	C	2
contig1	2837644	CDS		A>G	PROKKA_02659	*fhuA*	Ferrichrome-iron receptor precursor	Bmul_1594	*570R	nonsyn	SNP	A	A	A	A	A	A	A	A	A	G	A	A	A	A	G	A	2
contig1	2838056	CDS		C>T	PROKKA_02659	*fhuA*	Ferrichrome-iron receptor precursor	Bmul_1594	R433R	syn	SNP	C	C	C	C	C	C	C	C	C	T	C	C	C	C	T	C	2
contig2	356827	CDS		T>C	PROKKA_03372	-	Lactonase, 7-bladed beta-propeller	Bmul_6011	L149P	nonsyn	SNP	T	T	T	T	T	T	T	T	T	C	T	T	T	T	C	T	2
contig3	30510	CDS		C>T	PROKKA_03690	*rpoD_2*	RNA polymerase sigma factor RpoD	Bmul_4813	A666A	syn	SNP	C	C	C	C	C	C	C	C	C	T	C	C	C	C	T	C	2
contig3	796326	INT	2 bp up	C>T	PROKKA_04374	-	putative hemoglobin and hemoglobin-haptoglobin-binding protein 2 precursor	Bmul_3338	-	-	SNP	C	C	C	C	C	C	C	C	C	T	C	C	C	C	T	C	2
contig3	1192013	CDS		C>T	PROKKA_04715	*rstB_2*	Sensor protein RstB	Bmul_3678	R220C	nonsyn	SNP	C	C	C	T	C	C	C	T	C	C	C	C	C	C	C	C	2
contig3	1199449	CDS		C>G	PROKKA_04721	*rssA_4*	NTE family protein RssA	-	V498V	syn	SNP	C	C	C	C	C	C	C	C	G	C	C	C	C	G	C	C	2
contig3	1280318	CDS		C>A	PROKKA_04786	-	LigA protein	Bmul_3742	P68T	nonsyn	SNP	C	C	C	C	C	C	C	C	A	del	C	C	C	A	del	C	2
contig3	1290377	CDS		C>T	PROKKA_04796	-	peptidoglycan-binding LysM	Bmul_3752	E4220K	nonsyn	SNP	C	C	C	C	C	T	C	C	C	del	C	C	C	C	del	T	2
contig3	1517663	INT	41 bp down	C>G	PROKKA_04994	-	MerR family transcriptional regulator	Bmul_4022	-	-	SNP	C	C	C	C	C	C	C	C	C	G	C	C	C	C	G	C	2
contig1	1041735	CDS		T	PROKKA_00973	*recG*	ATP-dependent DNA helicase RecG	Bmul_2662	-	11 aa del	indel	cg cc gc cg	cg cc gc cg	cg cc gc cg	cg cc gc cg	cg cc gc cg	cg cc gc cg	cg cc gc cg	cg cc gc cg	cg cc gc cg	---- ---- ---- ----	cg cc gc cg	cg cc gc cg	cg cc gc cg	cg cc gc cg	---- ---- ---- ----	cg cc gc cg	2

												gcgcgcgcggccgatcgcgacgctg	gcgcgcgcggccgatcgcgacgctg	gcgcgcgcggccgatcgcgacgctg	gcgcgcgcggccgatcgcgacgctg	gcgcgcgcggccgatcgcgacgctg	gcgcgcgcggccgatcgcgacgctg	gcgcgcgcggccgatcgcgacgctg	gcgcgcgcggccgatcgcgacgctg	gcgcgcgcggccgatcgcgacgctg	-----------	gcgcgcgcggccgatcgcgacgctg	gcgcgcgcggccgatcgcgacgctg	gcgcgcgcggccgatcgcgacgctg	gcgcgcgcggccgatcgcgacgctg	----------	gcgcgcgcggccgatcgcgacgctg	
contig1	2839612				PROKKA_02660 to PROKKA_02663					5396 aa del	indel	C	C	C	C	C	C	C	C	C	16187-bp del	C	C	C	C	16187-bp del	C	2
contig3	1264028				PROKKA_04773 to PROKKA_04857					35323 aa del	indel	T	T	T	T	T	T	T	T	T	105968-bp del	T	T	T	T	105968-bp del	T	2
contig1	91926	CDS		C>T	PROKKA_00087	*birA_1*	Bifunctional ligase/repressor BirA	-	I114I	syn	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig1	148070	CDS		T>C	PROKKA_00132	*sbp_1*	sulfate ABC transporter periplasmic sulfate-binding protein	Bmul_2940	D157G	nonsyn	SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig1	184360	CDS		T>C	PROKKA_00166	-	hypothetical protein	Bmul_2972	T76A	nonsyn	SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig1	184660	CDS		A>G	PROKKA_00167	-	short chain dehydrogenase	Bmul_2973	D20G	nonsyn	SNP	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	1
contig1	225690	CDS		G>A	PROKKA_00206	*flgI*	Flagellar P-ring protein precursor	Bmul_3013	Q374*	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	1
contig1	230206	CDS		C>G	PROKKA_00210	*flgE*	Flagellar hook protein FlgE	Bmul_3017	G56A	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	1
contig1	257407	CDS		C>T	PROKKA_00236	*proP_1*	Proline/betaine transporter	Bmul_3041	V341M	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	1
contig1	297295	CDS		T>C	PROKKA_00276	*rlmI*	Ribosomal RNA large subunit methyltransferase I	Bmul_3083	Y213H	nonsyn	SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig1	319451	CDS		C>T	PROKKA_00297	*gatA_1*	Glutamyl-tRNA(Gln) amidotransferase subunit A	Bmul_3105	L67L	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	1
contig1	349983	INT	160 bp up	C>G	PROKKA_00327	-	O-Antigen ligase	-	-	-	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	1
contig1	353749	CDS		C>T	PROKKA_00330	*gapN_1*	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	Bmul_3133	R447Q	nonsyn	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1

contig1	356481	INT	80 bp up	G>T	PROKKA_00332	*pgtC*	Phosphoglycerate transport regulatory protein PgtC precursor	Bmul_3135	-	-	SNP	G	G	G	G	G	G	G	G	G	G	G	T	G	G	G	G	1
contig1	390082	CDS		G>A	PROKKA_00362	*srpA*	Catalase-related peroxidase precursor	Bmul_0006	P14P	syn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	1
contig1	414961	CDS		T>G	PROKKA_00388	*fliL*	Flagellar FliL protein	Bmul_0044	D95A	nonsyn	SNP	T	T	T	T	T	T	T	T	T	G	T	T	T	T	T	T	1
contig1	660737	CDS		G>C	PROKKA_00606	*rplD*	50S ribosomal protein L4	Bmul_0250	G70R	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	1
contig1	661934	CDS		G>A	PROKKA_00608	*rplB*	50S ribosomal protein L2	Bmul_0252	S157S	syn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	1
contig1	668615	CDS		T>C	PROKKA_00624	*rplO*	50S ribosomal protein L15	Bmul_0268	A10A	syn	SNP	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	1
contig1	846473	CDS		G>A	PROKKA_00794	*echA8_1*	enoyl-CoA hydratase	Bmul_2850	D46N	nonsyn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig1	866409	CDS		G>A	PROKKA_00810	*murF*	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D- alanine ligase	Bmul_2840	V36I	nonsyn	SNP	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	1
contig1	866580	CDS		C>A	PROKKA_00810	*murF*	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D- alanine ligase	Bmul_2840	L93I	nonsyn	SNP	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	1
contig1	872893	CDS		T>G	PROKKA_00815	*murC*	UDP-N-acetylmuramate-L-alanine ligase	Bmul_2835	N27K	nonsyn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	1
contig1	962108	CDS		C>T	PROKKA_00901	*bamE*	Outer membrane protein assembly factor BamE precursor	Bmul_2736	L120L	syn	SNP	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	1
contig1	1061087	CDS		C>T	PROKKA_00992	*aroG_1*	Phospho-2-dehydro-3-deoxyheptonate aldolase, Phe-sensitive	Bmul_2646	D146N	nonsyn	SNP	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig1	1112550	CDS		C>T	PROKKA_01040	*nrdR*	Transcriptional repressor NrdR	Bmul_2580	T11M	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	1
contig1	1136596	CDS		A>G	PROKKA_01065	*araC_1*	AraC family transcriptional regulator	Bmul_2555	W139R	nonsyn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig1	1161443	CDS		C>T	PROKKA_01084	-	hypothetical protein	-	G639G	syn	SNP	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	1
contig1	1198643	CDS		C>G	PROKKA_01119	*tagO_2*	glycosyl transferase family protein	Bmul_2502	A296P	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	1
contig1	1261838	CDS		C>G	PROKKA_01180	*ribBA_1*	Riboflavin biosynthesis protein RibBA	Bmul_2440	A173A	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	1
contig1	1265296	CDS		T>C	PROKKA_01184	*hemL*	Glutamate-1-semialdehyde 2,1-aminomutase	Bmul_2436	H362R	nonsyn	SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig1	1413257	CDS		A>T	PROKKA_01336	-	RmuC family protein	Bmul_2199	Q116L	nonsyn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	T	1

contig1	1484160	CDS		G>A	PROKKA_01407	-	hypothetical protein	Bmul_2131	R98R	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	1
contig1	1492366	INT	131 bp up	A>G	PROKKA_01415	*astC*	Succinylornithine transaminase	Bmul_2124	-	-	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig1	1524070	CDS		A>G	PROKKA_01437	-	phosphatase-like protein	Bmul_2102	L357P	nonsyn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A	1
contig1	1528294	CDS		G>T	PROKKA_01441	*ywnA*	BadM/Rrf2 family transcriptional regulator	Bmul_2098	G148C	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	T	G	G	G	G	G	1
contig1	1564272	INT	469 bp down	C>A	PROKKA_01467	-	hypothetical protein	-	-	-	SNP	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	C	1
contig1	1580768	CDS		C>T	PROKKA_01482	*trpD2*	glycosyl transferase family protein	Bmul_0490	P64P	syn	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig1	1592143	CDS		C>G	PROKKA_01495	-	tetratricopeptide repeat protein	Bmul_0516	G33R	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	1
contig1	1619024	CDS		C>T	PROKKA_01521	-	Phasin protein	Bmul_0542	S187S	syn	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig1	1674265	CDS		G>A	PROKKA_01569	*eamA_2*	putative amino-acid metabolite efflux pump	Bmul_0592	P111L	nonsyn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig1	1713494	CDS		G>C	PROKKA_01607	*rhlE_1*	DEAD/DEAH box helicase	Bmul_0627	S246C	nonsyn	SNP	G	G	G	G	G	G	C	G	G	G	G	G	G	G	G	G	1
contig1	1714335	INT	28 bp down	C>T	PROKKA_01608	*adhB_1*	gluconate 2-dehydrogenase	Bmul_0628	-	-	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	1
contig1	1719881	INT	22 bp down	G>A	PROKKA_01612	*lldR_2*	GntR family transcriptional regulator	Bmul_0632	-	-	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	1
contig1	1722689	CDS		C>T	PROKKA_01616	-	glycosyl transferase family protein	Bmul_0636	A955A	syn	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig1	1729086	CDS		A>G	PROKKA_01620	*dmlR_8*	LysR family transcriptional regulator	Bmul_0641	Q262R	nonsyn	SNP	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	1
contig1	1824059	CDS		C>T	PROKKA_01705	*yhhW_2*	Quercetin 2,3-dioxygenase	Bmul_0731	C24Y	nonsyn	SNP	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig1	1914967	CDS		G>A	PROKKA_01792	*cbl_1*	transcriptional regulator CysB-like protein	Bmul_0818	S138F	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	1
contig1	1922382	CDS		G>C	PROKKA_01800	*lptG*	LPS export system permease protein LptG	Bmul_0826	P358A	nonsyn	SNP	G	G	G	G	G	G	C	G	G	G	G	G	G	G	G	G	1
contig1	1956542	CDS		G>T	PROKKA_01831	*metG*	Methionine-tRNA ligase	Bmul_0857	G370G	syn	SNP	G	G	G	G	G	G	G	T	G	G	G	G	G	G	G	G	1
contig1	1995949	CDS		C>T	PROKKA_01864	*envZ_1*	Osmolarity sensor protein EnvZ	Bmul_0890	Q412*	nonsyn	SNP	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	1
contig1	1999848	CDS		C>A	PROKKA_01867	*cmpD_1*	Bicarbonate transport ATP-binding protein CmpD	Bmul_0893	R422R	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	1
contig1	2058424	CDS		G>C	PROKKA_01930	-	ATPase family associated with various cellular activities	Bmul_0945	S233W	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	1

contig1	2067154	CDS		A>G	PROKKA_01939	-	Trehalase	Bmul_0954	R290R	syn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig1	2152189	CDS		G>C	PROKKA_02018	*nuoF*	NADH-quinone oxidoreductase subunit F	Bmul_1033	G149A	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	C	G	G	G	G	G	1
contig1	2152190	CDS		C>A	PROKKA_02018	*nuoF*	NADH-quinone oxidoreductase subunit F	Bmul_1033	G149G	syn	SNP	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	1
contig1	2177740	CDS		C>G	PROKKA_02044	-	hypothetical protein	Bmul_1059	Q280H	nonsyn	SNP	C	C	C	C	G	C	C	C	C	C	C	C	C	C	C	C	1
contig1	2261768	CDS		C>T	PROKKA_02118	*glnG*	nitrogen metabolism transcriptional regulator NtrC	Bmul_1124	R10*	nonsyn	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig1	2282907	INT	145 bp up	C>G	PROKKA_02131	-	Phasin protein	Bmul_1136	-	-	SNP	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C	1
contig1	2316891	CDS		T>A	PROKKA_02165	-	hypothetical protein	Bmul_1172	H188Q	nonsyn	SNP	T	T	T	T	T	T	T	T	T	T	A	T	T	T	T	T	1
contig1	2341659	CDS		G>A	PROKKA_02189	*accA*	acetyl-CoA carboxylase carboxyltransferase subunit alpha	Bmul_1205	E238E	syn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig1	2368518	CDS		C>A	PROKKA_02219	*dmlR_13*	LysR family transcriptional regulator	Bmul_1227	P94T	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	1
contig1	2405787	INT	179 bp up	G>A	PROKKA_02248	*map_2*	Methionine aminopeptidase	Bmul_1256	-	-	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	1
contig1	2447646	CDS		C>G	PROKKA_02291	-	hypothetical protein	-	*80S	nonsyn	SNP	C	C	C	C	C	C	C	del	C	G	C	C	C	C	C	C	1
contig1	2457416	INT	5 bp down	C>G	PROKKA_02301	-	hypothetical protein	-	-	-	SNP	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	1
contig1	2543389	CDS		A>G	PROKKA_02387	-	Major Facilitator Superfamily transporter	Bmul_1342	C73R	nonsyn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig1	2551944	CDS		A>G	PROKKA_02395	*clpX*	ATP-dependent Clp protease ATP-binding subunit ClpX	Bmul_1350	S79S	syn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig1	2666055	CDS		G>C	PROKKA_02498	*surE_1*	stationary phase survival protein SurE	Bmul_1450	M246I	nonsyn	SNP	G	G	G	G	G	G	C	G	G	G	G	G	G	G	G	G	1
contig1	2710070	CDS		T>C	PROKKA_02538	*metC_1*	Cystathionine beta-lyase MetC	Bmul_1491	C179R	nonsyn	SNP	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	1
contig1	2755251	CDS		T>G	PROKKA_02577	*fis*	Fis family transcriptional regulator	Bmul_1516	A103A	syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	G	T	T	1
contig1	2755284	CDS		A>G	PROKKA_02577	*fis*	Fis family transcriptional regulator	Bmul_1516	A114A	syn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig1	2755289	CDS		G>C	PROKKA_02577	*fis*	Fis family transcriptional regulator	Bmul_1516	R116P	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	C	G	G	G	G	G	1
contig1	2771183	CDS		G>C	PROKKA_02589	*ghrB_2*	Glyoxylate/hydroxypyruvate reductase B	Bmul_1528	G180R	nonsyn	SNP	G	G	G	G	G	G	G	G	C	G	G	G	G	G	G	G	1

contig1	2803044	CDS		C>T	PROKKA_02627	*yghU*	Disulfide-bond oxidoreductase YghU	Bmul_1563	P104L	nonsyn	SNP	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	1
contig1	2857712	CDS		G>C	PROKKA_02664	*yojI*	ABC transporter ATP-binding protein YojI	Bmul_1599	G327A	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	1
contig1	2871189	INT	98 bp up	T>C	PROKKA_02677	*fimA_3*	Type-1 fimbrial protein, A chain precursor	Bmul_1612	-	-	SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig1	2881282	CDS		C>G	PROKKA_02684	*glnM_2*	putative glutamine ABC transporter permease protein GlnM	Bmul_1620	G69R	nonsyn	SNP	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C	1
contig1	2884313	CDS		C>T	PROKKA_02687	-	SpoVR family protein	Bmul_1623	R550H	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	1
contig1	2910949	CDS		C>A	PROKKA_02710	*gltC_4*	LysR family transcriptional regulator	Bmul_1646	C172F	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	1
contig1	2917249	CDS		G>A	PROKKA_02716	*pld_1*	Phospholipase D precursor	Bmul_1653	R61H	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	1
contig1	2970954	CDS		T>C	PROKKA_02771	*rbsC_3*	Ribose transport system permease protein RbsC	Bmul_1712	V272A	nonsyn	SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig1	2998460	CDS		T>C	PROKKA_02795	-	hypothetical protein	Bmul_1742	L120P	nonsyn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	1
contig1	3025430	CDS		G>C	PROKKA_02814	*gltC_5*	LysR family transcriptional regulator	Bmul_1761	R99G	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	C	G	G	G	G	G	1
contig1	3095500	CDS		G>A	PROKKA_02882	*mglA*	Galactose/methyl galactoside import ATP-binding protein MglA	Bmul_1890	P334S	nonsyn	SNP	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	1
contig1	3179448	CDS		C>T	PROKKA_02945	*fabL*	Enoyl-[acyl-carrier-protein] reductase [NADPH] FabL	Bmul_1953	A146A	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	1
contig1	3213080	CDS		T>A	PROKKA_02977	*shlB_2*	Hemolysin transporter protein ShlB precursor	Bmul_2070	*514C	nonsyn	SNP	T	T	T	T	T	T	T	T	T	T	A	T	T	T	T	T	1
contig1	3215368	CDS		T>C	PROKKA_02979	-	hypothetical protein	Bmul_1985	Y245C	nonsyn	SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig1	3256603	CDS		C>T	PROKKA_03013	*pstS*	Phosphate-binding protein PstS precursor	Bmul_2020	A60A	syn	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig1	3277742	CDS		C>T	PROKKA_03031	*proP_8*	Proline/betaine transporter	Bmul_2038	Q403*	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	1
contig2	79974	INT	567 bp down	G>A	PROKKA_03128	-	Sodium Bile acid symporter family protein	Bmul_5378	-	-	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig2	89242	CDS		T>C	PROKKA_03140	*bktB*	Beta-ketothiolase BktB	Bmul_5365	Y69C	nonsyn	SNP	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	1
contig2	95301	CDS		A>G	PROKKA_03144	*fadD_3*	Long-chain-fatty-acid-CoA ligase	Bmul_5361	V271A	nonsyn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1

contig2	98034	CDS		T>C	PROKKA_03146	*phoR_2*	Alkaline phosphatase synthesis sensor protein PhoR	Bmul_5359	H369R	nonsyn	SNP	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	1
contig2	98413	CDS		G>A	PROKKA_03146	*phoR_2*	Alkaline phosphatase synthesis sensor protein PhoR	Bmul_5359	L242F	nonsyn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig2	105414	CDS		T>A	PROKKA_03152	*dht*	D-hydantoinase/dihydropyrimidinase	Bmul_5353	Q233L	nonsyn	SNP	T	T	T	T	T	T	T	T	T	T	A	T	T	T	T	T	1
contig2	133339	CDS		G>A	PROKKA_03174	*ggt_2*	Gamma-glutamyltranspeptidase precursor	Bmul_5335	A524V	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	1
contig2	215194	CDS		G>C	PROKKA_03250	*pbuE_2*	Purine efflux pump PbuE	Bmul_6097	P319R	nonsyn	SNP	G	G	G	G	G	G	G	C	G	G	G	G	G	G	G	G	1
contig2	226932	CDS		A>G	PROKKA_03259	*ydcR_3*	putative HTH-type transcriptional regulator YdcR	-	V254A	nonsyn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig2	233899	CDS		G>A	PROKKA_03265	*eamA_4*	putative amino-acid metabolite efflux pump	-	V254I	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	1
contig2	238814	CDS		C>T	PROKKA_03270	-	FAD linked oxidase domain-containing protein	Bmul_6081	F256F	syn	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig2	272103	INT	9 bp up	C>T	PROKKA_03298	*ynfM_3*	Inner membrane transport protein YnfM	-	-	-	SNP	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	1
contig2	276039	CDS		G>C	PROKKA_03301	*nicA_1*	Nicotinate dehydrogenase subunit A	-	P118R	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	G	1
contig2	302358	CDS		T>C	PROKKA_03326	*pnpB_2*	p-benzoquinone reductase	Bmul_6057	T94A	nonsyn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	C	T	T	1
contig2	374350	CDS		G>C	PROKKA_03390	*gltC_8*	LysR family transcriptional regulator	Bmul_5993	G284R	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	G	1
contig2	379227	INT	22 bp down	C>T	PROKKA_03395	*dapA_3*	4-hydroxy-tetrahydrodipicolinate synthase	Bmul_5988	-	-	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig2	397662	CDS		G>A	PROKKA_03410	-	cytochrome c class I	Bmul_5973	C87C	syn	SNP	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	1
contig2	447808	CDS		G>A	PROKKA_03448	*ddpC_1*	putative D,D-dipeptide transport system permease protein DdpC	-	M228I	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	1
contig2	460944	CDS		A>T	PROKKA_03457	-	hypothetical protein	Bmul_6087	P59P	syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	T	1
contig2	473417	CDS		G>A	PROKKA_03469	*fadK*	Short-chain-fatty-acid--CoA ligase	-	T526I	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	1
contig2	539497	CDS		G>A	PROKKA_03526	-	MmgE/PrpD family protein	-	G467D	nonsyn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig2	546030	CDS		T>C	PROKKA_03532	-	Tropinesterase	-	W4R	nonsyn	SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1

contig2	554282	INT	209 bp down	G>A	PROKKA_03538	*mmgC_6*	Acyl-CoA dehydrogenase	-	-	-	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig2	567288	CDS		A>G	PROKKA_03548	*betB_4*	NAD/NADP-dependent betaine aldehyde dehydrogenase	-	Q83R	nonsyn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig2	574133	INT	40 bp down	C>A	PROKKA_03552	*yjjL_1*	L-galactonate transporter	-	-	-	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	1
contig2	581195	CDS		T>C	PROKKA_03557	-	Phage-related baseplate assembly protein	-	D450G	nonsyn	SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig2	586241	CDS		T>C	PROKKA_03561	*rcsC_1*	Sensor histidine kinase RcsC	-	I987V	nonsyn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	1
contig2	594801	CDS		C>G	PROKKA_03567	-	hypothetical protein	-	R133S	nonsyn	SNP	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C	1
contig2	688973	CDS		A>T	PROKKA_03642	*rpoN_2*	RNA polymerase sigma-54 factor	Bmul_5472	V363E	nonsyn	SNP	A	A	A	A	A	A	A	T	A	A	A	A	A	A	A	A	1
contig3	20627	CDS		T>C	PROKKA_03681	-	H-NS histone family protein	-	T15A	nonsyn	SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig3	34055	CDS		C>T	PROKKA_03691	*dnaG*	DNA primase	Bmul_4814	A259A	syn	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig3	118722	CDS		G>A	PROKKA_03772	-	Putative multidrug export ATP-binding/permease protein	Bmul_4939	G143E	nonsyn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig3	186903	CDS		C>A	PROKKA_03833	-	Cyclic nucleotide-binding domain protein	Bmul_5000	G228C	nonsyn	SNP	C	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig3	308608	CDS		C>T	PROKKA_03942	*cynR_8*	LysR family transcriptional regulator	Bmul_5103	N92N	syn	SNP	C	C	C	C	C	C	C	C	C	T	C	C	C	C	del	C	1
contig3	310473	CDS		G>A	PROKKA_03943	-	putative glutamate synthase subunit beta	Bmul_5104	A367T	nonsyn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	del	G	1
contig3	351175	CDS		G>A	PROKKA_03977	*dmlR_33*	LysR family transcriptional regulator	Bmul_5136	P249S	nonsyn	SNP	G	G	G	G	G	G	G	G	A	G	G	G	G	G	del	G	1
contig3	359042	CDS		G>A	PROKKA_03985	*fac-dex*	Fluoroacetate dehalogenase	Bmul_5144	Q313*	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	del	G	1
contig3	433969	CDS		G>A	PROKKA_04055	*aldA_2*	Lactaldehyde dehydrogenase	-	A382T	nonsyn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig3	510482	CDS		T>G	PROKKA_04129	*ydbD_2*	putative manganese catalase	Bmul_5251	S3A	nonsyn	SNP	T	T	T	T	G	T	T	T	T	T	T	T	T	T	T	T	1
contig3	521988	CDS		T>C	PROKKA_04138	*nodD2_6*	LysR family transcriptional regulator	Bmul_5270	Q203R	nonsyn	SNP	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	1
contig3	594714	CDS		A>G	PROKKA_04212	-	hypothetical protein	Bmul_3176	H3R	nonsyn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig3	597578	CDS		C>A	PROKKA_04213	-	type VI secretion protein	Bmul_3177	A606D	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	1
contig3	790722	CDS		G>A	PROKKA_04370	*hmuV*	Hemin import ATP-binding protein HmuV	Bmul_3334	R57W	nonsyn	SNP	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	1

contig3	799034	CDS		C>G	PROKKA_04378	-	Outer membrane porin protein precursor	Bmul_3342	G11A	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	1
contig3	838308	CDS		A>G	PROKKA_04413	*mdtE_2*	Multidrug resistance protein MdtE precursor	Bmul_3373	S119G	nonsyn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	1
contig3	866764	INT	9 bp up	T>G	PROKKA_04436	*gudP_4*	d-galactonate transporter	Bmul_3398	-	-	SNP	T	T	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	1
contig3	875608	INT	42 bp up	C>T	PROKKA_04445	*garD*	D-galactarate dehydratase	Bmul_3409	-	-	SNP	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	1
contig3	968898	CDS		T>C	PROKKA_04518	-	hypothetical protein	-	S563P	nonsyn	SNP	T	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	1
contig3	988084	CDS		C>T	PROKKA_04535	-	hypothetical protein	Bmul_3494	I55I	syn	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig3	994398	CDS		C>A	PROKKA_04541	*Hgd_4*	2-(hydroxymethyl)glutarate dehydrogenase	Bmul_3500	D11Y	nonsyn	SNP	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	1
contig3	1024236	CDS		G>C	PROKKA_04563	*yadA_2*	Adhesin YadA precursor	Bmul_3524	R1154R	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	1
contig3	1032266	CDS		C>T	PROKKA_04572	*naiP_6*	Putative niacin/nicotinamide transporter NaiP	Bmul_3533	D232N	nonsyn	SNP	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	1
contig3	1033674	CDS		G>C	PROKKA_04573	*yadH_3*	Inner membrane transport permease YadH	Bmul_3534	A68G	nonsyn	SNP	G	G	C	G	G	G	G	G	G	G	G	G	G	G	G	G	1
contig3	1055812	CDS		C>T	PROKKA_04590	*etfA_3*	Electron transfer flavoprotein subunit alpha	Bmul_3551	A183V	nonsyn	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig3	1118940	CDS		A>G	PROKKA_04648	*ampR_1*	LysR family transcriptional regulator	Bmul_3613	D40D	syn	SNP	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	1
contig3	1166703	CDS		C>T	PROKKA_04693	-	hypothetical protein	Bmul_3654	D87N	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	1
contig3	1182518	CDS		C>T	PROKKA_04706	*rna*	Ribonuclease I precursor	Bmul_3667	T142I	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	1
contig3	1205969	CDS		A>G	PROKKA_04725	*ampR_2*	LysR family transcriptional regulator	Bmul_3688	A126A	syn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig3	1240716	CDS		G>A	PROKKA_04750	-	hypothetical protein	Bmul_3713	D85D	syn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	1
contig3	1250719	CDS		A>G	PROKKA_04758	*puuA_3*	Gamma-glutamylputrescine synthetase PuuA	Bmul_3721	S433G	nonsyn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig3	1274224	CDS		C>T	PROKKA_04782	*ssaV*	type III secretion FHIPEP protein	Bmul_3739	P23S	nonsyn	SNP	C	C	C	C	C	C	C	C	T	del	C	C	C	C	del	C	1
contig3	1338410	CDS		G>A	PROKKA_04827	*nboR*	Nicotine blue oxidoreductase	Bmul_3790	Y36Y	syn	SNP	G	G	G	G	G	G	G	A	G	del	G	G	G	G	del	G	1
contig3	1341901	CDS		G>T	PROKKA_04831	-	PKHD-type hydroxylase	Bmul_3793	S207S	syn	SNP	G	G	G	G	G	G	G	G	G	del	G	G	G	T	del	G	1
contig3	1419415	CDS		A>G	PROKKA_04897	*rutR_2*	TetR family transcriptional regulator	Bmul_3926	L5L	syn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1

contig3	1437618	CDS		T>C	PROKKA_04912	-	Integrase core domain protein	Bmul_3939, Bmul_0288, Bmul_1339	S136S	syn	SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig3	1651094	INT	159 bp down	G>T	PROKKA_05088	-	Extracellular serine protease precursor	Bmul_4115	-	-	SNP	G	G	G	G	G	G	G	G	G	G	T	G	G	G	G	G	1
contig3	1657941	CDS		C>T	PROKKA_05093	*gudP_6*	putative glucarate transporter	Bmul_4121	A92A	syn	SNP	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	1
contig3	1680682	CDS		G>A	PROKKA_05113	*gabR_6*	HTH-type transcript-tional regulatory protein GabR	Bmul_4136	A38T	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	1
contig3	1723461	CDS		C>T	PROKKA_05142	-	FkbM family methyltransferase	Bmul_4165	T234M	nonsyn	SNP	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	1
contig3	1829583	CDS		G>A	PROKKA_05235	*mrcA_4*	peptidoglycan glycosyltransferase	Bmul_4259	F317F	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	1
contig3	1851820	CDS		C>T	PROKKA_05255	*atoE*	Short-chain fatty acids transporter	Bmul_4279	T3T	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	1
contig3	1921752	CDS		C>G	PROKKA_05316	*ssuA_5*	Putative aliphatic sulfonates-binding protein precursor	Bmul_4340	C241W	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	1
contig3	1934530	CDS		T>C	PROKKA_05329	-	Helix-turn-helix domain protein	Bmul_4367	S80S	syn	SNP	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	1
contig3	2026007	CDS		C>G	PROKKA_05409	*acg*	Putative NAD(P)H nitroreductase acg	-	P34R	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	1
contig3	2045464	CDS		C>G	PROKKA_05424	*dppA_3*	D-aminopeptidase	Bmul_4422	R171P	nonsyn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	1
contig3	2093195	CDS		G>C	PROKKA_05468	-	membrane protein	Bmul_4469	R22P	nonsyn	SNP	G	G	G	G	G	G	G	G	G	G	C	G	G	G	G	G	1
contig3	2132886	INT	119 bp up	G>T	PROKKA_05500	*glnQ_5*	Glutamine transport ATP-binding protein GlnQ	Bmul_4502	-	-	SNP	G	G	G	G	G	G	G	G	G	G	T	G	G	G	G	G	1
contig3	2215477	CDS		C>T	PROKKA_05573	*xanP*	Xanthine permease XanP	Bmul_4581	S272L	nonsyn	SNP	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig3	2287386	CDS		A>G	PROKKA_05630	-	hypothetical protein	-	W46R	nonsyn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig3	2304195	CDS		G>A	PROKKA_05648	-	putative bacteriophage protein	Bmul_1803	H300H	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	1
contig3	2322063	CDS		G>A	PROKKA_05671	-	hypothetical protein	Bmul_1828	L139L	syn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig3	2374729	CDS		A>G	PROKKA_05730	*accD*	acetyl-CoA carboxylase subunit beta	Bmul_4619	I253I	syn	SNP	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	1
contig3	2392109	CDS		G>A	PROKKA_05744	-	hypothetical protein	Bmul_4634	R34R	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	1
contig3	2398341	CDS		G>A	PROKKA_05750	-	Glyoxalase-like domain protein	Bmul_4641	R23R	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	1
contig3	2457867	CDS		G>T	PROKKA_05806	-	Amidohydrolase	Bmul_4701	T241N	nonsyn	SNP	G	T	G	G	G	G	G	G	G	G	G	G	G	G	G	G	1

contig3	2462576	CDS		A>G	PROKKA_05810	*serA*	D-3-phosphoglycerate dehydrogenase	Bmul_4705	V195A	nonsyn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig1	63816	CDS		T	PROKKA_00060	*frc_1*	Formyl-coenzyme A transferase	Bmul_0427	-	2 aa del	indel	tccgtc	tccgtc	tccgtc	tccgtc	tccgtc	tccgtc	tccgtc	tccgtc	tccgtc	tccgtc	tccgtc	tccgtc	tccgtc	------	tccgtc	tccgtc	1
contig1	268629	CDS		G	PROKKA_00247	*pucA*	putative xanthine dehydrogenase subunit A	Bmul_3054	-	frameshift	indel	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	-	1
contig1	388634	CDS		G	PROKKA_00361		cytochrome B561	Bmul_0005	-	17 aa del	indel	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	aacgtcaccggattcgccgcaacgatcgcgggcaactgcacacccgcaccc	---	1
contig1	1002688	CDS		C	PROKKA_00936	*dtd*	D-tyrosyl-tRNA(Tyr) deacylase	Bmul_2699	-	frameshift	indel	cg	cg	cg	cg	cg	cg	cg	--	cg	cg	cg	cg	cg	cg	cg	cg	1
contig1	1092287	CDS		A	PROKKA_01017	*mutL*	DNA mismatch repair protein MutL	Bmul_2621	-	16 aa del	indel	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	--	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	gcaccgcgcgcgcgagtgcctgcaggtcggcgtccttcagc	1

												agcgcag	agcgcag	agcgcag	agcgcag	agcgcag	agcgcag	agcgcag		agcgcag	agcgcag	agcgcag	agcgcag	agcgcag	agcgcag	agcgcag	agcgcag	
contig1	1140837	CDS		G	PROKKA_01068	*kdhB*	6-hydroxy-pseudooxynicotine dehydrogenase complex subunit beta	Bmul_2552	-	frameshift	indel	gc	gc	gc	gc	gc	gc	gc	--	gc	gc	gc	gc	gc	gc	gc	gc	1
contig1	1733619	CDS		T	PROKKA_01624	*tctD_3*	Transcriptional regulatory protein tctD	Bmul_0649	-	frameshift	indel	gc	gc	gc	gc	gc	gc	gc	--	gc	gc	gc	gc	gc	gc	gc	gc	1
contig1	1764172	CDS		C	PROKKA_01653	*dacB*	D-alanyl-D-alanine carboxypeptidase DacB precursor	Bmul_0679	-	4 aa del	indel	cgacgcgctcga	cgacgcgctcga	cgacgcgctcga	cgacgcgctcga	cgacgcgctcga	cgacgcgctcga	cgacgcgctcga	cgacgcgctcga	cgacgcgctcga	cgacgcgctcga	cgacgcgctcga	----------	cgacgcgctcga	cgacgcgctcga	cgacgcgctcga	cgacgcgctcga	1
contig1	1764306	CDS		G	PROKKA_01653	*dacB*	D-alanyl-D-alanine carboxypeptidase DacB precursor	Bmul_0679	-	frameshift	indel	-	-	-	-	t	-	-	-	-	-	-	-	-	-	-	-	1
contig1	1771192	INT	67 bp down	G	PROKKA_01661	*ttgR_1*	TetR family transcriptional regulator	Bmul_0687	-	-	indel	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	acgaacacgaa	----------	----------	----------	----------	----------	1
contig1	1903681	CDS		T	PROKKA_01778	*ascD*	CDP-6-deoxy-L-threo-D-glycero-4-hexulose-3-dehydrase reductase	Bmul_0805	-	2 aa del	indel	cgtcgg	cgtcgg	cgtcgg	cgtcgg	cgtcgg	cgtcgg	cgtcgg	-----	cgtcgg	cgtcgg	cgtcgg	cgtcgg	cgtcgg	cgtcgg	cgtcgg	cgtcgg	1
contig1	1945245	INT	159 bp up	A	PROKKA_01823	*mhqP*	Putative oxidoreductase MhqP	Bmul_0849	-	-	indel	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	ctcacgg	1
contig1	1982365	CDS		G	PROKKA_01851	-	hypothetical protein	Bmul_0877	-	frameshift	indel	-	-	-	-	-	-	-	c	-	-	-	-	-	-	-	-	1
contig1	2602189	INT	29 bp down	C	PROKKA_02438	-	hypothetical protein	Bmul_1390	-	-	indel	-	-	-	-	-	-	-	g	-	-	-	-	-	-	-	-	1
contig1	2868787	CDS		T	PROKKA_02676	*papC*	fimbrial biogenesis outer membrane usher protein	Bmul_1611	-	frameshift	indel	c	c	c	c	c	c	c	c	c	c	-	c	c	c	c	c	1
contig1	2878400	CDS		G	PROKKA_02681	*ttgR_3*	TetR family transcriptional regulator	Bmul_1617	-	frameshift	indel	ac	ac	ac	ac	ac	ac	ac	ac	ac	ac	ac	--	ac	ac	ac	ac	1
contig1	2878405	CDS		A	PROKKA_02681	*ttgR_3*	TetR family transcriptional regulator	Bmul_1617	-	2 aa insertion	indel	------	------	------	------	------	------	------	------	------	------	------	attttgt	------	------	------	------	1
contig1	2892952	CDS		G	PROKKA_02692	*rbsR*	LacI family transcriptional regulator	Bmul_1628	-	frameshift	indel	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	cgccgccgcgac	-----------	1

												ga ac	ga ac	ga ac	ga ac	ga ac	ga ac	ga ac	ga ac	ga ac	ga ac	ga ac	ga ac	ga ac	ga ac	ga ac		
contig1	3097524	INT	27 bp up	C	PROKKA_02883	*mglB*	D-galactose-binding periplasmic protein precursor	Bmul_1891	-	-	indel	-	-	-	-	-	-	-	g	-	-	-	-	-	-	-	-	1
contig2	638096	CDS		C	PROKKA_03603		hypothetical protein	Bmul_5510	-	4 aa del	indel	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	gc cg cg ca gg cg	---- ---- ----	1
contig3	269092	INT	23 bp down	C	PROKKA_03905	*proP_20*	Proline/betaine transporter	Bmul_5065	-	-	indel	---- ---- -	---- ---- -	---- ---- -	---- ---- -	---- ---- -	---- ---- -	---- ---- -	---- ---- -	---- ---- -	---- ---- -	---- ---- -	---- ---- -	---- ---- -	---- ---- -	---- ---- -	---- ---- -	gc cg cg ggt	1
contig3	286833	CDS		T	PROKKA_03922	*fixL_2*	Sensor protein FixL	Bmul_5082	-	5 aa del	indel	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	---- ---- ---- ---	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	cc gc gct tcc gc gc g	1
contig3	295434	INT	163 bp up	A	PROKKA_03931	-	Acetyltransferase (GNAT) family protein	Bmul_5091	-	-	indel	---- ----	---- ----	---- ----	---- ----	---- ----	---- ----	---- ----	---- ----	---- ----	---- ----	---- ----	---- ----	---- ----	---- ----	---- ----	---- ----	ctt gc ctg	1
contig3	582738	CDS		C	PROKKA_04201	-	hypothetical protein	-	-	framesh ift	indel	gct cg gt	gct cg gt	gct cg gt	gct cg gt	gct cg gt	gct cg gt	gct cg gt	gct cg gt	gct cg gt	gct cg gt	gct cg gt	gct cg gt	gct cg gt	gct cg gt	gct cg gt	---- ---	1
contig3	654495	CDS		C	PROKKA_04249	*fadD_4*	Long-chain-fatty-acid-CoA ligase	Bmul_3212	-	framesh ift	indel	a	a	a	a	a	a	a	-	a	a	a	a	a	a	a	a	1
contig3	856118	INT	39 bp up	T	PROKKA_04427	*dctB_2*	C4-dicarboxylate transport sensor protein DctB	Bmul_3389	-	-	indel	cg gc cgt cc aa c	cg gc cgt cc aa c	cg gc cgt cc aa c	cg gc cgt cc aa c	cg gc cgt cc aa c	cg gc cgt cc aa c	cg gc cgt cc aa c	cg gc cgt cc aa c	cg gc cgt cc aa c	cg gc cgt cc aa c	cg gc cgt cc aa c	cg gc cgt cc aa c	cg gc cgt cc aa c	cg gc cgt cc aa c	---- ---- ----	cg gc cgt cc aa c	1
contig3	1098459	CDS		T	PROKKA_04630	-	2-keto-3-deoxy-L-fuconate dehydrogenase	Bmul_3595	-	framesh ift	indel	cg	cg	cg	cg	cg	cg	cg	--	cg	cg	cg	cg	cg	cg	cg	cg	1
contig3	1340843	INT	38 bp down	C	PROKKA_04829	-	hypothetical protein	-	-	-	indel	-	-	-	-	-	-	-	g	-	-	-	-	-	-	-	-	1
contig3	1512900	CDS		T	PROKKA_04989	-	metal-dependent hydrolase	Bmul_4017	-	framesh ift	indel	cg	cg	cg	cg	cg	cg	cg	--	cg	cg	cg	cg	cg	cg	cg	cg	1
contig3	1653113	CDS		A	PROKKA_05089	*clcA_3*	chloride channel core	Bmul_4117	-	framesh ift	indel	gc	gc	gc	gc	gc	gc	gc	--	gc	gc	gc	gc	gc	gc	gc	gc	1
contig3	1886952	INT	35 bp up	G	PROKKA_05286	-	Long-chain-fatty-acid-CoA ligase	Bmul_4310	-	-	indel	-	-	-	-	-	-	-	c	-	-	-	-	-	-	-	-	1
contig3	2117637	CDS		T	PROKKA_05485	*glmS_4*	Glutamine-fructose-6-phosphate aminotransferase [isomerizing]	Bmul_4486	-	framesh ift	indel	gc	gc	gc	gc	gc	gc	gc	--	gc	gc	gc	gc	gc	gc	gc	gc	1

contig3	2324562	CDS		G	PROKKA_05674		hypothetical protein	Bmul_1831	-	frameshift	indel	-	-	-	-	-	-	-	c	-	-	-	-	-	-	-	-	1
contig3	2353647	INT	47 bp down	C	PROKKA_05713	-	3-demethylubiquinone-9 3-methyltransferase	Bmul_4601	-	-	indel	gcggct	gcggct	gcggct	gcggct	gcggct	gcggct	gcggct	gcggct	gcggct	gcggct	gcggct	gcggct	gcggct	gcggct	gcggct	------	1
contig1	1144822				PROKKA_01072	*syrM1_1*	HTH-type transcriptional regulator SyrM 1			123 aa del	indel	G	G	G	G	G	G	G	369-bp del	G	G	G	G	G	G	G	G	1
contig1	2444332				PROKKA_02289 to PROKKA_02299					3787 aa del	indel	C	C	C	C	C	C	C	11360-bp del	C	C	C	C	C	C	C	C	1
contig1	1576769				PROKKA_01479; PROKKA_01480					385 aa del	indel	T	T	T	T	T	T	T	1155-bp del	T	T	T	T	T	T	T	T	1
contig1	1577258				PROKKA_01479	*pagL*	Lipid A deacylase PagL precursor			49 aa del	indel	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	148-bp del	1
contig2	122606				PROKKA_03165; PROKKA_03166	*tRNA*				82 aa del	indel	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	246-bp del	1
contig3	275479				PROKKA_03911 to PROKKA_04047					50383 aa del	indel	C	C	C	C	C	C	C	C	C	C	C	C	C	C	151147-bp del	C	1
contig3	1447293				PROKKA_04923 to PROKKA_04945					6874 aa del	indel	G	G	G	G	G	G	G	20622-bp del	G	G	G	G	G	G	G	G	1
contig3	1643082				PROKKA_05088	-	Extracellular serine protease precursor			980 aa del	indel	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	2938-bp del	1
contig3	2418658				PROKKA_05769; PROKKA_05770					391 aa del	indel	G	G	G	G	G	G	G	1172-bp del	G	G	G	G	G	G	G	G	1

Table A2 - Mutations found among the 21 *B. multivorans* isolates of CF patient P426 (CDS = coding sequence, INT = intergenic, del = large deletion).

Contig	Position	Category	bp upstream or downstream CDS (if IG)	Mutation	Locus tag (bp upstream if IG)	Gene name	Annotation	Homologue in *B. multivorans* ATCC 17616	Mutations found (if CDS)	Type of mutation	Type of mutation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Nr of isolates with mutation
contig1	502108	CDS		T>G	PROKKA_00467	*ytnP*	putative quorum-quenching lactonase YtnP	-	S84R	non syn	SNP	T	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	20
contig1	536415	CDS		G>T	PROKKA_00501	*pbuE_1*	major facilitator transporter	Bmul_6097	L176M	non syn	SNP	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	20
contig2	970630	CDS		C>T	PROKKA_01972	*adrB*	cyclic-di-GMP phosphodiesterase	Bmul_5257	R145H	non syn	SNP	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	18
contig2	1264721	CDS		G>C	PROKKA_02243	*gcvA_11*	LysR family transcriptional regulator	Bmul_4992	P114A	non syn	SNP	G	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	18
contig3	1274062	CDS		G>C	PROKKA_04445	*czcA_2*	acriflavin resistance protein	Bmul_1519	G452A	non syn	SNP	G	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	18
contig3	1386691	CDS		C>A	PROKKA_04548	*hom*	Homoserine dehydrogenase	Bmul_1419	A121S	non syn	SNP	C	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	A	A	A	A	18
contig3	2578308	CDS		G>A	PROKKA_05665	*accC_1*	acetyl-CoA carboxylase biotin carboxylase subunit	Bmul_0367	D444N	non syn	SNP	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	18
contig3	2756613	INT	78 up	A>C	PROKKA_05835	*sfaG*	S-fimbrial protein subunit SfaG precursor	-			SNP	A	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	18
contig3	3197434	CDS		G>T	PROKKA_06251	*pilQ*	type IV pilus secretin PilQ	Bmul_0294	L387L	syn	SNP	G	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	T	T	T	T	18
contig2	61308	CDS		G>T	PROKKA_01182	*budC_2*	3-ketoacyl-ACP reductase	Bmul_3804	D219E	non syn	SNP	G	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	T	T	T	T	18
contig2	1503211	CDS		T>G	PROKKA_02456	*potH_2*	transport system permease protein	Bmul_5430	L225R	non syn	SNP	T	T	G	G	G	G	G	T	G	G	G	G	G	G	G	G	T	G	G	G	G	17

contig2	1503212	CDS		C>A	PROKKA_02456	*potH_2*	transport system permease protein	Bmul_5430	L225L	syn	SNP	C	C	A	A	A	A	A	C	A	A	A	A	A	A	A	A	C	A	A	A	A	17
contig3	1228781	CDS		G>C	PROKKA_04396	-	cobalt-precorrin-6A synthase	Bmul_1566	V311L	non syn	SNP	G	G	G	G	G	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	15
contig3	1620376	CDS		G>C	PROKKA_04768	*cyaA*	Ubiquinol oxidase subunit 1	Bmul_1210	A566G	non syn	SNP	G	G	G	G	G	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	15
contig3	1654573	CDS		G>A	PROKKA_04803	-	Phage late control gene D protein (GPD)	-	G183G	syn	SNP	G	G	G	G	G	A	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	15
contig1	645393	CDS		C>A	PROKKA_00596	*hcnA_1*	Hydrogen cyanide synthase subunit HcnA	Bmul_6004	G76C	non syn	SNP	C	C	C	C	C	A	A	C	A	A	A	A	A	A	A	A	C	A	A	A	A	14
contig3	1369080	CDS		C>A	PROKKA_04532	*acoB*	Acetoin 2,6-dichlorophenolindophenol oxidoreductase subunit beta	Bmul_1435	A233S	non syn	SNP	C	C	C	C	C	A	A	C	A	A	A	A	A	A	A	A	C	A	A	A	A	14
contig3	68863				PROKKA_03313; PROKKA_03314	*glnQ_4; flu*	Glutamine transport ATP-binding protein GlnQ; Antigen 43 precursor			275 aa deletion	indel	A	A	A	A	A	824-bp deletion	824-bp deletion	A	824-bp deletion	824-bp deletion	824-bp deletion	824-bp deletion	824-bp deletion	824-bp deletion	824-bp deletion	824-bp deletion	A	824-bp deletion	824-bp deletion	824-bp deletion	824-bp deletion	14
contig2	218150	CDS		G>A	PROKKA_01308	*rstB_1*	Sensor protein RstB	Bmul_3678	R119C	non syn	SNP	G	G	G	G	G	G	A	A	A	A	A	A	G	A	A	A	G	A	A	A	A	13
contig1	570779	INT	73 down	A>C	PROKKA_00530	-	Putative multidrug export ATP-binding/permease protein	Bmul_6066			SNP	A	A	A	A	A	A	C	A	C	C	C	C	A	C	C	C	A	C	C	C	C	12
contig3	1300130	CDS		G>A	PROKKA_04469	-	putative assembly protein	Bmul_1499	S254L	non syn	SNP	G	G	G	G	G	G	G	A	A	A	A	A	G	A	A	A	G	G	A	G	A	10
contig3	205812	CDS		G>A	PROKKA_03438	*mmgC_3*	Acyl-CoA dehydrogenase	Bmul_2570	G130G	syn	SNP	G	G	G	G	G	G	G	G	A	A	A	A	G	A	A	A	G	G	A	A	A	10
contig3	271374	CDS		C>G	PROKKA_03495	-	hypothetical protein	-	G471G	syn	SNP	C	C	C	C	C	C	C	G	G	G	G	G	del	G	G	G	del	C	G	del	G	10
contig3	278036	CDS		A>C	PROKKA_03500	*gmd_2*	GDP-mannose 4,6-dehydratase	-	L210R	non syn	SNP	A	A	A	A	A	A	A	C	C	C	C	C	del	C	C	C	del	A	C	del	C	10

contig3	281676	CDS		G>T	PROKKA_03501	*mfpsA*	Mannosylfructose-phosphate synthase	-	P251P	syn	SNP	G	G	G	G	G	G	G	T	T	T	T	T	del	T	T	T	del	G	T	del	T	10
contig2	278535	CDS		A>G	PROKKA_01370	-	hypothetical protein	-	K29R	non syn	SNP	A	A	A	A	A	A	A	G	G	G	G	G	del	G	G	G	A	A	G	A	G	10
contig2	1550406	CDS		C>A	PROKKA_02498	-	cell division protein DedD	Bmul_4617	P174Q	non syn	SNP	C	C	C	C	C	C	C	A	A	A	A	A	C	A	A	A	C	C	A	C	A	10
contig2	1699017	CDS		A>G	PROKKA_02625	*pgaC*	Poly-beta-1,6-N-acetyl-D-glucosamine synthase	Bmul_4484	L24P	non syn	SNP	A	A	A	A	A	A	A	G	G	G	G	G	A	G	G	G	A	A	G	A	G	10
contig3	63037	CDS		G>C	PROKKA_03307	*purB*	Adenylosuccinate lyase	Bmul_2717	A56G	non syn	SNP	G	G	G	G	G	G	G	C	C	C	C	C	G	C	C	C	G	G	C	G	C	10
contig3	923396	CDS		G>C	PROKKA_04097	*lplT_1*	phospholipid/glycerol acyltransferase	Bmul_1790	W590C	non syn	SNP	G	G	G	G	G	G	G	C	C	C	C	C	G	C	C	C	G	G	C	G	C	10
contig3	1070008	CDS		C>G	PROKKA_04259	-	putative assembly protein	Bmul_1703	G197G	syn	SNP	C	C	C	C	C	C	C	G	G	G	G	G	C	G	G	G	C	C	G	C	G	10
contig3	1745194	CDS		G>T	PROKKA_04896	*fixL_2*	Sensor protein FixL	Bmul_1132	G231C	non syn	SNP	G	G	G	G	G	G	G	T	T	T	T	T	G	T	T	T	G	G	T	G	T	10
contig3	2356074	CDS		A>T	PROKKA_05461	*lptD*	LPS-assembly protein LptD precursor / organic solvent tolerance protein	Bmul_0588	V754D	non syn	SNP	A	A	A	A	A	A	A	T	T	T	T	T	A	T	T	T	A	A	T	A	T	10
contig3	2396050	CDS		G>A	PROKKA_05494	*bmr3_2*	major facilitator transporter	Bmul_0554	A239V	non syn	SNP	G	G	G	G	G	G	G	A	A	A	A	A	G	A	A	A	G	G	A	G	A	10
contig2	144318	CDS		C>G	PROKKA_01249	-	LuxR family transcriptional regulator	Bmul_3738	T173T	syn	SNP	C	C	C	C	C	C	C	C	G	G	G	G	C	G	G	G	C	C	G	C	G	9
contig2	1429902	CDS		A>G	PROKKA_02397	-	hypothetical protein	Bmul_4347	T11A	non syn	SNP	A	A	A	A	A	A	A	A	G	G	G	G	A	G	G	G	A	A	G	A	G	9
contig3	2789456	CDS		T>G	PROKKA_05863	*zraR_8*	Fis family transcriptional regulator	Bmul_3117	E184D	non syn	SNP	T	T	T	T	T	T	T	T	G	G	G	G	T	G	G	G	T	T	G	T	G	9
contig2	1172281	CDS		G>C	PROKKA_02160	*pobB_2*	Phenoxybenzoate dioxygenase subunit beta	Bmul_5072	P256R	non syn	SNP	G	G	G	G	G	G	G	C	C	C	C	C	G	C	C	C	G	G	C	G	del	9
contig2	691600				PROKKA_01734	-	Outer membrane porin protein precursor			294 aa deletion	indel	T	T	T	T	T	T	T	T	880-bp del	880-bp del	880-bp del	880-bp del	T	880-bp del	880-bp del	880-bp del	T	T	880-bp del	T	880-bp del	9

																						etion	etion	etion	etion		etion	etion	etion			etion		etion	
contig3	23514	CDS		C>T	PROKKA_03270	*pncC*	Nicotinamide-nucleotide amidohydrolase PncC	Bmul_2757	A27T	non syn	SNP	C	C	C	C	C	C	C	T	T	C	T	T	C	T	T	T	C	C	C	C	T	8
contig1	16531				PROKKA_00017 to PROKKA_00225					72592 aa deletion	indel	T	T	T	T	T	T	T	T	2177774-bp deletion	T	2177774-bp deletion	2177774-bp deletion	T	2177774-bp deletion	2177774-bp deletion	2177774-bp deletion	T	T	T	2177774-bp deletion	2177774-bp deletion	8
contig3	127180	CDS		G>A	PROKKA_03364	-	hypothetical protein	-	M1I	non syn	SNP	G	G	G	G	G	G	G	G	A	G	A	A	G	A	A	A	G	G	G	G	A	7
contig3	419434	CDS		C>T	PROKKA_03648	*ydfH_3*	GntR family transcriptional regulator	Bmul_2372	A60T	non syn	SNP	C	C	C	C	C	C	C	C	T	C	T	T	C	T	C	T	C	C	C	C	T	6
contig3	1001405	CDS		C>T	PROKKA_04202	-	transcription elongation factor NusA	Bmul_1756	I117I	syn	SNP	C	C	C	C	C	C	C	C	T	C	T	T	C	T	C	T	C	C	C	C	T	6
contig3	1307725	CDS		T>C	PROKKA_04477	*metC*	Cystathionine beta-lyase MetC	Bmul_1491	Q207R	non syn	SNP	T	T	T	T	T	T	T	T	C	T	C	C	T	C	T	C	T	T	T	T	C	6
contig3	945339				PROKKA_04119 to PROKKA_04187					14633 aa deletion	indel	A	43898-bp deletion	A	A	A	A	A	A	A	A	43898-bp deletion	43898-bp deletion	A	A	A	43898-bp deletion	A	A	A	43898-bp deletion	43898-bp deletion	6
contig1	887178				PROKKA_00810 to PROKKA_00904					29820 aa deletion	indel	A	A	A	A	89458-bp deletion	A	A	A	A	89458-bp deletion	89458-bp deletion	A	A	A	A	A	A	A	89458-bp deletion	A	89458-bp deletion	5
contig1	1148967				PROKKA_01069 to PROKKA_01078					1668 aa deletion	indel	C	C	C	C	C	C	5002-bp deletion	C	5002-bp deletion	C	5002-bp deletion	C	C	C	C	5002-bp deletion	C	C	C	C	5002-bp deletion	5
contig1	1155187				PROKKA_01080 to PROKKA_01126					16318 aa deletion	indel	G	G	G	G	G	G	48952-bp del	G	48952-bp del	G	48952-bp del	G	G	G	G	48952-bp del	G	G	G	G	48952-bp del	5

																				eti on		eti on		eti on					eti on					eti on		
contig1	995236				PROKKA _00918 to PROKKA _00953					120 31 aa dele tion	indel	C	C	C	C	36 09 3- bp del eti on	C	C	C	C	36 09 3- bp del eti on	36 09 3- bp del eti on	C	C	C	C	C	C	C	36 09 3- bp del eti on	C	del	4
contig1	475138	INT	147 up	C>A	PROKKA _00443	-	Initiator Replication protein	Bmul_ 6162			SNP	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	A	C	C	C	C	A	3
contig1	806595	CDS		G>A	PROKKA _00740	-	Universal stress protein	Bmul_ 5859	I141I	syn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	A	G	G	G	G	A	3
contig2	43593	CDS		G>A	PROKKA _01167	*tpa_3*	Taurine-pyruvate aminotransferase	Bmul_ 3821	R367C	non syn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	A	G	G	G	G	A	3
contig2	1462431	INT	63 down	C>T	PROKKA _02418	*luxQ_ 4*	Autoinducer 2 sensor kinase/phosphatase LuxQ	Bmul_ 4695			SNP	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	T	C	C	C	C	T	3
contig3	812372	CDS		T>C	PROKKA _04003	-	Membrane-bound aldehyde dehydrogenase [pyrroloquinoline-quinone] precursor	Bmul_ 1961	F545L	non syn	SNP	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	C	T	T	T	T	C	3
contig3	1003769	CDS		A>G	PROKKA _04203	*infB*	Translation initiation factor IF-2	Bmul_ 1755	E383G	non syn	SNP	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	G	A	A	A	A	G	3
contig3	1099889	INT	113 up	C>T	PROKKA _04286	*pyrD*	Dihydroorotate dehydrogenase (quinone)	Bmul_ 1678			SNP	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	T	C	C	C	C	T	3
contig3	1193266	CDS		T>G	PROKKA _04365	*iucB_ 1*	N(6)-hydroxylysine O-acetyltransferase	Bmul_ 1596	V338G	non syn	SNP	T	T	T	T	T	T	T	T	T	T	G	T	T	T	T	G	T	T	T	T	G	3
contig3	1322782	CDS		T>G	PROKKA _04491	-	hypothetical protein	Bmul_ 1477	Q218P	non syn	SNP	T	T	T	T	T	T	T	T	T	T	G	T	T	T	T	G	T	T	T	T	G	3
contig3	1888759	INT	96 down	G>A	PROKKA _05029	*secG*	Protein-export	Bmul_ 1027			SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	A	G	G	G	G	A	3

							membrane protein SecG																										
contig3	2043737	CDS		G>C	PROKKA_05175	*nadE_2*	Glutamine-dependent NAD(+) synthetase	Bmul_0878	P49A	non syn	SNP	G	G	G	G	G	G	G	G	G	G	C	G	G	G	G	C	G	G	G	G	C	3
contig3	2711433	INT	17 down	T>G	PROKKA_05789	*fpr*	Ferredoxin-NADP reductase	Bmul_3050			SNP	T	T	T	T	T	T	T	T	T	T	G	T	T	T	T	G	T	T	T	T	G	3
contig1	1030237				PROKKA_00954 to PROKKA_00958					2505 aa deletion	indel	T	T	T	T	del	T	7513-bp deletion	T	7513-bp deletion	del	del	T	T	T	T	7513-bp deletion	T	T	del	T	del	3
contig1	1038936				PROKKA_00960 to PROKKA_01069					36393 aa deletion	indel	A	A	A	A	A	A	109177-bp deletion	A	109177-bp deletion	del	109177-bp deletion	A	A	A	A	A	A	A	del	A	A	3
contig3	824543	CDS		C>A	PROKKA_04013	*fecI_1*	putative RNA polymerase sigma factor FecI	Bmul_1935	R10R	syn	SNP	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	C	A	2
contig1	34573	CDS		A>T	PROKKA_00036	*ydaP_2*	Putative thiamine pyrophosphate-containing protein YdaP	Bmul_5505	K179M	non syn	SNP	A	A	A	A	A	A	del	T	del	A	del	del	A	del	del	del	T	A	A	del	del	2
contig1	216460	CDS		C>T	PROKKA_00213	*uvrA_1*	UvrABC system protein A	Bmul_5541	A1807A	syn	SNP	C	C	C	T	T	C	C	C	del	C	del	del	C	del	del	del	C	C	C	del	del	2
contig1	370924	CDS		T>A	PROKKA_00349	-	oxidoreductase alpha (molybdopterin) subunit	Bmul_5410	I657N	non syn	SNP	T	T	T	T	T	A	T	T	T	T	T	T	A	T	T	T	T	T	T	T	T	2
contig2	811057	CDS		A>G	PROKKA_01840	*aroG_1*	Phospho-2-dehydro-3-deoxyheptonate aldolase, Phe-sensitive	Bmul_3232	V28A	non syn	SNP	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	G	2
contig2	1401215	CDS		C>A	PROKKA_02371	*ispA*	polyprenyl synthetase	Bmul_4821	L139I	non syn	SNP	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	C	A	2
contig2	1418698	CDS		C>A	PROKKA_02386	-	Outer membrane	Bmul_4702	V310V	syn	SNP	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	C	C	A	C	C	C	2

							porin protein precursor																										
contig2	1487897	CDS		C>T	PROKKA_02442	-	Methyltransferase domain protein	Bmul_4677	D69D	syn	SNP	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	T	C	C	2
contig2	1493930	INT	44 up	G>C	PROKKA_02450	*hemR_2*	Hemin receptor precursor	Bmul_4669			SNP	G	G	G	G	G	G	C	G	G	G	G	G	G	G	G	G	G	C	G	G	G	2
contig2	1665876	INT	65 up	C>T	PROKKA_02592	*xsc*	Sulfoacetaldehyde acetyltransferase	Bmul_4516			SNP	C	C	C	C	C	T	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	2
contig2	1714157	CDS		C>T	PROKKA_02635	*pphA_2*	Phosphonopyruvate hydrolase	Bmul_4471	R103H	non syn	SNP	C	C	C	C	C	T	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	2
contig2	1737706	CDS		G>T	PROKKA_02656	*hcnC_3*	Hydrogen cyanide synthase subunit HcnC precursor	Bmul_4446	A371S	non syn	SNP	G	G	G	G	G	T	G	G	G	G	G	G	T	G	G	G	G	G	G	G	G	2
contig2	1909827	CDS		C>T	PROKKA_02798	-	chaperone protein HchA	Bmul_4323	A29V	non syn	SNP	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	T	C	C	2
contig2	2035288	CDS		G>A	PROKKA_02911	*araE*	Alpha-ketoglutaric semialdehyde dehydrogenase	Bmul_4215	R358Q	non syn	SNP	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	A	G	G	G	2
contig3	127177	INT	61 up	C>T	PROKKA_03364	-	hypothetical protein	-			SNP	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	T	C	C	2
contig3	353940	CDS		C>G	PROKKA_03572	*glk*	Glucokinase	Bmul_2433	G26A	non syn	SNP	C	C	C	G	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	2
contig3	390997	CDS		T>C	PROKKA_03618	*spoT*	Bifunctional (p)ppGpp synthase/hydrolase SpoT	Bmul_2399	S33P	non syn	SNP	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	C	T	T	T	2
contig3	452134	CDS		C>G	PROKKA_03676	*rfaE*	D-beta-D-heptose 7-phosphate kinase	Bmul_2252	G145G	syn	SNP	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C	C	C	G	C	C	C	2
contig3	1145720	CDS		C>T	PROKKA_04337	-	putative serine protein kinase PrkA	Bmul_1625	R308C	non syn	SNP	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	T	C	C	C	2
contig3	1612228	CDS		C>A	PROKKA_04760	*mepH_2*	Murein DD-endopeptidase MepH precursor	Bmul_1218	G70C	non syn	SNP	C	C	C	C	C	A	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	2

contig3	1731036	INT	38 up	A>T	PROKKA_04885	*iscR*	BadM/Rrf2 family transcriptional regulator	Bmul_1143			SNP	A	A	A	T	T	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	2
contig3	1756102	CDS		A>G	PROKKA_04904	*f1pep1*	Prolyl endopeptidase precursor	Bmul_1125	V146A	non syn	SNP	A	A	A	A	A	G	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	2
contig3	1823404	CDS		A>C	PROKKA_04961	-	hypothetical protein	-	K601T	non syn	SNP	A	A	A	A	A	A	C	A	A	A	A	A	A	A	A	A	A	C	A	A	A	2
contig3	1951341	CDS		G>A	PROKKA_05090	*rfaQ_3*	Lipopolysaccharide core heptosyltransferase RfaQ	-	A90V	non syn	SNP	G	G	G	G	G	A	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	2
contig3	2211372	CDS		A>G	PROKKA_05331	-	hypothetical protein	Bmul_0723	M32V	non syn	SNP	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	G	2
contig3	3373463	CDS		C>G	PROKKA_06417	*outO_2*	prepilin peptidase	Bmul_2820	W244C	non syn	SNP	C	C	C	C	C	G	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C	2
contig1	299121	CDS		G>A	PROKKA_00277	-	hypothetical protein	Bmul_5479	H89Y	non syn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	A	G	G	G	G	2
contig1	446262	CDS		G>A	PROKKA_00418	-	hypothetical protein	Bmul_5343	G16R	non syn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	A	G	G	G	G	2
contig1	558634	INT	100 up	G>A	PROKKA_00518	*braC_1*	Leucine-, isoleucine-, valine-, threonine-, and alanine-binding protein precursor	Bmul_6079			SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	A	G	G	G	G	2
contig2	1581994	INT	tRNA	C>G	PROKKA_02523	-	tRNA-Sec				SNP	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C	C	G	C	C	C	C	2
contig2	2293878	CDS		G>A	PROKKA_03128	-	Outer membrane porin protein precursor	Bmul_4039	N38N	syn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	A	G	G	G	G	2
contig3	635870	INT	74 up	T>C	PROKKA_03856	*proY*	Proline-specific permease ProY	Bmul_2091			SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	C	T	T	T	T	2
contig3	810899	CDS		G>A	PROKKA_04003	-	Membrane-bound aldehyde dehydrogenase [pyrroloquinoline-quinone] precursor	Bmul_1961	G54S	non syn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	A	G	G	G	G	2
contig3	1204221	INT	3 up	G>A	PROKKA_04374	*cobO*	Cob(I)yrinic acid a,c-diamide	Bmul_1587			SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	A	G	G	G	G	2

							adenosyltransferase																									
contig3	1746932	CDS		T>C	PROKKA_04896	*fixL_2*	Sensor protein FixL	Bmul_1132	M810T	non syn	SNP	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	C	T	T	T	T	2
contig3	2429786	CDS		G>T	PROKKA_05526	*hprK*	HPr kinase/phosphorylase	Bmul_0521	R200L	non syn	SNP	G	G	G	G	G	G	G	T	G	G	G	G	G	G	G	G	T	G	G	G	G	2
contig3	2496032	CDS		A>G	PROKKA_05592	*ptsH*	Phosphocarrier protein HPr	Bmul_0442	L77P	non syn	SNP	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	G	A	A	A	A	2
contig3	3332455	INT	304 down	G>A	PROKKA_06381	*echA8_4*	enoyl-CoA hydratase	Bmul_2850			SNP	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	A	G	G	G	G	2
contig2	1204840	INT	119 down	G	PROKKA_02188	-	Carboxymuconolactone decarboxylase family protein	Bmul_5046			indel	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	gccgacgcaa	-----------	-----------	gccgacgcaa	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	2
contig2	1906501	CDS		C	PROKKA_02794	-	Outer membrane porin protein precursor	Bmul_4327			indel	aa	aa	aa	aa	aa	aa	--	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	--	aa	aa	aa	2
contig3	271595	CDS		A	PROKKA_03495	-	hypothetical protein	-			indel	c	c	c	c	c	c	-	c	c	c	c	c	c	c	c	c	-	c	c	c	2
contig3	372634	CDS		G	PROKKA_03589	-	putative acetyltransferase	Bmul_2417			indel	-	-	-	-	-	c	-	-	-	-	-	-	c	-	-	-	-	-	-	-	-	2
contig3	1155926	CDS		A	PROKKA_04345	*ttgR_1*	TetR family transcriptional regulator	Bmul_1617			indel	---	---	---	---	---	---	---	gag	---	---	---	---	---	---	---	---	gag	---	---	---	---	2
contig3	1281297	CDS		G	PROKKA_04451	*vnfA*	Nitrogen fixation protein VnfA	Bmul_1516			indel	tc	tc	tc	tc	tc	tc	tc	--	tc	tc	tc	tc	tc	tc	tc	tc	--	tc	tc	tc	tc	2
contig1	1038909				PROKKA_00960; PROKKA_00961	*cbdA_1; naiP_2*	2-halobenzoate 1,2-dioxygenase large subunit; Putative niacin/nicotinamide transporter NaiP			826 aa deletion	indel	G	G	G	G	G	G	G	G	G	2476-bp deletion	G	G	G	G	G	G	G	G	2476-bp deletion	G	G	2
contig1	1127241				PROKKA_01048 to PROKKA_01069					6993 aa deletion	indel	T	T	T	T	T	T	del	T	del	T	del	T	T	T	T	20979-bp deletion	T	T	T	T	20979-bp deletion	2

contig3	269236				PROKKA_03494 to PROKKA_03506					6282 aa deletion	indel	A	A	A	A	A	A	A	A	A	A	A	A	del	A	A	A	1884 4-bp deletion	A	A	1884 4-bp deletion	A	2
contig1	12406	INT	84 down	G>A	PROKKA_00012	-	hypothetical protein	-			SNP	G	G	G	G	G	G	del	G	G	G	G	G	G	G	G	G	G	G	G	A	G	1
contig1	33042	CDS		G>A	PROKKA_00035	*ydaP_1*	Putative thiamine pyrophosphate-containing protein YdaP	Bmul_5505	G36S	non syn	SNP	G	G	G	G	G	G	del	G	del	G	del	del	A	del	del	del	G	G	G	del	del	1
contig1	43628	INT	325 down	C>T	PROKKA_00041	*adhB_1*	Alcohol dehydrogenase cytochrome c subunit precursor	-			SNP	C	C	C	C	C	C	del	C	del	C	del	del	T	del	del	del	C	C	C	del	del	1
contig1	99281	CDS		C>G	PROKKA_00098	-	hypothetical protein	-	R65R	syn	SNP	C	C	C	C	C	C	del	C	del	C	del	del	G	del	del	del	C	C	C	del	del	1
contig1	163483	CDS		T>G	PROKKA_00155	-	hypothetical protein	Bmul_5602	Q74P	non syn	SNP	T	T	T	T	T	T	G	T	del	T	del	del	T	del	del	del	T	T	T	del	del	1
contig1	163487	CDS		T>G	PROKKA_00155	-	hypothetical protein	Bmul_5602	T72P	non syn	SNP	T	T	T	T	T	T	G	T	del	T	del	del	T	del	del	del	T	T	T	del	del	1
contig1	242962	CDS		C>G	PROKKA_00231	*ttgC_1*	putative efflux pump outer membrane protein TtgC	Bmul_5523	A73P	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	1
contig1	316788	CDS		T>C	PROKKA_00300	*ydaD_1*	General stress protein 39	Bmul_5457	Q273R	non syn	SNP	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	1
contig1	461839	CDS		G>T	PROKKA_00431	*ggt_1*	Gamma-glutamyltranspeptidase precursor	Bmul_5335	S431R	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	G	G	G	G	G	G	1
contig1	465804	CDS		G>T	PROKKA_00435	-	metal dependent phosphohydrolase	Bmul_5331	L190L	syn	SNP	G	G	G	G	G	G	G	G	T	G	G	G	G	G	G	G	G	G	G	G	G	1
contig1	483288	INT	140 up	C>T	PROKKA_00450	*proQ*	ProQ activator of osmoprotectant transporter ProP	Bmul_6155			SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	1
contig1	502063	CDS		C>T	PROKKA_00467	*ytnP*	putative quorum-quenching	-	G99S	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1

							lactonase YtnP																												
contig1	621285	CDS		A>C	PROKKA_00575	*stp_1*	Multidrug resistance protein stp	-	Q292P	non syn	SNP	A	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1
contig1	721096	CDS		C>A	PROKKA_00656	*dxs_1*	1-deoxy-D-xylulose-5-phosphate synthase	Bmul_5943	L25I	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	1
contig1	771498	CDS		G>A	PROKKA_00708	-	Ribonuclease	Bmul_5891	H289H	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	1
contig1	792418	CDS		C>T	PROKKA_00726	-	hypothetical protein	Bmul_5873	A129V	non syn	SNP	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig1	931715	CDS		C>A	PROKKA_00852	-	AAA-like domain protein	-	T964K	non syn	SNP	C	C	C	C	del	C	C	C	A	del	del	C	C	C	C	C	C	C	del	C	del	1
contig1	933666	CDS		T>A	PROKKA_00856	-	hypothetical protein	-	T85T	syn	SNP	T	T	T	T	del	T	T	T	T	del	del	T	A	T	T	T	T	T	del	T	del	1
contig1	941673	CDS		C>T	PROKKA_00859	-	hypothetical protein	-	P1456S	non syn	SNP	C	C	C	C	del	C	C	C	T	del	del	C	C	C	C	C	C	C	del	C	del	1
contig1	1058786	INT	125 up	A>T	PROKKA_00978	-	hypothetical protein	-			SNP	A	A	A	A	A	A	del	A	del	A	A	A	A	A	A	A	A	T	A	A	A	1
contig1	1082099	CDS		C>G	PROKKA_01008	-	hypothetical protein	-	T121S	non syn	SNP	C	C	C	C	G	C	del	C	del	C	C	C	C	C	C	C	C	C	C	C	C	1
contig1	1170556	CDS		G>C	PROKKA_01098	-	hypothetical protein	-	P111P	syn	SNP	G	G	G	G	G	G	del	G	del	G	del	G	G	G	G	del	G	C	G	G	del	1
contig2	13420	CDS		C>T	PROKKA_01140	-	type VI secretion protein	Bmul_3913	Q25Q	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig2	17994	CDS		C>G	PROKKA_01143	*clpB1*	type VI secretion ATPase	Bmul_3906	G289R	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	1
contig2	24055	CDS		G>T	PROKKA_01149	-	type VI secretion protein	Bmul_3902	P65P	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	G	G	G	1
contig2	24056	CDS		A>C	PROKKA_01149	-	type VI secretion protein	Bmul_3902	S66R	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	A	A	A	1
contig2	33596	CDS		G>A	PROKKA_01157	*ynfM_1*	major facilitator transporter	Bmul_3894	V164I	non syn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	1
contig2	136719	CDS		T>C	PROKKA_01241	*nolT*	YscJ/HrcJ family type III secretion apparatus lipoprotein	Bmul_3744	S57G	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	T	1
contig2	144529	CDS		A>C	PROKKA_01249	-	LuxR family transcriptional regulator	Bmul_3738	L103R	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	C	A	A	A	A	A	A	A	A	1

contig2	172331	CDS		G>C	PROKKA_01273	-	parallel beta-helix repeat-containing protein	Bmul_3715	G147G	syn	SNP	G	G	G	G	C	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	1
contig2	184934	CDS		C>A	PROKKA_01281	-	hypothetical protein	Bmul_3707	P37Q	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	1
contig2	199879	CDS		A>T	PROKKA_01295	*basR*	Transcriptional regulatory protein BasR	Bmul_3691	E17V	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	T	A	A	A	1
contig2	244659	CDS		C>G	PROKKA_01333	*virS_3*	AraC family transcriptional regulator	Bmul_3655	P47R	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	1
contig2	285020	CDS		G>C	PROKKA_01376	-	Integrase core domain protein	-	C113S	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	del	G	G	G	G	C	G	G	G	1
contig2	336988	CDS		G>A	PROKKA_01434	*lptB_1*	Lipopolysaccharide export system ATP-binding protein LptB	Bmul_3626	S223L	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	1
contig2	429927	CDS		C>T	PROKKA_01517	*fdhA_3*	Glutathione-independent formaldehyde dehydrogenase	Bmul_3539	P293P	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	1
contig2	444339	CDS		G>A	PROKKA_01530	*cmpR_3*	LysR family transcriptional regulator	Bmul_3526	D197N	non syn	SNP	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	1
contig2	468353	CDS		A>G	PROKKA_01552	-	Hrp-dependent type III effector protein	Bmul_3499	E244G	non syn	SNP	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	1
contig2	495662	CDS		T>G	PROKKA_01578	-	hypothetical protein	-	E216D	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	T	1
contig2	495668	CDS		G>C	PROKKA_01578	-	hypothetical protein	-	D214E	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	1
contig2	495701	CDS		A>C	PROKKA_01578	-	hypothetical protein	-	F203L	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	A	1
contig2	504117	INT	85 up	C>T	PROKKA_01585	*yjcC_1*	diguanylate phosphodiesterase	Bmul_3471			SNP	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig2	510715	CDS		G>A	PROKKA_01590	*mmgC_2*	Acyl-CoA dehydrogenase	Bmul_3466	S377S	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	1
contig2	555119	INT	303 up	T>C	PROKKA_01619	*nadE_1*	NH(3)-dependent NAD(+) synthetase	Bmul_3443			SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	1
contig2	555131	INT	291 up	T>G	PROKKA_01619	*nadE_1*	NH(3)-dependent	Bmul_3443			SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	1

							NAD(+) synthetase																											
contig2	691485	INT	9 up	G>A	PROKKA_01734	-	Outer membrane porin protein precursor	Bmul_3342			SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	1
contig2	692133	CDS		A>G	PROKKA_01734	-	Outer membrane porin protein precursor	Bmul_3342	N214D	non syn	SNP	A	A	A	A	A	A	A	A	del	del	del	del	G	del	del	del	A	A	del	A	del	1
contig2	692531	CDS		C>G	PROKKA_01734	-	Outer membrane porin protein precursor	Bmul_3342	Y346*	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C	1
contig2	707733	INT	56 up	G>A	PROKKA_01750	*nylB*	6-aminohexanoate-dimer hydrolase	Bmul_3326			SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	1
contig2	774027	CDS		G>T	PROKKA_01804	-	hypothetical protein	Bmul_3270	*90L	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	G	1
contig2	810557	CDS		G>T	PROKKA_01840	*aroG_1*	Phospho-2-dehydro-3-deoxyheptonate aldolase, Phe-sensitive	Bmul_3232	L194I	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	G	1
contig2	843174	CDS		G>C	PROKKA_01867	*araA*	L-arabinose 1-dehydrogenase	Bmul_3205	R149G	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	G	1
contig2	848053	CDS		C>G	PROKKA_01871	*gbpR_3*	LysR family transcriptional regulator	Bmul_3201	R32R	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	1
contig2	849735	CDS		C>T	PROKKA_01872	*fhuA_1*	Ferrichrome-iron receptor precursor	Bmul_3200	T119I	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	1
contig2	866790	CDS		G>T	PROKKA_01882	-	hypothetical protein	Bmul_3190	T3T	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	T	G	G	G	G	G	G	G	G	1
contig2	890004	CDS		G>A	PROKKA_01895	*rhsB*	putative deoxyribonuclease RhsB	Bmul_3178	Q829*	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig2	914780	CDS		C>T	PROKKA_01913	*smc_1*	Chromosome partition protein Smc	Bmul_3159	R195C	non syn	SNP	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig2	928096	CDS		C>T	PROKKA_01926	*bdhA_1*	D-beta-hydroxybutyrate dehydrogenase	Bmul_5312	R190H	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	1
contig2	937804	CDS		A>G	PROKKA_01937	*nphR_2*	AraC family transcriptional regulator	Bmul_5301	R176R	syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A	1

contig2	945601	CDS		A>C	PROKKA_01946	*fabG_3*	3-oxoacyl-[acyl-carrier-protein] reductase FabG	Bmul_5292	L126R	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	A	1
contig2	952691	CDS		C>T	PROKKA_01953	*catD_1*	3-oxoadipate enol-lactonase 2	Bmul_5285	P230L	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	1
contig2	974235	CDS		C>T	PROKKA_01976	*groL_1*	molecular chaperone GroEL	Bmul_5253	I292I	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	1
contig2	982700	CDS		C>T	PROKKA_01984	*btr_1*	AraC family transcriptional regulator	Bmul_5245	P133S	non syn	SNP	C	C	C	C	C	C	C	C	del	C	C	C	T	C	C	C	C	C	C	C	C	1
contig2	1114946	CDS		T>A	PROKKA_02107	*naiP_4*	major facilitator transporter	Bmul_5124	L419Q	non syn	SNP	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	del	1
contig2	1203065	CDS		G>A	PROKKA_02187	*bag*	IgA FC receptor precursor	-	E51E	syn	SNP	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	1
contig2	1224230	CDS		G>C	PROKKA_02206	*qseF*	Transcriptional regulatory protein QseF	-	D378H	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	1
contig2	1319913	CDS		C>T	PROKKA_02295	*cheB_1*	Chemotaxis response regulator protein-glutamate methylesterase	Bmul_4941	S88L	non syn	SNP	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig2	1381942	CDS		C>A	PROKKA_02351	*modA*	Molybdate-binding periplasmic protein precursor	Bmul_4885	A48A	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	1
contig2	1503720	CDS		C>T	PROKKA_02456	*potH_2*	transport system permease protein	Bmul_5430	L395F	non syn	SNP	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	1
contig2	1576568	CDS		C>A	PROKKA_02518	*fdoH*	Formate dehydrogenase-O iron-sulfur subunit	Bmul_4596	I246I	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	1
contig2	1579862	CDS		T>A	PROKKA_02521	*selA*	L-seryl-tRNA(Sec) selenium transferase	Bmul_4593	L436*	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	A	T	T	1
contig2	1603020	CDS		C>G	PROKKA_02540	*dppA_1*	Periplasmic dipeptide transport	Bmul_4568	H372D	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	1

							protein precursor																										
contig2	1692876	CDS		A>G	PROKKA_02620	-	Cytochrome c	Bmul_4489	L40P	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	1
contig2	1707649	CDS		C>G	PROKKA_02632	*recD2*	ATP-dependent RecD-like DNA helicase	-	A174A	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	1
contig2	1813969	CDS		G>C	PROKKA_02719	*mshA_4*	D-inositol 3-phosphate glycosyltransferase	Bmul_4402	R346G	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	C	G	G	G	G	G	G	G	1
contig2	1908658	CDS		C>A	PROKKA_02797	*gabR_6*	HTH-type transcriptional regulatory protein GabR	Bmul_4324	G161V	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	1
contig2	1943657	CDS		C>A	PROKKA_02827	-	Omega-amino acid-pyruvate aminotransferase	Bmul_4294	P363T	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	1
contig2	2107835	CDS		G>A	PROKKA_02991	*otsA_1*	Alpha,alpha-trehalose-phosphate synthase [UDP-forming]	Bmul_4183	R349Q	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	1
contig2	2121924	CDS		G>C	PROKKA_03002	-	hypothetical protein	-	I254M	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	G	G	G	G	1
contig2	2185194	CDS		T>C	PROKKA_03052	*chaA*	Sodium/proton antiporter ChaA	Bmul_4124	V173A	non syn	SNP	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	1
contig2	2191033	CDS		G>C	PROKKA_03057	*opuCA*	Carnitine transport ATP-binding protein OpuCA	Bmul_4119	A115P	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	G	G	1
contig2	2209935	CDS		G>A	PROKKA_03062	*iorB_4*	Isoquinoline 1-oxidoreductase subunit beta	Bmul_4113	G177S	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig2	2220511	CDS		C>T	PROKKA_03066	-	Transglutaminase-like superfamily protein	Bmul_4109	G30R	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	1
contig2	2349776	CDS		C>G	PROKKA_03178	*gbsA_3*	Betaine aldehyde	Bmul_3992	A171A	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	1

							dehydrogenase																										
contig2	2381824	CDS		C>A	PROKKA_03210	-	putative ATP-dependent helicase Lhr	-	T153T	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	1
contig2	2399978	CDS		G>T	PROKKA_03226	*bepE_2*	Efflux pump membrane transporter BepE	Bmul_3931	R770S	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	G	1
contig3	28033	CDS		G>A	PROKKA_03274	*rnaSA*	Guanyl-specific ribonuclease Sa	Bmul_2753	R7H	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	1
contig3	229867	CDS		A>G	PROKKA_03457	*xdhA_2*	Xanthine dehydrogenase molybdenum-binding subunit	Bmul_2551	Q666R	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A	1
contig3	274762	CDS		G>A	PROKKA_03497	*pimB*	GDP-mannose-dependent alpha-(1-6)-phosphatidylinositol monomannoside mannosyltransferase	-	S267L	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	del	G	G	G	del	A	G	del	G	1
contig3	275439	CDS		G>A	PROKKA_03497	*pimB*	GDP-mannose-dependent alpha-(1-6)-phosphatidylinositol monomannoside mannosyltransferase	-	S41S	syn	SNP	G	G	G	G	G	G	G	A	G	G	G	G	del	G	G	G	del	G	G	del	G	1
contig3	308463	CDS		C>G	PROKKA_03524	*waaA*	3-deoxy-D-manno-octulosonic acid transferase	Bmul_2494	A172G	non syn	SNP	C	C	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig3	420538	CDS		C>T	PROKKA_03649	-	Putative glutamine amidotransferase	Bmul_2371	L192L	syn	SNP	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	1
contig3	426450	CDS		T>C	PROKKA_03652	*pcpB*	monooxygenase FAD-binding	Bmul_2368	S527P	non syn	SNP	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	1
contig3	442326	CDS		A>C	PROKKA_03667	*pheA*	chorismate mutase	Bmul_2261	S85R	non syn	SNP	A	A	A	A	A	A	A	A	C	A	A	A	A	A	A	A	A	A	A	A	A	A	1

contig3	448164	CDS		C>A	PROKKA_03672	*ihfB*	Integration host factor subunit beta	Bmul_2256	T2N	non syn	SNP	C	C	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig3	479095	CDS		T>G	PROKKA_03706	*dfrA*	Dihydrofolate reductase	Bmul_2221	I98L	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	1
contig3	504128	CDS		C>A	PROKKA_03733	*moeA_2*	Molybdopterin molybdenum transferase	Bmul_2196	A51S	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	1
contig3	549968	INT	73 up	G>T	PROKKA_03783	-	hypothetical protein	-			SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	G	G	G	G	1
contig3	580803	INT	92 up	C>T	PROKKA_03811	*limB_3*	Limonene 1,2-monooxygenase	Bmul_2136			SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	1
contig3	581529	CDS		A>G	PROKKA_03812	-	Glycosyl transferase family 2	Bmul_2134	N63D	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A	1
contig3	604243	CDS		C>T	PROKKA_03831	*cph2_5*	PAS/PAC sensor(s)-containing diguanylate cyclase/phosphodiesterase	Bmul_2116	T199T	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	1
contig3	619128	CDS		G>A	PROKKA_03842	*mdtB*	acriflavin resistance protein	Bmul_2105	Q565*	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig3	758407	CDS		C>G	PROKKA_03963	*nlpD_3*	Murein hydrolase activator NlpD precursor	Bmul_1990	G217R	non syn	SNP	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig3	794664	CDS		G>C	PROKKA_03986	-	NAD-dependent epimerase/dehydratase	Bmul_1981	A295G	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	G	G	G	G	G	G	G	1
contig3	882169	CDS		C>T	PROKKA_04060	*rbsC_6*	Ribose transport system permease protein RbsC	Bmul_1889	A94V	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	1
contig3	901264	INT	234 up	T>C	PROKKA_04077	*uvrD_2*	DNA helicase II	Bmul_1873			SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	T	1
contig3	925543	CDS		C>G	PROKKA_04099	-	Electron transfer flavoprotein-ubiquinone oxidoreductase	Bmul_1788	A359A	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C	1

contig3	925885	CDS		C>G	PROKKA_04099	-	Electron transfer flavoprotein-ubiquinone oxidoreductase	Bmul_1788	A245A	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C	1
contig3	925915	CDS		C>T	PROKKA_04099	-	Electron transfer flavoprotein-ubiquinone oxidoreductase	Bmul_1788	K235K	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	1
contig3	989003	INT	421 up	C>T	PROKKA_04188	-	lipoprotein	Bmul_1770			SNP	C	del	C	C	C	C	C	C	C	C	del	del	C	C	T	del	C	C	C	del	del	1
contig3	1087274	CDS		G>C	PROKKA_04275	*panE_4*	2-dehydropantoate 2-reductase	Bmul_1688	V71V	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	G	G	1
contig3	1096783	CDS		C>T	PROKKA_04283	*afr*	1,5-anhydro-D-fructose reductase	Bmul_1681	R263R	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	1
contig3	1097086	CDS		C>A	PROKKA_04284	*yecS_2*	Inner membrane amino-acid ABC transporter permease protein	Bmul_1680	C210F	non syn	SNP	C	C	C	C	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig3	1104485	CDS		G>A	PROKKA_04293	*echA8_3*	putative enoyl-CoA hydratase	Bmul_1672	A360T	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	1
contig3	1113816	CDS		T>C	PROKKA_04305	-	hypothetical protein	Bmul_1658	Q118R	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig3	1114149	CDS		G>C	PROKKA_04305	-	hypothetical protein	Bmul_1658	T7S	non syn	SNP	G	G	G	C	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	1
contig3	1187838	CDS		A>G	PROKKA_04364	*grsB*	Gramicidin S synthase 2	Bmul_1597	H184R	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A	1
contig3	1212137	CDS		A>G	PROKKA_04379	*cobN*	Aerobic cobaltochelatase subunit CobN	Bmul_1582	H1222R	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A	1
contig3	1250451	CDS		G>T	PROKKA_04421	*ramA*	(R)-stereoselective amidase	-	M186I	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	G	1
contig3	1266312	CDS		C>G	PROKKA_04439	-	Xylose isomerase-like TIM barrel	Bmul_1525	A138A	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	1
contig3	1402723	CDS		G>T	PROKKA_04562	-	ISXO2-like transposase domain protein	-	S41R	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	G	G	G	1

contig3	1406972	CDS		G>A	PROKKA_04567	*rpsR*	30S ribosomal protein S18	Bmul_1402	L18F	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig3	1439231	CDS		G>C	PROKKA_04601	-	nicotinamidase/pyrazinamidase	Bmul_1369	H55D	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	1
contig3	1523859	CDS		C>G	PROKKA_04673	*sdhA_3*	Succinate dehydrogenase flavoprotein subunit	Bmul_1298	A218A	syn	SNP	C	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig3	1584517	CDS		C>A	PROKKA_04728	*clsA*	Major cardiolipin synthase ClsA	Bmul_1243	L166I	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	1
contig3	1611930	CDS		G>T	PROKKA_04760	*mepH_2*	Murein DD-endopeptidase MepH precursor	Bmul_1218	P170Q	non syn	SNP	G	G	T	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	1
contig3	1638814	CDS		T>C	PROKKA_04784	-	hypothetical protein	-	I67V	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	T	1
contig3	1664485	CDS		G>C	PROKKA_04816	-	Phage-related baseplate assembly protein	-	P221R	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	1
contig3	1725353	CDS		T>C	PROKKA_04879	*fdx_1*	2Fe-2S ferredoxin	Bmul_1149	H108R	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	T	1
contig3	1729484	CDS		T>A	PROKKA_04884	*iscS_2*	Cysteine desulfurase	Bmul_1144	N305I	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	A	T	T	T	T	T	T	T	T	1
contig3	1731488	CDS		G>T	PROKKA_04886	*yfkJ*	protein tyrosine phosphatase	Bmul_1142	I14I	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	G	G	G	1
contig3	1747755	CDS		T>C	PROKKA_04897	*tmoT*	two component LuxR family transcriptional regulator	Bmul_1131	*213R	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	T	1
contig3	1859312	CDS		G>A	PROKKA_04996	-	hypothetical protein	Bmul_1059	V358I	non syn	SNP	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	1
contig3	1950433	CDS		A>G	PROKKA_05090	*rfaQ_3*	Lipopolysaccharide core heptosyltransferase RfaQ	-	*392R	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A	A	A	1
contig3	2010489	CDS		G>C	PROKKA_05139	-	Rubredoxin	Bmul_0916	W17S	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	G	G	1
contig3	2037980	CDS		T>G	PROKKA_05170	*hisP_3*	Histidine transport ATP-binding protein HisP	Bmul_0883	L155R	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	T	1

contig3	2128370	CDS		T>C	PROKKA_05256	*livH_5*	High-affinity branched-chain amino acid transport system permease	Bmul_0798	Q347R	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig3	2128371	CDS		G>T	PROKKA_05256	*livH_5*	High-affinity branched-chain amino acid transport system permease	Bmul_0798	Q346K	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	T	G	G	G	G	G	G	G	G	1
contig3	2149307	CDS		C>A	PROKKA_05275	*dut*	Deoxyuridine 5'-triphosphate nucleotidohydrolase	Bmul_0780	R23S	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	1
contig3	2161705	CDS		G>C	PROKKA_05284	*icd_1*	Isocitrate dehydrogenase [NADP]	Bmul_0771	A389G	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	1
contig3	2171106	CDS		G>A	PROKKA_05291	*yegW_1*	GntR family transcriptional regulator	Bmul_0763	A67T	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	1
contig3	2206844	CDS		A>T	PROKKA_05327	*cmaA1*	Cyclopropane mycolic acid synthase 1	Bmul_0727	D159V	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	T	A	A	A	A	A	A	A	A	1
contig3	2207131	CDS		T>C	PROKKA_05327	*cmaA1*	Cyclopropane mycolic acid synthase 1	Bmul_0727	Y255H	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	T	1
contig3	2229130	CDS		G>C	PROKKA_05345	*estB_2*	Esterase EstB	Bmul_0708	R249G	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	1
contig3	2229131	CDS		C>G	PROKKA_05345	*estB_2*	Esterase EstB	Bmul_0708	P249P	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	1
contig3	2235706	CDS		T>A	PROKKA_05351	-	putative ABC transporter-binding protein precursor	Bmul_0702	I87F	non syn	SNP	T	T	T	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	1
contig3	2278315	CDS		A>T	PROKKA_05388	*rfaF_2*	ADP-heptose-LPS heptosyltransferase 2	Bmul_0665	H272L	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	T	A	A	A	1
contig3	2286655	CDS		G>C	PROKKA_05397	*fimA_3*	Fimbrial protein precursor	Bmul_0656	G92G	syn	SNP	G	G	G	G	G	G	G	G	G	C	G	G	G	G	G	G	G	G	G	G	G	1
contig3	2287870	INT	23 down	C>T	PROKKA_05399	*sucD*	Succinyl-CoA ligase [ADP-	Bmul_0654			SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	1

							forming] subunit alpha																												
contig3	2297765	CDS		T>C	PROKKA_05408	*dmlR_37*	LysR family transcriptional regulator	Bmul_0641	N260D	non syn	SNP	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	1
contig3	2311096	CDS		G>A	PROKKA_05419	*cc4_2*	Cytochrome c4 precursor	Bmul_0629	A223T	non syn	SNP	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	1
contig3	2356176	CDS		C>G	PROKKA_05461	*lptD*	LPS-assembly protein LptD precursor	Bmul_0588	G720A	non syn	SNP	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig3	2356177	CDS		C>T	PROKKA_05461	*lptD*	LPS-assembly protein LptD precursor	Bmul_0588	G719S	non syn	SNP	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig3	2424376	CDS		T>C	PROKKA_05520	-	Lipopolysaccharide-assembly, LptC-related	Bmul_0527	F92L	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	1
contig3	2427933	INT	60 up	G>A	PROKKA_05524	*hpf*	Ribosome hibernation promoting factor	Bmul_0523			SNP	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig3	2428767	CDS		G>C	PROKKA_05525	*ptsN*	Nitrogen regulatory protein PtsN	Bmul_0522	G55A	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	1
contig3	2430037	CDS		G>A	PROKKA_05526	*hprK*	HPr kinase/phosphorylase	Bmul_0521	A284T	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig3	2448847	CDS		A>G	PROKKA_05547	*pagL*	Lipid A deacylase PagL precursor	Bmul_0487	Y156H	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	1
contig3	2465205	CDS		C>G	PROKKA_05559	-	impB/mucB/samB family protein	Bmul_0476	P483P	syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	1
contig3	2537284	CDS		G>A	PROKKA_05629	*mlaE_1*	putative phospholipid ABC transporter permease protein MlaE	Bmul_0403	R55R	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig3	2558710	CDS		C>A	PROKKA_05648	-	cell division protein FtsN	Bmul_0384	A2E	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	1
contig3	2579639	CDS		G>C	PROKKA_05667	*kipA_1*	urea amidolyase-like protein	Bmul_0365	A103A	syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	C	G	G	G	G	G	G	G	G	1
contig3	2602857	INT	10 up	C>A	PROKKA_05682	*gcvA_18*	LysR family transcriptional regulator	Bmul_2946			SNP	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	1

contig3	2620498	CDS		G>C	PROKKA_05699	-	Outer membrane porin protein 32 precursor	Bmul_2963	H128D	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	1
contig3	2627381	CDS		G>A	PROKKA_05706	*yihS*	N-acylglucosamine 2-epimerase	Bmul_2968	C9Y	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	1
contig3	2724952	CDS		C>A	PROKKA_05803	*fliI*	flagellar protein export ATPase FliI	Bmul_3064	T350K	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	1
contig3	2725926	CDS		A>T	PROKKA_05804	*fliJ*	Flagellar FliJ protein	Bmul_3065	E130V	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	T	A	A	A	A	A	1
contig3	2778397	INT	28 up	G>T	PROKKA_05851	*gatC*	Glutamyl-tRNA(Gln) amidotransferase subunit C	Bmul_3106			SNP	G	G	G	G	T	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	1
contig3	2789704	CDS		C>G	PROKKA_05863	*zraR_8*	Fis family transcriptional regulator	Bmul_3117	E101Q	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	1
contig3	2932992	CDS		G>A	PROKKA_06008	*gspD*	Putative type II secretion system protein D	Bmul_0063	A181V	non syn	SNP	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	1
contig3	3011968	CDS		G>A	PROKKA_06073	*stp_9*	Multidrug resistance protein stp	Bmul_0130	G306D	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G	1
contig3	3028855	CDS		C>T	PROKKA_06087	-	hypothetical protein	Bmul_0144	A37T	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	1
contig3	3031354	CDS		C>T	PROKKA_06089	*rep*	ATP-dependent DNA helicase Rep	Bmul_0146	G260G	syn	SNP	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1
contig3	3060211	INT	16 down	G>A	PROKKA_06111	*flhC*	Flagellar transcriptional regulator FlhC	Bmul_0161			SNP	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	1
contig3	3064330	CDS		G>C	PROKKA_06115	*cheA*	Chemotaxis protein CheA	Bmul_0165	G513R	non syn	SNP	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	G	1
contig3	3104312	CDS		T>A	PROKKA_06154	*macB_2*	Macrolide export ATP-binding/permease protein MacB	Bmul_0204	Q199L	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	A	T	T	T	T	T	1
contig3	3105522	CDS		A>C	PROKKA_06155	-	FtsX-like permease family protein	Bmul_0205	F184C	non syn	SNP	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	A	1

contig3	3187132	CDS		C>A	PROKKA_06241	*ccsB*	Cytochrome c biogenesis protein CcsB	Bmul_0282	L646I	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	C	1
contig3	3253851	CDS		T>C	PROKKA_06310	-	hypothetical protein	Bmul_2934	E63G	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	1
contig3	3343234	CDS		T>C	PROKKA_06390	*fadD_4*	Long-chain-fatty-acid-CoA ligase	Bmul_2848	*558Q	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	1
contig3	3348779	CDS		T>C	PROKKA_06396	*penA_3*	peptidoglycan glycosyltransferase	Bmul_2842	V354A	non syn	SNP	T	T	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	1
contig3	3365910	INT	82 up	C>G	PROKKA_06411	-	preprotein translocase subunit SecA	Bmul_2827			SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	C	1
contig3	3411751	CDS		C>G	PROKKA_06453	*nrdB*	Ribonucleoside-diphosphate reductase subunit beta	Bmul_2788	P37R	non syn	SNP	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	C	1
contig1	342152	INT	214 up	G	PROKKA_00322	-	hypothetical protein	Bmul_5435			indel	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	cgcagc	-----------	1
contig1	899298	CDS		G	PROKKA_00823	*parM*	Plasmid segregation protein ParM	-			indel	-	-	-	-	-	-	-	-	c	-	-	-	-	-	-	-	-	-	-	-	-	1
contig1	1096803	CDS		C	PROKKA_01020	*oatA_1*	O-acetyltransferase OatA	-			indel	-	-	-	-	-	-	-	-	-	-	-	-	-	g	-	-	-	-	-	-	-	1
contig1	1119377	CDS		C	PROKKA_01043	-	hypothetical protein	-			indel	ga	ga	ga	ga	ga	ga	ga	--	ga	ga	ga	ga	ga	ga	ga	ga	ga	ga	ga	ga	ga	1
contig2	712936	INT	67 up	G	PROKKA_01754	*ttr*	Acetyltransferase	Bmul_3322			indel	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	tcgcgttcgcgt	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	1
contig2	712936	INT	67 up	G	PROKKA_01754	*ttr*	Acetyltransferase	Bmul_3322			indel	tcgcgt	tcgcgt	tcgcgt	tcgcgt	tcgcgt	------	tcgcgt	tcgcgt	tcgcgt	tcgcgt	tcgcgt	tcgcgt	tcgcgt	tcgcgt	tcgcgt	tcgcgt	tcgcgt	tcgcgt	tcgcgt	tcgcgt	tcgcgt	1
contig2	803135	CDS		C	PROKKA_01832	*lvr_2*	Levodione reductase	Bmul_3241			indel	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	----------	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	cgcgcgcgcgcgagcgcgac	1

												cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg		cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	cg cg gc cg atc cc ctg	
contig2	956880	CDS		T	PROKKA _01959	*czcD _1*	Cadmium, cobalt and zinc/H(+)-K(+) antiporter	Bmul_ 5279			indel	ca cg ag	ca cg ag	ca cg ag	ca cg ag	ca cg ag	ca cg ag	ca cg ag	ca cg ag	ca cg ag	ca cg ag	ca cg ag	ca cg ag	ca cg ag	ca cg ag	ca cg ag	ca cg ag	------	ca cg ag	ca cg ag	ca cg ag	ca cg ag	1
contig2	1087172	CDS		G	PROKKA _02084	*gcvA _10*	LysR family transcription al regulator	Bmul_ 5146			indel	c	c	c	c	c	c	c	c	del	c	-	c	c	c	c	c	c	c	c	c	c	1
contig2	1203055	CDS		C	PROKKA _02187	*bag*	IgA FC receptor precursor	-			indel	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	---------	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	gg tgc cc ga	1
contig2	1593682	CDS		G	PROKKA _02532	-	Fumarylacet oacetate (FAA) hydrolase family protein	Bmul_ 4583			indel	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	a	1
contig2	1767378	INT	89 down	T	PROKKA _02683	*gbpR _7*	LysR family transcription al regulator	Bmul_ 4418			indel	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	ac ga gc c	-------	ac ga gc c	ac ga gc c	1
contig2	2033547	CDS		G	PROKKA _02910	-	Gluconate 2-dehydrogena se cytochrome c subunit precursor	Bmul_ 4216			indel	-	-	-	-	-	-	-	c	-	-	-	-	-	-	-	-	-	-	-	-	-	1
contig3	133669	INT	140 up	T	PROKKA _03371	-	alanine racemase domain-containing protein	Bmul_ 2655			indel	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	----------------------------	gg cg cg cg gc ca gc gg cg cg cg gc ca gc	1
contig3	420883	CDS		C	PROKKA _03649	-	Putative glutamine	Bmul_ 2371			indel	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	-	a	1

							amidotransferase																									
contig3	515849	INT	34 up	G	PROKKA_03745	*plsX*	putative glycerol-3-phosphate acyltransferase PlsX	Bmul_2184			indel	c	c	c	c	c	c	c	c	c	c	c	c	-	c	c	c	c	c	c	c	c	1
contig3	519948	CDS		C	PROKKA_03749	*acpP_2*	Acyl carrier protein	Bmul_2180			indel	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	-	a	a	a	1
contig3	544783	CDS		T	PROKKA_03776	-	hypothetical protein	-			indel	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	a	1
contig3	1091841	CDS		T	PROKKA_04278	*rnr*	Ribonuclease R	Bmul_1686			indel	--	--	--	--	--	--	--	--	--	--	gc	--	--	--	--	--	--	--	--	--	--	--	1
contig3	1092429	CDS		C	PROKKA_04279	*rlmB*	RNA methyltransferase	Bmul_1685			indel	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	-----------	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	tgaaggttcttacggtt	1
contig3	1155929	CDS		G	PROKKA_04345	*ttgR_1*	TetR family transcriptional regulator	Bmul_1617			indel	----	----	----	----	----	----	----	----	----	----	----	----	----	gatc	----	----	----	----	----	----	----	----	1
contig3	1322758	CDS		T	PROKKA_04491	-	hypothetical protein	Bmul_1477			indel	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	--	gc	1
contig3	1402890	INT	45 up	C	PROKKA_04562	-	ISXO2-like transposase domain protein	-			indel	-	-	-	a	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
contig3	1618995	INT	137 up	G	PROKKA_04765	-	hypothetical protein	Bmul_1213			indel	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	-	c	1
contig3	1724563	CDS		G	PROKKA_04877	-	17 kDa surface antigen	Bmul_1151			indel	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	-----	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	cgccgcagc	1
contig3	1948468	CDS		G	PROKKA_05088	-	Glycosyl transferase family 2	-			indel	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	cgt	---	cgt	cgt	1
contig3	2448851	CDS		C	PROKKA_05547	*pagL*	Lipid A deacylase PagL precursor	Bmul_0487			indel	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	t	-	1
contig3	2449067	CDS		C	PROKKA_05547	*pagL*	Lipid A deacylase PagL precursor	Bmul_0487			indel	ca	ca	ca	ca	ca	ca	ca	ca	ca	ca	ca	ca	ca	ca	ca	ca	ca	--	ca	ca	ca	ca	ca	ca	1
contig3	2472098	INT	80 down	G	PROKKA_05566	*tar_8*	Methyl-accepting chemotaxis protein II	Bmul_0469			indel	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	gccgtgtc	----------	1

contig3	2536493	CDS		A	PROKKA_05629	*mlaE_1*	putative phospholipid ABC transporter permease protein MlaE	Bmul_0403			indel	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	---------------------------	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	tcacgcgcatcgcgtcgagctcctcgg	1
contig3	2625971	CDS		G	PROKKA_05704	-	hypothetical protein	-			indel	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	------------------	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	acccgctcgaccgctca	1
contig3	2884572	CDS		G	PROKKA_05960	-	Reverse transcriptase (RNA-dependent DNA polymerase)	-			indel	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	a	-	1
contig3	3188635	INT	27 down	A	PROKKA_06242	*ccsA_1*	Cytochrome c biogenesis protein CcsA	Bmul_0283			indel	ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	---	ccg	1
contig3	3196331	CDS		G	PROKKA_06251	*pilQ*	type IV pilus secretin PilQ	Bmul_0294			indel	------	------	------	------	------	------	------	------	tcggcc	------	------	------	------	------	------	------	------	------	------	------	------	1
contig1	7203				PROKKA_00008 to PROKKA_00155					52091 aa deletion	indel	A	A	A	A	A	A	156271 bp deletion	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1
contig1	450595				PROKKA_00422; PROKKA_00423	-	tRNA-Ala; tRNA-Ile			170 aa deletion	indel	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	508-bp deletion	C	1
contig1	855545				PROKKA_00785 to PROKKA_00800					5999 aa deletion	indel	G	G	G	G	G	G	17997-bp deletion	G	G	G	G	G	G	G	G	G	G	G	G	G	G	1

contig1	995236				PROKKA_00918 to PROKKA_00958					14169 aa deletion	indel	C	C	C	C	del	C	C	C	C	del	del	C	C	C	C	C	C	C	C	del	C	42505-bp deletion	1
contig2	262950				PROKKA_01353 to PROKKA_01422					21216 aa deletion	indel	T	T	T	T	T	T	T	T	T	T	T	T	63646-bp deletion	T	T	T	T	T	T	T	T	1
contig2	979219				PROKKA_01981 to PROKKA_02095					39847 aa deletion	indel	G	G	G	G	G	G	G	G	119541 bp deletion	G	G	G	G	G	G	G	G	G	G	G	G	1
contig2	1100890				PROKKA_02097 to PROKKA_02178					30832 aa deletion	indel	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	92495-bp deletion	1
contig2	1355298				PROKKA_02325 to PROKKA_02331					2031 aa deletion	indel	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	6093 bp deletion	C	C	C	1
contig3	269236				PROKKA_03494 to PROKKA_03505					5884 aa deletion	indel	A	A	A	A	A	A	A	A	A	A	A	A	17652-bp deletion	A	A	A	del	A	A	del	A	1