Identification of genetic determinants involved in multicellular aggregation in *Burkholderia multivorans*

Catarina Reis Fernandes Cabeções

Thesis to obtain the Master of Science Degree in

Microbiology

Supervisor: Doctor Leonilde de Fátima Morais Moreira

Examination Committee:

Chairperson: Doctor Jorge Humberto Gomes Leitão

Supervisor: Doctor Leonilde de Fátima Morais Moreira

Member of the Committee: Doctor Dalila Madeira Mil-Homens

November 2021

Preface

The work presented in this thesis was performed at Institute for Bioengineering and Biosciences (iBB) of Instituto Superior Técnico (Lisbon, Portugal), during the period September 2020-September 2021, under the supervision of Professor Leonilde de Fátima Morais Moreira.

Declaration

I declare that this document is an original work of my own authorship and that it fulfils all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Acknowledgements

First of all, I would like to thank my supervisor Professor Leonilde Moreira for the constant support and guidance during the development of this work.

I also thank every member of the Biological Sciences Research Group (BSRG) that helped me throughout my experimental work, in particular Mirela Ferreira and Sara Gomes, for all the teachings and for always showing great availability, as well as for making the hardest days at the lab more fun. Finally, I would like to thank my family and friends, especially my mother, for her unconditional support and invaluable motivation. A special thanks to my dear friend Sara, who has been with me since the beginning and also Beatriz and Mariana, who have been by my side throughout this journey.

This work was financed by national funds from Fundação para a Ciência e a Tecnologia, in the scope of the project UIDB/04565/2020 and UIDP/04565/2020 of the research unit Institute for Bioengineering and Biosciences, and the project LA/P/0140/2020 of the Associate Laboratory institute for Health and Bioeconomy i4HB.

Resumo

Os microrganismos podem viver como células livres ou formando agregados. Esses agregados multicelulares são de grande importância, especialmente em ambientes clínicos, devido à maior capacidade de prevalecerem em diversas condições ambientais. Muitas bactérias do complexo *Burkholderia cepacia* são capazes de formar agregados, permitindo-lhes estabelecer infeções crónicas em pacientes com de fibrose quística (FQ), levando muitas vezes à diminuição da função pulmonar. No entanto, os mecanismos moleculares envolvidos na formação destes agregados são praticamente desconhecidos. Neste trabalho, caracterizaram-se dois isolados sequenciais de *Burkholderia multivorans* obtidos de um paciente com FQ, nos quais P0426-1 forma agregados e P0426-2, não. A comparação da sequência do genoma identificou 4 mutações pontuais em regiões codificantes, 2 *indels* e uma grande deleção em P0426-2. Uma das mutações é num gene que codifica a lactonase YtnP. Este tipo de enzimas é conhecida por regular a comunicação por quorum sensing através da degradação de N-acil-homoserina lactonas (AHLs). A expressão de YtnP-85R do isolado P0426-2 no isolado P0426-1 diminuiu a biomassa dos agregados e afetou a sua estrutura. A expressão de YtnP-85R em duas outras estirpes de *B. multivorans* formadoras de agregados, P0213-1 e P0148-1, causou alterações na estrutura dos agregados em P0213-1, mas não em P0148-1. Para determinar a variedade de AHLs possivelmente degradados por esta lactonase, a proteína recombinante His_6YtnP-85S foi produzida e purificada com sucesso em *E. coli*, mas não se determinou a atividade enzimática. No futuro dever-se-à caracterizar esta nova lactonase, nomeadamente o seu envolvimento na regulação da expressão génica e patogénese em *Burkholderia*.

Palavras-chave: *Burkholderia multivorans*, Agregados celulares, Lactonase, Fibrose quística.

Abstract

Microorganisms can live freely or as multicellular aggregates. These aggregates are of great importance, especially in clinical settings, due to the increased ability to prevail in diverse environmental conditions. *Burkholderia cepacia* complex are some of the bacteria capable of forming multicellular aggregates, contributing to the establishment of chronic infection in cystic fibrosis (CF) patients, often leading to decreased lung function. However, the molecular mechanisms involved in the formation of cellular aggregates by these bacteria are mostly unknown. Here, we have characterized two *Burkholderia multivorans* sequential isolates from a CF patient, in which P0426-1 was able to form aggregates and P0426-2, not. Comparison of their genome sequence identified 4 single nucleotide polymorphisms (SNP) mutations in coding sequences, 2 indels, and a large deletion in P0426-2. One of the SNPs was in a gene encoding the putative lactonase YtnP. These enzymes are known to regulate quorum sensing communication by degrading N-acyl-homoserine lactones (AHLs). Expression of YtnP-85R from isolate P0426-2 into isolate P0426-1 resulted in the reduction of aggregates biomass and affected aggregates´ structure. Similarly, expression of YtnP-85R in two different *B. multivorans* strains, P0213-1 and P0148-1, which form aggregates, caused alterations in aggregates structure in P0213-1, but not in P0148-1. To determine the range of AHLs degraded by this putative lactonase, the recombinant His_6YtnP-85S was successfully overexpressed and purified in *E. coli*, but no enzyme activities were done. Further work needs to be performed in order to characterize this novel lactonase enzyme, namely its possible role in regulating gene expression and pathogenesis in *Burkholderia*.

Keywords: *Burkholderia multivorans,* Planktonic cellular aggregates, Lactonase, Cystic fibrosis

List of contents

Acknowledgments ..... **iii**

Resumo ..... **iv**

Abstract ..... **v**

List of contents ..... **vi**

List of figures ..... **viii**

List of tables ..... **x**

List of abbreviations ..... **xi**

1. Introduction **1**

1.1. Cystic Fibrosis 2

1.1.1. Epidemiology 3

1.1.2. CF in the lungs 4

1.1.3. Microorganisms affecting CF lungs 6

1.2. *Burkholderia cepacia* complex 7

1.2.1. Epidemiology 8

1.3. Bacterial biofilms 10

1.3.1. Surface-attached biofilms 11

1.3.2. Planktonic cellular aggregates 12

1.3.3. Elements controlling biofilm formation 13

1.3.3.1. Quorum Sensing and its role in controlling biofilm formation 13

1.3.3.2. Quorum quenching 15

1.3.3.3. Efflux pumps 15

1.4. *Burkholderia multivorans* aggregates 17

1.5. Aim of this thesis 18

2. Materials and Methods **19**

2.1. Bacterial strains and growth conditions 19

2.2. Genomic DNA extraction 20

2.3. DNA manipulation techniques 20

2.4. Cloning of *B. multivorans* genes for phenotype complementation 20

2.5. Complementation assays 22

2.6. RNA manipulation techniques 22

2.6.1. RNA extraction and purification 22

2.6.2. Real-Time qPCR 23

2.7. Growth curves 23
2.8. Quantification of cellular aggregates and free cells 24
2.9. Microscopy analysis 24
2.10. Protein manipulation techniques 24
2.10.1. Protein expression 24
2.10.2. Purification of His6-tagged proteins by affinity chromatography 25
2.10.3. Denaturing polyacrylamide gel electrophoresis (SDS-PAGE) 26
2.11. Sequence alignment and phylogenetic analysis 27
2.12. Statistical analyses 27
3. Results and Discussion 28
3.1. Cellular aggregation in *B. multivorans* isolates P0426-1 and P0426-2 28
3.1.1. *In vitro* growth analysis of *B. multivorans* P0426-1 and P0426-2 30
3.1.2. Genetic alterations between *B. multivorans* isolates P0426-1 and P0426-2 30
3.2. Analysis of the putative quorum quenching lactonase (YtnP) from isolates P0426-1 and P0426-2 32
3.2.1. Comparative sequence analysis and phylogenetic distribution of the YtnP protein 32
3.2.2. Genetic complementation with the *ytnP* gene 35
3.2.3. Overexpression and purification of YtnP-85S protein 40
3.3. Analysis of the outer membrane protein (TtgC) from isolates P0426-1 and P0426-2 42
3.3.1. Comparative sequence analysis and phylogenetic distribution of the TtgC protein . 42
3.3.2. Genetic complementation with the *ttgC* gene 44
3.4. Comparative sequence analysis of the PbuE and Stp proteins 45
3.5. Gene expression analysis by real-time quantitative PCR 47
4. Concluding remarks and future perspectives 48
5. References 50
6. Supplementary material 61

List of figures

Figure 1: CFTR Structure 2

Figure 2: Age at diagnosis of all the CF patients screened in 2019 4

Figure 3: Mucociliary clearance defect in CF epithelial airways 5

Figure 4: Microbial diversity by age 7

Figure 5: Biofilm structure 10

Figure 6: Schematic representation of biofilm formation 12

Figure 7: Schematic representation of planktonic cellular aggregate formation 13

Figure 8: Quorum sensing system in gram-negative bacteria. 14

Figure 9: Four potential roles of efflux pumps in biofilm formation 16

Figure 10: pBBR1MCS and pET-23a(+) vectors used on the different constructs of *B. multivorans* genes 22

Figure 11: Screening for aggregate formation by *B. multivorans* clinical isolate P0426-1 and P0426-2 in both SM (A) and SCFM medium (B) 28

Figure 12: Quantification of cellular aggregates and free cells of P0426-1 and P0426-2 2 grown in SM medium at 37˚C, 180 rpm for 48h 29

Figure 13: Growth curve of *B. multivorans* P0426-1 and P0426-2 at 37ºC for 48 hours in SM medium 30

Figure 14: Amino acid sequence alignment generated by ClustalX of a representative sequence of YtnP protein present in different species 33

Figure 15: Sequence alignment presenting the relationship between *B. multivorans* P0426-1 and P0426-2 lactonases YtnP analysed in our study and other experimentally validated AHL lactonases 33

Figure 16: Phylogeny of YtnP protein 35

Figure 17: Screening for aggregate formation by P0426-1, P0426-2 expressing a different lactonase 36

Figure 18: Quantification of cellular aggregates and free cells 37

Figure 19: Screening for aggregate formation by P0213-1, P0148-1 and respective complemented mutants 38

Figure 20: Quantification of cellular aggregates and free cells 39

Figure 21: 15% SDS-PAGE gel with YtnP overexpression assay samples 41

Figure 22: Analysis by SDS-PAGE of His-tagged YtnP protein purification by nickel affinity chromatography 42

Figure 23: Amino acid sequence alignment generated by ClustalX of a representative sequence of TtgC protein present in different species 43

Figure 24: Phylogeny of TtgC protein 43

Figure 25: Screening for aggregate formation by P0426-1, P0426-2 expressing a different outer membrane protein TtgC 44

Figure 26: Quantification of cellular aggregates and free cells 45

Figure 27: Amino acid sequence alignment generated by ClustalX of a representative sequence of Stp protein present in different species 46

Figure 28: Amino acid sequence alignment generated by ClustalX of a representative sequence of Stp protein present in different species 46

Figure 39: Real-time PCR 47

List of tables

Table 1: Strains and plasmids used in this study 9
Table 2: Primers used to perform Real-Time quantitative PCR 23
Table 3: Elution buffers used during the process of purification of His-tagged proteins 25
Table 4: Composition of denaturing polyacrylamide gel..... 26
Table 5: Mutations identified between *B. multivorans* P0426-1 and P0426-2..... 31

List of abbreviations

CF - Cystic Fibrosis

CFTR – Cystic Fibrosis Transmembrane Conductance Regulator

cAMP - cyclic adenosine monophosphate

ABC - ATP-Binding Cassette

NBD - Nucleotide Binding Domain

ASL - Airway Surface Liquid

AMPs - Antimicrobial Peptides

Bcc – *Burkholderia cepacia* complex

MLST – Multilocus sequence typing

PCR - Polymerase chain reaction

BAL - Bronchoalveolar lavage

QS – Quorum sensing

EPS - Extracellular Polymeric Substances

AHL – Acyl-Homoserine Lactone

LTTR - LysR-Type Transcriptional Regulator

QQ – Quorum quenching

eDNA - Extracellular DNA

MFS - major facilitator superfamily

MATE - multidrug and toxin extrusion

RND - resistance-nodulation-division

SMR - small multidrug resistance

mRNA – messenger RNA

MST - Monosaccharide Transporter

1. Introduction

Cystic fibrosis is a recessive disease caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene. This is one of the most common hereditary diseases in the population with European descent, being believed that 1 in every 25 Caucasians are carriers of the disease (Polgreen *et al.* 2018; Cant *et al.* 2014). CFTR protein transports chloride ions through the cell membrane, helping maintain an equilibrium in exocrine organs. Mutations in this regulator can affect multiple organ systems like the pancreas, lungs, liver, intestine, and reproductive organs. Infection of the lungs with microorganisms like *Pseudomonas aeruginosa* and *Burkholderia cepacia* complex make pulmonary diseases the major causes of mortality and morbidity (Naehrig *et al.* 2017; Fanen *et al*. 2014; Sheppard *et al.* 2002).

It is known that microorganisms can accumulate on many types of surfaces, forming sessile communities. Biofilms are composed of microbial communities, water, and polysaccharides, among other macromolecules, and can be found in a wide range of surfaces and conditions, protecting bacteria from external factors, such as antimicrobial compounds, extreme temperature, etc (Yin *et al*. 2019). Bacteria can adhere to surfaces and/or form planktonic cellular aggregates, which happens when they adhere to each other instead of a surface (Schleheck *et al.* 2009). Biofilms are frequently related with a lot of diseases caused by pathogens. *Pseudomonas* and *Burkholderia* are some of the bacteria that can form either attached biofilms or planktonic cellular aggregates (also known as mobile biofilms), allowing them to cause persistent infections that can often lead to severe damage on the lungs of patients suffering from cystic fibrosis (Schwab *et al.* 2014).

To this day, the knowledge of the molecular mechanisms involved in the formation of these multicellular aggregates is still very little. Thereby, this introduction will give an overview of cystic fibrosis condition, followed by a short introduction to *Burkholderia* biology, and then focus on literature related to the molecular mechanisms involved in the formation of theses aggregates by different bacteria, but giving emphasis on the ones that can cause chronic infection in CF patients.

1.1. Cystic fibrosis

Cystic fibrosis (CF) identified in 1989 is an autosomal recessive hereditary disease that affects mainly populations of European descent and arises from mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene located in the human chromosome 7 (Lyczak *et al.* 2002). This protein is regulated by cyclic adenosine monophosphate (cAMP) and is a part of the ATP-binding cassette (ABC) family of membrane proteins working as a channel across the membrane of cells producing mucus, tears, saliva, sweat, and digestive enzymes. This protein carries negatively charged particles called chloride ions into and out of cells, which helps to maintain an equilibrium in certain tissues (Cant *et al.* 2014; Ehre *et al*. 2014). When this regulator is not working properly, chloride becomes trapped in cells which will cause a lack of hydration of the cellular surface influencing exocrine organ systems like the pancreas, lungs, liver, intestine, and reproductive organs. Due to this, the most common test used for the diagnosis of CF is the Sweat Chloride Test (Ratjen *et al.* 2015).

Despite the effect on other organs, the major cause of morbidity and mortality in CF patients is lung disease, triggered by the clogging of the bronchial passage which causes difficulty breathing, while bacterial infections will progressively destroy the lungs (Welsh and Smith, 1995).

As shown in figure 1 the CFTR protein comprises two transmembrane domains (MSD-1 and MSD-2) each followed by a nucleotide binding domain (NBD1 and NBD2), being these last two the ones that bind and hydrolyse ATP. CFTR also includes a unique regulatory domain (R).

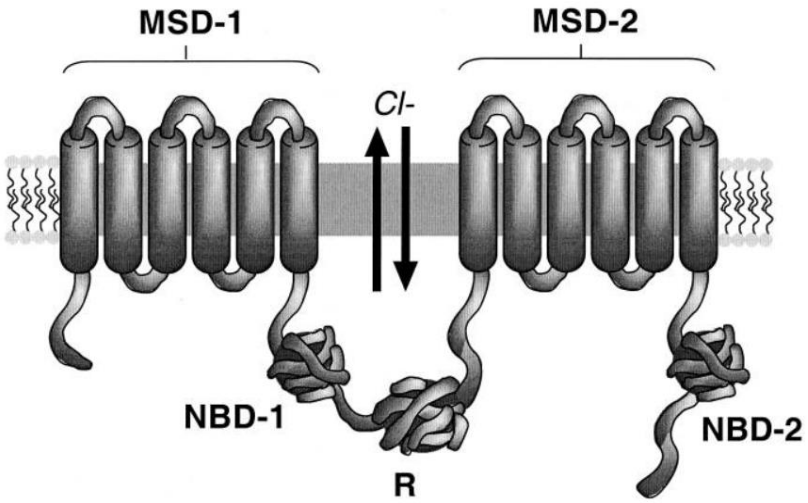

Figure 1: CFTR Structure. Protein channel member of the ABC family, being its function the transport of ions through the cell membrane. It has two nucleotide binding domains (NBD1 and NBD2) and a regulatory domain (R). (Retrieved from Lyczak *et al. (*2002).

Most mutations occur in NBD1, being the first described and most common, the loss of a phenylalanine at position 508 in the CFTR protein. The result of this deletion is a folding defect that causes not only problems in CFTR processing and intracellular transport to the plasma membrane, but also destabilizes the ones that try to reach it (Bose *et al.* 2019).

There have been developed many treatments for the many signs and symptoms affecting CF patients, which have led to improvements in the quality of life and also life expectancy. In recent times, CFTR modulators became accessible only for patients with certain types of mutations, with the aim to improve the function of CFTR protein. There are four main types of CFTR modulators all with different approaches: Potentiators that increase open channel probability by holding open the channel to allow chlorine to flow through; Correctors that help the protein with folding increasing the traffic to the cell surface; Stabilizers that extend the half-life of the CFTR protein at the plasma membrane and, lastly, Amplifiers that increase the quantity of CFTR available in the cell. The type of modulator used will depend on the class of mutation causing the disease (Clancy et al. 2019; Fukuda and Okiyoneda, 2018).

1.1.1. Epidemiology

There were 31.199 individuals with CF in the Patient registry - Annual Data Report of 2019, which contains data from Cystic fibrosis patients under care at CF Foundation accredited care centers in the United States, who consented to have their data entered, being more than 90.0 percent of them Caucasians (Marshall *et al.* 2020). The frequency of CF, in all European countries, is thought to be around 0.0074 percent. The number of infants diagnosed with CF continues stable while the number of adults keeps going up, representing now 56.0 percent of the CF population making this no longer a Pediatric disease. As a result of neonatal screening that is taking place in certain countries and also new treatments, the prevalence of CF has been decreasing, while the disease complexity has increased over the last few years. Even with all the advances in diagnosis and treatment the life expectancy of individuals with this disease is only around 46 years (Quintana-Gallego *et al.* 2014; Bhagirath *et al.* 2016).

Nowadays, there is both pre-conceptional carrier screening and neonatal screening. It is available, in the USA, a program of pre-conceptional carrier screening for 25 of the most common CF mutations mostly for couples that are Caucasian or have a family history of CF. As can be seen in figure 2 most diagnoses were done at young ages which allows for higher life expectancy due to early treatment. Neonatal screening usually has a sensitivity of 85.0-90.0 percent starting with a biochemical test for immunoreactive trypsin (IRT) in plasma. For children with elevated IRT a genetic test is performed to detect a mutation and then a sweat chloride test is done to confirm the diagnosis. Having DNA based testing indicates that unaffected heterozygous carriers are discovered (Brice *et al.* 2007).

In Portugal, the average age of patients with CF was 31.23 ± 9 years, being 54.0 percent of them female and 46.0 percent male. As can be seen in figure 2 most diagnoses were done at young ages which allows for higher life expectancy due to early treatment (Silva A *et al.* 2016).

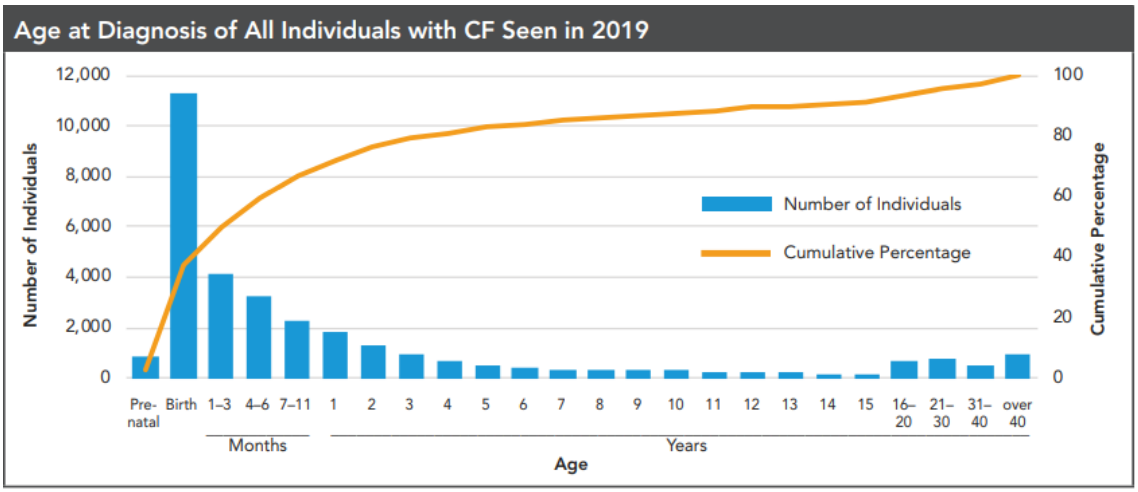

Figure 2: Age at diagnosis of all the CF patients screened in 2019. (Retrieved from Marshall *et al.* (2020))

1.1.2. CF in the lungs

The main function of the respiratory system is to provide oxygen and remove carbon dioxide from the vascular system, consequently, giving oxygen to all the organs in the body. The human respiratory system is very complex, being divided in the upper respiratory tract, consisting of the nasal cavity, oral cavity, paranasal sinuses, pharynx and larynx, and the lower respiratory tract which is also divided in conductive and respiratory zones (Bhagirath *et al.* 2016).

It is known that the morphology of the airways plays a particularly important role in the good function of the lungs. This way the shape and size of the airways can be a guide for the severity of the respiratory disease (Montaudon *et al.* 2007).

The function of the conducting airways is to conduct and warm the air that comes from the upper respiratory system, while the respiratory airways is where the gas exchange occurs. The epithelial of both these zones is composed of goblet cells, ciliated cells, and basal cells, which are all covered by the airway surface liquid consisting of a periciliary layer and a mucus layer (Fig. 3a). To remove the possible dangerous matter that can be trapped in the mucus layer, the periciliary layer is struck by the cilia from the ciliated cells. This process is called mucolarity clearance and is considered essential for the proper functioning of the respiratory system. For this process to be efficient is especially important that the regulation of the airway surface liquid (ASL) hydration, by the transport of water and ions, is working properly, which as was previously mentioned does not occur in CF patients (Bhagirath *et al.* 2016; Saint-Criq and Gray, 2016).

Most ASL originates from the submucosal glands but, its final composition and volume are also regulated by the surface epithelia, this is because the airway epithelial cells express not only CFTR but also epithelial sodium channel (ENaC) (Saint-Criq and Gray, 2016). ENaC is a sodium channel bound to the membrane that allows the transport of sodium ions (Na^+) (Hanukoglu and Hanukoglu, 2016). The absorption of Na^+ and secretion of Cl^- is what allows the surface airway epithelial cells to regulate the

volume of ASL, which, as previously said, is extremely important to the efficiency of mucus clearance (Saint-Criq and Gray, 2016; Tarran *et al.* 2002). Some mutations have been found to affect CFTR and HCO_3^- transport, which is thought to suggest that CFTR can mediate HCO_3^- transport. HCO_3^- is a particularly important ion with many functions, including serving as a pH buffer and increasing the ion and protein solubility in fluids. Seeing that pH has an effect in mucin viscosity it is safe to say that HCO_3^- is of extreme importance in CF (Choi *et al.* 2001).

What happens in CF patients is that the mucus clearance is affected due to the depletion of ASL, which will not allow for the elimination of bacterial infections. It is expected that if there is any defect in the CFTR of submucosal glands there will be a decrease of ASL depth and periciliary layer. Submucosal glands contain both serous cells, which produce a watery secretion, and mucus cells, that are responsible for secreting mucins (main component of mucus). If there is a defect in the fluid secretion by the serous cells, mucin secretion will also be affected, even if mucus cells do not express CFTR, which can lead to abnormally adherent mucus that will become fixed to the gland orifice or the originating goblet cells (Fig. 3b) (Inglis *et al.* 1998; Engelhardt *et al.* 1992).

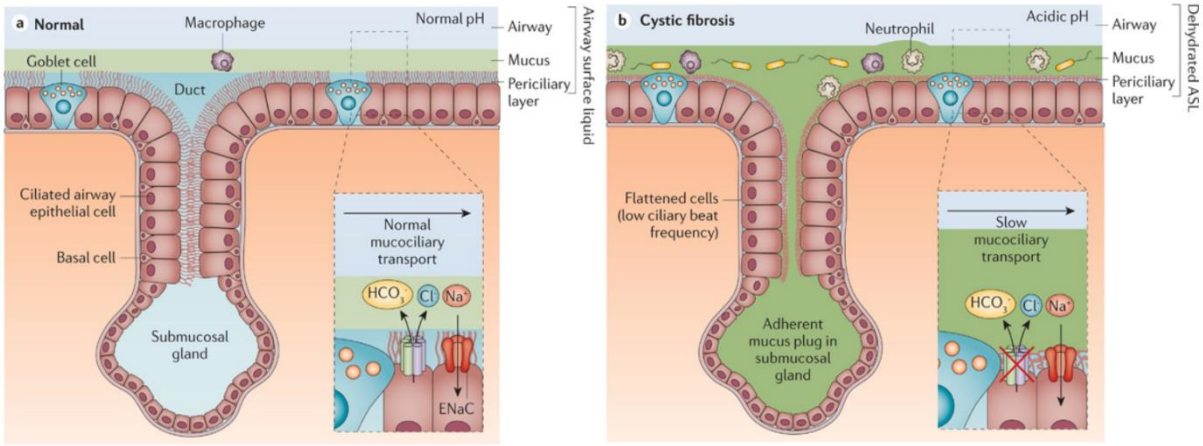

Figure 3: Mucociliary clearance defect in CF epithelial airways. (a) Airway epithelium of a healthy individual; (b) Airway epithelium of CF patient with dehydrated ASL. (Adapted from Ratjen et al. (2015))

Previously, it was thought that ASL became depleted due to the lack of Cl^- secretion, which would affect the secretion of fluids in surface epithelia. It is known that CFTR can also regulate other ion channels like ENaC, having an effect on the secretion of various molecules that can affect ASL pH regulation, which it has been shown to be important for bacterial killing (Hoegger *et al.* 2014; Pezzulo *et al.* 2012). The short palate lung and nasal epithelial clone 1 (SPLUNC1) is a protease inhibitor that serves as example of molecules that depend on optimal ASL pH, which function is to regulate antimicrobial peptides (AMPs) (Garland *et al.* 2013; Ahmad *et al.* 2016). It is extremely important that the airways preserve the ability to eliminate pathogens and keep an almost sterile environment in the lungs. The inappropriate functioning of this ability can lead to respiratory system complications that will eventually cause respiratory failure (Rafeeq and Murad, 2017).

The immune system also plays an important role in eliminating pathogens from the airways, relying on the action of innate immune cells, such as, neutrophils and macrophages. Considered the first line of defence against pathogens, neutrophils, release both oxidants and proteases that have been proven to affect CFTR function. These cells form extracellular traps (NETs) can prevent the spread and eliminate all types of pathogens, like bacteria, yeasts, viruses, and protozoans (Martínez-Alemán *et al.* 2017). However, these traps are thought to increase the viscosity of secretions, inhibiting pathogen clearance, and potentially enhancing biofilm development, which cause damaging effects in the airway of CF patients (Young *et al*. 2011).

As previously mentioned, the thick mucus produced in the airways of CF patients will alter mucociliary clearance increasing the chances of an infection. The presence of pathogens will trigger a massive neutrophils recruitment to the site of infection. Neutrophils kill pathogens indirectly through releasing inflammatory mediators. The recruitment of more immune cells will be prompted by the increase of these inflammatory mediators. Eventually, this will form a cycle of inflammation that prolongs itself. This type of inflammation usually happens very early in CF patients and can determine the severity of the disease (Khan *et al.* 2019; Cantin *et al.* 2015).

1.1.3. Microorganisms affecting the CF lungs

Chronic infection of the lungs is a huge problem affecting patients with CF and can be caused by various opportunistic pathogens like *Pseudomonas aeruginosa* and members of the *Burkholderia cepacia* complex (Bcc). As shown in figure 4, *Staphylococcus aureus* is the most common pathogen in patients with CF and as the individuals get older it is usually seen in simultaneous with *P. aeruginosa*. Over 60.0 percent of young individuals with CF are colonized by a least one microorganism, and as these individuals get older the percentages can go up to 80.0 percent. Eventually, chronic infection and associated inflammation of the airway will cause respiratory failure in the majority of patients with CF (Marshall *et al.* 2020).

In most CF patients from the Patient registry - Annual Data Report of 2019 , the early colonizers are microorganisms, like *S. aureus* and *Haemophilus influenzae*, at an early age. This early colonization can lead to damages in the epithelial surfaces and later colonization by *P. aeruginosa*. It is proven that the main cause of mortality, by lung function decline, in CF patients is the chronic infection by *P. aeruginosa* (Lyczak *et al.* 2002). Some bacteria are emerging as a threat to CF patients, such as methicillin-resistant *S. aureus* (MRSA), which are found colonizing around 30.0 percent of patients, and *Burkholderia cepacia* complex that are a challenge because of inherent antibiotic resistance and potential for patient-to-patient spread (Marshall *et al.* 2020).

Different genera of anaerobic bacteria can also be identified in around 64.0 percent of sputum samples collected from adults with CF. It is believed that the presence of *Pseudomonas aeruginosa* increases the chance of colonization by anaerobic bacteria (Tunney *et al.* 2008). Formerly, the focus was directed to characterising bacteria colonizing the lung of these patients, but recently fungal species, such as *Aspergillus*, and viral communities, like influenza A and B viruses have been identified (Amin *et al.* 2010; Wat *et al.* 2008).

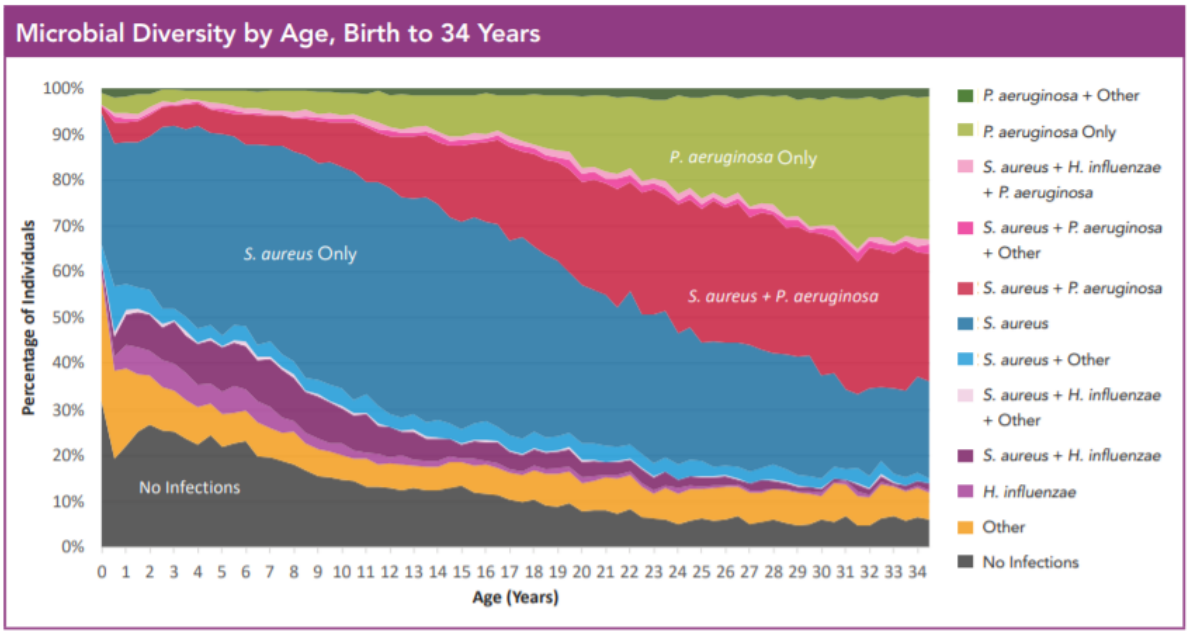

Figure 4: Microbial diversity by age. (Retrieved from Marshall *et al.* (2020))

1.2. *Burkholderia cepacia* complex (Bcc)

A phytopathogenic bacterium was first described, in 1950 by William Burkholder, as the agent causing bacterial rot of onion bulbs in New York (Burkholder, 1950). The species previously known as *Pseudomonas cepacia* were reclassified, in 1992, to the new genus *Burkholderia* (Yabuuchi *et al.* 1992).

B. cepacia was later noted to be not just a single species, but multiple ones, more specifically, 5 genomovars. *B. cepacia* genomovar I maintained the name of the type species of this complex, *B. cepacia*, for genomovar II it was given the name *B. multivorans*, and genomovars III, IV and V correspond, respectively, to *B. cenocepacia*, *B. stabilis* and *B. vietnamiensis*. For this reason, *B. cepacia* is now named *B. cepacia* complex, that consists of many microorganisms that even though are phenotypically similar, are genetically different (Coenye *et al.* 2001). Nowadays this complex is composed of more than 22 species, all very versatile in different environments due to their genome size, which ranges between 6 and 9 Mb, and DNA separation in two or more chromosomes and several plasmids (Devanga Ragupathi and Veeraraghavan, 2019).

These Gram-negative bacilli are believed to be opportunistic pathogens in immunocompromised people, such as patients suffering from cystic fibrosis (CF), causing a pneumonic illness, respiratory failure, and bacteremia, known as “cepacia syndrome”. These seem to be the only CF pathogens that can cross over the epithelial barrier and enter the blood stream causing bacteremia (Mahenthiralingam *et al.* 2005).

Bcc species can adapt to changing conditions being able to live in a variety of environments, but most species were isolated from sputum of CF patients. Some Bcc species are considered animal or plant pathogens, but others can bring beneficial effects to plant hosts, such as releasing iron and fixing nitrogen. Due to the benefits these bacteria bring to host plants their economic value has been increasing, although the increased use of *Burkholderia* spp. in agriculture would also increase human exposure to a potential danger (Devanga Ragupathi and Veeraraghavan, 2019). Infections caused by Bcc species are usually extremely hard to treat due to the antimicrobial resistance characteristic of them. Efflux pumps, biofilm formation, and reduced permeability offer resistance to many antimicrobials including, aminoglycosides, quinolones, and β-lactams (Leitão *et al.* 2010; George *et al.* 2009).

The identification of Bcc species is often difficult due to the many similarities between them. For example, the 16S rRNA method cannot be used because the sequence is identical in some of the species. For the correct identification of Bcc species genetic methods are usually the ones applied. The most accurate method is a Multilocus sequence typing (MLST) method that identifies the different species by comparing several house-keeping genes (Baldwin *et al.* 2005). A new method was then developed, as an adaptation from the MLST, that can be done directly from a sputum sample, and is performed by a polymerase chain reaction (PCR), separating the many isolates into the specific species, using genes such as the *recA* gene. However, using the *recA* gene to identify the species can also have some complications due to *B. cenocepacia* being divided in 4 phylotypes (IIIA-D) (Drevinek *et al.* 2010).

Bcc strains have the capacity for interpatient transmission as confirmed by the decrease of infections when infection control measures were implemented (Muhdi *et al.* 1996). The infection with some strains, such as, ET12 and SLC6 was basically eliminated with the implementation of these measures, but this did not occur for other Bcc strains, which suggests acquisition form other sources (LiPuma, 2010). Some strains, like the PHDC are considered widely distributed in the natural environment, while others, such as the Midwest strain is not usually found outside of CF patients. The widespread of Bcc strains is due to a variety of factors, like the distribution, infection control measures and other uncharacterized factors (Springman *et al.* 2009; Coenye, 2004).

1.2.1. Epidemiology

Despite the presence of Bcc species usually causing chronic infection in CF patient, it is also possible that it leads to only a transient infection. Even though Bcc species are closely related genetically, some species are not even associated with human infection, and the most commonly found in CF patients is

Burkholderia cenocepacia followed by *Burkholderia multivorans*. This phenomenon has yet to be explained since this predominance does not occur in their natural environment (Reik *et al.* 2005).

Bcc are usually present in 3.0 percent of CF patients by the age of 18. In the USA, *Burkholderia* species were isolated from 1218 CF patients for 7 years (1997-2004), 45.0 percent of these patients were infected with *B. cenocepacia* and 39.0 percent by *B. multivorans* (Reik *et al.* 2005). A somewhat similar distribution occurred in France where 45.0 percent of isolates were from *B. cenocepacia* and 52.0 percent from *B. multivorans* (Brisse *et al.* 2004). In the case of Italy and Portugal, the most predominant species was also *B. cenocepacia*, with 87.0 and 52.0 percent, respectively, but *B. multivorans* was at very low numbers (Bevivino *et al.* 2002; Cunha *et al.* 2003). The rarely found species *B. stabilis* was found in around 18.0 percent of the 23 Portuguese CF patients and is thought to be the cause of "cepacia syndrome" in one of the patients (Cunha *et al.* 2003).

Although most Bcc infections occur with only one strain, it is possible that a coinfection occurs and even a replacement of the initial strain with a different one (Yang *et al.* 2006). There have been described cases where the strain that is initially causing the infection is replaced by a different strain. For example, this happened in 6 patients that were infected by *B. multivorans* and later were infected with *B. cenocepacia*. One of these patients maintained a coinfection with both *B. multivorans* and *B. cenocepacia* for 2 years until their death (Ledson *et al.* 1998). In another case, 24 patients infected with Bcc strains and 3 patients with *B. gladioli* were all replaced with *B. cenocepacia* strains (Mahenthiralingam *et al.* 2001). A replacement of a *B. cenocepacia* strains with other Bcc strains was not described in any of these studies.

Bacteria from the Bcc do not exclusively infect immunocompromised patients. Cases of infection of non-CF patients have been described along the years, being most of them nosocomial. These infections are usually caused via contaminated hospital equipment, but patient-to-patient transmission from CF patients has been reported. The infections caused by intravenous catheters can cause bacteremia in non-CF patients, but the infection subsides when the contaminated equipment is removed (Ledson *et al.* 1998). Also, in hospital settings, contaminated water or mechanical ventilation can cause pulmonary infection in immunocompetent patients (Bernhardt *et al.* 2003). In a study done in the UK, Bcc species were isolated from 112 non-CF patients, and *B. cepacia* was considered the most common species, being found in 37 of the 112 patients, followed by *B. cenocepacia*, which was present in 25 of the 112 non-CF patients. A very rare case of transmission from CF children to their immunocompetent mothers was also reported (Ledson *et al.* 1998).

In 2015, Hauser et al., reported what is thought to be the first case of Cepacia syndrome in a patient without cystic fibrosis. The 64-year-old patient had a history of systemic arterial hypertension and diabetes, and during the hospital stay he received both non-invasive and invasive ventilation, which is thought to be the main cause of the Bcc infection. Cultures from both blood and bronchoalveolar lavage (BAL) were done and in both cases, there was Bcc growth. The patient died after 30 days in the intensive care unit (Hauser and Orsini, 2015).

1.3. Bacterial biofilms

It is known that bacteria have the ability to live in both planktonic and biofilm form. When in biofilm form the bacterial cells are sheltered by a self-produced extracellular polymeric matrix, composed of extracellular polymeric substances (EPS), such as, polysaccharides, proteins, lipids, nucleic acids, and other biopolymers (Costerton, 1999). The basic structure of a biofilm was described in 1994, as being a heterogenic structure composed of microcolonies enclosed in an extracellular polymeric matrix and water channels (Fig. 5). The purpose of these water channels is thought to be oxygen and nutrient transport to the microcolonies and efflux of metabolic products (Costerton *et al.* 1995).

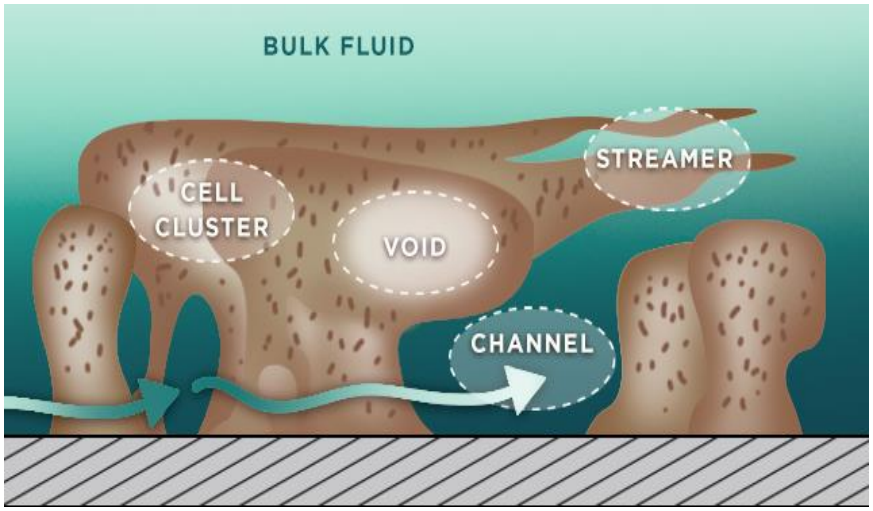

Figure 5: Biofilm structure. Water channels between the bacterial microcolonies enclosed in an extracellular polymeric matrix. The arrows indicate the water flow in the channels. (Retrieved from (www.technologynetworks.com/immunology/articles/care-in-the-community-how-biofilms-improve-bacterial-survival-311685))

Even though in the planktonic form the growth rate of bacteria seems to be higher, their preferred form is usually biofilm (Jefferson, 2004). The predominance of bacteria in biofilm form could be due to many different reasons, being the main ones, the protection that they provide against environmental changes, making them 1000 times more resistant to antimicrobials provided by the EPS layer of the biofilm than when in planktonic form. The cells density in mixed biofilms allows for a higher rate of horizontal gene transfer, which could mean that genes encoding antibiotic resistance could easily be transferred between the present bacteria species (Rasmussen and Givskov, 2006).

Biofilms can be an important tool at an industrial level, such as the use of biofilm forming bacteria in the treatment of municipal sewage and industrial wastewater, but when it comes to the clinical level, biofilms can be a major problem because of the ability of certain bacteria to form biofilms in medical implants and tissue (Yu *et al.* 2007; Haaber *et al.* 2012).

When it comes to the human health, biofilms can have a positive or negative impact. Commensal bacteria, like *Staphylococcus epidermis*, stimulate the host immune response, preventing the colonization and adhesion of potential pathogenic bacteria. Nevertheless, many human diseases are associated with biofilms, a common example is dental plaque. The composition of dental plaque can be

altered according to local environmental changes, such as acidification caused by increased consumption of fermentable carbohydrates. This acidification will end up killing the bacterial species in the healthy biofilm that have low acid tolerance, and allowing the growth of acid-tolerant bacteria, like *Streptococcus mutans* and *Streptococcus sobrinus*. This change in the composition of the biofilm can lead to oral cavity disease caused by the demineralization of the enamel (Haaber *et al.* 2012; Sbordone *et al.* 2003).

Another example of biofilms related to human disease is cystic fibrosis. As previously mentioned, CF causes the formation of thick mucus that blocks the airway, allowing for bacteria like *P. aeruginosa* to colonize the lungs. *P. aeruginosa* is one of the most important pathogens colonizing the lungs of CF patients, able to persist for decades because of the overproduction of alginate, a matrix polysaccharide (Romling *et al.* 1994). This offers the bacteria resistance to the immune system, antibiotics, and phagocytosis, triggering chronic inflammation that leads to severe damage to the lungs (Pedersen *et al.* 1992).

Two types of biofilms will be considered in this introduction, surface-attached biofilms, and planktonic cellular aggregates.

1.3.1. Surface-attached biofilms

Biofilms are considered sessile communities of bacteria that can accumulate on solid biotic and abiotic surfaces, including, contact lenses, implants, catheters and animal and plant tissue (Haaber *et al.* 2012).

There are 5 stages of biofilm formation (Fig. 6). The first stage is the initial reversible attachment (1), which is thought to be regulated by the nutrient availability in the environment. In this stage flagella and type IV pili-mediated motilities are extremely important for interaction and attachment to the surface. The irreversible attachment includes two stages (2-3) in which the EPS matrix starts to form and the bacteria replicate forming a microcolony. In the maturation stage (4) the biofilm has a mushroom-like form where bacteria are able to communicate with each other and are arranged according to the different needs and metabolism, i.e., anaerobic bacteria tend to inhabit deeper layer of the biofilm for less contact with oxygen. The last stage of biofilm formation is dispersion (5), as shown in Figure 6, where a detachment occurs allowing the now free bacteria or cellular aggregate to start a new cycle (Rabin *et al.* 2015).

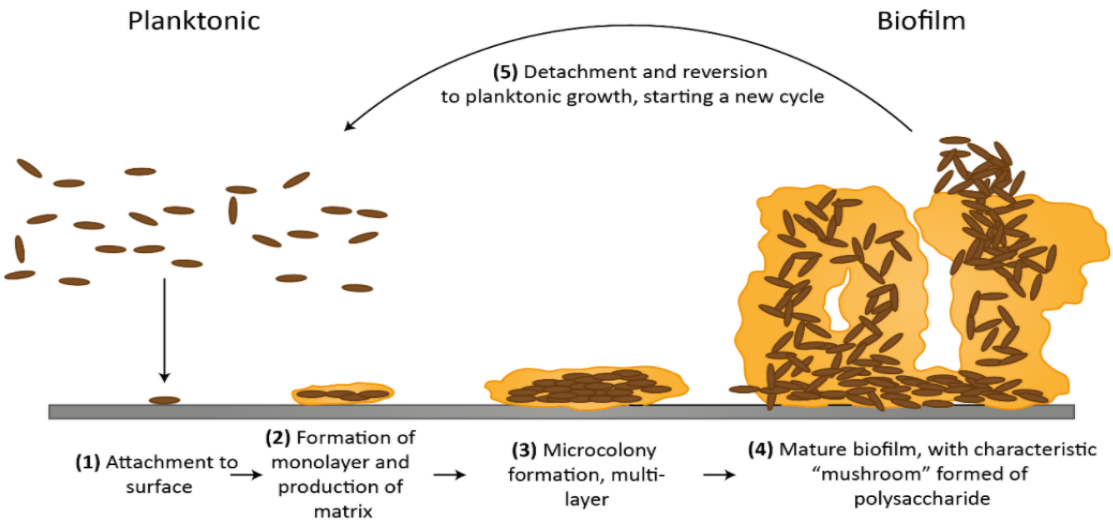

Figure 6: Schematic representation of biofilm formation. Five steps of biofilm formation according to the current model: (1) Attachment of bacterial cells to surface; (2) Formation of the matrix; (3) Cell growth and division forming a microcolony; (4) Mature biofilm with a mushroom-like appearance; (5) Detachment of cells from the biofilm and beginning of new cycle. (Retrieved from (https://www.immunology.org/public-information/bitesized-immunology/pathogens-and-disease/biofilms-and-their-role-in))

1.3.2. Planktonic cellular aggregates

When bacteria adhere to each other instead of a surface they form surface-independent biofilms, usually called planktonic cellular aggregates. These aggregates present the same physiological characteristic as other forms of biofilm and their formation can be a response to environmental stresses, being thought to be mediated by the same mechanisms as surface-attached biofilms (Klebensberger *et al.* 2006).

Some bacteria are able to form cellular aggregates in wastewater, being usually called clumps or flocs and are a strategy for bacteria to be able to survive in bodies of water. These floc-forming bacteria, like *Aquincola tertiaricarbonis*, can be used in wastewater treatment. Clumps and flocs are both different types of surface-independent biofilms. Clumps are formed by cells that adhere to each other by their non-flagellated poles, being able to later separate due to the reversibility of the process, while flocs are formed by bacteria flocculation that may occur after clumping. In the flocculation mode cells are encased in an EPS matrix making this irreversible (Bible *et al.* 2008).

Just like in surface-attached biofilms, the formation of planktonic cellular aggregates has 5 stages (Fig. 7): the first one in which bacteria approach each other by swimming motility or Brownian motions, the second and third stage where bacteria adhere to one another and begin the production of the EPS matrix, rendering the attachment irreversible. In the fourth stage in which the bacteria start to grow and replicate forming microcolonies, lastly the fifth stage is where the mature biofilm starts to disperse through either planktonic cells ou small aggregates to possibly start a new cycle (Dienerowitz *et al.* 2014).

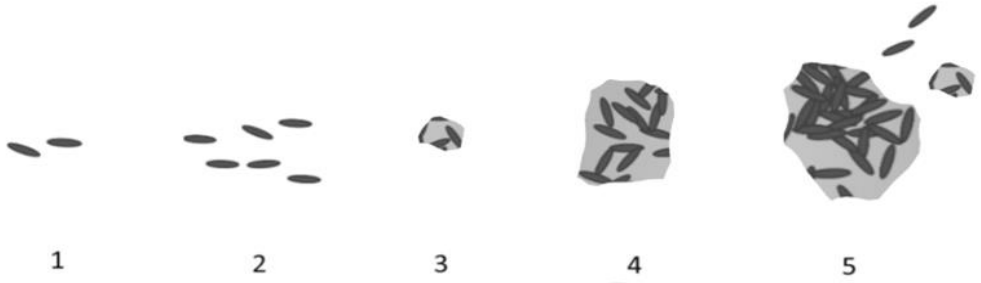

Figure 7: Schematic representation of planktonic cellular aggregate formation. (1) Cell-to-cell approach through motility or Brownian motion; **(2-3)** Interaction between single cells and EPS matrix production forming the aggregate; **(4)** Bacterial growth and replication; **(5)** Dispersion of single cells or small aggregates. (Adapted from (https://www.immunology.org/public-information/bitesized-immunology/pathogens-and-disease/biofilms-and-their-role-in))

The study of these planktonic cellular aggregates is of great importance due to their relation with human disease. It has been proven that important pathogens in cystic fibrosis also can form these aggregates, such as bacteria from the *Burkholderia cepacia* complex (Silva *et al.* 2017).

When studying *P. aeruginosa*, in log fase the population appears to prefer the aggregate form, while in the stationary fase the single cells were more predominant. Even though in log fase most of the population preferred grow as aggregates, in certain cases of environmental stress, such as oxygen limitation, the aggregates tend to disperse releasing planktonic cells (Schleheck *et al*. 2009). These planktonic aggregates display similar responses and behaviour as surface-attached biofilms, allowing for the same benefits but maintaining the motility of single free cells, makes the infections more difficult to eliminate (Alhede *et al.* 2011).

1.3.3. Elements controlling biofilm formation

1.3.3.1. Quorum Sensing and its role in controlling biofilm formation

In 1979, Nielsen et al., described Quorum sensing (QS) for the first time, in two marine bacterial species. These species were able to produce light due to enzymes encoded by a specific operon. They discovered that the light was only produced when the bacterial concentration was high (Nealson and Hastings, 1979). Other examples of QS were later described, i.e., the antibiotic production by *Streptomyces* spp., but in most cases, it was not believed that bacteria could communicate because it was thought to be exclusive to eukaryotes. Nowadays, evidence has been established that most bacteria have the ability to communicate with each other through chemical signals, and that a wide variety of these signals are involved, allowing bacteria to even distinguish different species (Miller and Bassler, 2001).

The QS mechanism allows to monitor cell density and synchronize gene expression, for example, to form biofilms. Bacteria are able to communicate with each other by releasing an array of signals usually

called “autoinducers”. When the cell density is low the concentration of autoinducer in the media will be small, causing it to not be detected by bacteria. With higher cells density the concentration of autoinducers will reach a threshold, allowing bacteria to detect them, inducing, or repressing the expression of specific genes. QS is mediated by acyl-homoserine lactones (AHLs) in Gram-negative bacteria and by small peptides in Gram-positive bacteria (Lyon and Muir, 2003).

Proteins belonging to the LuxR family of regulators will sense AHL when they are in high concentration. The AHL will bind to the LuxR regulator altering its conformation and allowing it to bind its other domain to the DNA, regulating the transcription of the target gene (Fig. 8). However, some LuxR homologues can act as repressors of gene expression when AHLs are absent. The type of *luxI-luxR* homologous QS system can vary between the different species (Fuqua *et al.* 1994).

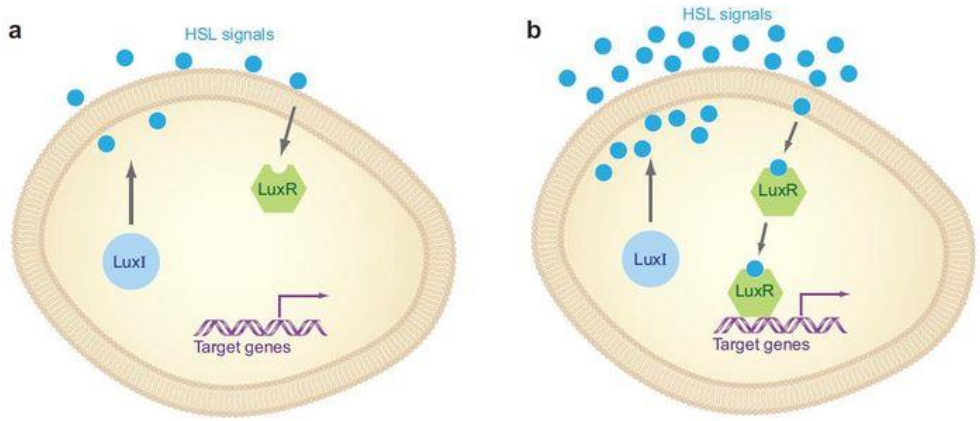

Figure 8: Quorum sensing system in gram-negative bacteria. The bacteria secrete AHLs (blue circles) that in high concentrations activate the AHL receptor and induce the expression of QS regulated genes. (a) Low cell density; (b) High cell density. (Retrieved from (http://2013.igem.org/Team:NTU-Taida/Project/Background))

A lot of the physiological processes of bacteria are regulated by QS, being one of these processes’ biofilm formation. In the case of *S. aureus*, AgrA regulates the transcription of genes involved in the dispersion of biofilm (Boles and Horswill, 2008). In *Bacillus subitlis* the production of the EPS matrix is regulated by a QS molecule named surfactin (Lopez *et al.* 2008). In *P. aeruginosa* there are 2 different AHL-based QS systems controlling biofilm formation, Las and Rhl. Some extracellular virulence factors and the RhIR-RhII system are both controlled by the LasR-LasI system, being the production of a variety of secondary metabolites regulated by the RhIR-RhII system. A *Pseudomonas* Quinolone Signal (PQS) is a third QS related molecule that stimulates certain virulence factors by connecting the LysR-type transcriptional regulator PqsR, although the precise role of this molecule is still poorly understood (Lee and Zhang, 2014).

Lewenza et al, discovered, in 1999, a QS system belonging to *B. cepacia*, consisting of CepI and CepR, two LuxR homologs responsible for regulating a variety of virulence factors, such as, motility and biofilm production. CepR is the regulator that can be found in all Bcc bacteria. The CepRI system was proven to not be involved in adhesion to surfaces, because mutants for cepI and cepR were able to

adhere to surfaces but could not create a mature biofilm. The QS system CciIR was also found to have an effect in biofilm formation in these bacteria, as well as the transcriptional regulator ShyR (Lewenza *et al.* 1999).

1.3.3.2. Quorum Quenching

Quorum quenching (QQ) was the name given to all processes disrupting QS. This disruption can be done in a variety of ways, such as, inhibition, mimicking, degradation, or modification of QS signaling molecules (Alagarasan *et al.* 2017). In the case of inhibition of the signaling molecules, closantel and triclosan are known inhibitors of QS by disrupting AHL synthesis. Furanones are molecules able to mimic AHL molecules and can also shorten their half-life, decreasing the concentration of these signaling molecules in the cell. An enzyme capable of degrading AHL have been describes in *Bacillus cereus*, due to the toxicity of its products against some bacteria. QQ can also be done by modification of signaling molecules, such as the oxidoreductase activity in AHL that modifies the acyl chain, inactivating QS (Basavaraju *et al.* 2016; Chan *et al.* 2010).

These inhibitory molecules can be found in a variety of phyla, including, bacteria, plants, fungi and even mammals. There are two groups of QS inhibitory molecules distinguished by the target and mode of action, the quorum quenching enzymes (QQ), and the quorum sensing inhibitors (QSIs). Lactonases, oxidoreductases, and acylases are all considered QQ enzymes (Hong *et al.* 2012).

Lactonases capable of degrading AHLs and affecting biofilm formation have been studied both *in vivo* and *in vitro*, concluding that they should be considered candidates for antivirulence therapies. *B. cepacia*, according to Chan et al., possesses both QS and QQ systems, describing an oxidoreductase capable of resisting *Erwinia carotovora* virulence factors by degrading AHLs (Chan *et al.* 2011). Also, more recently, two *B. cepacia* lactonases, YtnP and Y2-aiiA, were able to inhibit virulence from *P. aeruginosa* by interfering with its QS systems, Las, Rhl, and Pqs (Malešević *et al.* 2020).

1.3.3.3. Efflux pumps

Efflux pumps are used by bacteria to pump out waste products and toxic substances. They can be either single or multiple component systems, being the multicomponent systems only found in Gram-negative bacteria, such as the AcrAB-TolC (*Escherichia coli*) and MexAB-OprM (*Pseudomonas aeruginosa*) systems that transport the waste directly to the exterior of the cell. Efflux pumps are divided in five superfamilies: major facilitator superfamily (MFS), multidrug and toxin extrusion (MATE), ATP-binding cassette (ABC), resistance-nodulation-division (RND), and small multidrug resistance (SMR). Many studies have suggested the role of these pumps in biofilm formation and QS (Fig. 9) (Alav *et al.* 2018).

All bacteria possess multiple efflux pumps, which makes it extremely difficult for one type of efflux pump inhibitor (EPI) to prevent biofilm formation. Nevertheless, in studies done using two complementary EPIs for the different efflux pumps, biofilm formation was completely inhibited, and the antibiotic resistance that bacteria acquire when in biofilms can also be blocked by these inhibitors. This was also confirmed with knockout mutans for two efflux pump encoding genes, being biofilm formation very reduced when compared with the wild-type strain (Kvist *et al.* 2008).

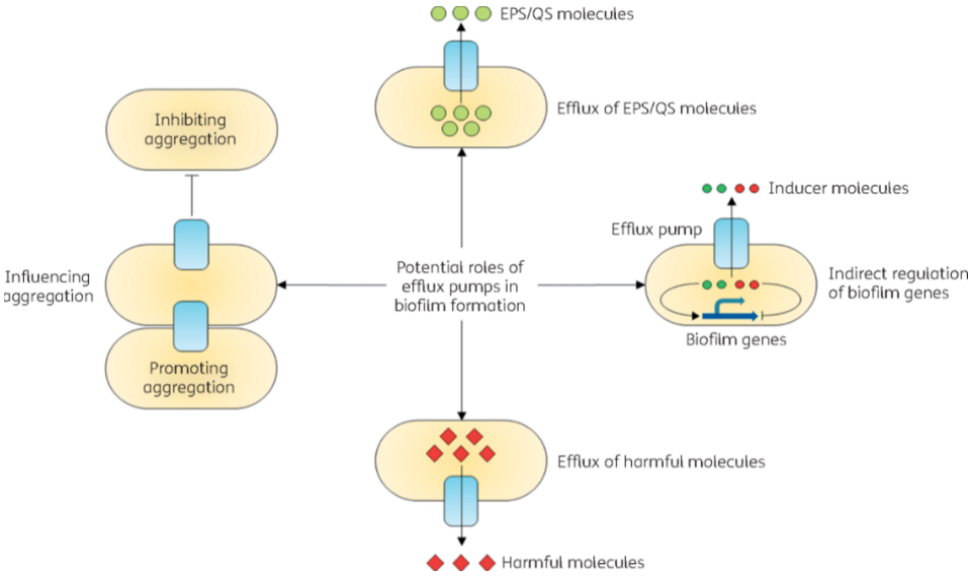

Figure 9: Four potential roles of efflux pumps in biofilm formation. (Retrieved from Alav *et al.* (2018))

In *P. aeruginosa*, many studies have proven the involvement of efflux pumps in biofilm formation, through AHL transport. As previously mentioned, AHLs are signalling molecules involves in QS of Gram-negative bacteria. When *P. aeruginosa* cells were incubated with an EPI, the concentration of the specific AHL increased inside the cells, proving the involvement of these pumps in AHL transport. This increase in concentration also occurred in mutants lacking genes that encode efflux pumps (Pearson *et al*. 1999).

Efflux pumps are also thought to be involved in biofilm formation by *Burkholderia* spp., also by AHL transport. In 2009, Buroni et al., described two genes, BCAL1675 and BCAL2821, that encode for efflux pumps in *B. cenocepacia*. When mutants lacking these genes were analysed, the concentration of AHLs in the growth media was much lower than the wild-type strain. Similar intracellular accumulation of AHLs occurred in *B. pseudomallei* mutants lacking the *bpeAB* gene, which encodes for an RND-efflux pump (Buroni *et al.* 2009; Chan *et al.* 2007).

1.4. *Burkholderia multivorans* aggregates

The knowledge concerning the formation of planktonic aggregates by *Burkholderia multivorans* is still scarce. However, a study by Silva et al. involving several Bcc species, described the formation of planktonic aggregates by both clinical and environmental isolates in salts media containing carbon sources, i.e., fructose or glucose. Also, in the same study, the analysis of *B. multivorans* ATCC 17616 deletion mutant for the gene encoding the transcription regulator LdhR presented a decrease in the ability to form multicellular aggregates (Silva et al. 2017).

More recently, a study done by Ferreira, 2018, concerning the genes involved in the formation of planktonic aggregates by *B. multivorans* was performed. In this study, a genetic library created through random insertional mutagenesis was screened and the different phenotypes were analysed. Some of the mutants analysed showed a connection between the gene and the phenotype, being one example, the mutant named E70. This mutant has a disruption in the *ppsA* gene encoding a phosphoenolpyruvate synthase involved in the synthesis of AHL in *B. cenocepacia*, that displayed lower capacity to form aggregates. The second example is the mutant G38 that also presented a reduced ability to form aggregates, with a disruption in the *gltB* gene that encodes a glutamate synthase, an enzyme involved in glutamate synthesis. Without this enzyme glutamine will accumulate in the cell inactivating the NtrBC signaling system, causing a decrease in swarming motility and exopolysaccharide production. The last example is B63, a mutant producing larger aggregates and with decreased virulence. This mutant had a disruption in the *xdhA_3* encoding a xanthine dehydrogenase molybdenum-binding subunit XdhA, an enzyme involved in purine salvage pathways important for the GMP/GTP maintenance and, eventually, c-di-GMP, evidencing the involvement of c-di-GMP in biofilm formation by these bacteria.

Another study on the formation of planktonic cellular aggregates by clinical isolates of *B. multivorans*, tested the impact of different stresses on the ability to aggregate, being osmotic and nitrogen stresses strain specific. In isolates that did not display the ability to form aggregates, ciprofloxacin was found to induce their formation, which resulted on a great increase on aggregate biomass percentage. Also, the impact of the LdhA enzyme on aggregation was tested by creating a *ldhA* mutant and evaluating its ability to aggregate. The role of this D-lactate dehydrogenase LdhA in aggregation was confirmed by the overall decrease in aggregation presented by mutant, while the overexpression of LdhA resulted in a strong increase (Gomes, 2018).

1.5. Aim of this thesis

Organisms of the *Burkholderia cepacia* complex can cause infections in CF patients and are considered a risk factor for their survival and quality of life. A way of trying to avoid chronic infection with these organisms is through antibiotic treatment which can cause an increase in resistance by the colonizing species and jeopardize future treatments. Another way these species increase their resistance to treatments is by forming multicellular aggregates that protect the bacteria from external stress such as the immune system, oxidative stress and antimicrobials.

Although the number of studies assessing multicellular aggregate formation by *Burkholderia multivorans* is scarce, the knowledge of the mechanisms involved in their formation could possibly improve the chances of bacterial eradication. Therefore, the aim of this thesis is to identify molecular players that may be involved in the formation of planktonic cellular aggregates in *Burkholderia multivorans*. The approach comprised the identification of mutations between two *B. multivorans* isolates from the same CF patient in which one has the ability to form multicellular aggregates (P0426-1) while the other is not able to (P0426-2), and their study to identify if any of these mutations could be involved in the loss of this ability by the second isolate.

2. Materials and Methods

2.1. Bacterial strains and growth conditions

Bacterial strains and plasmids used in this study are listed in table 1. *E. coli* was grown at 37ºC in Lennox Broth (LB) with or without agar, supplemented with Ampicillin (100 µg/ml) or chloramphenicol (25 µg/ml) when required to maintain selective pressure. *Burkholderia multivorans* clinical isolates from a cystic fibrosis patient, and their derivative strains were grown in LB, SCFM (synthetic cystic fibrosis medium) (Palmer *et al.* 2007) or in SM medium (12.5 g/l $Na_2HPO_4.2H_2O$, 3 g/l KH_2PO_4, 1 g/l K_2SO_4, 1 g/l NaCl, 0.2 g/l $MgSO_4.7H_2O$, 0.01 g/l $CaCl_2.2H_2O$, 0.001 g/l $FeSO_4.7H_2O$, 1 g/l yeast extract, 1 g/l casamino acids, pH 7.2), supplemented with 20 g/l of D-mannitol, at 37ºC with 180 rpm of orbital agitation.

Table 1: Strains and plasmids used in this study.

Bacterial strains or plasmids	Description	Reference or source
Bacterial strains – *Burkholderia multivorans*		
P0426-1	Cystic fibrosis isolate, Canada Isolation date: 11/02/1997	D.P. Speert University of British Columbia
P0426-2	Cystic fibrosis isolate, Canada Isolation date: 17/03/1997	D.P. Speert University of British Columbia
P0213-1	Cystic fibrosis isolate, Canada Isolation date: 13/02/1996	D.P. Speert University of British Columbia
P0148-1	Cystic fibrosis isolate, Canada Isolation date: 21/02/1989	D.P. Speert University of British Columbia
Bacterial strains – *Escherichia coli*		
DH5-α	DH5α recA1 Δ(lacZYA-argF)U169 ϕ80dlacZΔM15	Gibco BRL
BL21 (DE3)	F− ompT hsdSB (rB−mB−) dcm gal λ(DE3)	Stratagene
Plasmids		
pBBR1MCS	4,707-bp broad-host-range cloning vector, Cm^r	(Kovach *et al.* 1994)
pRK2013	Tra^+ Mob^+ (RK2) Km::Tn7 ColE1 origin, helper plasmid, Km^r	(Figurski *et al.* 1979)
pET-23a (+)	Cloning/expression vector, T7 promoter, C-terminal 6× His-Tag, Ap^r	Novagen
pCR20-1	pBBR1MCS derivative containing a fragment with the *ytnP* gene coding sequence from P0426-1	This work
pCR20-5	pBBR1MCS derivative containing a fragment with the *ytnP* gene coding sequence from P0426-2	This work

pCR21-1	pBBR1MCS derivative containing a fragment with the *ttgC* promotor region from P0426-1	This work
pCR21-2	pBBR1MCS derivative containing a fragment with the *ttgC* gene coding sequence from P0426-1	This work
pCR21-3	pBBR1MCS derivative containing a fragment with the *ttgC* gene coding sequence from P0426-2	This work
pCR21-4	pET-23a (+) derivative containing a fragment with the *ytnP* gene coding sequence from P0426-1	This work

Abbreviations: Cm^r, chloramphenicol resistance; Km^r, kanamycin resistance; Ap^r, ampicillin resistance

2.2. Genomic DNA extraction

A small mass of cells grown in solid medium was resuspended in 1 mL of solution K (10 mM Tris, pH= 7.8; 5 mM EDTA, pH = 8; SDS 0.5%; proteinase K at 50 µg/mL) and incubated for 3 hours at 40 °C. 600 µL of potassium acetate 3M, pH = 5.5 was then added and mixed. This preparation was centrifuged at 16000 xg for 10 minutes. The supernatant was collected onto a new tube and 0.7 volumes of isopropanol added. Another centrifugation was performed in the same conditions. The supernatant was then discarded with a syringe and the pellet washed with 200 µL of ethanol 70% (v/v). This was followed by another centrifugation at 16000 xg for 5 minutes. The supernatant was discarded with the syringe and the pellet was dried in the Speed Vacuum for 20 minutes at 45°C in V-AL mode. 50 µL of sterile water was added and the solution was left overnight at room temperature.

2.3. DNA manipulation techniques

DNA purification, DNA restriction, agarose gel electrophoresis, DNA amplification by PCR, and *E. coli* transformation were performed using standard procedures (Russel, 2001). NZYMiniprep kit from NZYTech® was used to perform the extraction of plasmid DNA. The protocol followed during the procedure was the one provided by this specific kit.

2.4. Cloning of *B. multivorans* genes for phenotype complementation

To perform complementation assays, four plasmids expressing different genes were constructed. The first two plasmids were constructed by cloning the *ytnP* gene belonging to both *B. multivorans* isolates P0426-1 and P0426-2. For that, the 1.3-Kb *ytnP* gene from each isolate was cloned into the pBBR1MCS broad-host-range vector (Fig. 10A). Each gene was amplified from the genomic DNA of *B. multivorans* P0426-1 and P0426-2 by PCR, using a set of forward and reverse primers that included engineered restriction endonuclease sites. The primers used for amplification of the *ytnP* gene were YtnP_Fw 5'- GGGTACCAGATCGAACGTGAACTGCTGGA - 3', and YtnP_Rev 5' – CTCTAGAGGCGACCTTTCACCAATTGTCA – 3'. Following PCR procedure, purification of the resulting DNA was performed using NZYGelpure kit from NZYTech®. After DNA purification both amplified fragments and pBBR1MCS vector were doubled digested with KpnI/XbaI followed by ligation.

E. coli electrocompetent cells were transformed by electroporation, using Bio-Rad Gene Pulser II system (400 Ω, 25 µF, 2.5 kV), grown for 1 hour before being plated onto LB plates supplemented with 25 µg/ml chloramphenicol. Plasmids were then recovered and digested with the restriction enzymes used before, followed by agarose gel electrophoresis to verify the presence of fragments corresponding to the gene. The clones from the two constructs were sequenced through Sanger sequencing to confirm whether the sequences were correct, originating vector pCR20-1 with *ytnP* gene from *B. multivorans* P0426-1 and pCR20-5 with *ytnP* gene from *B. multivorans* P0426-2.

The two other plasmids were constructed by cloning the promotor region and the *ttgC* gene belonging to both *B. multivorans* isolates P0426-1 and P0426-2. The 1.6-Kb *ttgC* gene and its own promotor region from each isolate were cloned into the pBBR1MCS broad-host-range vector. Like previously described the promotor and each of the genes were amplified by PCR, using a set of forward and reverse primers that included engineered restriction endonuclease sites. The primers used for amplification of the promotor region were PromTtgC_Fw 5'-GGGTACCTCGTGAAGCTGCTCGTGTCGA - 3', and PromTtgC_Rev 5' – CTCTAGAGGCAGTGCGGATTTCAGGA – 3'. The amplified fragments and pBBR1MCS vector were doubled digested with XbaI/KpnI followed by ligation. *E. coli* electrocompetent cells were transformed by electroporation, originating vector pCR21-1. The primers used for the amplification of the *ttgC* gene from each isolate were TtgC_Fw 5'- GGATATCGTGTTCTACGTCGTGCTGC - 3', and TtgC_Rev 5' – CTCTAGAACTTCATGTGCGCAGCGGC – 3'. The amplified fragment and pCR21-1 vector were doubled digested with EcoRV/XbaI followed by ligation. *E. coli* electrocompetent cells were transformed by electroporation, originating vector pCR21-2 with the *ttgC* gene from *B. multivorans* P0426-1 and pCR21-3 with the *ttgC* gene from *B. multivorans* P0426-2.

A last construct was performed to purify the YtnP protein from the P0426-1 isolate. For this construct the *ytnP* gene was amplified using the following primers pairs: pET_YtnP_Fw 5'-CGCTAGCAGCGCAACCATCCACAATGT-3', and pET_YtnP_Rev 5'–CCTCGAGGTAATCCCACGTCACGGGTA – 3'. FW primer has the restriction sequence for NheI, and REV primer has the restriction sequence for XhoI. The DNA fragment was cloned into pET-23a (+) (Fig. 10B). Ligation mix was transformed in competent *E. coli* DH5-α through electroporation resulting in plasmid pCR21-4. After confirmation, the plasmid was extracted using the NZYMiniprep kit from NZYTech® and transformed into electrocompetent *E. coli* BL21(DE3) through electroporation.

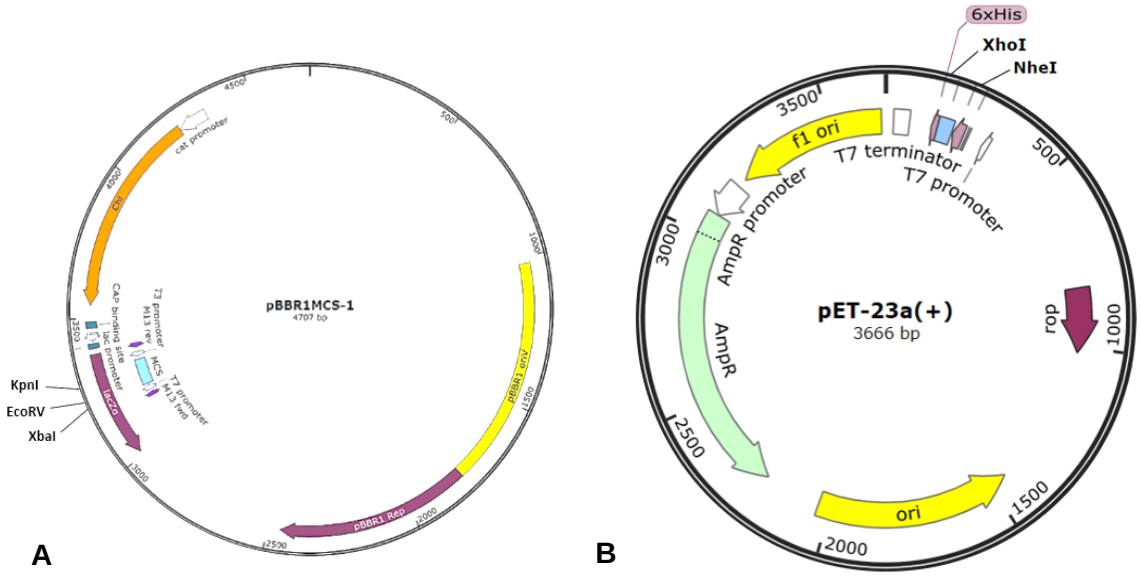

Figure 10: pBBR1MCS and pET-23a(+) vectors used on the different constructs of *B. multivorans* genes. pBBR1MCS (A) is the broad-host-range cloning vector used for phenotype complementation analysis and, pET-23a(+) (B) the vector used for his-tagged protein expression and purification.

2.5. Complementation assays

Complementation assays were performed using triparental conjugation. The receptor strains were *B. multivorans* P0426-1, P0426-2, P0213-1, and P0148-1; the helper strain contained plasmid pRK2013; and *E. coli* the donor strain would depend on the assay that was being performed. The selection occurred in LB medium supplemented with 250 mg/mL of chloramphenicol and 40 mg/mL of gentamicin.

2.6. RNA manipulation techniques

2.6.1. RNA extraction and purification

Isolates under study were inoculated in SM medium with an initial OD_{640nm} 0.1, and incubated at 37ºC with 180 rpm of orbital agitation. Samples were collected at 4, 8 and 24 hours and RNA was extracted with RNeasy Mini Kit by Qiagen, following an optimized procedure, based on manufacturer recommendations. At the time points chosen 1 volume of culture was taken and two volumes of RNAprotect® Bacteria Reagent were added, promoting mRNA stabilization. The mixture was centrifuged, and the pellet was resuspended in 200 µl of TE buffer (30 mM Tris-HCL, 1 mM EDTA, pH 8.0) containing 15 mg/ml of lysozyme and 20 µl of proteinase K (Qiagen). After 10 minutes of enzymatic lysis 700 µl of RLT buffer were added, the mixture was vortexed for 10s and 500 µl of 96% ethanol were added, mixed and transferred to a RNeasy Mini spin column. The columns containing the mixture were then centrifuged and washed with 700 µl of RW1. After being washed 500 µl of RPE were added to the column and these were centrifuged. The membrane of the columns was dried by centrifuging at full speed for 1 min, the RNeasy spin columns were placed in new 1.5 ml collection tubes and the RNA was

eluted with 40 µl of RNA-free water. To avoid contamination of genomic DNA, a step of DNA digestion was performed using Thermo Scientific DNase I, RNase-free, an endonuclease that digests single- and double-stranded DNA.

All steps described above were performed using RNase-free material. RNA concentration was estimated using a UV spectrophotometer (ND-1000 UV-Vis, NanoDrop Technologies, USA). RNA samples were stored immediately after extraction at -80ºC.

2.6.2. Real-Time qPCR

Real-Time qPCR was performed using a relative quantification method based on a one-step protocol. A dilution of the RNA samples was performed in order to obtain 12.5 ng of RNA in 50 µl, as well as 1:10 dilutions of the primers. For each well of the 96-well plate, 5.6 µl of DEPC-treated H2O, 1.6 µl of the primer solution and 2 µl of RNA were added. A mixture with One-step NZY qPCR Green master mix (2x), ROX and NZYRT mix from NZYTech® was prepared assuming that for each well there needs to be 10 µl One-step NZY qPCR Green master mix (2x), ROX of and 0.8 µl of NZYRT mix.

The relative expression levels of the target genes were normalized to those of the housekeeping gene *rpoD*, an RNA polymerase sigma factor, and the target genes from P0426-1 isolate were compared with the corresponding target genes from isolate P0426-2. The relative oligonucleotide primers for the genes under study (Table 2) were designed in Benchling website. The method used to treat the data was the ΔΔCT method (ΔΔCT = ΔCT(a target sample)−ΔCT(a reference sample) = (CTD − CTB)−(CTC − CTA)).

Table 2: Primers used to perform Real-Time quantitative PCR.

Primer	Sequence
GalU_FW	5’ ATGATCCTCGATCGCGCGCTTG’3
GalU_RV	5’ GTCGTCGACAAGCCGCTGATCC ’3
OatA_FW	5’ ACTGGATCGGCGTCGTGAAGGA ’3
OatA_RV	5’ CACGAGCGGCGCGATCAGATAG ’3
GltR_FW	5’ GCTCGCGATCTTCTCGGCTTCC ’3
GltR_RV	5’ TGAGCACTTTCCCGAACGCACC ’3
AlaS_FW	5’CGAGATGGAGTCCGATGCACGC’3
AlaS_RV	5’ GTAGACGGCCTTGCGCTCGATC ’3

2.7. Growth curves

Cells from an overnight culture were inoculated in 50 mL of fresh SM medium for a OD640nm of 0.1. The culture was incubated at 37ºC with 180 rpm of orbital agitation. OD640nm readings were taken over time for 7 hours and then at 24 and 48 hours. Growth rates were calculated from the exponential phase of growth.

2.8. Quantification of cellular aggregates and free cells

The quantification of cellular aggregates and free cells was made based on the protocol previously described (Haaber *et al.* 2012), with some adjustments. Bacterial cultures grown in SM medium for 48 hours at 180 rpm of orbital agitation were transferred to a 50 mL Falcon tube and centrifuged at 200 xg for 30 seconds at 25ºC. After centrifugation, cell suspensions rested for 10 minutes, in order to settle down the aggregates still in suspension. The supernatant was removed by pipetting and placed in a new 50 mL Falcon tube. Suspensions containing aggregates were transferred to 2 mL tube with snap cap and then, after several quick-spins, a 2 mL tube was obtained with all cellular aggregates of the same mutant strain. The 50 mL Falcon tube with free cells and growth medium was centrifuged at 1700 xg for 15 minutes at 25ºC. The resulting pellet was resuspended in 5 mL of growth medium and then, several centrifugations at 16000 xg of orbital agitation for 2 minutes were performed in 2 mL Eppendorf tubes, in order to collect all free cells of the isolate.

In the end of the procedure, two 2 mL Eppendorf tubes were obtained for each strain, one with all the free cells and one with the aggregates. These tubes were placed at 60ºC during at least 72 hours until all cellular aggregates and free cells were dried, presenting a brown color. All 2 mL Eppendorf tubes were weighted before the collection of samples and after samples were dried.

2.9. Microscopy analysis

B. multivorans strains grown in SM or SCFM medium for 48 hours were visualized on Zeiss Axioplan microscope, equipped with a Axiocam 503 color Zeiss camera, using a 10x 0.3 NA objective, and controlled with the Zen software.

2.10. Protein manipulation techniques

Protein manipulation techniques were mainly used to overexpress and purify the YtnP protein fused with a His_6-tag to allow easy purification by affinity chromatography. In order to overexpress the protein, the gene coding sequence was previously cloned into pET-23a(+) as described in Section 2.4 and the best conditions were optimized.

2.10.1. Protein overexpression

The pET-23a(+) derived vector pCR21-4 was inserted into *E. coli* BL21 (DE3) cells for overexpression of the His_6-tagged protein. After transformation the cells were inoculated in 50 ml of LB liquid medium containing ampicillin (100 µg/ml) at an initial OD_{640nm} of 0.1 and incubated at 37˚C with 250 rpm of orbital agitation. When the cultures reached an OD_{640nm} of 0.7, protein overexpression was

induced by the addition of 0.4 mM of IPTG to the growth medium and cultures were incubated for 4 hours, under the same conditions. The cells were harvested by centrifugation at 7000 xg for 5 minutes at 4˚C. The supernatant was discarded, and the pellet was stored at -80˚C until further use.

2.10.2. Purification of His_6-tagged proteins by affinity chromatography

The recombinant His_6-tagged protein overexpressed was purified using a HisTrap FF columns (Amersham Biosciences). These columns retain histidine tagged fusion proteins when charged with Ni^{2+} ions. The protein can be desorbed from the column with buffers containing imidazole, which preserve the antigenicity and functionality of the proteins.

The overexpressed cell extracts were defreezed on ice and then resuspended in 6 ml of Sonication buffer. The cells were disrupted by ultrasonic vibration (six cycles for 30 seconds each, 70 Watts) with a Branson sonifier 250 (Branson) and the envelopes and unbroken bacteria were removed by centrifugation at 12000xg for 30 minutes at 4ºC. The clear supernatant was carefully transferred to a clean Falcon.

The HisTrap FF column was washed with 5 ml distilled water and then regenerated with 10 ml of Solution C. After being regenerated, the column was charged with 0.5 ml of 0.1 M nickel sulphate salt solution and then washed again with 5 ml of distilled water. To purify the protein the column was balanced with 10 ml Buffer I (Table 3) and then the sample was fully applied. A constant increasing gradient of imidazole concentration was used to elute the protein, with imidazole concentrations varying form 10 mM to 500 mM. One by one, 5 ml of each buffer were applied to the column and 1ml aliquots were recovered during the elution. To later determine the fractions containing the protein of interest, 20 µl of each collected aliquot were analysed by SDS-PAGE (Section 2.10.3).

Phosphate buffer 8x – 80 mM $Na_2HPO_4.2H_2O$; 80 mM $NaH_2PO_4.H_2O$; 4 mM NaCl (pH 7.4)

Sonication Buffer – Phosphate Buffer 1x (pH 7.4); 10 mM Imidazole

Solution C – 20 mM $Na_2HPO_4.2H_2O$; 0.5 mM NaCl; 0.05 M EDTA (pH 7.4)

Table 3: Elution buffers used during the process of purification of His-tagged proteins.

Buffer	[Imidazole]	Phosphate buffer 8x pH 7.4	Imidazole 2 mM pH 7.4	Distilled H_20
I	10 mM	3 ml	0.12 ml	Until 24 ml
II	20 mM	1 ml	0.08 ml	Until 8 ml
III	40 mM	1 ml	0.16 ml	Until 8 ml

IV	60 mM	1 ml	0.24 ml	Until 8 ml
V	100 mM	1 ml	0.40 ml	Until 8 ml
VI	300 mM	1 ml	1.20 ml	Until 8 ml
VII	500 mM	1 ml	2.00 ml	Until 8 ml

2.10.3. Denaturing polyacrylamide gel electrophoresis (SDS-PAGE)

The discontinuous gel system was described by Laemmli (Laemmli, 1970) and is used to separate proteins on polyacrylamide gels. This system is composed of two different gels: the running gel (15% polyacrylamide) that allows the proteins to separate according to their relative molecular mass, and the stacking gel (4% polyacrylamide), where the proteins are applied and concentrated. The composition of each gel is detailed in Table 4. The *N*,*N*,*N'*,*N'*–tetramethylethylenediamine (TEMED) and the ammonium persulphate (APS) solution are added just before pouring the gel to prevent premature polymerization of the acrylamide.

Table 4: Composition of denaturing polyacrylamide gel.

Stock solution	Running gel (15%)	Stacking gel (4%)
Separating buffer	1.875 ml	-
Stacking buffer	-	375 µl
H2O	505 µl	1.305 ml
SDS 10%	50 µl	20 µl
Acrylamide	2.5 ml	270 µl
APS 10%	50 µl	30 µl
TEMED	2.5 µl	2 µl

After polymerization, the gel is immersed in running buffer 1X. The protein samples previously resuspended in gel loading buffer and incubated for 10 minutes at 100 ºC were applied (20 µl) in each well of the gel. A molecular mass standard PageRuler™ Plus prestained protein ladder (Thermo Scientific) was used as reference. Separation of proteins is achieved by applying 150 V, until the bromophenol blue contained in the loading buffer reached the end of the gel. The SDS-PAGE gels were stained by emersion in BlueSafe (NZYTech) for 20 minutes. During incubation, the gel was gently agitated to distribute the dye uniformly over the gel. After staining, the gel was washed with distilled water.

Separating buffer – 1.5 M Tris base, 0,4% (v/v) SDS. pH is adjusted to 8.8 with HCl.

Stacking buffer – 0.5 M Tris base, 0.4% (v/v) SDS. pH is adjusted to 6.8 with HCl.

Gel loading buffer – 100 mM Tris base (pH 6.8), 4% (wt/v) SDS, 20% (v/v) glycerol, 0.2% (wt/v) bromophenol blue, 200 mM DTT.

Running buffer 10X – 0.25 M Tris base, 1.92 M Glycin, 1% (wt/v) SDS.

2.11. Sequence alignment and phylogenetic analysis

The amino acid sequences of the proteins being studied were obtained by BLASTP search at the National Center for Biotechnology (NCBI) and the *Burkholderia* genome database. Protein alignments were performed using CLUSTALW and the phylogenetic trees were built using the neighbor-joining method both in MEGA-X v.10.5.0.

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 8.0.1 for Windows (GraphPad Software, San Diego California USA, www.graphpad.com). Differences were considered statistically significant for P-values lower than 0.05.

3. Results and Discussion

3.1. Cellular aggregation in *B. multivorans* isolates P0426-1 and P0426-2

A longitudinal series of 21 isolates recovered from the cystic fibrosis patient P0426 between 1997 and 2004 and available from the Canadian *Burkholderia cepacia* complex research and referral repository (CBCCRRR) was available for this study. Out of these isolates, the first one has the ability to form large multicellular aggregates while the remaining ones were incapable of forming these aggregates or formed very small ones (Gomes, 2018). The first isolate P0426-1 and the second, P0426-2, are sequential isolates recovered a month apart of each other, being the first able to form large aggregates whilst the second could not form any. To quantify the planktonic cellular aggregates formed by P0426-1 and P0426-2, both isolates were grown in SM medium, which has high carbon to nitrogen ratio and in SCFM medium that mimetizes nutrients available in the CF lung environment.

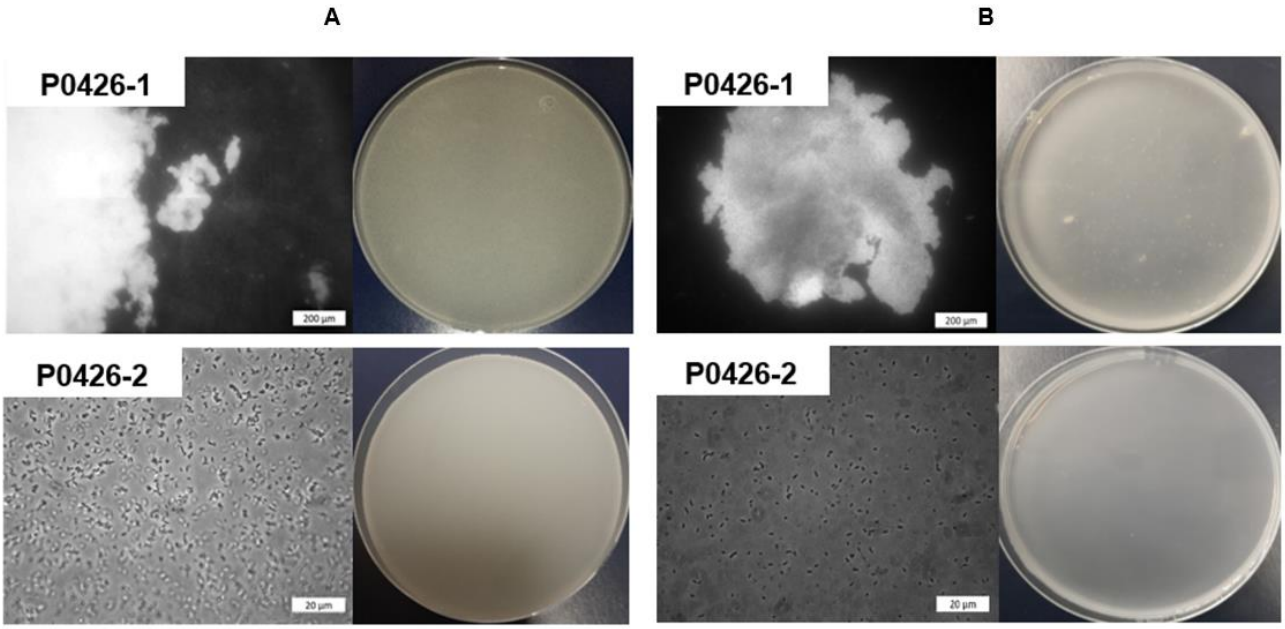

Figure 11: Screening for aggregate formation by *B. multivorans* clinical isolate P0426-1 and P0426-2 in both SM (A) and SCFM medium (B). For each strain is shown a microscopy image (on the left) and the liquid culture on a petri dish (on the right). Each culture was grown at 37ºC, 180 rpm of orbital agitation, for 48 hours.

As can be seen in Figure 11A, *B. multivorans* P0426-1 grows both as free cells and planktonic aggregates in SM medium, while from *B. multivorans* P0426-2 no cellular aggregates were seen. Since the sputum layer of the CF lung is a complex substrate, a synthetic cystic fibrosis medium (SCFM) was used to simulate the composition of CF sputum. Tests performed on other bacteria, such as *Pseudomonas aeruginosa,* revealed similar phenotypes during growth in sputum from patients suffering from CF and in SCFM (Palmer *et al.* 2007), validating this medium. Although using SCFM might provide a better depiction of the formation of multicellular aggregates by isolate P0426-1 in CF sputum, the results were very similar to the ones obtained in SM medium (Fig. 11B).

To obtain a better perception of the differences in aggregate formation ability by both strains, the planktonic cellular aggregates were quantified by weighting the dry biomass of both planktonic cells and aggregates separately.

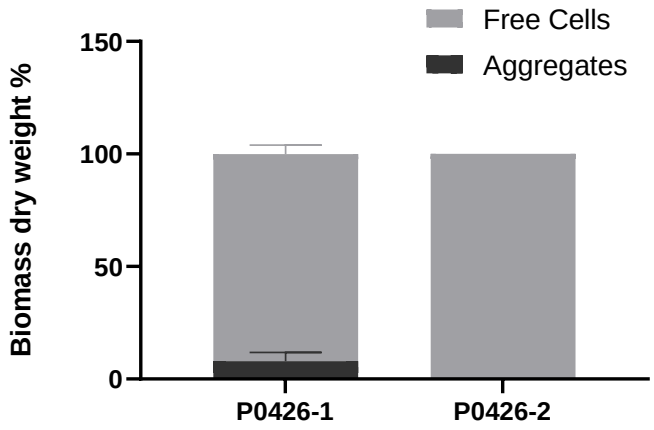

Figure 12: Quantification of cellular aggregates and free cells of isolates P0426-1 and P0426-2 grown in SM medium at 37˚C, 180 rpm for 48h. Error bars correspond to the standard deviations of the mean values of three independent experiments.

As it can be seen in Figure 12, isolate P0426-1 produces around 8% of the biomass as planktonic aggregates, being the majority of the biomass produced as free cells. From the isolate P0426-2 we were unable to obtain aggregates, having 100% of the biomass in the form of free cells.

This result confirms previous observations on the ability of the first isolate P0426-1 to form multicellular aggregates while isolate P0426-2 seem to grow only as free cells.

3.1.1. *In vitro* growth analysis of *B. multivorans* P0426-1 and P0426-2

It is important to determine if the inability to form planktonic cellular aggregates by isolate P0426-2 is caused by an impairment of growth or not. With that purpose, isolate P0426-1 and P0426-2 were grown in SM medium for 48 hours with 180 rpm of orbital agitation at 37ºC.

Growth curves shown in Figure 13 confirm a very similar growth behaviour, both in exponential and stationary phase. According to data from figures 12 and 13, it was concluded that, under the tested conditions, the two strains display differences in the formation of planktonic cellular aggregates without interfering with the growth behaviour of the bacteria.

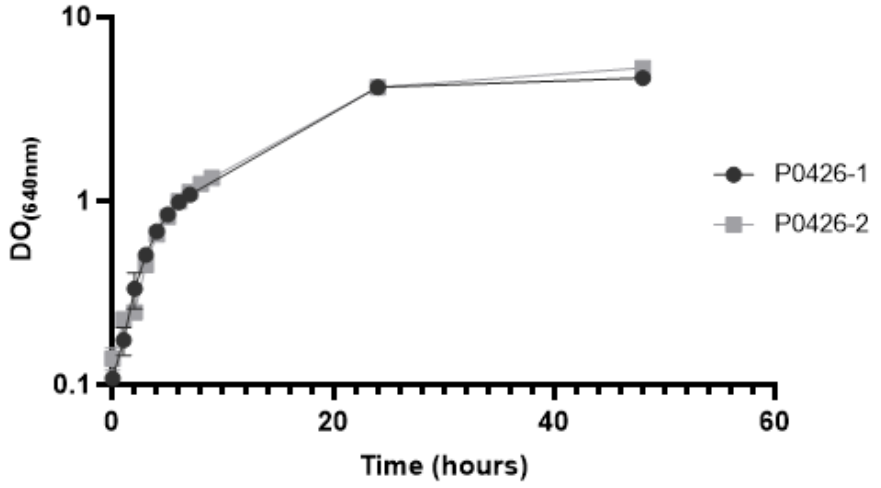

Figure 13: Growth curve of *B. multivorans* P0426-1 and P0426-2 at 37ºC for 48 hours in SM medium. Error bars indicate the standard deviation.

3.1.2. Genomic alterations between *B. multivorans* isolates P0426-1 and P0426-2

To understand this inability to form planktonic aggregates by *B. multivorans* P0426-2 and since the sequence of the genome of both isolates was available, the differences between them were analysed. In comparison to isolate P0426-1, isolate P0426-2 has four single nucleotide polymorphisms (SNP) mutations in the coding region of: a gene encoding a putative quorum-quenching lactonase (YtnP), a purine efflux pump (PbuE), a multidrug resistance protein (Stp), and an outer membrane protein (TtgC). This caused nonsynonymous mutations and except for protein Stp they are present in all isolates of the longitudinal series (Table 5). Other differences between the two isolates are two small indel mutations in intergenic regions, being the genes that could be affected by the 7 base-pairs insertion the *bceT* gene encoding the UTP-glucose-1-phosphate uridylyltransferase (BceT) and/or the acyltransferase (BceU), both involved in cepacian biosynthesis. The 10 nucleotides deletion might affect a LysR-type transcriptional regulator (GltR) and/or an alanyl-tRNA synthetase (AlaS). A large deletion

of approximately 44 kb is also present in P0426-2 as well as 6 other isolates, and includes mainly genes encoding hypothetical proteins.

Table 5: Mutations identified between *B. multivorans* P0426-1 and P0426-2. (Coding region – CDS; Intergenic region – IG; Single-nucleotide substitution – SNP; Synonymous mutation – Syn; Non-synonymous – Nonsyn).

Annotation	Location	Locus tag	Amino acid	Effect in protein	Type of mutation	P426-1	P426-2	Nº of isolates with mutation
Putative quorum-quenching lactonase (YtnP)	CDS	FEP59_00467	S>R	Nonsyn	SNP	T	G	20
Purine efflux pump (PbuE)	CDS	FEP59_00501	L>M	Nonsyn	SNP	G	T	20
Multidrug resistance protein (Stp)	CDS	FEP59_00575	Q>P	Nonsyn	SNP	A	C	1
Toluene efflux pump outer membrane protein (TtgC)	CDS	FEP59_01688	T>A	Nonsyn	SNP	T	C	20
Between FEP59_02506 and FEP59_02507	IG	-	-	-	Indel	-	+7cgaacgc	10
Between FEP59_04039 and FEP59_04040	IG	-	-	-	Indel	-	-10cacgcaggct	20
Several hypothetical proteins	-	-	-	-	indel	-	del 44 kb	7

Lactonases are known quorum quenching enzymes that can affect biofilm formation by degrading AHLs, which disrupts bacterial quorum sensing-based communication. Taking into consideration that many microbial species regulate a number of processes through QS, the disruption of this mode of communication is considered of a great potential to a variety of different fields such as medical or agricultural (Chan *et al.* 2010).

The role of efflux pumps in biofilm formation has been described previously for different bacteria, including *Escherichia coli*, *Salmonella enterica and Klebsiella pneumoniae*. TtgC is the outer membrane channel of the TtgABC efflux system, which is a member of the RND family of efflux systems. These systems are believed to be present in all major kingdoms and play an important role in antimicrobial resistance, as well as resistance to heavy metals (Tseng *et al.* 1999). This specific efflux system has been proven to have an effect in biofilm formation in *Pseudomonas putida* by Tettmann *et al*., which showed that when the *ttgABC* genes were overexpressed there was an increase in biofilm formation (Tettmann *et al.* 2014).

PbuE is a putative purine efflux pump involved in the efflux of purine ribonucleosides, being its role to control the intracellular concentration of purine bases. This protein has been proven to modulate the expression of two regulons, PurR and G-box. When studying a mutant that overexpresses PbuE, it could be seen that there was also an increase in the expression of PurR (Sause *et al.* 2019, Goncheva *et al.* 2019). As this HTH-type transcriptional repressor (PurR) is proven to be involved in biofilm formation in *Staphylococcus aureus*, a mutation on the PbuE protein could possibly have an effect in multicellular aggregate formation in our bacterial system.

The Stp protein is a sugar transport protein part of the MST (Monosaccharide Transporter) superfamily which belongs to the major facilitator superfamily (MFS) of efflux pumps. The MFS superfamily is ubiquitous in living organisms and is considered the largest group of transporters. Various studies have been successful in proving the involvement of this efflux pump in the communication of certain bacteria with the host cells, mainly in stages such as adhesion, invasion and possibly biofilm formation (Pasqua *et al.* 2021).

The two indel mutation in intergenic region could be affecting the expression of the genes nearby. Since UTP-glucose-1-phosphate uridylyltransferase (BceT) enzyme is involved in the synthesis of the activated sugar-nucleotide UDP-D-Glucose, which is then the precursor for UDP-D-glucuronic acid and UDP-D-galactose, changes in the levels of these sugar nucleotides could affect the amount of cepacian being produced. The acetyltransferase BceU is involved in the decoration of the cepacian repeat-unit and could contribute to differences at the acetylation level of the final polymer (Ferreira *et al.* 2011). Both changes could have an effect on biofilm formation as documented in several publications (Ferreira *et al.* 2007; Ferreira *et al.* 2013). Some LysR-type transcriptional regulators have been implicated in quorum sensing regulation and to have an effect in biofilm formation (Lee and Zhang, 2014). Therefore, *gltR* might be important for biofilm formation. No involvement of alanyl-tRNA synthetase in biofilm formation has yet been reported.

The large deletion that occurred in isolate P0426-2 eliminates 62 genes encoding hypothetical proteins, but since some isolates having this region are also unable to produce aggregates, this mutation will not be considered in further studies.

From the mutation analysis, genes mutated in all isolates are our primary targets for being implicated in aggregate formation. In particular, we will investigate genes *ytnP* and *ttgC*, as described in the next sections.

3.2. Analysis of the putative quorum quenching lactonase (YtnP) from isolates P0426-1 and P0426-2

3.2.1. Comparative sequence analysis and phylogenetic distribution of the YtnP protein

When the sequences of the YtnP protein homologs from different species were aligned, it is possible to see that in all, except for isolate P0426-1, the amino acid at position 85 is an arginine, while in P0426-1 is replaced by a serine (Fig. 14). The fact that this amino acid is conserved in all the other species opens the possibility that this might be a conserved area which may eventually have relevance in functional terms.

The putative quorum quenching lactonase (YtnP) where the mutation occurred belongs to the metallo-β-lactamase family. The enzymes from this family depend on a dinuclear zinc site in order to be able to catalyse the reaction of hydrolyzation of N-acyl-l-homoserine lactones (AHLs). The most well-known is the autoinducer inactivator A (AiiA) from *Bacillus thuringiensis,* but others have also been

proven capable of degrading AHLs and affecting biofilm formation, such as the two *B. cepacia* lactonases, YtnP and Y2-aiiA (Malešević *et al.* 2020).

Search of the *ytnP* gene in the *Burkholderia* genus (strains with sequenced genomes), showed that it is not present in many of the strains from different species. Some of the strains lacking the *ytnP* gene are *B. multivorans* ATCC 17616, *B. stabilis* FERMP-21014 or *B. vietnamiensis* G4.

	70aa															85 aa												97aa
B. multivorans P0426-1	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	S	S	G	E	R	L	I	L	I	D	S	G	V
B. multivorans P0426-2	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	G	E	R	L	I	L	I	D	S	G	V
B. multivorans R-20526	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	G	E	R	L	I	L	I	D	S	G	V
B. cepacia LO6	L	Q	R	D	M	F	D	W	A	L	N	V	A	L	V	R	S	G	D	R	L	I	L	I	D	S	G	V
B. dolosa AU0158	L	Q	R	D	M	F	D	W	A	L	N	V	A	L	V	R	S	G	D	R	L	I	L	I	D	S	G	V
B. territorii MSMB2203WGS	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	A	D	R	L	I	L	I	D	S	G	V
B. lata 383	L	Q	R	D	M	F	D	W	A	L	N	V	A	L	V	R	S	G	D	R	L	I	L	I	D	S	G	V
B. ambifaria AMMD	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	G	E	R	L	I	L	I	D	S	G	V
B. contaminans MS14	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	G	E	R	L	I	L	I	D	S	G	V
B. diffusa RF2-non-BP9	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	A	D	R	L	I	L	I	D	S	G	V
B. anthina AZ-4-2-10-S1-D7	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	G	D	R	L	I	L	I	D	S	G	V
B. cenocepacia J2315	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	A	D	R	L	I	L	I	D	S	G	V
B. seminalis FL-5-4-10-S1-D7	L	Q	R	D	R	F	D	W	A	L	N	I	A	L	V	R	S	A	D	R	L	I	L	I	D	S	G	V
B. ubonensis MSMB1600WGS	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	G	E	Q	L	I	L	I	D	S	G	V
Paraburkholderia phytofirmans	L	Q	R	D	M	F	D	W	A	L	N	V	A	L	V	R	S	G	E	R	L	I	L	I	D	S	G	V
Cupriavidus sp. YR651	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	G	E	R	L	I	L	I	D	S	G	V
Caballeronia glathei	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	G	E	R	L	I	L	I	D	S	G	V
Ralstonia sp. 25mfcol4.1	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	G	D	R	L	I	L	I	D	S	G	V
Bordetella petrii ATCC BAA-461	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	G	D	H	L	I	L	V	D	S	G	V
Pseudomonas acidophila	L	Q	R	D	T	F	D	W	A	L	N	I	A	L	I	R	S	G	E	R	L	I	L	V	D	S	G	V
Rhodanobacter sp. T12-5	L	Q	R	D	M	F	D	W	A	L	N	I	A	L	V	R	S	G	D	R	L	I	L	I	D	S	G	V
Sphingomonas haloaromaticamans	L	G	P	D	M	F	D	W	A	L	N	V	L	V	V	R	S	G	E	Q	I	I	L	V	D	A	G	L
Sphingobium baderi	L	G	P	D	M	F	D	W	A	L	N	V	L	V	V	R	S	G	E	Q	V	I	L	V	D	A	G	L
Pseudomonas citronellolis	L	G	P	E	A	F	D	W	A	L	N	V	L	V	V	R	S	G	E	Q	T	I	L	V	D	A	G	L
Parvibaculum lavamentivorans	L	G	P	D	M	F	D	W	A	L	N	V	L	V	V	R	S	G	D	K	T	I	L	I	D	A	G	L
Luteibacter sp. OK325	L	P	-	E	S	F	D	W	A	V	N	V	V	V	V	R	S	G	E	Q	I	I	L	I	D	S	G	L
Massilia sp. Root1485	L	P	P	D	K	F	D	W	P	L	N	V	L	V	V	R	S	G	D	Q	T	I	L	V	D	A	G	V
Achromobacter sp. Root565	M	P	P	D	A	F	D	W	P	L	N	V	M	V	A	R	S	G	D	Q	T	I	L	I	D	A	G	L
Variovorax paradoxus	M	P	P	D	A	F	D	W	P	L	N	V	M	V	A	R	S	G	E	Q	T	I	L	I	D	A	G	L
Pandoraea fibrosis	M	P	P	D	A	F	D	W	P	L	N	V	M	V	V	R	S	G	S	Q	T	I	L	I	D	A	G	L
Ensifer sp. Root278	L	P	Q	D	A	F	D	W	S	L	N	V	V	V	V	R	S	G	A	Q	T	I	L	I	D	A	G	L
Polyangium fumosum	L	P	-	D	K	F	D	W	A	L	N	V	I	V	V	R	S	G	G	R	T	I	L	V	D	S	G	L
Phenylobacterium sp.	L	G	P	D	A	F	D	W	P	L	N	V	L	V	V	R	S	G	E	Q	I	I	L	V	D	A	G	L
Bradyrhizobium icense	L	P	P	D	A	Y	D	W	A	L	N	V	V	V	V	R	S	G	D	R	T	I	L	V	D	A	G	L

Figure 14: Amino acid sequence alignment generated by ClustalX of a representative sequence of YtnP protein present in different species.

Protein YtnP is 310 amino acids long and does not show a leader peptide, being most likely a cytoplasmatic protein. Alignment of YtnP-85S and YtnP-85R to biochemically characterized lactonases confirms the presence of the conserved metallo-β-lactamase dinuclear zinc-binding domain HxHxDM required for AHL degradation (Fig. 15). By this alignment it is possible to observe that the amino acid at position 85 is not in a conserved region. Nevertheless, all amino acids at this position are charged (D, R, Q) except the serine found in YtnP-85S of isolate P0426-1.

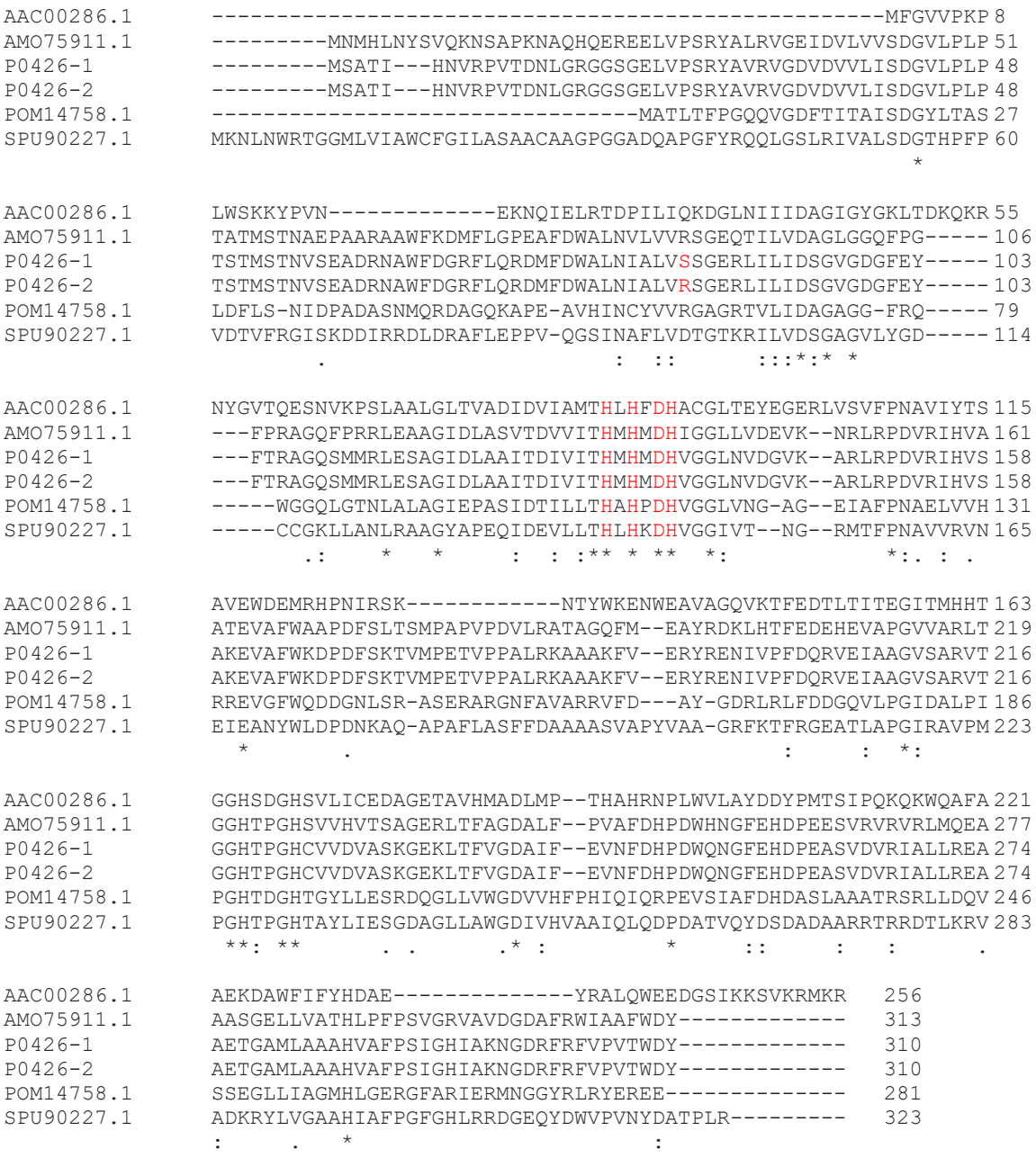

Figure 15: Sequence alignment presenting the relationship between *B. multivorans* P0426-1 and P0426-2 lactonases YtnP analysed in this study and other experimentally validated AHL lactonases. Multiple sequence alignment was performed using Clustal omega algorithm by the neighbor-joining (NJ) method. Letters in red represent the highly conserved amino acids of the metallo-β-lactamase superfamily-HXHXDH dinuclear zinc-binding motif in active site with aspartate and histidine residues required for AHL-degrading activity. AACOO286.1, YtnP of *Bacillus subtilis*; AMO75911.1, YtnP of *Pseudomonas citronellolis*; POM14758.1, YtnP of *Burkholderia cepacia*; SPU90227.1, Y2-aiiA of *Burkholderia cepacia*.

To better analyse the similarities among these sequences and their distribution, the neighbour joining method was used to construct a phylogenetic tree based on the ClustalX alignment (Fig. 16). The YtnP protein of *B. multivorans* seem to form an independent group within the majority of

Burkholderia species. Nevertheless, it is interesting to notice the wide distribution of this lactonase enzyme in the Proteobacteria.

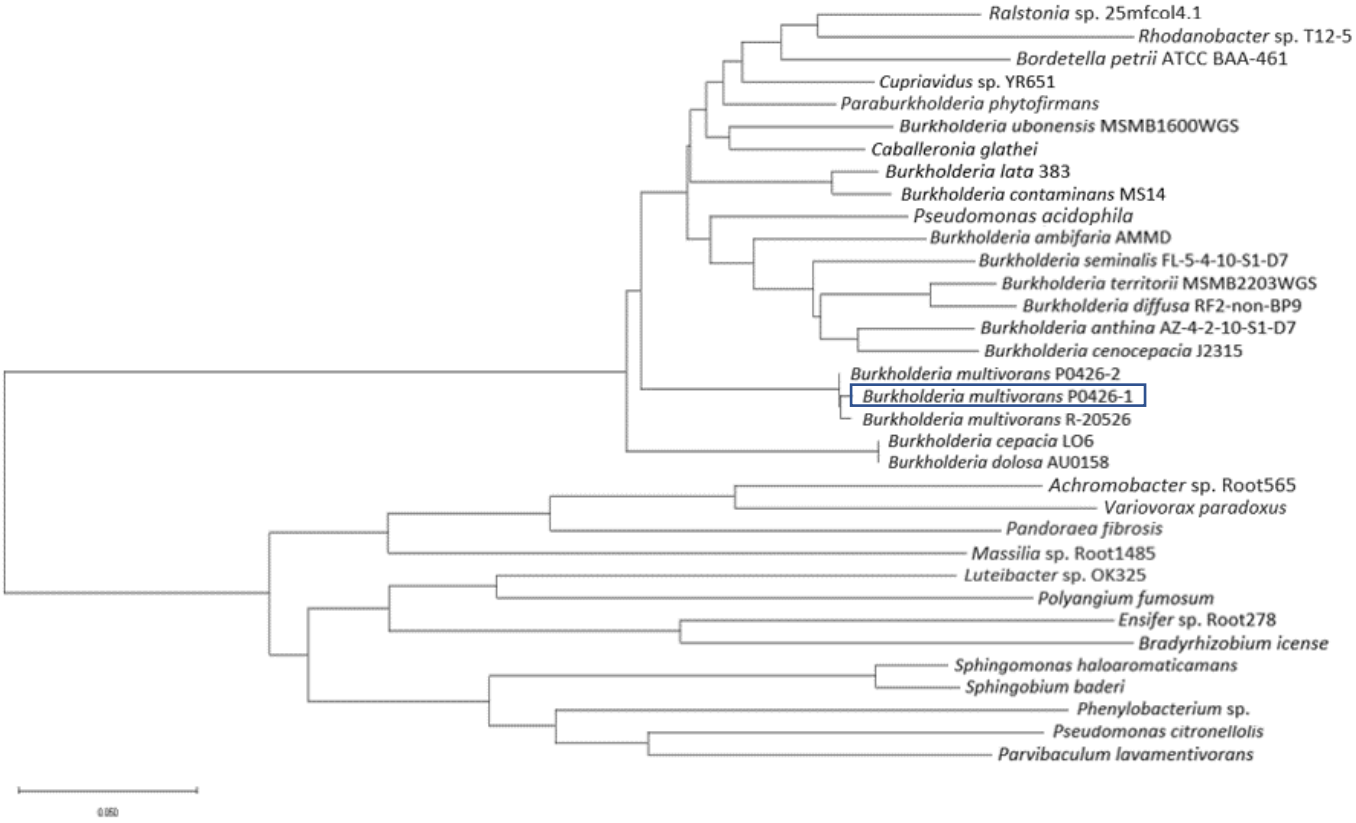

Figure 16: Phylogeny of YtnP protein. Neighbor-joining phylogenetic tree of the deduced amino acid sequences of YtnP predicted proteins from 12 *Burkholderia* species and other species from different genera. Scale: number of substitutions per site.

3.2.2. Genetic complementation with the *ytnP* gene

As the YtnP protein of isolate P0426-2 has an arginine residue at position 85 and that is conserved in all its homologues of Fig. 14 we hypothesized this would be the wild-type protein. Then, YtnP of isolate P0426-1 would perhaps be affected in its lactonase activity. In the absence or reduction of this activity the intracellular levels of AHLs in P0426-1 would be higher, leading perhaps to aggregate formation of a larger size due to the expression of genes implicated in this trait. Contrastly, in isolate P0426-2, the levels intracellular of AHLs would be lower and no multicellular aggregates would be formed. To test our hypothesis, the 1.3-Kb *ytnP* gene of isolate P0426-1 and P0426-2 was cloned into the pBBR1MCS broad-host-range vector, leading to the formation of pCR20-1 and pCR20-5 expressing the *ytnP* gene of P0426-1 and P0426-2, respectively. A complementation assay was performed in which the plasmid pCR20-5 and the empty vector were introduced into isolate P0426-1 and pCR20-1 and the empty vector into P0426- 2. Cells were grown in SM medium at 37ºC with 180 rpm of orbital agitation for 48 hours followed by inspection of aggregates formation.

The microscopic and macroscopic appearance of the planktonic cellular aggregates can be seen in Figure 17. When compared with the parental isolate P0426-1, the introduction of the empty pBBR1MCS vector stimulated aggregates formation of larger size (compare Fig. 17A vs Fig. 17B). The expression of YtnP-85R in P0426-1 did not prevent aggregate formation, but the structure of the aggregates seems to be more compact and with more ramifications (compare Fig.17B vs Fig. 17C). Regarding isolate P0426-2, the introduction of the empty vector or expressing Ytnp-85S, as expected, did not have any impact since cells remained as free-cells (Fig. 17D-F).

To determine whether there were quantitative differences in aggregates formation by the complementation of the isolates, the percentage of biomass dry weight recovered from both aggregates and free cells is shown in Figure 18.

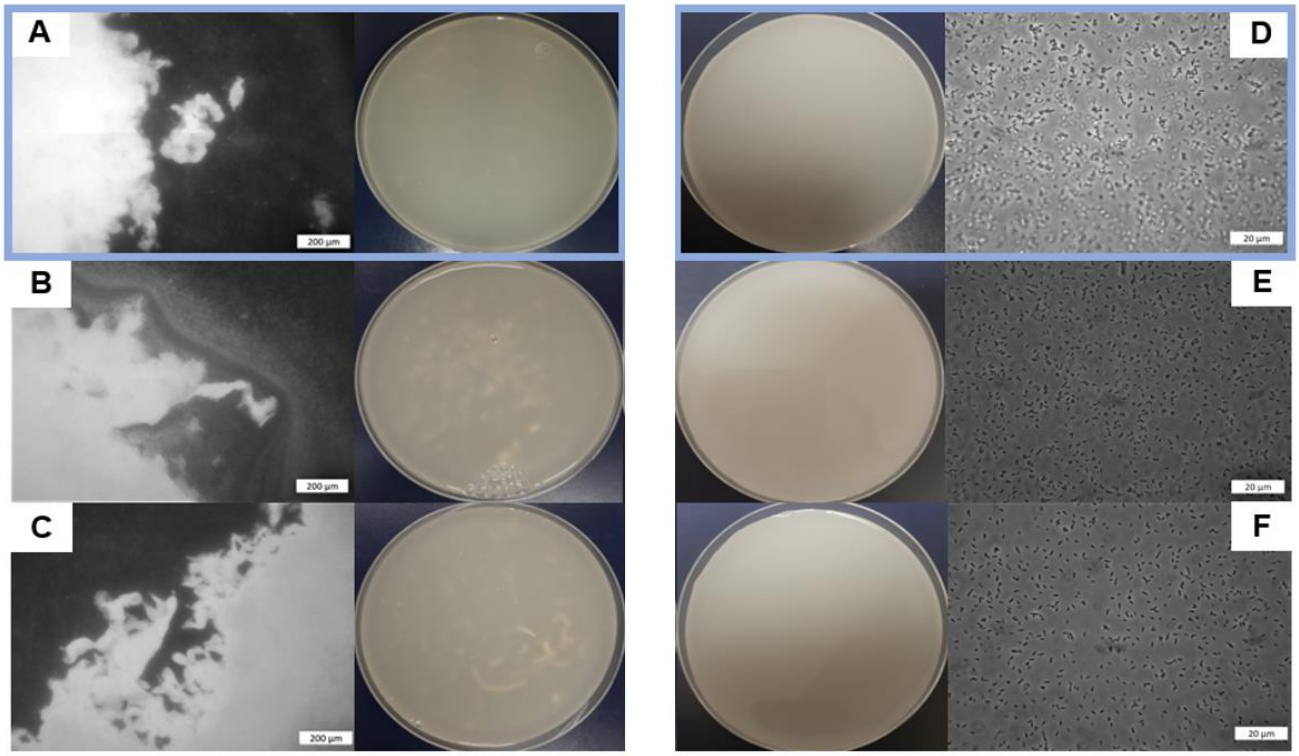

Figure 17: Screening for aggregate formation by P0426-1, P0426-2 expressing a different lactonase. A- P0426-1; **B-** P0426-1/pBBR1MCS; **C-** P0426-1/pCR20-5 (YtnP-85R); **D-** P0426-2; **E-** P0426-2/pBBR1MCS; **F-** P0426-2/pCR20-1 (YtnP-85S).

Confirming the macroscopic observation, the parental isolate P0426-1 produces 8% of its biomass in the form of cellular aggregates, while with the empty vector this value is considerably higher (53%). The comparison of the aggregates´ biomass between P0426-1/ pBBR1MSC and P0426-1-YtnP-85R showed a statistically significant reduction (53% to 22%) for the isolate expressing YtnP-85R (Fig. 18A). As expected, no aggregates were obtained from isolate P0426-2 with or without the expression of YtnP-85S (Fig. 18B). This result suggests that YtnP-85R expression in P0426-1 might have affected both the structure of the aggregates and their relative amount.

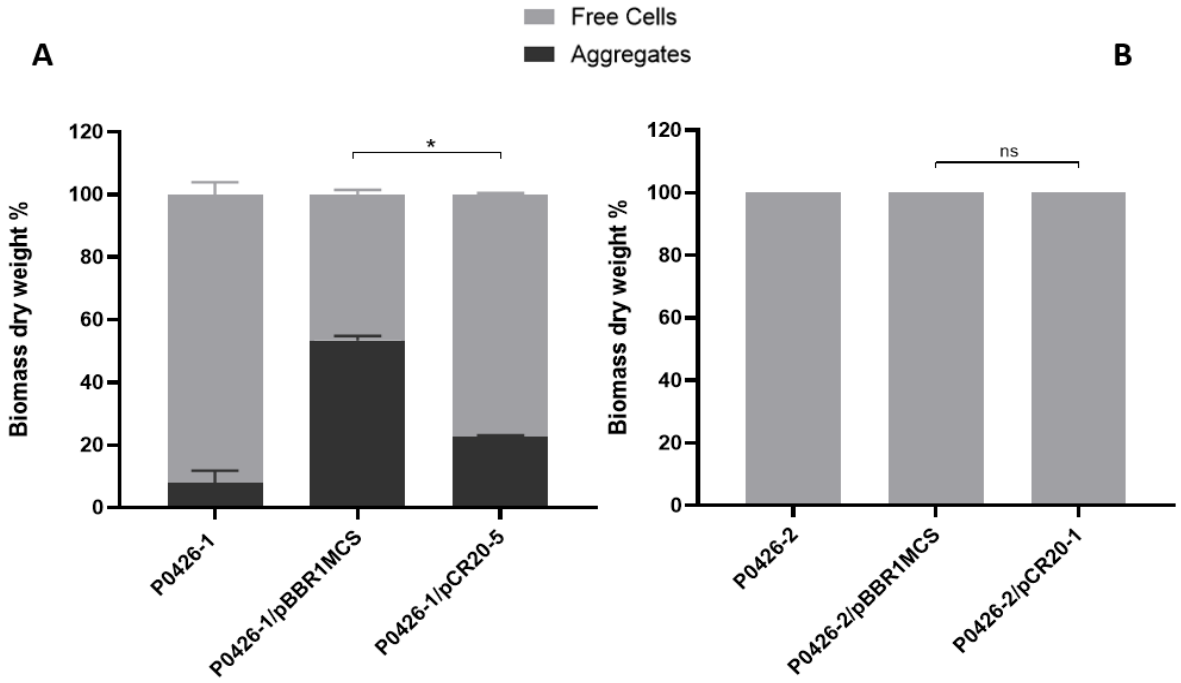

Figure 18: Quantification of cellular aggregates and free cells. A- P0426-1 parental and complemented with the empty vector or expressing YtnP-85R; **B-** P0426-2 parental and complemented with the empty vector or expressing YtnP-85S. Error bars correspond to the standard deviation. Significance level (one-way ANOVA followed by Dunnett's multiple comparison test) between cellular aggregates of P0426-1 and P0426-2 with the empty vector and cellular aggregates of the complemented isolates expressing the YtnP-85R or YtnP-85S gene variants: *, P <0.05; ns, not statistically significant.

To determine whether the lactonase enzymes under study would have impact in aggregates formation in other *Burkholderia*, we have chosen isolate P0213-1 and P0148-1 recovered from two different CF patients. Isolates P0213-1 and P148-1, which lack this gene in their genome, were grown in SM medium at 37ºC with 180 rpm of orbital agitation for 48 hours. Aggregates were visualized macroscopically in the broth and at the microscope (Fig. 19 A and E, respectively). Then, the empty vector, as well as the plasmids expressing YtnP-85S or YtnP-85R were mobilized to each of the isolates.

As shown in Figure 19, all strains were able to form planktonic cellular aggregates. When comparing the parental strain P0213-1, with the respective complemented form, some phenotypic/morphologic changes can be seen. The P0213-1 isolate expressing YtnP-85R (Fig. 19D) presented the strongest phenotypic difference in comparison with the other 2 strains. Instead of the large aggregates seen in the other strain (Fig. 19A-C), P0213-1 expressing YtnP-85R produced higher number of small aggregates. The microscopic shape of the planktonic cellular aggregates formed by the expression of the YtnP-85R lactonase was significantly different from the other strains, appearing to have lost the ability to form the ramifications that can be seen in the other P0213-1 strains.

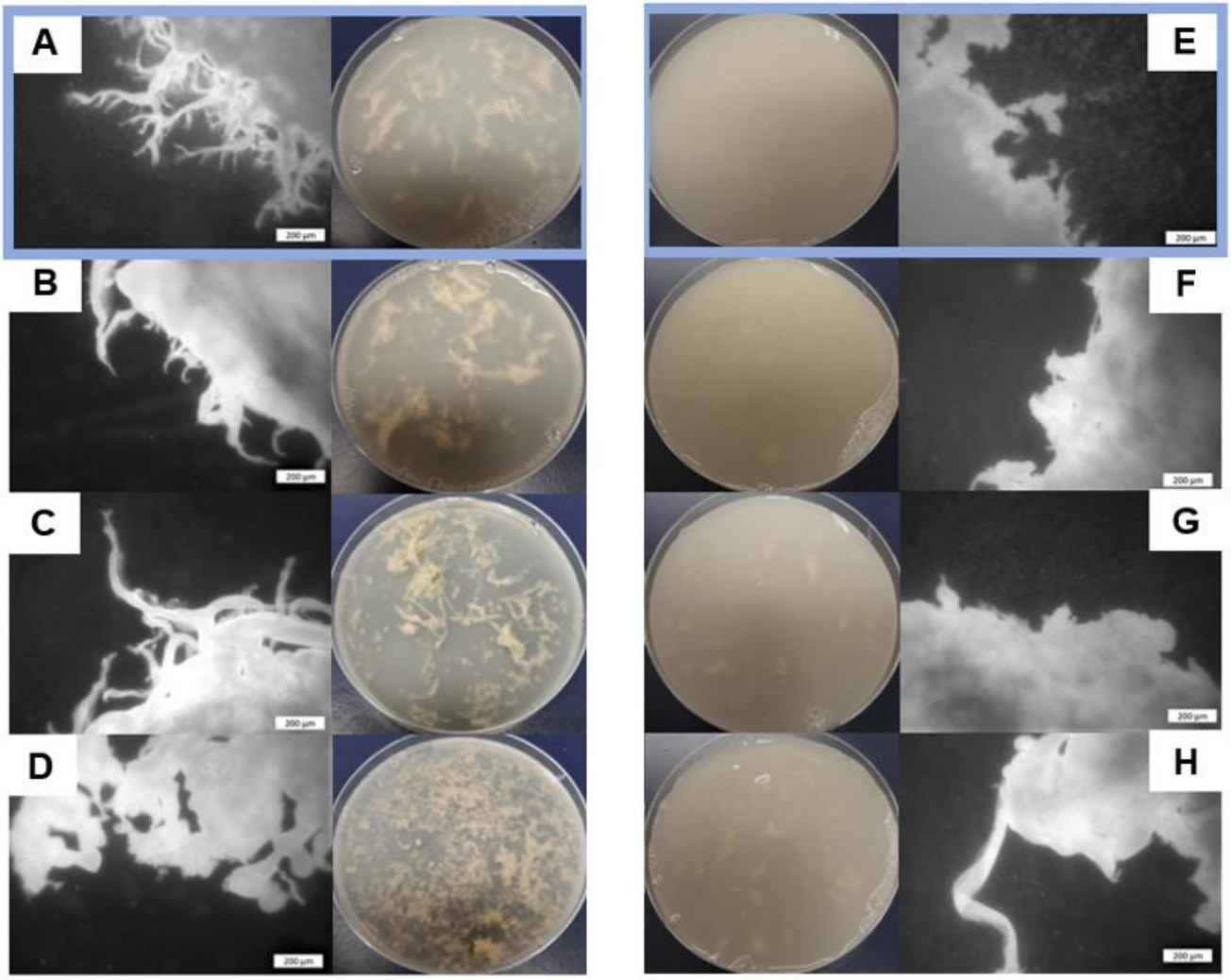

Figure 19: Screening for aggregate formation by P0213-1, P0148-1 and respective complemented mutants. **A-** P0213-1; **B-** P0213-1/pBBR1MCS; **C-** P0213-1/pCR20-1 (YtnP-85S); **D-** P0213-1/pCR20-5 (YtnP-85R); **E-** P0148-1; **F-** P0148-1/pBBR1MCS **G-** P0148-1/pCR20-1 (YtnP-85S); **H-** P0148-1/pCR20-5 (YtnP-85R).

A similar analysis of isolate P0148-1 expressing both lactonase variants did not show significant difference in the structure of the aggregates (Fig. 19E-H).

The planktonic cellular aggregates and free cells were quantified by weighing the dry biomass, in order to better visualise the differences between the isolates.

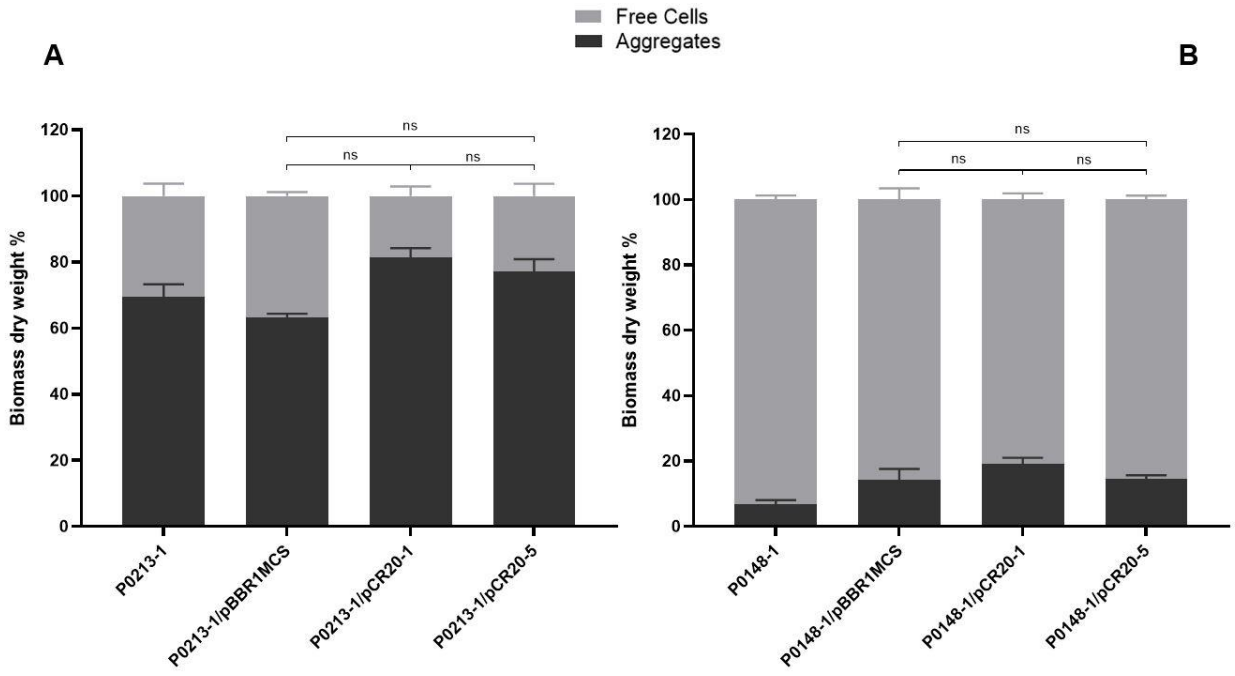

Figure 20: Quantification of cellular aggregates and free cells. A- P0213-1 and complemented with the lactonase gene; **B-** P0148-1 and complemented with the lactonase gene. Error bars correspond to the standard deviation. Significance level (one-way ANOVA followed by Dunnett's multiple comparison test) between cellular aggregates and free cells of the parental (P0213-1 and P0148-1) and cellular aggregates and free cells of the complemented isolates was determined: ns, not statistically significant.

Analysis of Figure 20A allows to visualize an increase of aggregates biomass in the P0213-1 strain expressing both lactonase variants, although this increase was not statistically significant when compared to the presence of the empty vector. Expression of the two lactonase variants in isolate P0148-1 also did not have significant impact in aggregate formation when compared to the strain with the empty vector (Fig. 20B).

Altogether, our data suggests that expression of YtnP-85R affected aggregate morphology in P0213-1, but no difference was observed in strain P0148-1.

Lactonases, as previously mentioned, have been proven to be able to disrupt QS in several bacterial species. However, lactonase efficiency in inhibiting QS depends on several factors, such as AHL hydrolytic activity levels and the specificity of the lactonase being studied. Some lactonases are able to completely inhibit biofilm formation while other only slightly decrease the formation and/or the size of the aggregates (Rémy *et al.* 2020). This could explain the phenotype that can be seen in P0213-1 expressing YtnP-85R (Fig. 19D). Although there was not a decrease in multicellular aggregate formation, the size of the aggregates decreased significantly which could be suggest YtnP-85R is active but the specificity to this AHLs is low.

Lactonases can have different specificity and different biological roles. Some lactonases act intracellularly playing a role in self-regulation of AHL-mediated quorum sensing. It has already been

shown that the fine-tunning of AHL concentration in *P. aeruginosa* is achieved by AHL acylase activity (Sio *et al.* 2006). Contrastly, extracellular lactonases respond to QS signals produced by competitors living within the same niche giving the quorum quenching producer a competition advantage. Such an example of this type of lactonase was recently characterized in *Burkholderia cepacia* BCC4135 (Malešević *et al.* 2020). This strain produces the lactonases YtnP and Y2-aiiA, being the first intracellular and the second extracellular. Analysis of the enzymatic potential showed that YtnP has a higher preference for short and middle-long chain AHLs, while the Y2-aiiA was efficient against both short- and long-chained AHLs. Therefore, the authors proposed that the intracellular YtnP might be involved in self-regulation of AHL-mediated QS while Y2-aiiA, by displaying broader substrate specificity may have the ability to respond to different QS signals produced by competitors. The YtnP homolog of *B. multivorans* characterized in this study shows low homology to Y2-aiiA (29% similarity) and YtnP (30% similarity) from *B. cepacia* and is a new putative lactonase.

3.2.3. Overexpression and purification of YtnP-85S protein

With the aim of characterizing the biochemical activity and AHLs substrate specificity of the YtnP lactonase enzyme, we envisaged the cloning of the two gene alleles, and recombinant protein expression and purification. Nevertheless, due to time limitations only one of the genes was cloned. To this purpose, the 1.3-Kb PCR fragment from P0426-1 isolate was amplified and cloned into the expression vector pET-23a(+) under the control of the T7 promoter, creating pCR21-4 (Table 1). The overexpression of the protein as an His-tagged derivative was accomplished by introduction of the plasmid pCR21-4 into *E. coli* BL21 (DE3) cells. The transformed *E. coli* BL21 (DE3) cells were grown in LB medium and various conditions (such as temperature and IPTG concentration) were tested to improve the amount of protein expressed and the ones that showed higher levels of the His-tagged protein were selected. Although a very high overexpression of His_6YtnP protein was not achieved, the best condition for protein expression in *E. coli* BL21 (DE3) was at 37 ºC with 100 µg/L of ampicillin, with an IPTG concentration of 0.4 mM and an induction time of 4 hours.

The overexpressed His_6-tagged protein was analysed by SDS-PAGE (Figure 21). Although a clear band corresponding to the predicted protein size (31 kDa) was not clearly visible in lanes 4 and 5, its expected location according to the molecular mass is indicated with and arrow.

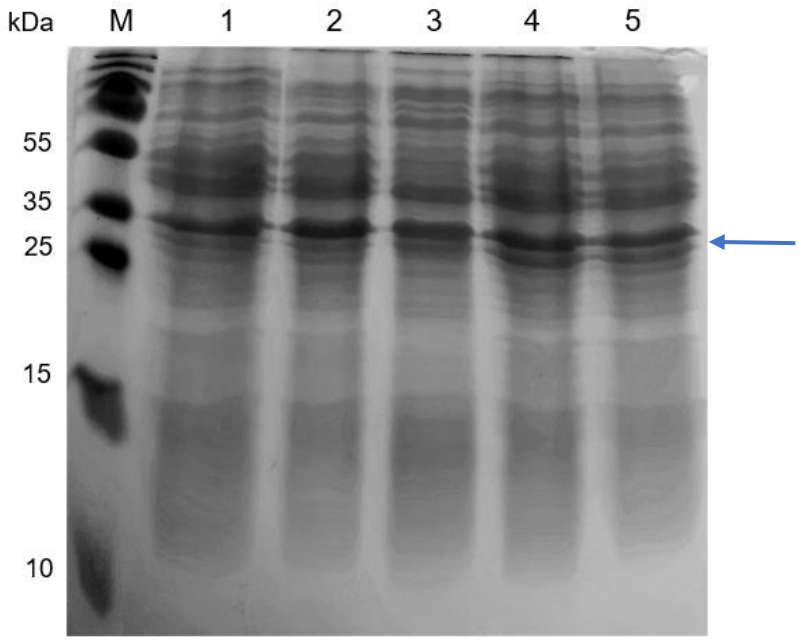

Figure 21: 15% SDS-PAGE gel with YtnP overexpression assay samples. Lane M - PageRuler™ Plus prestained protein ladder (Thermo Scientific); Lane 1 – *E. coli* BL21 (DE3); Lane 2 - pET-23a(+); Lane 3 – pCR21-4 with 0mM IPTG (t0); Lane 4 – pCR21-4 0.4 mM IPTG (t4); Lane 5 – pCR21-4 0.4 mM IPTG (t24).

The purification of the recombinant protein was performed using nickel affinity chromatography. The protein extract was fully applied to the column and a constant increasing gradient varying form 10 mM to 500 mM of imidazole concentration was used to elute the protein. The eluted fraction from the column were analysed through SDS-PAGE (Figure 22), with the protein present in 7 of the fractions. The fraction containing higher concentration of the His_6YtnP protein was the third fraction eluted with the buffer containing 100 mM of imidazole. Small amounts of contaminants were present in most of the fractions recovered with the protein, being the fifth fraction eluted with the buffer containing 100 mM of imidazole the only one where no contaminants were visible.

The purified YtnP protein could be used in further studies to better understand the function of this lactonase and the importance of these types of enzymes in inhibiting virulence in some bacterial species by interfering with their QS systems. But, before that, there is still room for improvement of its expression and further purification. For example, gel filtration chromatography, should be assayed.

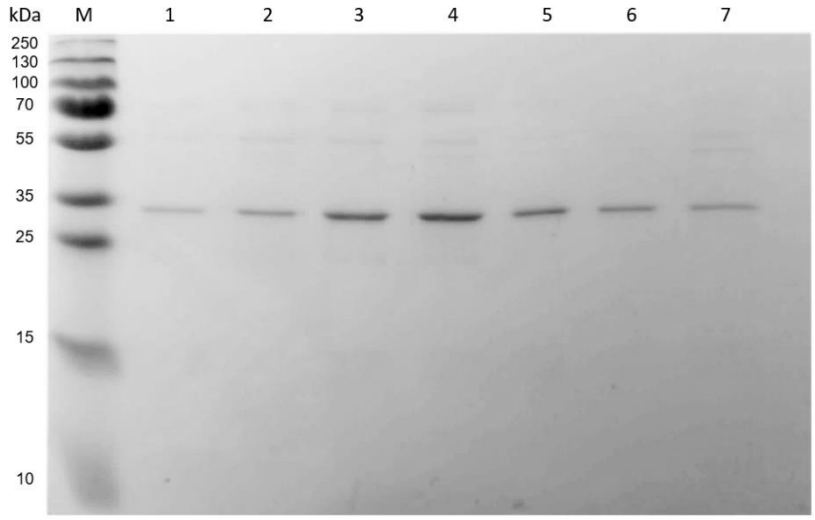

Figure 22: Analysis by SDS-PAGE of His-tagged YtnP protein purification by nickel affinity chromatography. Lane M - PageRuler^{TM} Plus prestained protein ladder (Thermo Scientific); Lane 1 - 50(3) mM imidazole; Lane 2- 100(1) mM imidazole; Lane 3- 100(2) mM imidazole; Lane 4- 100(3) mM imidazole; Lane 5- 100(4) mM imidazole; Lane 6- 100(5) mM imidazole; Lane 7- 300(1) mM imidazole.

3.3. Analysis of the outer membrane protein (TtgC) from isolates P0426-1 and P0426-2

3.3.1. Comparative sequence analysis and phylogenetic distribution of the TtgC protein

The mutation in the TtgC protein occurred on the position 503 of the amino acid sequence which in total has 511 amino acids (Figure 23). An adenine was replaced by a guanine (ACG>GCG), which caused a change in the resulting amino acid from a threonine to an alanine. When the sequences of the different species were aligned it is possible to see that this amino acid is not conserved. Like the second isolate, four of the sequences retrieved from the different species had an alanine at position 503, being one of them the sequence from *B. multivorans* ATCC 17616. The fact that this amino acid is not conserved in all the other species opens the possibility that this might not be a conserved area and could possibly be irrelevant in functional terms.

Even though the amino acid in which the mutation occurred is not conserved, this protein was chosen for further studies since this mutation is also present in all longitudinal isolates that were not able to form planktonic cellular aggregates.

	491 aa												503 aa								511 aa
Burkholderia multivorans P0426-1	A	P	T	A	V	G	G	-	-	-	A	A	T	A	K	A	E	I	A	G	R
Burkholderia multivorans P0426-2	A	P	T	A	V	G	G	-	-	-	A	A	A	A	K	A	E	I	A	G	R
Burkholderia multivorans ATCC 17616	A	P	T	A	V	D	G	-	-	-	A	A	A	A	K	A	E	I	A	G	R
Burkholderia cepacia LO6	T	P	T	A	V	G	G	-	-	-	A	A	S	G	K	E	N	V	A	A	R
Burkholderia dolosa AU0158	T	P	T	A	V	G	G	-	-	-	A	A	S	G	K	E	N	V	A	A	R
Burkholderia anthina AZ-4-2-10-S1-D7	A	P	T	A	I	G	D	-	-	-	A	G	A	G	K	A	E	V	A	G	R
Burkholderia pseudomultivorans MSMB2199WGS	A	P	T	A	V	G	D	-	-	-	A	A	T	A	K	A	E	I	A	G	R
Burkholderia vietnamiensis G4	A	P	T	A	V	G	D	-	-	-	A	A	T	G	K	A	E	V	A	A	R
Burkholderia territorii MSMB2203WGS	A	P	A	A	V	G	D	-	-	-	A	G	T	G	K	Q	D	V	A	V	R
Burkholderia pyrrocinia DSM 10685	A	P	T	A	V	G	D	-	-	-	A	A	S	V	K	A	E	V	A	A	R
Burkholderia contaminans MS14	A	P	T	A	V	G	D	-	-	-	A	A	T	G	K	A	E	V	A	A	R
Burkholderia cenocepacia J2315	A	P	T	A	V	G	D	-	-	-	A	A	S	G	K	A	D	V	A	A	R
Burkholderia seminalis FL-5-4-10-S1-D7	A	P	T	A	V	G	D	-	-	-	A	A	S	V	K	A	E	V	A	A	R
Burkholderia lata 383	A	P	T	A	V	G	D	-	-	-	A	A	S	V	K	A	E	V	A	A	R
Burkholderia diffusa RF2-non-BP9	A	P	A	A	I	G	D	-	-	-	A	A	H	-	K	E	D	V	A	A	R
Burkholderia ambifaria AMMD	A	P	T	A	V	G	D	-	-	-	A	A	T	G	K	A	D	V	A	A	R
Burkholderia stabilis ATCC BAA-67	A	P	T	A	V	G	D	-	-	-	A	A	S	G	K	E	S	V	A	A	R
Burkholderia ubonensis MSMB1600WGS	V	P	A	A	V	G	D	-	-	-	A	A	R	-	K	E	D	V	A	A	R
Burkholderia oklahomensis C6786	A	D	V	A	A	G	A	-	-	-	R	E	P	A	Q	Q	D	V	A	A	R
Paraburkholderia phenazinium	V	T	T	A	S	A	A	-	-	-	A	K	D	A	P	Q	Q	V	A	K	Q
Paraburkholderia megapolitana	T	A	A	A	D	V	A	-	-	-	N	S	D	A	Q	Q	Q	I	A	K	R
Trinickia fusca	A	I	D	L	V	D	S	-	-	-	D	S	S	G	T	S	K	V	T	K	Q
Pandoraea sputorum	A	A	P	A	V	P	T	-	-	-	A	N	A	A	G	A	T	V	A	A	R
Caballeronia mineralivorans PML1	A	A	P	A	A	P	A	-	-	-	A	D	A	N	-	-	-	I	A	S	R
Bordetella hinzii	P	A	V	A	A	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Pseudomonas sp. SST3	A	K	D	C	L	P	N	G	R	K	A	V	I	T	Q	P	E	T	A	L	N
Pararobbsia silviterrae	A	P	K	Q	D	P	A	-	-	-	A	A	V	P	-	-	P	V	A	S	R
Achromobacter sp. AONIH1	P	G	A	G	A	G	Q	V	A	V	K	-	-	-	-	-	-	-	-	-	-
Herbaspirillum sp. CF444	P	T	E	T	Q	A	K	N	D	A	A	G	-	-	-	-	-	-	N	K	M
Methyloversatilis sp. RAC08	A	E	A	I	N	-	-	T	S	D	A	G	P	V	Q	P	L	H	A	R	H
Pigmentiphaga humi	P	S	T	A	M	V	G	T	S	M	A	T	-	-	-	-	I	I	T	T	P
Pusillimonas sp. YR330	-	-	-	-	-	-	-	T	G	-	-	-	-	-	-	-	S	M	A	R	L
Oryzomicrobium terrae	P	S	A	V	S	S	A	T	A	Q	A	G	V	A	A	G	P	V	A	T	A
Azoarcus communis SWub3	P	T	L	G	H	G	P	T	P	D	A	G	-	-	-	-	L	I	A	R	Q
Ralstonia solanacearum GMI1000	D	P	P	A	P	Q	T	-	-	-	-	-	-	-	-	-	T	V	S	Q	R

Figure 23: Amino acid sequence alignment generated by ClustalX of a representative sequence of TtgC protein present in different species.

In order to better analyse the similarities among these sequences, neighbour joining method was used to construct a phylogenetic tree based on the ClustalX alignment (Fig. 24). The TtgC protein is conserved among the *Burkholderia* genus, but is also present in other Proteobacteria.

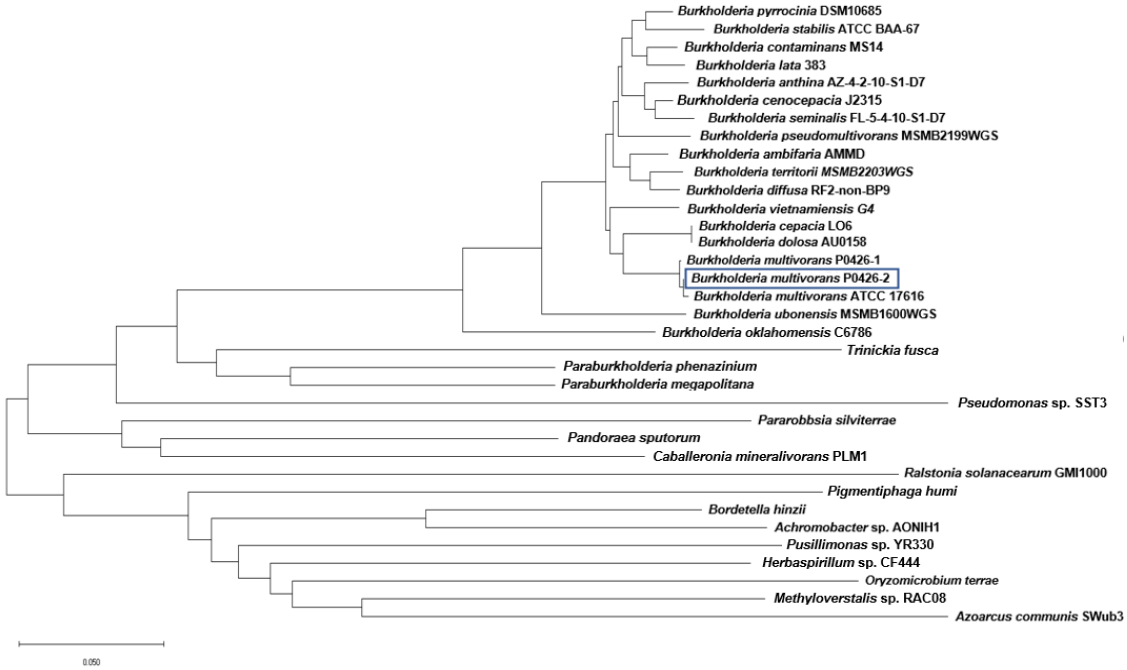

Figure 24: Phylogeny of TtgC protein. Neighbor-joining phylogenetic tree of the deduced amino acid sequences of TtgC predicted proteins from 19 *Burkholderia* species and other species from different genera. Scale: number of substitutions per site.

3.3.2. Genetic complementation with the *ttgC* gene

The 1.6-Kb containing the *ttgC* gene and its own promotor from each clinical isolate was cloned into the pBBR1MCS broad-host-range vector, in order to study the possible role of this outer membrane protein in the formation of multicellular aggregates. A complementation assay was performed to test whether this change in amino acid from a threonine to an alanine might have caused the impairment of multicellular aggregate formation in P0426-2. The plasmid pCR21-3, together with the empty vector was introduced into P0426-1 and pCR21-2 into P0426- 2. Cells were grown in SM medium at 37ºC with 180 rpm of orbital agitation for 48 hours followed by inspection of aggregates formation.

The microscopic and macroscopic appearance of the planktonic cellular aggregates can be seen in Figure 25. In the complementation of P0426-1 none of the strains showed significant differences in the structure of the aggregates analysed, but the number of aggregates formed seems very different. In the complementation of P0426-2 there were also no changes when comparing to the respective parental strain since cells remained freely. For a better understanding of the differences in aggregate formation by the complementation of the isolates, the percentage of biomass dry weight recovered from both aggregates and free cells was determined.

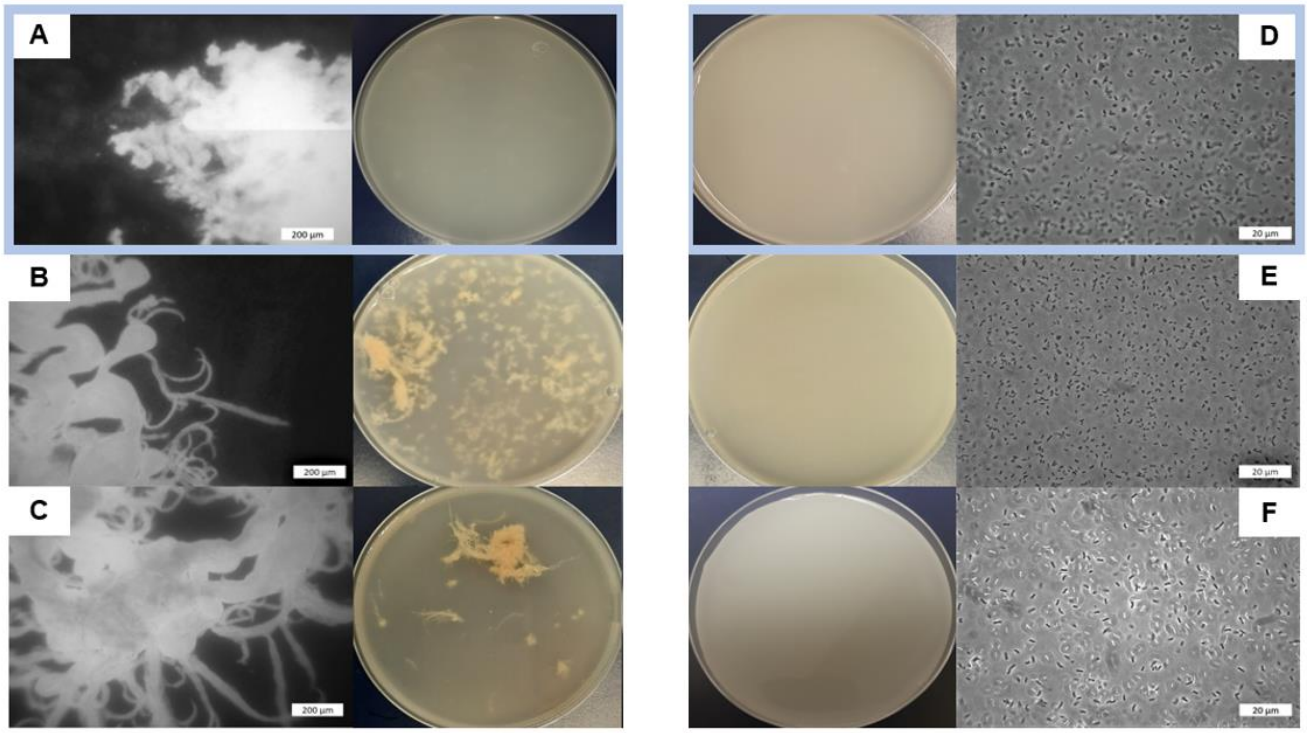

Figure 25: Screening for aggregate formation by P0426-1, P0426-2 expressing a different outer membrane protein TtgC. A- P0426-1; **B-** P0426-1/pBBR1MCS; **C-** P0426-1/pCR21-3 (TtgC-503A); **D-** P0426-2; **E-** P0426-2/pBBR1MCS; **F-** P0426-2/pCR21-2 (TtgC-503T).

As depicted in Figure 26A, the presence of the empty vector in isolate P0426-1 increased in the formation of planktonic cellular aggregates when compared the parental strain, but the expression of TtgC-503A led to a decrease of aggregate formation. Nevertheless, this difference is not statistically significant. Complementation of P0426-2 with the *ttgC* gene from P0426-1 (TtgC-503T) also did not alter

the proportion between free cells and aggregates (Fig. 26B), since neither the parental strain nor the complemented were able to form aggregates.

In conclusion, expression of the TtgC-503A in P0426-1 seems to have a slight effect in aggregates biomass and structure since the aggregates were generally of a larger size, suggesting a possible (but small) role of this protein in aggregate formation. The more likely function of this protein and genes in the vicinity is multidrug resistance, but it could not be immediately excluded a possible role in biofilm/aggregates formation.

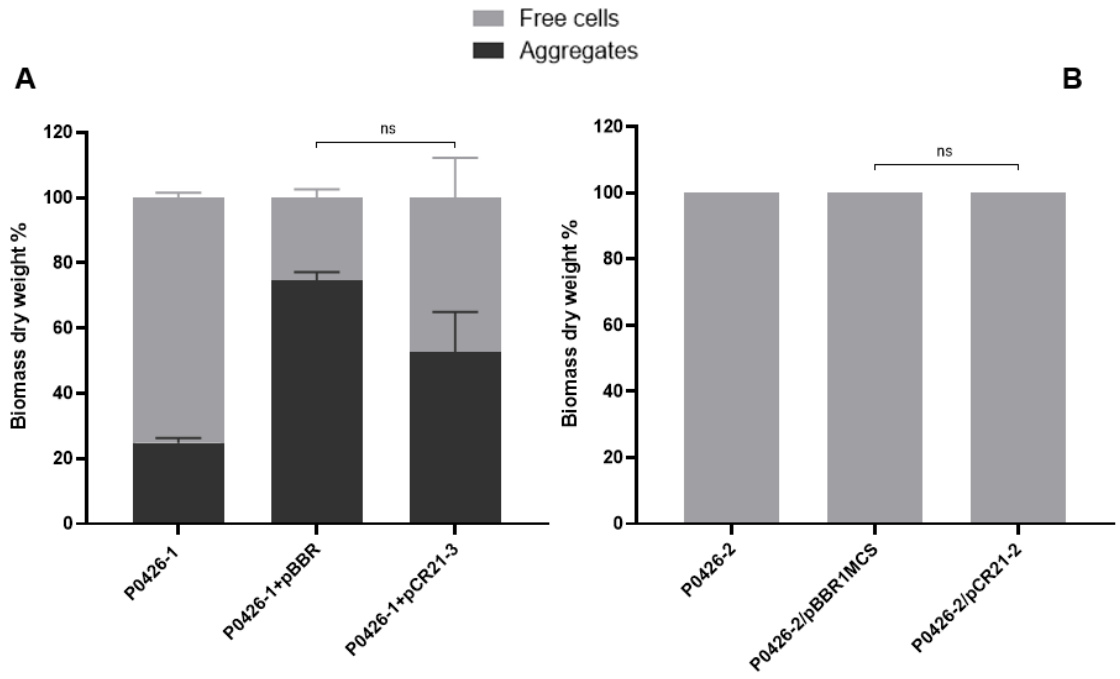

Figure 26: Quantification of cellular aggregates and free cells. A- P0426-1 parental and complemented; **B-** P0426-2 parental and complemented. Error bars correspond to the standard deviation. Significance level (one-way ANOVA followed by Dunnett's multiple comparison test) between cellular aggregates and free cells of the parental (P0426-1 and P0426-2) and cellular aggregates and free cells of the complemented isolates was determined: ns, not statistically significant.

3.4. Comparative sequence analysis of the Stp and PbuE proteins

The mutation in the Stp protein occurred on the position 292 of the 471 amino acid sequence (Figure 27). An adenine was replaced by a cytosine (CAG>CCG), which caused a change in the resulting amino acid from a glutamine to a proline. When the sequences of the different species were aligned it is possible to see that this amino acid is conserved in all strains except the isolate P0426-2 and maps between two transmembrane domains, perhaps with no significant effect on protein function. Since this mutation was only present in P0426-2 of the longitudinal series, no further experiments were carried out.

Figure 27: Amino acid sequence alignment generated by ClustalX of a representative sequence of Stp protein present in different species.

The mutation in the major facilitator superfamily PbuE protein occurred on the position 237 of the amino acid sequence which in total has 406 amino acids (Figure 28). A cytosine was replaced by a adenine (CTG>ATG), which caused a change in the resulting amino acid from a leucine to a methionine. Just like for the Stp protein, when the sequences of the different species were aligned it is possible to see that this amino acid is conserved in all the identified proteins except in P0426-2. This mutation is present in all isolates of the longitudinal series (except P0426-1), being an interesting protein, but due to lack of time no further studies were performed.

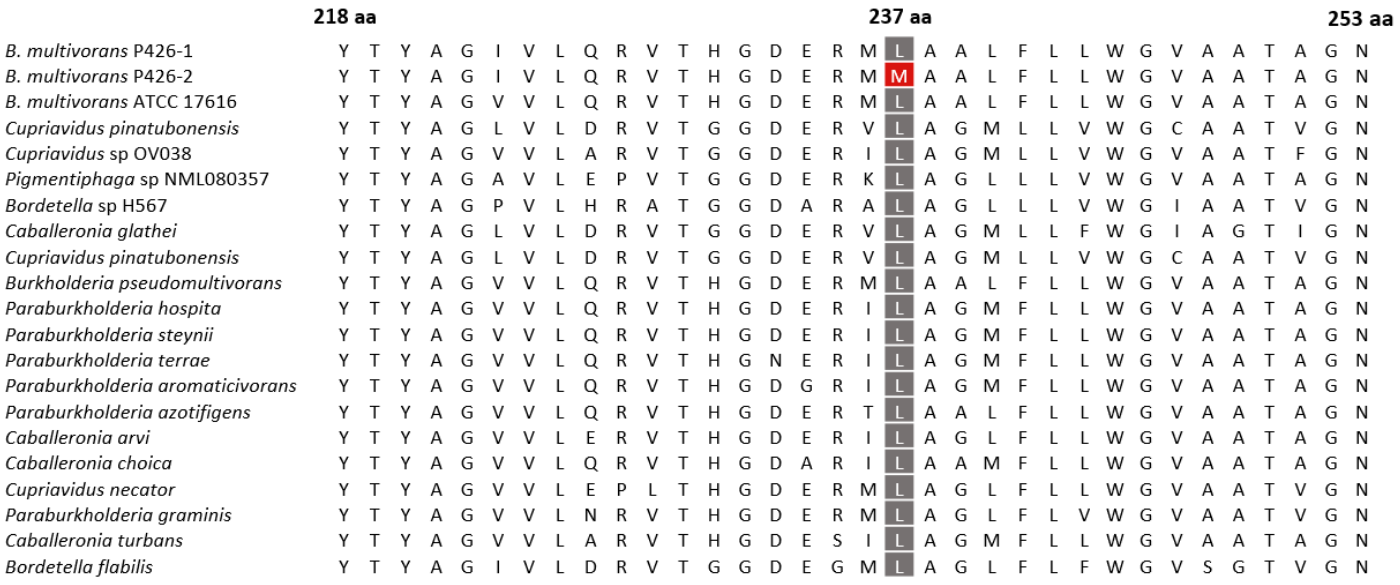

Figure 28: Amino acid sequence alignment generated by ClustalX of a representative sequence of PbuE protein present in different species.

3.5. Gene expression analysis by real-time quantitative PCR

A RT-qPCR assay based on SYBR Green detection was performed to assess the expression of the four genes that might be affected by the two mutations in intergenic regions (possibly promotor regions) of the two isolates (Table 5). Each reaction was performed once for the three different times (4, 8 and 24 hours of growth), and three biological replicates analysed.

The method used to analyse the data was the comparative $\Delta\Delta C_T$ method (Rao *et al.* 2013) . The C_T is described as the number of cycles that are necessary for the fluorescent signal emitted by the SYBR Green to cross the threshold line and is inversely proportional to the quantity of target mRNA present in the sample.

As previously mentioned, the RNA was extracted from 3 different points (4, 8 hours corresponding to exponential and late exponential phase and 24 hours corresponding to stationary phase), followed by digestion with DNase to avoid possible contaminations and the quantification of RNA through a UV spectrophotometer (ND-1000 UV-Vis, NanoDrop Technologies, USA). To assess if the samples were contaminated with DNA, reactions without the reverse transcriptase were made. The quality was evaluated. In this case the housekeeping gene *rpoD* was used, and only the samples with a C_T value higher than 30 (low DNA contamination) would be used for further studies.

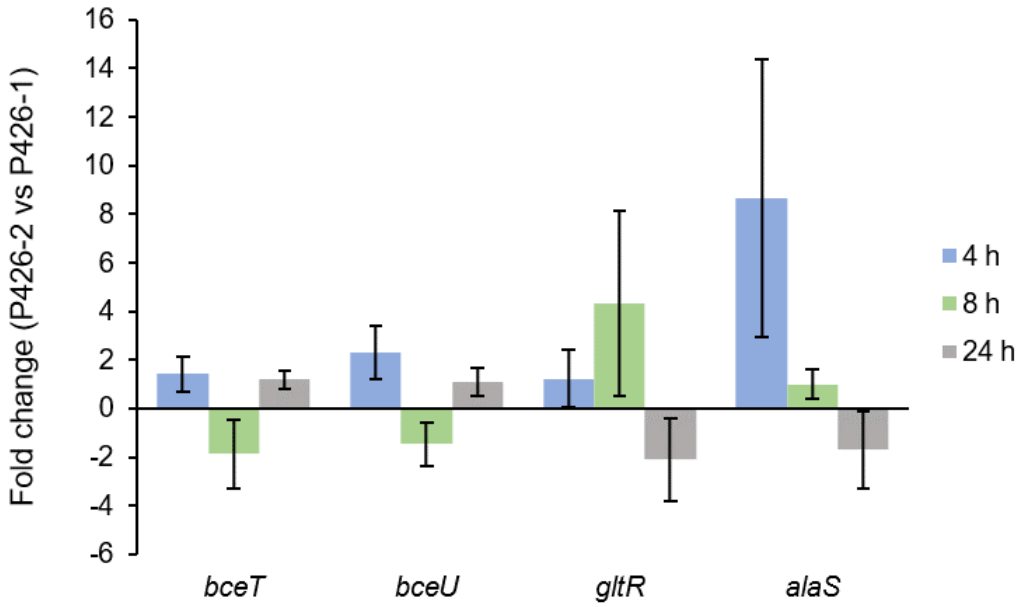

Figure 29: Expression levels of the target genes. Values are given in the form of C_T values) The columns represent mean $2^{-\Delta\Delta CT}$ values, the bars standard deviations.

The first mutation to be addressed is a 7-nucleotide insertion in the intergenic region between the two divergently transcribed genes, namely *bceT* and *bceU*. Considering a 2-fold cuttoff for significantly different expression, we can conclude that none of the genes seem to be affected. Previous analysis of cepacian biosynthesis by the two different isolates indicates that both are able to present the mucoid phenotype due to cepacian biosynthesis in yeast extract mannitol medium (Pessoa, 2017). This result

is in line with the expression data of *bceT* and *bceU* genes here reported, although a quantitative analysis for exopolysaccharide should be done to confirm that they produced a similar amount.

Regarding the second region, a 10-nucleotide deletion that occurred between the divergently expressed genes *alaS* and *gltR,* could affect the expression of both genes at 8 hours of growth. Nevertheless, the standard error is very high, and no definite conclusion can be taken without repeating this experiment.

4. Concluding remarks and future perspectives

Some microorganisms are able to form multicellular aggregates, which could either be surface-attached biofilms and planktonic cellular aggregates. These planktonic cellular aggregates are of great importance, especially in clinical settings, due to the increased ability to prevail in diverse environmental conditions. It has been proven that bacteria in biofilms are more resistant to treatment with antimicrobial compounds when compared to free cells, and, that biofilm development is related to many aspects of human pathogenesis within a clinical context. *Burkholderia cepacia* complex are some of the bacteria capable of forming these aggregates, allowing them to establish chronic infection in patients suffering from cystic fibrosis, often leading to decreased lung function. However, the understanding of the molecular mechanisms involved in the formation of planktonic cellular aggregates by these bacteria still requires further study.

The aim of this study was to identify some molecular players that may be involved in the formation of planktonic cellular aggregates in *Burkholderia multivorans*. The approach comprised the identification of mutations between two sequential *B. multivorans* isolates from the same CF patient in which one has the ability to form aggregates (P0426-1) while the other is not able to (P0426-2), and their study in order to identify if any of these mutations could be involved in the loss of aggregate formation by the second isolate.

There were 7 mutations distinguishing the two clinical isolates, four SNP mutations in coding regions, which were in a gene encoding a putative quorum-quenching lactonase (YtnP), an efflux pump (PbuE), a multidrug resistance protein (Stp) and an outer membrane protein (TtgC). Between the two isolates there are also a large deletion of hypothetical proteins and two small indel mutations in intergenic regions, being the genes that could be affected by these mutations the UTP-glucose-1-phosphate uridylyltransferase (BceT) and acetyltransferase (BceU), and a LysR-type transcriptional regulator (GltR) and an alanyl-tRNA synthetase (AlaS).

Previous phenotypic characterization of P0426-1 and P426-2 identified a few differences between them, namely a decreased swimming and swarming motility and increased resistance to ciprofloxacin by the P0426-2 isolate (Pessoa, 2017). Other phenotypes such as the presence of O antigen in the lipopolysaccharide, antimicrobial resistance against several antibiotics, the mucoid phenotype, adhesin to CF lung epithelial cells and virulence in *Galleria mellonella* showed no significant differences. The increased resistance to ciprofloxacin by P0426-2 isolate could be caused by mutation in genes such as

pbuE, *stp* or *ttgC*. But differences in motility and aggregate formation might be more likely caused by quorum sensing regulation. For that reason, the most evident target would be gene *ytnP* encoding a putative lactonase.

A complementation assay was performed to test whether expression of a different form of the lactonase in *B. multivorans* P0426-1 would affect aggregates formation. The expression of YtnP-85R in P0426-1 did not prevent aggregate formation, but the structure of the aggregates was more compact and with more ramifications. Furthermore, the comparison of the aggregates´ biomass between P0426-1/ pBBR1MCS and P0426-1-YtnP-85R showed a statistically significant reduction (53% to 22%) for the isolate expressing YtnP-85R. In addition, the complementation of isolate P0213-1 with YtnP-85R also showed macroscopic and microscopic changes. Instead of the large aggregates seen in the parental isolate, P0213-1 expressing YtnP-85R produced higher number of small aggregates that microscopically appeared to have lost the ability to form ramifications. This result suggests that YtnP-85R expression might have affected both the structure of the aggregates and the relative amount. Although the expression of YtnP-85R in P0426-1 and P0213-1 did not abolish aggregation of cells, this preliminary data provides good indications that this novel lactonase might be relevant in the regulation of *Burkholderia* pathogenesis. Future studies will involve the construction of a YtnP deletion mutant; studies on expression of this gene during the growth phases; determination of the range of AHLs being degraded by this lactonase; and measure the intracellular levels of AHLs in the presence and absence of YtnP. As a final remark, is important to say that despite the observed mutations, their effect on aggregates formation might be an indirect one, for example through regulation of gene expression. In that case, perhaps determining the transcriptome of the two isolates would give us some clues on this virulence mechanism.

5. References

Ahmad, S., Tyrrell, J., Walton, W. G., Tripathy, A., Redinbo, M. R., & Tarran, R. (2016). Short palate, lung, and nasal epithelial clone 1 has antimicrobial and antibiofilm activities against the *Burkholderia cepacia* Complex. Antimicrobial Agents and Chemotherapy, 60(10), 6003–6012.

Alagarasan, G., Aswathy, K. S., & Madhaiyan, M. (2017). Shoot the message, not the messenger—combating pathogenic virulence in plants by inhibiting quorum sensing mediated signaling molecules. In Frontiers in Plant Science (Vol. 8). Frontiers Media SA.

Alav, I., Sutton, J. M., & Rahman, K. M. (2018). Role of bacterial efflux pumps in biofilm formation. Journal of Antimicrobial Chemotherapy, 73(8), 2003–2020.

Alhede, M., Kragh, K. N., Qvortrup, K., Allesen-Holm, M., van Gennip, M., Christensen, L. D., Jensen, P. Ø., Nielsen, A. K., Parsek, M., Wozniak, D., Molin, S., Tolker-Nielsen, T., Høiby, N., Givskov, M., & Bjarnsholt, T. (2011). Phenotypes of Non-Attached Pseudomonas aeruginosa Aggregates Resemble Surface Attached Biofilm. In M. A. Webber (Ed.), PLoS ONE (Vol. 6, Issue 11, p. e27943). Public Library of Science (PLoS).

Amin, R., Dupuis, A., Aaron, S. D., & Ratjen, F. (2010). The effect of chronic infection with A*spergillus fumigatus* on lung function and hospitalization in patients with cystic fibrosis. Chest, 137(1), 171–176.

B. Hollmann, M. Perkins, D. Walsh. Biofilms and their role in pathogenesis. Centre of Biomolecular Sciences, University of Nottingham. https://www.immunology.org/public-information/bitesized-immunology/pathogens-and-disease/biofilms-and-their-role-in#:~:text=The%20role%20of%20biofilms%20in%20pathogenesis&text=One%20example%20of%20a%20positive,and%20the%20prevention%20of%20adhesion.

Baldwin, A., Mahenthiralingam, E., Thickett, K., Honeybourne, D., Maiden, M. C. J., Govan, J. R., Speert, D., LiPuma, J. J., Vandamme, P., Dowson, C. G. (2005). Multilocus sequence typing scheme that provides both species and strain differentiation for the *Burkholderia cepacia* Complex. Journal of Clinical Microbiology, p. 4665–4673.

Basavaraju, M., Sisnity, V. S., Palaparthy, R., & Addanki, P. K. (2016). Quorum quenching: Signal jamming in dental plaque biofilms. Journal of Dental Sciences, 11(4), 349–352.

Basics of the CFTR Protein. https://www.cff.org/Research/Research-Into-the-Disease/Restore-CFTR-Function/Basics-of-the-CFTR-Protein/. Accessed 20/12/2020.

Bernhardt, S. A., Spilker, T., Coffey, T., LiPuma, J. J. (2003). *Burkholderia cepacia* complex in cystic fibrosis: frequency of strain replacement during chronic infection, Clinical Infectious Diseases, Volume 37, Issue 6, Pages 780–785.

Bevivino, A., Dalmastri, C., Tabacchioni, S., Chiarini, L., Belli, M. L., Piana, S., Materazzo, A., Vandamme, P., & Manno, G. (2002). *Burkholderia cepacia* complex bacteria from clinical and

environmental sources in italy: genomovar status and distribution of traits related to virulence and transmissibility. Journal of Clinical Microbiology, 40(3), 846–851.

Bhagirath, A. Y., Li, Y., Somayajula, D., Dadashi, M., Badr, S., & Duan, K. (2016). Cystic fibrosis lung environment and *Pseudomonas aeruginosa* infection. BMC Pulmonary Medicine, 16(1).

Bible, A. N., Stephens, B. B., Ortega, D. R., Xie, Z., & Alexandre, G. (2008). Function of a chemotaxis-like signal transduction pathway in modulating motility, cell clumping, and cell length in the Alphaproteobacterium *Azospirillum brasilense*. Journal of Bacteriology, 190(19), 6365–6375.

Boles, B. R., & Horswill, A. R. (2008). AGR-mediated dispersal of *Staphylococcus aureus* biofilms. PLoS Pathogens, 4(4), e1000052.

Bose, S. J., Bijvelds, M. J. C., Wang, Y., Liu, J., Cai, Z., Bot, A. G. M., de Jonge, H. R., & Sheppard, D. N. (2019). Differential thermostability and response to cystic fibrosis transmembrane conductance regulator potentiators of human and mouse F508del-CFTR. American Journal of Physiology-Lung Cellular and Molecular Physiology, 317(1), L71–L86.

Brice, P., Jarrett, J., & Mugford, M. (2007). Genetic screening for cystic fibrosis: An overview of the science and the economics. Journal of Cystic Fibrosis, 6(4), 255–261.

Brisse, S., Cordevant, C., Vandamme, P., Bidet, P., Loukil, C., Chabanon, G., Lange, M., & Bingen, E. (2004). Species distribution and ribotype diversity of *Burkholderia cepacia* complex isolates from french patients with cystic fibrosis. Journal of Clinical Microbiology, 42(10), 4824–4827.

Burkholder, W. H. 1950. Sour skin, a bacterial rot of onion bulbs. Phytopathology 40:115-118.

Buroni, S., Pasca, M.R., Flannagan, R.S. (2009). Assessment of three resistance-nodulation-cell division drug efflux transporters of *Burkholderia cenocepacia* in intrinsic antibiotic resistance. BMC Microbiol; 9: 200.

CFTR Modulator Types. https://www.cff.org/Research/Developing-New-Treatments/CFTR-Modulator-Types/. Accessed 15/12/2020.

Cant, N., Pollock, N., & Ford, R. C. (2014). CFTR structure and cystic fibrosis. The International Journal of Biochemistry & Cell Biology, 52, 15–25.

Cantin, A. M., Hartl, D., Konstan, M. W., & Chmiel, J. F. (2015). Inflammation in cystic fibrosis lung disease: Pathogenesis and therapy. Journal of Cystic Fibrosis, 14(4), 419–430.

Chan, K.-G., Atkinson, S., Mathee, K., Sam, C.-K., Chhabra, S., Cámara, M., Koh, C.-L., & Williams, P. (2011). Characterization of N-acylhomoserine lactone-degrading bacteria associated with the *Zingiber officinale* (ginger) rhizosphere: Co-existence of quorum quenching and quorum sensing in *Acinetobacter* and *Burkholderia*. BMC Microbiology, 11(1), 51.

Chan, K.-G., Wong, C.-S., Yin, W.-F., Sam, C.-K., & Koh, C.-L. (2010). Rapid degradation of N-3-oxo-acylhomoserine lactones by a Bacillus cereus isolate from Malaysian rainforest soil. Antonie van Leeuwenhoek, 98(3), 299–305.

Chan, Y.Y., Bian, H.S., Tan, T.M.C. (2007). Control of quorum sensing by a *Burkholderia pseudomallei* multidrug efflux pump. J Bacteriol; 189: 4320–4.

Choi, J. Y., Lee, M. G., Ko, S., Muallem, S. (2001). Cl^- -Dependent HCO_3 - Transport by cystic fibrosis. JOP. J. Pancreas, 2(4 Suppl):243-246.

Clancy, J. P., Cotton, C. U., Donaldson, S. H., Solomon, G. M., VanDevanter, D. R., Boyle, M. P., Gentzsch, M., Nick, J. A., Illek, B., Wallenburg, J. C., Sorscher, E. J., Amaral, M. D., Beekman, J. M., Naren, A. P., Bridges, R. J., Thomas, P. J., Cutting, G., Rowe, S., Durmowicz, A. G., … Tuggle, K. L. (2019). CFTR modulator theratyping: Current status, gaps and future directions. Journal of Cystic Fibrosis, 18(1), 22–34.

Coenye, T. (2004). Recovery of *Burkholderia cenocepacia* strain PHDC from cystic fibrosis patients in Europe. Thorax, 59(11), 952–954.

Coenye, T., LiPuma, J. J., Henry, D., Hoste, B., Vandemeulebroecke, K., Gillis, M., Speert, D. P., & Vandamme, P. (2001). *Burkholderia cepacia* genomovar VI, a new member of the *Burkholderia cepacia* complex isolated from cystic fibrosis patients. International Journal of Systematic and Evolutionary Microbiology, 51(2), 271–279.

Cools, F., Torfs, E., Vanhoutte, B., de Macedo, M. B., Bonofiglio, L., Mollerach, M., Maes, L., Caljon, G., Delputte, P., Cappoen, D., & Cos, P. (2018). *Streptococcus pneumoniae* galU gene mutation has a direct effect on biofilm growth, adherence and phagocytosis in vitro and pathogenicity in vivo. Pathogens and Disease, 76(7).

Costerton, J. W. (1999). Bacterial biofilms: a common cause of persistent infections. Science, 284(5418), 1318–1322.

Costerton, J. W., Lewandowski, Z., DeBeer, D., Caldwell, D., Korber, D., & James, G. (1994). Biofilms, the customized microniche. Journal of Bacteriology, 176(8), 2137–2142.

Cunha, M. V., Leitao, J. H., Mahenthiralingam, E., Vandamme, P., Lito, L., Barreto, C., Salgado, M. J., & Sa-Correia, I. (2003). Molecular analysis of *Burkholderia cepacia* complex isolates from a portuguese cystic fibrosis center: a 7-year study. Journal of Clinical Microbiology, 41(9), 4113–4120.

Cystic Fibrosis Mutation Database. http://www.genet.sickkids.on.ca/cftr/StatisticsPage.html, Accessed 15/12/2020.

De Boeck, K., & Amaral, M. D. (2016). Progress in therapies for cystic fibrosis. The Lancet Respiratory Medicine, 4(8), 662–674.

Dechecchi, M. C., Tamanini, A., & Cabrini, G. (2018). Molecular basis of cystic fibrosis: from bench to bedside. Annals of Translational Medicine, 6(17), 334–334.

Devanga Ragupathi, N. K., & Veeraraghavan, B. (2019). Accurate identification and epidemiological characterization of *Burkholderia cepacia* complex: an update. Annals of Clinical Microbiology and Antimicrobials, 18(1).

Devanga Ragupathi, N. K., & Veeraraghavan, B. (2019). Accurate identification and epidemiological characterization of *Burkholderia cepacia* complex: an update. Annals of Clinical Microbiology and Antimicrobials, 18(1).

Dienerowitz, M., Cowan, L. V., Gibson, G. M., Hay, R., Padgett, M. J., & Phoenix, V. R. (2014). Optically trapped bacteria pairs reveal discrete motile response to control aggregation upon cell–cell approach. Current Microbiology, 69(5), 669–674.

Drevinek, P., Vosahlikova, S., Dedeckova, K., Cinek, O., & Mahenthiralingam, E. (2010). Direct culture-independent strain typing of *Burkholderia cepacia* complex in sputum samples from patients with cystic fibrosis. Journal of Clinical Microbiology, 48(5), 1888–1891.

Ehre, C., Ridley, C., & Thornton, D. J. (2014). Cystic fibrosis: An inherited disease affecting mucin-producing organs. The International Journal of Biochemistry & Cell Biology, 52, 136–145.

Engelhardt, J. F., Yankaskas, J. R., Ernst, S. A., Yang, Y., Marino, C. R., Boucher, R. C., Cohn, J. A., & Wilson, J. M. (1992). Submucosal glands are the predominant site of CFTR expression in the human bronchus. Nature Genetics, 2(3), 240–248.

Fanen, P., Wohlhuter-Haddad, A., & Hinzpeter, A. (2014). Genetics of cystic fibrosis: CFTR mutation classifications toward genotype-based CF therapies. The International Journal of Biochemistry & Cell Biology, 52, 94–102.

Ferreira, A. S. (2011). Insights into the role of extracellular polysaccharides in *Burkholderia* adaptation to different environments. In Frontiers in Cellular and Infection Microbiology (Vol. 1). Frontiers Media SA.

Ferreira, A. S., Leitão, J. H., Sousa, S. A., Cosme, A. M., Sá-Correia, I., & Moreira, L. M. (2007). Functional analysis of *Burkholderia cepacia* genes *bceD* and *bceF* , Encoding a phosphotyrosine phosphatase and a tyrosine autokinase, respectively: role in exopolysaccharide biosynthesis and biofilm formation. In Applied and Environmental Microbiology (Vol. 73, Issue 2, pp. 524–534). American Society for Microbiology.

Ferreira, A. S., Silva, I. N., Oliveira, V. H., Becker, J. D., Givskov, M., Ryan, R. P., Fernandes, F., & Moreira, L. M. (2013). Comparative transcriptomic analysis of the *Burkholderia cepacia* tyrosine kinase *bcef* mutant reveals a role in tolerance to stress, biofilm formation, and virulence. In Applied and Environmental Microbiology (Vol. 79, Issue 9, pp. 3009–3020). American Society for Microbiology.

Ferreira, M. R. (2018). Molecular mechanisms of planktonic cellular aggregation in *Burkholderia multivorans*. Master Thesis from Instituto Superior Técnico.

Figurski, D. H., & Helinski, D. R. (1979). Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proceedings of the National Academy of Sciences, 76(4), 1648–1652.

Fukuda, R., & Okiyoneda, T. (2018). Peripheral protein quality control as a novel drug target for cftr stabilizer. Frontiers in Pharmacology, 9.

Fuqua, W. C., Winans, S. C., & Greenberg, E. P. (1994). Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. Journal of Bacteriology, 176(2), 269–275.

Garland, A. L., Walton, W. G., Coakley, R. D., Tan, C. D., Gilmore, R. C., Hobbs, C. A., Tripathy, A., Clunes, L. A., Bencharit, S., Stutts, M. J., Betts, L., Redinbo, M. R., & Tarran, R. (2013). Molecular basis for pH-dependent mucosal dehydration in cystic fibrosis airways. Proceedings of the National Academy of Sciences, 110(40), 15973–15978.

George, A. M., Jones, P. M., & Middleton, P. G. (2009). Cystic fibrosis infections: treatment strategies and prospects. FEMS Microbiology Letters, 300(2), 153–164.

Gilchrist, F. J., Webb, A. K., Bright-Thomas, R. J., & Jones, A. M. (2012). Successful treatment of cepacia syndrome with a combination of intravenous cyclosporin, antibiotics and oral corticosteroids. Journal of Cystic Fibrosis, 11(5), 458–460.

Gomes, S. C. (2018). Genetic and environmental conditions influencing cellular aggregates formation in *Burkholderia multivorans*. Master Thesis from Instituto Superior Técnico.

Goncheva, M. I., Flannagan, R. S., Sterling, B. E., Laakso, H. A., Friedrich, N. C., Kaiser, J. C., Watson, D. W., Wilson, C. H., Sheldon, J. R., McGavin, M. J., Kiser, P. K., & Heinrichs, D. E. (2019). Stress-induced inactivation of the Staphylococcus aureus purine biosynthesis repressor leads to hypervirulence. Nature Communications, 10(1).

Haaber, J., Cohn, M. T., Frees, D., Andersen, T. J., & Ingmer, H. (2012). Planktonic aggregates of *Staphylococcus aureus* protect against common antibiotics. PLoS ONE, 7(7), e41075.

Hanukoglu, I., & Hanukoglu, A. (2016). Epithelial sodium channel (ENaC) family: Phylogeny, structure–function, tissue distribution, and associated inherited diseases. Gene, 579(2), 95–132.

Hauser, N., & Orsini, J. (2015). Cepacia syndrome in a non-cystic fibrosis patient. Case Reports in Infectious Diseases, 2015, 1–4.

Hoegger, M. J., Fischer, A. J., McMenimen, J. D., Ostedgaard, L. S., Tucker, A. J., Awadalla, M. A., Moninger, T. O., Michalski, A. S., Hoffman, E. A., Zabner, J., Stoltz, D. A., & Welsh, M. J. (2014). Impaired mucus detachment disrupts mucociliary transport in a piglet model of cystic fibrosis. Science, 345(6198), 818–822.

Hollmann, B., Walsh, D. and Perkins, M., n.d. Biofilms and their role in pathogenesis.

Hong, K.-W., Koh, C.-L., Sam, C.-K., Yin, W.-F., & Chan, K.-G. (2012). Quorum quenching revisited—from signal decays to signalling confusion. Sensors, 12(4), 4661–4696.

Inglis, S. K., Corboz, M. R., & Ballard, S. T. (1998). Effect of anion secretion inhibitors on mucin content of airway submucosal gland ducts. American Journal of Physiology-Lung Cellular and Molecular Physiology, 274(5), L762–L766.

Jefferson, K. K. (2004). What drives bacteria to produce a biofilm? FEMS Microbiology Letters, 236(2), 163–173.

Khan, M. A., Ali, Z. S., Sweezey, N., Grasemann, H., & Palaniyar, N. (2019). Progression of Cystic Fibrosis Lung Disease from Childhood to Adulthood: Neutrophils, Neutrophil Extracellular Trap (NET) Formation, and NET Degradation. Genes, 10(3), 183.

Klebensberger, J., Rui, O., Fritz, E., Schink, B., & Philipp, B. (2006). Cell aggregation of *Pseudomonas aeruginosa* strain PAO1 as an energy-dependent stress response during growth with sodium dodecyl sulfate. Archives of Microbiology, 185(6), 417–427.

Kovach, M. E., Phillips, R. W., Elzer, P.H., Roop, R. M. II, Peterson, K. M. (1994). pBBR1MCS: a broad-host-range cloning vector. Biotechniques 16:800–802.

Kragh, K. N., Hutchison, J. B., Melaugh, G., Rodesney, C., Roberts, A. E. L., Irie, Y., Jensen, P. Ø., Diggle, S. P., Allen, R. J., Gordon, V., & Bjarnsholt, T. (2016). Role of multicellular aggregates in biofilm formation. MBio, 7(2).

Kvist, M., Hancock, V., & Klemm, P. (2008). Inactivation of Efflux Pumps Abolishes Bacterial Biofilm Formation. Applied and Environmental Microbiology, 74(23), 7376–7382.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 227(5259), 680–685.

Ledson, M. J., Gallagher, M. J., & Walshaw, M. J. (1998). Chronic *Burkholderia cepacia* bronchiectasis in a non-cystic fibrosis individual. Thorax, 53(5), 430–432.

Lee, J., & Zhang, L. (2014). The hierarchy quorum sensing network in *Pseudomonas aeruginosa*. Protein & Cell, 6(1), 26–41.

Leitão, J. H., Sousa, S. A., Ferreira, A. S., Ramos, C. G., Silva, I. N., & Moreira, L. M. (2010). Pathogenicity, virulence factors, and strategies to fight against *Burkholderia cepacia* complex pathogens and related species. Applied Microbiology and Biotechnology, 87(1), 31–40.

Lewenza, S., Conway, B., Greenberg, E. P., & Sokol, P. A. (1999). Quorum sensing in *Burkholderia cepacia*: identification of the LuxRI homologs CepRI. Journal of bacteriology, 181(3), 748–756.

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

Liu, D., Momb, J., Thomas, P. W., Moulin, A., Petsko, G. A., Fast, W., & Ringe, D. (2008). Mechanism of the quorum-quenching lactonase (AiiA) from *Bacillus thuringiensis*. 1. Product-Bound Structures. Biochemistry, 47(29), 7706–7714.

Lopez, D., Fischbach, M. A., Chu, F., Losick, R., & Kolter, R. (2008). Structurally diverse natural products that cause potassium leakage trigger multicellularity in *Bacillus subtilis*. Proceedings of the National Academy of Sciences, 106(1), 280–285.

Lyczak, J. B., Cannon, C. L., & Pier, G. B. (2002). Lung infections associated with cystic fibrosis. Clinical Microbiology Reviews, 15(2), 194–222.

Lyon, G. J., & Muir, T. W. (2003). Chemical signaling among bacteria and its inhibition. Chemistry & Biology, 10(11), 1007–1021.

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

Mahenthiralingam, E., Vandamme, P., Campbell, M. E., Henry, D. A., Gravelle, A. M., Wong, L. T. K., Davidson, A. G. F., Wilcox, P. G., Nakielna, B., & Speert, D. P. (2001). Infection with *Burkholderia cepacia* complex genomovars in patients with cystic fibrosis: virulent transmissible strains of genomovar iii can replace Burkholderia multivorans. Clinical Infectious Diseases, 33(9), 1469–1475.

Malešević, M., Stanisavljević, N., Novović, K., Polović, N., Vasiljević, Z., Kojić, M., & Jovčić, B. (2020). *Burkholderia cepacia* YtnP and Y2-aiiA lactonases inhibit virulence of *Pseudomonas aeruginosa* via quorum quenching activity. Microbial Pathogenesis, 149, 104561.

Marshall, B., Faro, A., Elbert, A., Fink, A., Cho, J., Rizvi, S., Petren, K., Dorokhina, O., Ostrenga, J., Wu, R., O'Neil, T. D. 2020. Patient registry - Annual data report 2019 cystic fibrosis foundation patient registry.

Martínez-Alemán, S. R., Campos-García, L., Palma-Nicolas, J. P., Hernández-Bello, R., González, G. M., & Sánchez-González, A. (2017). Understanding the entanglement: neutrophil extracellular traps (NETs) in cystic fibrosis. Frontiers in Cellular and Infection Microbiology, 7.

Melaugh, G., Hutchison, J., Kragh, K. N., Irie, Y., Roberts, A., Bjarnsholt, T., Diggle, S. P., Gordon, V. D., & Allen, R. J. (2016). Shaping the growth behaviour of biofilms initiated from bacterial aggregates. PLOS ONE, 11(3), e0149683.

Miller, M. B., & Bassler, B. L. (2001). Quorum sensing in bacteria. Annual review of microbiology, 55(1), 165–199.

Montaudon, M., Berger, P., Cangini-Sacher, A., de Dietrich, G., Tunon-de-Lara, J. M., Marthan, R., & Laurent, F. (2007). Bronchial measurement with three-dimensional quantitative thin-section CT in patients with cystic fibrosis. Radiology, 242(2), 573–581.

Muhdi, K., Edenborough, F. P., Gumery, L., O'Hickey, S., Smith, E. G., Smith, D. L., & Stableforth, D. E. (1996). Outcome for patients colonised with *Burkholderia cepacia* in a Birmingham adult cystic fibrosis clinic and the end of an epidemic. Thorax, 51(4), 374–377.

Naehrig, S., Chao, C.-M., & Naehrlich, L. (2017). Cystic fibrosis [JD]. Dtsch Arztebl Int., 114(33-34): 564–574.

Nealson, K. H., & Hastings, J. W. (1979). Bacterial bioluminescence: its control and ecological significance. Microbiological reviews, 43(4), 496–518.

O'Sullivan, B. P., Freedman, S. D. (2009). Cystic fibrosis. Lancet 373:1891–1904.

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), 8079–8087. https://doi.org/10.1128/jb.01138-07

Pasqua, M., Bonaccorsi di Patti, M. C., Fanelli, G., Utsumi, R., Eguchi, Y., Trirocco, R., Prosseda, G., Grossi, M., & Colonna, B. (2021). Host - bacterial pathogen communication: The wily role of the multidrug efflux pumps of the MFS family. Frontiers in Molecular Biosciences, 8.

Pearson, J.P.; Van Delden, C.; Iglewski, B.H. (1999). Active efflux and diffusion are involved in transport of *Pseudomonas aeruginosa* cell-to-cell signals. J Bacteriol; 181: 1203–10.

Pedersen, S. S., Hoiby, N., Espersen, F., & Koch, C. (1992). Role of alginate in infection with mucoid *Pseudomonas aeruginosa* in cystic fibrosis. Thorax, 47(1), 6–13.

Percival, S. L., Malic, S., Cruz, H., & Williams, D. W. (2011). Introduction to Biofilms. In Springer Series on Biofilms (pp. 41–68). Springer Berlin Heidelberg.

Pessoa, Filipa Duarte. 2017. "Long-term evolution of *Burkholderia multivorans* bacteria during chronic respiratory infections of cystic fibrosis patients," Master Thesis, Instituto Superior Técnico.

Pezzulo, A. A., Tang, X. X., Hoegger, M. J., Abou Alaiwa, M. H., Ramachandran, S., Moninger, T. O., Karp, P. H., Wohlford-Lenane, C. L., Haagsman, H. P., van Eijk, M., Bánfi, B., Horswill, A. R., Stoltz, D. A., McCray, P. B., Welsh, M. J., & Zabner, J. (2012). Reduced airway surface pH impairs bacterial killing in the porcine cystic fibrosis lung. Nature, 487(7405), 109–113.

Polgreen, P. M., Brown, G. D., Hornick, D. B., Ahmad, F., London, B., Stoltz, D. A., & Comellas, A. P. (2018). CFTR heterozygotes are at increased risk of respiratory infections: A population-based study. Open Forum Infectious Diseases, 5(11).

Proesmans, M., Vermeulen, F., & De Boeck, K. (2008). What's new in cystic fibrosis? From treating symptoms to correction of the basic defect. European Journal of Pediatrics, 167(8), 839–849.

Quintana-Gallego, E., Delgado-Pecellín, I., & Calero Acuña, C. (2014). CFTR Protein repair therapy in cystic fibrosis. Archivos de Bronconeumología (English Edition), 50(4), 146–150.

Rabin, N., Zheng, Y., Opoku-Temeng, C., Du, Y., Bonsu, E., & Sintim, H. O. (2015). Biofilm formation mechanisms and targets for developing antibiofilm agents. Future Medicinal Chemistry, 7(4), 493–512.

Rafeeq, M. M., & Murad, H. A. S. (2017). Cystic fibrosis: current therapeutic targets and future approaches. Journal of Translational Medicine, 15(1).

Rao, X., Huang, X., Zhou, Z., & Lin, X. (2013). An improvement of the 2ˆ(-delta delta CT) method for quantitative real-time polymerase chain reaction data analysis. Biostatistics, bioinformatics and biomathematics, 3(3), 71–85.

Rasmussen, T. B., & Givskov, M. (2006). Quorum-sensing inhibitors as anti-pathogenic drugs. International Journal of Medical Microbiology, 296(2–3), 149–161.

Ratjen, F., Bell, S. C., Rowe, S. M., Goss, C. H., Quittner, A. L., & Bush, A. (2015). Cystic fibrosis. Nature Reviews Disease Primers, 1(1).

Reik, R., Spilker, T., & LiPuma, J. J. (2005). Distribution of *Burkholderia cepacia* complex species among isolates recovered from persons with or without cystic fibrosis. Journal of Clinical Microbiology, 43(6), 2926–2928.

Remans, K., Vercammen, K., Bodilis, J., & Cornelis, P. (2010). Genome-wide analysis and literature-based survey of lipoproteins in *Pseudomonas aeruginosa*. Microbiology, 156(9), 2597–2607.

Rémy, B., Plener, L., Decloquement, P., Armstrong, N., Elias, M., Daudé, D., & Chabrière, É. (2020). Lactonase specificity is key to quorum quenching in *Pseudomonas aeruginosa*. In Frontiers in Microbiology (Vol. 11). Frontiers Media SA.

Romling, U., Fiedler, B., Bosshammer, J., Grothues, D., Greipel, J., von der Hardt, H., & Tummler, B. (1994). Epidemiology of chronic *Pseudomonas aeruginosa* infections in cystic fibrosis. Journal of Infectious Diseases, 170(6), 1616–1621.

Russel, J. S. 2001. Molecular cloning: A laboratory manual. New York.

Saint-Criq, V., & Gray, M. A. (2016). Role of CFTR in epithelial physiology. Cellular and Molecular Life Sciences, 74(1), 93–115.

Sause, W. E., Balasubramanian, D., Irnov, I., Copin, R., Sullivan, M. J., Sommerfield, A., Chan, R., Dhabaria, A., Askenazi, M., Ueberheide, B., Shopsin, B., van Bakel, H., & Torres, V. J. (2019). The purine biosynthesis regulator PurR moonlights as a virulence regulator in *Staphylococcus aureus*. Proceedings of the National Academy of Sciences, 116(27), 13563–13572.

Sbordone, L., & Bortolaia, C. (2003). Oral microbial biofilms and plaque-related diseases: microbial communities and their role in the shift from oral health to disease. Clinical Oral Investigations, 7(4), 181–188.

Schleheck, D., Barraud, N., Klebensberger, J., Webb, J. S., McDougald, D., Rice, S. A., & Kjelleberg, S. (2009). *Pseudomonas aeruginosa* PAO1 preferentially grows as aggregates in liquid batch cultures and disperses upon starvation. PLoS ONE, 4(5), e5513.

Schwab, U., Abdullah, L. H., Perlmutt, O. S., Albert, D., Davis, C. W., Arnold, R. R., Yankaskas, J. R., Gilligan, P., Neubauer, H., Randell, S. H., & Boucher, R. C. (2014). Localization of *Burkholderia cepacia* complex bacteria in cystic fibrosis lungs and interactions with *Pseudomonas aeruginosa* in hypoxic mucus. Infection and Immunity, 82(11), 4729–4745.

Sheppard, M. N., & Nicholson, A. G. (2002). The pathology of cystic fibrosis. Current Diagnostic Pathology, 8(1), 50–59.

Silva, A., Amorim, A., Azevedo, P., Lopes, C., & Gamboa, F. (2016). Cystic fibrosis – characterization of the adult population in Portugal. Revista Portuguesa de Pneumologia (English Edition), 22(3), 141–145.

Silva, I. N., Ramires, M. J., Azevedo, L. A., Guerreiro, A. R., Tavares, A. C., Becker, J. D., & Moreira, L. M. (2017). Regulator LdhR and d-lactate dehydrogenase LdhA of *Burkholderia multivorans* play roles in carbon overflow and in planktonic cellular aggregate formation. Applied and Environmental Microbiology, 83(19).

Silva, I. N., Santos, P. M., Santos, M. R., Zlosnik, J. E. A., Speert, D. P., Buskirk, S. W., Bruger, E. L., Waters, C. M., Cooper, V. S., & Moreira, L. M. (2016). Long-term evolution of *Burkholderia multivorans* during a chronic cystic fibrosis infection reveals shifting forces of selection. MSystems, 1(3).

Sio, C. F., Otten, L. G., Cool, R. H., Diggle, S. P., Braun, P. G., Bos, R., Daykin, M., Cámara, M., Williams, P., & Quax, W. J. (2006). Quorum quenching by an n -acyl-homoserine lactone acylase from *Pseudomonas aeruginosa* PAO1. In Infection and Immunity (Vol. 74, Issue 3, pp. 1673–1682). American Society for Microbiology.

Springman, A. C., Jacobs, J. L., Somvanshi, V. S., Sundin, G. W., Mulks, M. H., Whittam, T. S., Viswanathan, P., Gray, R. L., LiPuma, J. J., & Ciche, T. A. (2009). Genetic diversity and multihost pathogenicity of clinical and environmental strains of *Burkholderia cenocepacia*. Applied and Environmental Microbiology, 75(16), 5250–5260.

Tarran, R., Loewen, M. E., Paradiso, A. M., Olsen, J. C., Gray, M. A., Argent, B. E., Boucher, R. C., & Gabriel, S. E. (2002). Regulation of murine airway surface liquid volume by CFTR and Ca2+-activated Cl− conductances. Journal of General Physiology, 120(3), 407–418.

Tettmann, B., Dötsch, A., Armant, O., Fjell, C. D., & Overhage, J. (2014). Knockout of extracytoplasmic function sigma factor ecf-10 affects stress resistance and biofilm formation in *Pseudomonas putida* KT2440. Applied and Environmental Microbiology, 80(16), 4911–4919.

Tseng, T.T., Gratwick, K.S., Kollman, J., Park, D., Nies, D.H., Goffeau, A., Saier, M.H. Jr. (1999). The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human

disease and development proteins. J Mol Microbiol Biotechnol. Aug;1(1):107-25. PMID: 10941792.

Tunney, M. M., Field, T. R., Moriarty, T. F., Patrick, S., Doering, G., Muhlebach, M. S., Wolfgang, M. C., Boucher, R., Gilpin, D. F., McDowell, A., & Elborn, J. S. (2008). Detection of anaerobic bacteria in high numbers in sputum from patients with cystic fibrosis. American Journal of Respiratory and Critical Care Medicine, 177(9), 995–1001.

Wat, D., Gelder, C., Hibbitts, S., Cafferty, F., Bowler, I., Pierrepoint, M., Evans, R., & Doull, I. (2008). The role of respiratory viruses in cystic fibrosis. Journal of Cystic Fibrosis, 7(4), 320–328.

Welsh, M. and Smith, A., 1995. Cystic Fibrosis the genetic defects underlying this lethal disease have now been shown to eliminate or hobble a critical channel through which a constituent of salt enters and leaves cells. Scientific American, Inc.

Yabuuchi, E., Kosako, Y., Oyaizu, H., Yano, I., Hotta, H., Hashimoto, Y., Ezaki, T., Arakawa, M. (1992). Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* Palleroni and Holmes 1981 comb. nov. Microbiol Immunol., 36(12):1251–75.

Yang, J. H., Spilker, T., & LiPuma, J. J. (2006). Simultaneous coinfection by multiple strains during *Burkholderia cepacia* complex infection in cystic fibrosis. Diagnostic Microbiology and Infectious Disease, 54(2), 95–98.

Yeom, D. H., Kim, S.-K., Lee, M.-N., & Lee, J.-H. (2013). Pleiotropic effects of acyltransferases on various virulence related phenotypes of *Pseudomonas aeruginosa*. Genes to Cells, 18(8), 682–693.

Yin, W., Wang, Y., Liu, L., & He, J. (2019). Biofilms: The microbial “protective clothing” in extreme environments. International Journal of Molecular Sciences, 20(14), 3423.

Young, R. L., Malcolm, K. C., Kret, J. E., Caceres, S. M., Poch, K. R., Nichols, D. P., Taylor-Cousar, J. L., Saavedra, M. T., Randell, S. H., Vasil, M. L., Burns, J. L., Moskowitz, S. M., & Nick, J. A. (2011). Neutrophil extracellular trap (NET)-mediated killing of *Pseudomonas aeruginosa*: evidence of acquired resistance within the CF airway, independent of CFTR. PLoS ONE, 6(9), e23637.

Yu, D., Xia, M., Zhang, L., Song, Y., Duan, Y., Yuan, T., Yao, M., Wu, L., Tian, C., Wu, Z., Li, X., Zhou, J., & Qiu, D. (2017). RpoN (σ54) is required for floc formation but not for extracellular polysaccharide biosynthesis in a floc-forming *Aquincola tertiaricarbonis* strain. Applied and Environmental Microbiology, 83(14).

https://www.technologynetworks.com/immunology/articles/care-in-the-community-how-biofilms-improve-bacterial-survival-311685, accessed 13/12/2020.

http://2013.igem.org/Team:NTU-Taida/Project/Background, retrieved 20/12/1010.

https://www.therespiratorysystem.com/, accessed 12/12/2020.

6. Supplementary material

Table S1: Amino acid sequences used in the alignment of representative sequences of YtnP protein present in different species.

	Sequences
***B. multivorans* P0426-1**	MSATIHNVRPVTDNLGRGGSGELVPSRYAVRVGDVDVVLISDGVLPLPTSTMSTNVSEADRNAWFD GRFLQRDMFDWALNIALVSSGERLILIDSGVGDGFEYFTRAGQSMMRLESAGIDLAAITDIVITHMHMD HVGGLNVDGVKARLRPDVRIHVSAKEVAFWKDPDFSKTVMPETVPPALRKAAAKFVERYRENIVPFD QRVEIAAGVSARVTGGHTPGHCVVDVASKGEKLTFVGDAIFEVNFDHPDWQNGFEHDPEASVDVRIA LLREAAETGAMLAAAHVAFPSIGHIAKNGDRFRFVPVTWDY
***B. multivorans* P0426-2**	MSATIHNVRPVTDNLGRGGSGELVPSRYAVRVGDVDVVLISDGVLPLPTSTMSTNVSEADRNAWFD GRFLQRDMFDWALNIALVRSGERLILIDSGVGDGFEYFTRAGQSMMRLESAGIDLAAITDIVITHMHMD HVGGLNVDGVKARLRPDVRIHVSAKEVAFWKDPDFSKTVMPETVPPALRKAAAKFVERYRENIVPFD QRVEIAAGVSARVTGGHTPGHCVVDVASKGEKLTFVGDAIFEVNFDHPDWQNGFEHDPEASVDVRIA LLREAAETGAMLAAAHVAFPSIGHIAKNGDRFRFVPVTWDY
***B. multivorans* R-20526**	MSATIHNVRPVTDNLGRGGSGELVPSRYAVRVGDVDVVLISDGVLPLPTSTMSTNVSEADRNAWFD GRFLQRDMFDWALNIALVRSGERLILIDSGVGDGFEYFTRAGQSMMRLESAGIDLAAITDIVITHMHMD HVGGLNVDGVKARLRPDVRIHVSAKEVAFWKDPDFSKTVMPETVPPALRKAAAKFVERYRENIVPFD QRVDIAAGVSARVTGGHTPGHCVVDVASKGEKLTFVGDAIFEVNFDHPDWQNGFEHDPEASVDVRIA LLREAAETGAMLAAAHVAFPSIGHIAKNGDRFRFVPVTWDY
***B. cepacia* LO6**	MSETIRNVQPVVVNLGKEGAGELVPSRYAVRIGDVDVVLISDGVLPLPTSTMSTNVSEADRNAWFDG RFLQRDMFDWALNVALVRSGDRLILIDSGVGDGFEYFTRAGQSMMRLESAGIDLAAITDIVITHMHMD HVGGLNLDGVQGKLRPDVRIHVSAAEVEFWKNPDFSKTVMPEAVPPALRDAAARFAKRYAENIVQFD RTVQVAPGVSARVTGGHTPGHCVVDVASNGEKLTFVGDAIFEVNFDHPDWQNGFEHDPQASVDVRI ALLEEAADTGALLAAAHVAFPSIGHIARNGDGFRFVPVLWDY
***B. dolosa* AU0158**	MSETIRNVQPVVVNLGKEGAGELVPSRYAVRIGDVDVVLISDGVLPLPTSTMSTNVSEADRNAWFDG RFLQRDMFDWALNVALVRSGDRLILIDSGVGDGFEYFTRAGQSMMRLESAGIDLAAITDIVITHMHMD HVGGLNLDGVQGKLRPDVRIHVSAAEVEFWKNPDFSKTVMPEAVPPALRDAAARFAKRYAENIVQFD RTVQVAPGVSARVTGGHTPGHCVVDVASNGEKLTFVGDAIFEVNFDHPDWQNGFEHDPQASVDVRI ALLEEAADTGALLAAAHVAFPSIGHIARNGDGFRFVPVLWDY

B. territorii **MSMB2203W**	MNGSIRNVRPVTARLGNEGAGELVPSRYAVRIGDVDVVLISDGVLPLPTSTMSTNVSEADRNEWFDG RFLQRDMFDWALNIALVRSADRLILIDSGVGDGFEYFTRAGRSVMRVESAGIDLAAITDIVITHMHMDH VGGLNVDGVKARLNPDVRIHVSAAELAFWKNPDFSKTVMPETVPPALRKAAARFVELYREHIVPFDRT VEVAAGVSARVTGGHTPGHCVVDIASNGERLTFAGDAIFEVNFDHPDWQNGFEHDPQAATDVRLAL FNEAADTGAMLAAAHVAFPSIGHIARDGDGFRFVPVVWDY
B. lata **383**	MSETTRNVQPVISDLGKEGSGELVPSRYAVRVGDVDVVLISDGILPLPTSTMSTNVSEADRNAWFDG RFLQRDMFDWALNVALVRSGDRLILIDSGVGDGFEYFTRAGKSVMRLESAGIDLSAITDIVITHMHMDH VGGLNVDGVRARLRPDVRIHVSAAEVAFWKNPDFSKTVMPETVPPALRKAAEKFVALYGENIVQFDQ TVEVAAGVSARVTGGHTPGHCVVDIASNGEKLTFAGDAIFEVNFDHPEWQNGFEHDPETAAAVRIAL FNEAAETGALLAAAHVAFPSIGHIARNGDGFRFVPVLWDY
B. ambifaria **AMMD**	MNKSIRNVQPVTAPLGNEEAGELVSSRYAVRIGDVDVVLISDGVLPLPTSTMSTNVSEADRNEWFDG RFLQRDMFDWALNIALVRSGERLILIDSGVGDGFEYFTRAGRSVMRLESAGIDLAAITDIVITHMHMDH VGGLNTDGVKARLSPDVRIHVAAAEVAFWKNPDFSKTVMPETVPPALRKAAAKFVELYGENIVQFDR TVEVAPGVSARVTGGHTPGHCVVDIASKGEKLTFVGDAVFEVNFDNPDWQNGFEHDPEGAVDVRIA LFNEAAETGAMLAAAHVAFPSIGHVAKNGDGFRFVPVTWDY
B. contaminans **MS14**	MSETTRNVQPVISDLGKKGSGELVPSRYAVRVGDVDVVLISDGILPLPTSTMSTNVSEADRNAWFDG RFLQRDMFDWALNIALVRSGERLILIDSGVGDGFEYFTRAGKSVMRLESAGIDLSALTDIVITHMHMDH VGGLNVDGVKAKLRPDVRIHVSAAEVAFWKNPDFSKTVMPETVPPALRKAAEKFVTLYGENIVQFDQ TVEVAAGVSARVTGGHTPGHCVVDIASNGEKLTFAGDAIFEVNFDHPEWQNGFEHDPETAAAVRIAL FNEAAETGALLAAAHVAFPSIGHIAKTGDGFRFVPVLWDY
B. diffusa **RF2-non-BP9**	MNGPIRNVRPVTTRLGKEAAGELVPSRYAVRIGDVDVVLISDGVLPLPTSTMSTNVSEADRNEWFDG RFLQRDMFDWALNIALVRSADRLILIDSGVGDGFEYFTRAGRSVMRLESAGIDLAAITDIVITHMHMDH VGGLNVDGVKARLSPDVRIHVSAAELAFWKNPDFSKTVMPETVPPALRKAAARFVELYREHIVPFDHT VEVAAGVSARVTGGHTPGHCVVDIASDGERLTFSGDAIFEVNFDHPDWQNGFEHDPQAATNVRIALF NEAVETGAMLAAAHVAFPSIGHIARDGDGFRFVPVVWDY
B. anthina **AZ-4-2-10-S1-D7**	MNAPIRNVRPVTARLGNESSGELVSSRYAVRIGDVDVVLISDGVLPLPTSTMSTNVSEADRNAWFDG RFLQRDMFDWALNIALVRSGDRLILIDSGVGDGFEYFTRAGRSVMRLESAGIDLAAITDIVITHMHMDH VGGLNVDGVKAKLRPDVRIHVSAAEVAFWKHPDFSKTVMPETVPPALRKAAARFMELYGEHVVQFD RTVEVAAGVSARVTGGHTPGHCVVDIASNGEKLTFAGDAIFEVNFDNPDWQNGFEHDPEAAVDVRL ALFNEAADTGSMLAAAHVAFPSIGHVARDGDGFRFVPVVWDY
B. cenocepacia **J2315**	MNAPIRNVQPVVACLGHEGSGELVSSRYAVRIGAVDVVLISDGVLPLPTSTMSTNVSETDRNAWFDG RFLQRDMFDWALNIALVRSADRLILIDSGVGDGFEYFTRAGRSVMRLESAGIDLAAITDIVITHMHMDH VGGLNVDGVKAKLNPDVRIHVSAAEVAFWENPDFSKTVMPETVPPALRKAAARFMALYGENIVQFDR TVEVAAGVSARVTGGHTPGHCVVDVASNGEKLTFAGDAIFEVNFDHPDWQNGFEHDPDAAVDVRLA LFNEAADTGALLAAAHVAFPSIGHVARNGDGFRFVPVVWDY
B. seminalis **FL-5-4-10-S1-D7**	MNAPLRNVQPVTDRLGKEGSGELVPSRYAVRIGDVDVVLISDGVLPLPTSTMSTNVSEAERNAWFDG RFLQRDRFDWALNIALVRSADRLILIDSGVGDGFEYFTRAGRSVMRLESAGIDLAAITDIVITHMHMDH VGGLNTDGVKARLRPDVRIHVSAAEMAFWKNPDFSKTVMPETIPPALRKAAARFVTLYGENIVTFDRT VEVATGVTARVTGGHTPGHCVVDVASNGEKLTFAGDAIFEVNFDHPDWQNGFEHDPEAAVDVRIALF NEAADTGALLAAAHVAFPSIGHVARDGSGFRFVPVVWDY

***B. ubonensis* MSMB1600WGS**	MSETTRNVQPVTANLGKEGSGELVPSRYAVRVGDIDVVLISDGVLPLPTSTMSTNVSEADRNEWFDG RFLQRDMFDWALNIALVRSGEQLILIDSGVGDGFEYFTRAGQSVMRLESAGIDLAAITDIVITHMHMDH VGGLNVDGVKARLRPDVRIHVSAAEVAFWKNPDFSKTVMPETVPPALRKAAEKFVKLYSENIVQFDQ VTEVAAGVSARVTGGHTPGHCVVDIASNDEKLTFVGDAIFEVNFDNPAWQNGFEHDPEAAVDVRIAL FDDAARTGALLAAAHVAFPSIGHIAKNGESYRFVPVMWDY
Paraburkholderi a phytofirmans	MSETTRNVQPVTANLGKEGSGELVPSRYAVRVGDVDVVLISDGVLPLPTSTMSTNVSEADRNAWFD GRFLQRDMFDWALNVALVRSGERLILIDSGVGDGFEYFARAGRSVMRLESAGIDLAAITDIVITHMHM DHVGGLNVDGVKARLRPDVRIHVSAAEVEFWKNPDFSKTVMPETVPPALRKAAAKFVELYGENIVRF DQTVEVAAGVSARVTGGHTPGHCVVDVASKGEKLTFVGDAIFEVNFDNPRWQNGFEHEPEIAADVRI ALLNEAAETGALLAAAHVAFPSIGHIAKNGESFRFVPVQWDY
***Cupriavidus* sp. YR651**	MSQTTRNVQPVTANLGKEGSAELVPSRYAVRVGDVDVVLISDGILPLPTSTMSTNVSEADRNKWFDG RFLQRDMFDWALNIALVRSGERLILIDSGVGDGFEYFTRAGQSVMRLESAGIDLAAITDIVITHMHMDH VGGLNVDGVKAKLHPDVRIHVSATEVAFWENPDFSKTVMPETVPPALRKAAAKFVELYSENIVQFDR EVEVAAGVSARVTGGHTPGHCVVDVASNGEKLTFVGDAIFEVGFDKPEWQNGFEHDPEVAVDVRIA LLNEAAETGALLAAAHVAFPSIGHIAKNGEGFRFVPVQWDY
Caballeronia glathei	MSATTRNVQPVTAKIGKEGSGELVPSRYAVRVGDVDVVLISDGVLPLPTSTMSTNVSESDRNEWFDG RFLQRDMFDWALNIALVRSGERLILIDSGVGDGFEYFSRAGQSVMRLESAGIDLTAITDIVVTHMHMD HVGGLNVDGVKARLRPDVRIHVSAAEVEFWKNPDFSKTVMPETVPPALREAAAKFVKLYSENIVQFD QSVEVAEGVSARVTGGHTPGHCVVDVASNGEKLTFVGDAIFEVNFDNPDWQNGFEHDPEAAVDVRI ALFNEAAETGALLAAAHVAFPSIGHVAKNGEGFRFVPVQWDY
***Ralstonia* sp. 25mfcol4.1**	MSETIRNVQPVIANLGRDGPGELVPSRYAVRVGDVDVVLISDGILPLPTSTMSTNVSAADRNEWFDGR FLQRDMFDWALNIALVRSGDRLILIDSGVGDGFEYFTRAGRSVMRLETAGIDLAAITDIVITHMHMDHV GGLNVDGVKARLRPDVRIHVAAAEVEFWKNPDFSKTVMPETVPPALRKAAARFVELYSEQIVQFDQA VHVAAGVTARVTGGHTPGHCVVDVASNGEKLTFVGDAIFEVGFDHPEWQNGFEHEPEVATNVRIALL NEAAETGAMLAAAHIAFPSIGHIAKNGNGFRFVPVQWDY
Bordetella petrii ATCC BAA-461	MSQTTRNVQPVTATLGKENAGELVPSRYAVRVGDIDVVLISDGILPLPTSTMSTNVKESDRNEWFDG RFLQRDMFDWALNIALVRSGDHLILVDSGVGDGFEYFSRAGRSVMRLESAGIDLAAITDIVITHMHMD HVGGLNVDGVKAKLRPDVRIHVAASEVEFWQNPDFSKTVMPEAVPPALRKAAAKFVELYRDNIVPFG QTVEVAAGVSARVTGGHTPGHCVVDVASQGEKLTFVGDAIFEVGFDNPEWQNGFEHEPEKAVDVRI SLLNEAAETGALLAAAHVAFPSIGHIAKNGKSFRFVPVQWDY
Pseudomonas acidophila	MTATTRNVRPVTEKLGKEGSGELVPSRYAVRIGDLDVVLISDGVLPLPTSTMSTNVSEADRNEWFDG RFLQRDTFDWALNIALIRSGERLILVDSGVGDGFEYFTRAGRSVMRLETAGIDLSAITDIVVTHMHMDH VGGLNTDGVKARLRPDVRIHVAAAEVAFWKNPDFSKTVMPETVPPALRRAAAKFVALYGENIVQFEQ TVEVAAGVSARVTGGHTPGHCVVDVASNGEKLTFVGDAIFEVNFDNPAWQNGFEHDPEGAVDVRIA LFNEAAETGALLAAAHVAFPSIGHIAKHGDGFRFVPVQWDY

***Rhodanobacter* sp. T12-5**	MSATARNVQAVTANIGKDDSGELVSSRYAVRIGDVDVVLISDGILPLPTSTMSTNVSEAERNQWFDHR FLQRDMFDWALNIALVRSGDRLILIDSGVGDGFEYFSRAGRSVLRLETAGIDLAAITDIVVTHMHMDHV GGLNVDGIKAKLRPDVRIHVAAAEVEFWKHPDFSRTVMPEAVPPALRKAAARFVELYSDNIVPFDRTV EVAAGVTARVTGGHTPGHCVVDVASQGEKLTFVGDMIFEVGFDHPTWQNGFEHEPEVAATARIALLN EAAETGSLLAAAHIAFPSIGHVAKNGESFRFVPVQWDY
Sphingomonas haloaromaticamans	MKLETIAKPGRTDADDLVPSRYALSVGSLDVLIISDGVLPLPTETMSTNVTPDARAAWFNGMFLGPDM FDWALNVLVVRSGEQIILVDAGLGGQFPGFPRAGQFPRRLEAAGIPLEAVTDVVITHMHMDHVGGLLV DEVRKRLRPDVRIHVTSTEVEFWKDPDFSLTSMPSPVPDVLRSTAAQFVEAYGDKIRTFGNELKVAP GVTAKLTGGHTPGHCVVRVESNGERLTFAGDALFPVGFDHPDWHNGFEHDPEESVRVRVDLLREAA ATGELLVATHLAFPSLGRVALDGDAFRWVASIWDY
Sphingobium baderi	MKLETIAKPGRTDADDLVPSRYALSVGGLDVLIVSDGVLPLPTETMSTNVTPDARAAWFNDMFLGPD MFDWALNVLVVRSGEQVILVDAGLGGQFPGFPRAGQFPKRLEAAGIPLEAVTDVVITHMHMDHVGGL LVDEVRSRLRPDVRIHVTSTEVEFWKDPDFSLTSMPSPVPDVLRSTAAEFVQAYGDKISTFENELEVA PGVTAKLTGGHTPGHCVVRVESNGERLTFAGDALFPVGFDHPDWHNGFEHDPEEAVRVRVDLLREA AQSGELLVATHLAFPSLGRVALDGDAFRWVASIWDY
Pseudomonas citronellolis	MNMHLNYSVQKNSAPKNAQHQEREELVPSRYALRVGEIDVLVVSDGVLPLPTATMSTNAEPAARAA WFKDMFLGPEAFDWALNVLVVRSGEQTILVDAGLGGQFPGFPRAGQFPRRLEAAGIDLASVTDVVIT HMHMDHIGGLLVDEVKNRLRPDVRIHVAATEVAFWAAPDFSLTSMPAPVPDVLRATAGQFMEAYRD KLHTFEDEHEVAPGVVARLTGGHTPGHSVVHVTSAGERLTFAGDALFPVAFDHPDWHNGFEHDPEE SVRVRVRLMQEAAASGELLVATHLPFPSVGRVAVDGDAFRWIAAFWDY
Parvibaculum lavamentivorans	MTTHLSYVPQKSGAEFEELVPSRYALRIGEIDVLVVSDGVLPLPTETMATNVAPAVRAAWFKDMFLGP DMFDWALNVLVVRSGDKTILIDAGLGGQFPGFPRAGLFPKRLEAAGIDLASVTDIIVTHMHMDHVGGL LVDDVKKRLRPDVRIHVSATEVDFWGTPNFSQTAMPTPVPDVLRTTATEFLKVYGANLQTFAEEREV APGVVARLTGGHTPGHSVVHVTSGGERLTFAGDALFPVGFDHPDWHNGFEHDPEESVRVRVRLLQ EAAATGELLVATHLSFPSVGRVALDGEAFRWVAASWDF
***Luteibacter* sp. OK325**	MSAVHSIPHTVEATSRDLVPSRYALKIGEIDVMVISDGVLPIPFEVMSTNVAPAGRAAWLDNMLLPESF DWAVNVVVVRSGEQIILIDSGLGGEFEGFPRAGQLVKRLEAAGIDLASLTDVVLTHLHMDHVGGLLGD GVKERLRPDVRIHVAAAEVAFWESPDFSRTVMPPPIPDVLRATAKQFLADYRGQLRQFDEEYEVAPG VVARRTGGHTPGHCVVRIASGGEALTFAGDAVFPVAFDHPDWQNGFEHDPEEAVNVRLRLFRELAA SGELLVATHLSFPSVGRVAIDGDAFRWVPGYWDY
***Massilia* sp. Root1485**	MLTVDSPLRPASHGPEELVPSRYALKVGDIDVMVVSDGVLPLPATTMATNADPADLADWLDHMFLPP DKFDWPLNVLVVRSGDQTILVDAGVGTQFPGFPRAGQFPARLEAVGIDLASVTDIVITHMHMDHVGG LLAPGVKERLRPDVRIHVTDAEVKFWASPDFSLTVMPEAVPPVLRSTAARFLAEYRDRLLVFKERREV APGVVARRTGGHTPGHCVVDLASGGKRLMFAGDAVFPVNFEHPGWQNGFEHDPEEAAKVRDRLFR ELMIDGGLLVAAHLPFPSIGRVMADGDAYRYVPVFWDY

***Achromobacter* sp. Root565**	MFTADTLDPITHQAIDLVPSRYAVQVGDIDALVVSDGVLPLPTSTMATNADPADLANWLDHMFMPPDA FDWPLNVMVARSGDQTILIDAGLGGQFSGFPRAGQFPQRLQSAGIPLEAVTDVIITHMHMDHVGGLLV PGVKERLRKDVRIHVNATEVAFWTSPDFSQTVMPKPVPAVLRSTATSFYETYRDQIQVFEDRKEVAP GVIVRLTGGHTPGHCVVDLISGGERLMFAGDAIFPVGFDHPEWHNGFEHDPEESARVRVGLFRELAE NKGLLVAAHLPFPSVGRVAVDGDAFRWVPVIWDF
Variovorax paradoxus	MFSVNNNNLPPGCQPSEIVPSRYAVEVGDIDALVVSDGVLPLPTKTLATNADPVDLANWLDHMFMPP DAFDWPLNVMVARSGEQTILIDAGLGGQFSGFPRAGQFPQRLASAGIEIESVTDVIITHMHMDHVGGL LVEGVKQRLRPDVRIHVSAAEVAFWTAPDFTQTVMPNPVPEVLRSTARSFYEVYRDKLRIFEDKCEVA PGVHVRLTGGHTPGHSVVDLVSRGERLTFAGDAIFPVGFDHPDWHNGFEHDPEEAARVRMDLFREL AENGGLLVAAHLPFPSVGHVAADGDVFRWVPVIWDF
Pandoraea fibrosis	MPTSDNILFASGNRPVDLVPSRYALRVGDIDAMVISDGVLPLPTATMATNVAPSDLAAWLDSMCMPP DAFDWPLNVMVVRSGSQTILIDAGLGGQFPGFPRAGQFPSRLQSAGIDLESVTDIIVTHMHMDHVGG LLVPEVKARLSPNVRIHVSATEVAFWTSPDFSHTVMPTPVPAVLRTTATSFYNEYREQLHVFEETREV APGVIARITGGHTPGHCVVDIVSGDERLTFAGDAMFPVGFDHPDWQNGFEHDPEESTRVRIGLFEEL AENRGLLVAAHLPFPSVGRIARNGDAYRWVPVLWDF
***Ensifer* sp. Root278**	MSPLEARSHTNTTLEPAPPGRSAPEELVPSRYAVRVGDIEVLVVSDGVLPLPTTMLGHNADPADRAA WLGDMFLPQDAFDWSLNVVVVRSGAQTILIDAGLGMDPDLNLPRAGQLIRRLEAAGIDLAAVTDVVLT HMHMDHIGGLLVDGVKEQLRPDLRIHVAAAEIKFWESPDFSHAVMPQGFPDALRATAKRFANEYRNH LRPFEDECEVAPGVVAQRTGGHTPGHCVVRVASGGDRLMFAGDALFAVAFDHPDWYNGFEHDPEE AARVRVRLMKELAETGSSLVATHLPFPSVGHVAIDGDRFRWVPVFWDY
Polyangium fumosum	MSVDNTPHRGKPASDELVPSRYALRIGEIDVLVISDGVLPLPFATMATNVEPAVRTAWLDNLFLPDKF DWALNVIVVRSGGRTILVDSGLGKEFSGFPRAGQLALRLEAAGIDLADLTDVVLTHLHMDHIGGLLGD GIRERLRPDLRIHVAAAEVNFWAAPDFSRTTMPSPIPDVLRATAQRFLKEYHSLLRPFEEEHEVAAGV VVRRTGGHTPGHSVVRLASGGERLTFAGDAVFPVSFEHPEWHNGFEHDPEASVNVRIRLFRELSAT RELLVATHLSFPSVGRVAVDGDAFRFVPASWDY
***Phenylobacterium* sp.**	MKLDKISPAGGTASADLVPSRYALRIGEIDVLVISDGILPLATSIMSSNVEPAARTAWFKDMFLGPDAFD WPLNVLVVRSGEQIILVDAGLGGQFPGFPRAGQFPQRLEAAGIDLGSVTDVVITHMHMDHVGGLLVD DVKSRLRPDLRIHVSAAEVDFWKSPDFSRTVMPPPIPEVLRSTATEFAQMYGDRLQTFEQEHEVAPG VVVQRTGGHTPGHSIVRVTAGGERLTFAGDALFPVAFDHPDWQNGFEHDPEEAVRVRVRLMQDAA ATREILVATHLPFPSVGRVAIEGDAFRWVPAVWDY
Bradyrhizobium icense	MKQETNMDVDTSLANTSHPGGPGSNELVPSRYAVQVGDIDVLVVSDGVLPLPTAMLAHNIDPAVRAT WLKDMFLPPDAYDWALNVVVVRSGDRTILVDAGLGLDPDLHLPRAGQLIKRLEAAGIDLASVTDVVLT HMHMDHIGGLLVEGVKERLRPDLRIHVAAAEVKFWEAPDFSHVSMPPGFPDALRSTAKRFVKEYRSQ LRPFEEEYEVAPGVVVTRTGGHTPGHSVVRVASGGDRLTFAGDLVFAVGFEHPEWYNGFEHDPEES ARVRISLLRELAETGGLLVATHLPFPSVGHVAVDGDAFRWVPVFWDY