

**Adaptation of *Burkholderia cepacia* complex bacteria to the cystic fibrosis lung: focus on lipopolysaccharide O-antigen and cell wall physical properties**

**Ahmed Amir Elsayed Hassan**

**Supervisor:** Doctor Isabel Maria de Sá Correia Leite de Almeida

**Thesis approved in public session to obtain the PhD Degree in Biotechnology and Biosciences**

**Jury final classification: Pass with Distinction**

**2020**



**UNIVERSIDADE DE LISBOA**  
**INSTITUTO SUPERIOR TÉCNICO**

**Adaptation of *Burkholderia cepacia* complex bacteria to  
the cystic fibrosis lung: focus on  
lipopolysaccharide O-antigen and cell wall physical properties**

**Ahmed Amir Elsayed Hassan**

**Supervisor:** Doctor Isabel Maria de Sá Correia Leite de Almeida

**Thesis approved in public session to obtain the PhD Degree in  
Biotechnology and Biosciences**

**Jury final classification: Pass with Distinction**

**Jury**

**Chairperson:** Doctor Duarte Miguel de França Teixeira dos Prazeres, Instituto Superior Técnico, Universidade de Lisboa

**Members of the committee:**

Doctor Isabel Maria de Sá Correia Leite de Almeida, Instituto Superior Técnico, Universidade de Lisboa

Doctor Arsénio do Carmo Sales Mendes Fialho, Instituto Superior Técnico, Universidade de Lisboa

Doctor Paula Alexandra Quintela Videira, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

Doctor Mário Manuel Silveira Rodrigues, Faculdade de Ciências, Universidade de Lisboa

Doctor João Paulo dos Santos Gomes, Instituto Nacional de Saúde Doutor Ricardo Jorge

**Funding Institution - Fundação para a Ciência e Tecnologia**

BIOTECnico doctoral program PD/BD/113629/2015

**2020**



To my parents, Amir and Ihsan  
To my wife "Kaotar" and our daughter "Lyna"  
To my sister, my brothers and our family members



## Acknowledgment

This PhD has been a truly life-changing experience for me and it would not have been possible to do without the support and guidance that I received from many people. First and foremost, I would like to express my sincere gratitude to Professor Isabel Sá-Correia, my supervisor, coordinator of the PhD program “BIOTECnico” and head of the Biological Sciences Research Group (BSRG) of the Institute for Bioengineering and Biosciences (iBB), Instituto Superior Técnico. I acknowledge her for accepting me as a nominated PhD student under her irreplaceable supervision, guidance, support and patience. Her continuous commitment got me to be who I am today. I am literally most thankful for her help in writing as well as the scientific formation.

I am also grateful to my thesis committee members “CAT – members”, Professor Arsenio Fialho and Professor Jorge Leitão, for their suggestions and their interest in the presented work. Additionally, I acknowledge all the collaboration of all co-authors of the papers presented in this thesis, in particular Professor Miguel Valvano, Professor Vaughn Cooper and Professor Antonio Molinaro. I would like to express my profound thanks to Professor Mário Rodrigues, not only for being collaborator, but for having shared with me ideas, knowledge and some skills of scientific writing. My gratefulness to all doctors of the Bacteriology Laboratory, Santa Maria Hospital, for the contribution in the isolation and culture of the strains under study and those of the Adult Cystic Fibrosis Centre, where the patients were followed, for making available their clinical data. I would like to thank Dr. Sandra Santos and Dr. Carla Coutinho not only for their contributions to several works here presented but for being there most of the time needed.

The financial support from the Portuguese Foundation for Science and Technology (FCT) is gratefully acknowledged, especially for the BIOTECnico PhD fellowship (PD/BD/113629/2015). Funding received by iBB – Institute for Bioengineering and Biosciences from FCT (UID/BIO/04565/2019 and UID/BIO/04565/2013) and from Programa Operacional Regional de Lisboa 2020 (Project N. 007317) are also acknowledged. Part of this work was financially supported by FCT project contract: PTDC/ FIS-NAN/6101/2014. Another part of this research was also performed under the COST Action BM1003 “Microbial cell surface determinants of virulence as targets for new therapeutics in Cystic Fibrosis”.

Without doubts, this journey has revealed the real ones who stands next to you despite when and where, namely Claudia, Pedro, Sara, Rui and Rita. I should thank Ana and Joana, the survivors of the former PhDs on *Burkholderia*. I would like to thank everyone from BSRG, former and present members, particularly Rita, Mariana, Carina, Mónica, Nuno, Dalila, Luis, Margarida, Ricardo, Marta, Nuno, Miguel, Joana and Silvia. The memory is full of individuals who deserve to be thanked. Many

thanks to the aforementioned individuals and all others who I met (as Miguel, Tiago and Pedro from the AFM lab) for their significant inputs, motivational words and constant support.

Finally, given that the PhD is possible to be accomplished, my life lesson in this journey was that, to succeed, you have to work, harder and smarter, to overcome the obstacles on your path. To do so, I have been away of every one of my family who I would like to apologize and thank for their understanding; my parents (Ihsan and Amir), my love (Kaotar), my sister (Maha), my brothers (Mohammed and Maged) and our family members. They have been the support that I have needed.

**A Amir Hassan**

## Abstract

The *Burkholderia cepacia* complex (Bcc) comprises 24 closely related opportunistic pathogenic species with high potential to cause serious chronic respiratory infections in cystic fibrosis (CF) patients. The CF lung is a hostile environment due to host immune responses, continuous antibiotic therapy, high levels of oxidative stress and low oxygen concentration. This thesis work contributed to the understanding of the adaptive strategies of Bcc bacteria to the CF lung at a genome-wide level and with a focus on the cell envelope. The genome sequences of twenty *B. cenocepacia* and *B. multivorans* clonal variants that co-inhabited the same CF patient-lung environment, for at least 3 years, were compared. Mutation rates were in the range of 2.08-2.27 SNPs/year leading to rapid genotypic and phenotypic diversification. The comparative genomic analysis indicated that the lipopolysaccharide O-antigen locus undergoes alterations during chronic infection. A novel hybrid O-antigen locus, encoding a novel O-antigen structure only present in the first *B. cenocepacia* isolate, was compared with the loci of 10 subsequent clonal variants that lack the O-antigen. A systematic retrospective longitudinal screening, including 357 isolates from 19 chronically infected patients, representing 21 strains of six Bcc species, was performed to know if the loss of the O-antigen during chronic infection can be considered a general phenomenon. Only isolates from the two most prevalent and feared species, *B. cenocepacia* and *B. multivorans*, showed the referred tendency. To understand the role of long-term infection in *B. cenocepacia* cell surface properties, atomic force microscopy was used to study cell adhesion, cell morphology and mechanical properties. The ability of the early isolate to adhere to AFM Si<sub>3</sub>N<sub>4</sub>-tip was significantly higher compared with late variants that lack the O-antigen and exhibit identical adhesion values. A reduction of cell size and cell shape evolution from the rod form typical of the species to a coccoid-like form was also observed.

**Keywords:** *Burkholderia cepacia* complex; Chronic infection; Cystic fibrosis; Adaptation and evolution; Cell envelope.



## Resumo

O complexo *Burkholderia cepacia* (Bcc) inclui 24 espécies patogênicas oportunistas muito relacionadas, com elevado potencial de causar infecções respiratórias crônicas graves em pacientes com fibrose quística (CF). O pulmão de CF é um ambiente hostil devido às respostas imunes do hospedeiro, à terapia continuada com antibióticos, elevados níveis de stresse oxidativo e baixa concentração de oxigênio. Este trabalho de tese vem contribuir para a compreensão das estratégias adaptativas daquelas bactérias ao pulmão de CF, envolvendo todo o genoma e foco no envelope celular. Foram comparadas as sequências genômicas de vinte variantes clonais de *B. cenocepacia* e *B. multivorans* que co-habitaram o mesmo ambiente do pulmão de um doente com FQ, durante pelo menos três anos. As frequências de mutação encontravam-se na gama 2,08-2,27 SNPs/ano, levando à rápida diversificação genotípica e fenotípica. A análise genômica comparativa indicou que o locus do O-antígeno (OAg) do lipopolissacárido (LPS) sofre alterações durante a infecção crônica. Um novo locus híbrido para o OAg, codificando uma nova estrutura presente apenas no primeiro isolado de *B. cenocepacia*, foi comparado com os loci de 10 variantes clonais subsequentes que não possuem o OAg. Foi realizado um estudo longitudinal e retrospectivo que incluiu 357 isolados de 19 pacientes infectados cronicamente, representando 21 estirpes de seis espécies de Bcc, para saber se a perda do OAg durante a infecção crônica pode ser considerada um fenômeno geral. Somente os isolados das duas espécies mais prevalentes e temidas, *B. cenocepacia* e *B. multivorans*, apresentaram a tendência referida. Para entender o papel que a infecção crônica exerce nas propriedades da superfície celular de *B. cenocepacia*, foi usada a microscopia de força atômica (AFM) para estudar a adesão celular, morfologia celular e propriedades mecânicas. A capacidade de adesão do isolado inicial é significativamente superior à dos variantes tardios que não possuem o OAg e exibem valores de adesão idênticos. Também foi observada uma redução do tamanho das células e a sua evolução da forma bastonete típica da espécie para a forma cocóide.

**Palavras-chave:** Complexo *Burkholderia cepacia*; Infecção crônica; Fibrose quística; Adaptação e evolução; Invólucro celular.



## Resumo Alargado

As infecções respiratórias crônicas, incluindo as causadas por bactérias do complexo *Burkholderia cepacia* (Bcc), continuam a ser das mais importantes causas de morte prematura dos doentes com fibrose quística (CF). Os pulmões dos doentes com CF constituem ambientes dinâmicos caracterizados pela ação de uma forte resposta imunitária do hospedeiro, por elevadas e continuadas concentrações de antibióticos e elevados níveis de stresse oxidativo. É possível que estirpes de diferentes espécies do Bcc apresentem diferentes trajetos de evolução adaptativa em resposta aos vários stresses ambientais a que se encontram expostas no pulmão de CF. Até ao presente, os estudos realizados para compreender a diversificação de bactérias do Bcc nos pulmões de CF e o seu potencial para evoluir em resposta a pressão seletiva exercida por fármacos e a resposta imune são limitados quando comparados com o conhecimento que existe para *Pseudomonas aeruginosa*. Consequentemente, o trabalho desta tese veio contribuir para o avanço do conhecimento no que respeita às estratégias adaptativas usadas por bactérias do Bcc durante infecções respiratórias crônicas, quer ao nível de todo o genoma quer focado no invólucro celular.

Apesar dos vários estudos de genómica comparativa realizados anteriormente, utilizando variantes clonais de isolados clínicos de Bcc obtidos durante a infeção crónica, não há informação quanto aos padrões evolutivos que ocorrem em estirpes de diferentes espécies de Bcc quando são expostas a um mesmo ambiente pulmonar de CF durante uma co-infeção. Este trabalho vem contribuir para colmatar essa falha já que compara sequências do genoma de variantes clonais de *B. cenocepacia* e *B. multivorans* (as espécies de Bcc mais prevalentes mundialmente em pacientes de CF) que co-habitaram o mesmo ambiente seletivo do hospedeiro durante um período de pelo menos 3 anos. Neste caso de estudo retrospectivo, foram comparados vinte variantes clonais derivados de duas estirpes ancestrais (9 isolados de *B. multivorans* e 11 isolados de *B. cenocepacia*). Ambas as espécies apresentaram taxas de mutações/ano de 2.08 (*B. cenocepacia*) e 2.27 (*B. multivorans*) durante a co-infeção. Estas taxas são consideradas associadas à rápida diversificação da população bacteriana que ocorre nos pulmões de CF, que resulta na divergência de cada estirpe em sub-linhagens, cada uma com a sua própria assinatura genómica, o que é consistente com os resultados aqui obtidos. Os resultados também sugerem que a infeção por *B. cenocepacia* envolva períodos de diversificação dominados por seleção positiva, enquanto a infeção por *B. multivorans* envolve períodos de diversificação dominados por seleção positiva seguidos de períodos de evolução relativa neutra. Este diferente padrão de evolução adaptativa das duas espécies durante a co-infeção poderá ter causas genética e os padrões evolutivos registados parecem ser específicos para cada espécie, confirmando publicações anteriores onde cada espécie foi estudada em condições independentes.

A análise genómica comparativa realizada revelou ainda que de entre os genes mutados em *B. cenocepacia* e *B. multivorans*, estão genes que codificam proteínas reguladoras de resposta a stress oxidativo e proteínas de deteção de metais pesados, indicando que o sistema imunitário do hospedeiro possa ter um papel fundamental na evolução de *B. cenocepacia* e *B. multivorans* durante estas infeções crónicas, protegendo de bactérias contra espécies reativas de oxigénio produzidas por células fagocitárias. Outros genes mutados encontrados em ambas as espécies incluem genes envolvidos em mecanismos de defesa e a sua regulação, sugerindo resposta a terapêutica prolongada com antibióticos. A resistência a antibióticos é também uma característica evolutiva amplamente descrita para vários patogénicos associados à CF e genes relacionados com essa resistência encontram-se sujeitos a pressão seletiva. Uma resposta evolutiva comumente descrita, envolvendo as alterações genéticas que ocorrem em genes presentes no agrupamento de genes da biossíntese do OAg do lipossacárido, foi também observada neste estudo, indicando que aquele agrupamento de genes se encontra sob forte pressão seletiva durante a infeção crónica.

Assim, quer os resultados obtidos da análise de genómica comparativa realizada quer de outros estudos de biologia celular, indicam que a fração do lipopolissacárido (LPS) correspondente ao OAg sofre alterações durante infeção crónica, o que pode estar associado à adaptação e evasão do sistema imunitário e à pressão seletiva por parte da terapia com antibióticos. Para uma melhor compreensão deste tema, durante o trabalho de dissertação foi caracterizado um novo *locus* genético envolvido na biossíntese do OAg do LPS nos onze variantes clonais de *B. cenocepacia*, que foram analisado na abordagem de genómica comparativa mencionada anteriormente. Este novo *locus* biossintético híbrido do AgO (pois possui genes homólogos em *B. multivorans* ATCC 17616) fornece evidência para a existência de uma nova estrutura química do OAg que foi durante este trabalho determinada em colaboração com o grupo de A. Molinaro, Universidade de Nápoles Federico II, Complexo Universitário Monte Sant'angelo, Nápoles, Itália. A origem da região homóloga em *B. multivorans*, no *locus* do OAg, é desconhecida, mas apresenta um conteúdo GC significativamente diferente dos restantes genes do locus para a sua biossíntese, assim como do resto do genoma. Todos os subsequentes 10 variantes clonais de *B. cenocepacia* recolhidos do mesmo paciente até pouco antes da sua morte com a síndrome de cepacia não apresentam OAg, indicando a sua perda durante a adaptação de *B. cenocepacia* envolvido na infeção crónica do pulmão. No entanto, apesar do Bcc ser constituído de 24 espécies relacionadas, permanece por entender se a perda OAg confere ou não alguma vantagem às diferentes espécies durante o curso da infeção. No presente trabalho foi realizada uma triagem retrospectiva sistemática e longitudinal para a perceber se a perda de expressão do OAg, durante infeção no contexto de CF pode ser considerada um fenómeno geral, observado em diferentes espécies do Bcc e favorecendo a infeção persistência. Para isso, foram utilizados 357 isolados obtidos de 19 pacientes com infeção crónica que estiveram em tratamento num hospital central de Lisboa. Estes isolados representam 21 estirpes de seis/sete espécies/linhagens de Bcc que são, a nível mundial, frequentemente

ou raramente isolados destes pacientes: *B. cenocepacia* recA linhagens IIIA e IIIB, *B. multivorans*, *B. dolosa*, *B. stabilis*, *B. cepacia* e *B. contaminans*. Entre estas seis espécies, as duas mais prevalentes e recedadas, *B. cenocepacia* e *B. multivorans*, mostraram tendência para perder o OAg durante a infeção crónica. As estirpes de *B. cenocepacia* recA linhagem IIIA que levam normalmente a infeções particularmente destrutivas mostraram uma perda mais frequente do OAg, comparado com as estirpes da linhagem IIIB. Pela primeira vez, foi mostrado que *B. cenocepacia* e *B. contaminans* não perdem o OAg, mesmo durante infeções prolongadas com duração até 9.6 e 15.2 anos, respetivamente. Além disso, estirpes de *B. stabilis* apresentam um fenótipo estável relativamente à presença ou ausência do OAg durante o período de infeção. No caso do único paciente infetado pela espécie mais distante do Bcc, *B. dolosa*, a fração OAg esteve ausente desde o início da infeção até à morte do paciente após 5.5 anos, consistente com a estrutura do LPS estabelecida anteriormente para o primeiro isolado.

A perda ou modificação do OAg parece ter um papel importante durante o processo de infeção, especialmente no passo de colonização (adesão). Neste contexto, a microscopia de força atómica (AFM) tornou-se uma ferramenta essencial para o entendimento de adesão bacteriana e a nanomecânica de sistemas vivos. A AFM foi utilizada para estudar a morfologia de superfície e mapear as propriedades mecânicas dos variantes clonais sequenciais de *B. cenocepacia*. Esta parte do trabalho permitiu constatar que a capacidade do isolado inicial para aderir à ponta de Si<sub>3</sub>N<sub>4</sub> é significativamente maior quando em ambiente líquido do que a capacidade dos variantes mais tardios, que não possuem AgO e cujos valores de adesão são semelhantes entre eles. Estes resultados são consistentes com a ideia de que a variabilidade no OAg do LPS afeta a adesão bacteriana e colonização, e que a capacidade de evasão aos mecanismos de defesa do hospedeiro é um fator importante na adaptação de *B. cenocepacia* durante infeção crónica. De facto, o módulo de elasticidade de superfície diminuiu significativamente desde o isolado inicial até ao último variante clonal. No entanto, o isolado intermédio IST4113, altamente resistente a antibióticos, apresenta valores máximos. A elevada rigidez e aumento de elasticidade foram recentemente associadas a uma menor permeabilidade da membrana externa, o que resulta numa redução da entrada por difusão de antibióticos nas células. Pela primeira vez foi observado nesta tese um padrão consistente e progressivo de diminuição da altura e o aumento do rácio largura/comprimento de células de *B. cenocepacia* durante infeção prolongada, o que resulta na evolução da estrutura celular desde a típica forma de bastonete para uma morfologia mais próxima da forma esférica dos *cocci*. Consistente com esta hipótese, vários estudos mostraram a relevância do tamanho da superfície celular quando as células bacterianas são confrontadas com um ataque do sistema imunológico, sendo que um reduzido tamanho permite uma evasão mais eficiente às defesas do hospedeiro. É provável que a evolução adaptativa observada possa levar a um melhor desempenho em condições limitantes de oxigénio, uma aquisição mais eficiente de nutrientes e mais fácil evasão; favorecendo infeções persistentes e a patogenicidade.

Em conclusão, esta tese vem reforçar a ideia de que bactérias do Bcc são agentes patogénicos altamente adaptáveis ao pulmão de CF. Nesse nicho, estas bactérias encontram múltiplas condições de stresse que favorecem a diversificação genética e fenotípica durante a infeção crónica, resultando na emergência de populações heterogéneas mais adaptadas para sobreviver nesses ambientes, em particular ao nível do invólucro celular. Este conhecimento é crucial para o desenvolvimento de novos alvos a estratégias terapêuticas que resultem num melhor tratamento das infeções por bactérias daquele complexo.

**Palavras-chave:** Complexo Burkholderia cepacia; Infeção crónica; Fibrose quística; Adaptação e evolução; Invólucro celular.

## Table of contents

---

Acknowledgment .....	i
Abstract .....	iii
Resumo .....	v
Resumo Alargado.....	vii
Table of contents.....	xi
List of Figures .....	xv
List of Tables .....	xix
List of Abbreviations .....	xxi

### 1. Introduction

1.1. <i>Burkholderia cepacia</i> complex bacteria (Bcc) .....	1
1.2. Bcc bacteria as opportunistic human pathogens in cystic fibrosis.....	4
1.2.1. Bcc bacteria and cystic fibrosis .....	4
1.2.2. Epidemiology of Bcc respiratory infections in cystic fibrosis patients .....	5
1.2.2.1. <i>Epidemiology of Bcc respiratory infections in CF patients and epidemic strains</i> .....	5
1.2.2.2. <i>Epidemiology of Bcc respiratory infections in the major Portuguese CF centre</i> .....	6
1.2.3. Within-host adaptation of Bcc bacteria during chronic respiratory infections.....	7
1.2.3.1. <i>Bcc bacterial factors involved in host-pathogen interaction and recognition</i> .....	7
1.2.3.2. <i>Bcc phenotypic adaptation during chronic infection</i> .....	9
1.2.3.3. <i>Mechanisms underlying genome-wide Bcc adaptation to the CF lung</i> .....	11
1.2.4. Bcc cell surface in host-pathogen interaction and adaptation to the CF lung .....	18
1.2.4.1. <i>The cell surface of Bcc bacteria</i> .....	18
1.2.4.2. <i>LPS biosynthesis in Bcc</i> .....	19
1.2.4.3. <i>LPS modification in Bcc during chronic respiratory infection in CF and pathogenicity role</i> .....	20
1.3. Atomic Force Microscopy (AFM) as an advanced technique to study bacteria at the cell surface level .....	23
1.3.1. Fundamentals of AFM .....	23
1.3.2. Experimental considerations for the AFM imaging of the bacterial cells.....	25
1.3.3. Mechanical and adhesion properties of Gram-negative bacteria studied by AFM .....	26
1.4. Thesis outline and research objectives.....	29

### 2. Comparative evolutionary patterns of *Burkholderia cenocepacia* and *Burkholderia multivorans* during chronic co-infection of a cystic fibrosis patient lung

2.1. Abstract .....	33
2.2. Introduction.....	34

2.3. Materials and Methods.....	36
2.3.1. Bcc bacterial isolates and growth conditions .....	36
2.3.2. Genomic DNA sequencing, assembly and annotation .....	37
2.3.3. Variant calling and SNP/INDEL detections .....	38
2.3.4. Assessment of population structure and phylogeny.....	39
2.3.5. Examining the co-infecting <i>B. cenocepacia</i> and <i>B. multivorans</i> evolving populations for Selection.....	39
2.3.6. Genomic features and structural analysis of the ancestor <i>B. cenocepacia</i> and <i>B. multivorans</i> strains genomes .....	39
2.3.7. Ethics statement.....	40
2.3.8. Data and nucleotide sequence accession numbers .....	40
2.4. Results.....	41
2.4.1. Genomic analysis of <i>B. cenocepacia</i> and <i>B. multivorans</i> ancestor strains .....	41
2.4.2. Comparative analysis of the genome sequences of <i>B. cenocepacia</i> and <i>B. multivorans</i> sequential clonal isolates .....	45
2.4.3. Evolutionary dynamics of <i>B. cenocepacia</i> and <i>B. multivorans</i> co-infecting populations inside the CF lung .....	46
2.4.4. <i>B. cenocepacia</i> and <i>B. multivorans</i> genes under convergent evolution in the CF lung.....	49
2.4.5. Orthologous Genes, under selection in the co-evolved <i>B. cenocepacia</i> and <i>B. multivorans</i> populations .....	51
2.5. Discussion.....	55
<b>3. Structure of O-antigen and hybrid biosynthetic locus in <i>Burkholderia cenocepacia</i> clonal variants recovered from a cystic fibrosis patient</b>	
3.1. Abstract.....	61
3.2. Introduction.....	62
3.3. Materials and Methods.....	63
3.3.1. Bacterial strains and growth conditions.....	63
3.3.2. LPS extraction, purification, and compositional analyses .....	64
3.3.3. NMR spectroscopy .....	65
3.3.4. MALDI mass spectrometry .....	66
3.3.5. Small-scale LPS extraction for SDS-polyacrylamide gel electrophoresis analysis.....	66
3.3.6. Genomic DNA sequencing, assembly, and annotation .....	66
3.3.7. <i>In silico</i> characterization of O-antigen biosynthetic gene clusters.....	67
3.3.8. Molecular cloning of <i>wbil</i> and <i>bmul_2510</i> genes in <i>B. cenocepacia</i> clonal variants lacking O-antigen .....	67
3.3.9. Plasmid conjugation into <i>B. cenocepacia</i> .....	68
3.3.10. Western blot .....	68
3.4. Results.....	68
3.4.1. Only the first of the 11 sequential clonal variants of <i>B. cenocepacia</i> produce OAg.....	68
3.4.2. Chemical structure of the LPS of <i>B. cenocepacia</i> isolates with or without O-antigen .....	69

3.4.3. Genomic sequence of the <i>B. cenocepacia</i> clonal isolates reveals of a novel hybrid O-antigen biosynthetic cluster .....	74
3.4.4. Mutations in <i>wbil</i> and <i>Bmul_2510</i> are not involved in loss of OAg .....	76
3.4.5. Analysis of the OAg cluster of a co-infecting <i>B. multivorans</i> isolate rules out that the hybrid cluster in IST439 arose from gene transfer in the lung .....	78

3.5. Discussion .....	80
-----------------------	----

#### **4. *Burkholderia cepacia* complex species differ in the frequency of variation of the lipopolysaccharide O-antigen expression during cystic fibrosis chronic respiratory infection**

4.1. Abstract .....	83
4.2. Introduction .....	84
4.3. Materials and methods .....	86
4.3.1. Bacterial isolates and growth conditions .....	86
4.3.2. Species identification and genotyping of Bcc isolates .....	87
4.3.3. LPS extraction and SDS-PAGE analysis .....	87
4.3.4. Ethics .....	88
4.4. Results .....	88
4.4.1. During chronic infections, <i>B. cenocepacia recA</i> lineage IIIA and <i>B. multivorans</i> isolates were more prone to lose the OAg present in early isolates, compared with <i>B. cenocepacia recA</i> lineage IIIB. ....	88
4.4.2. The OAg is stably expressed in <i>B. cepacia</i> , <i>B. contaminans</i> and <i>B. stabilis</i> .....	94
4.4.3. The LPS from early and late <i>B. dolosa</i> isolates lack the OAg .....	94
4.5. Discussion .....	95

#### **5. Variation of *Burkholderia cenocepacia* cell wall morphology and mechanical properties during cystic fibrosis lung infection, assessed by atomic force microscopy**

5.1. Abstract .....	99
5.2. Introduction .....	100
5.3. Methods .....	102
5.3.1. Bacterial Strains and growth conditions. ....	102
5.3.2. Preparation of the AFM samples .....	102
5.3.3. AFM observations and measurements .....	103
5.3.4. Growth curves .....	104
5.3.5. Biofilm formation assays .....	104
5.3.6. Statistics .....	105
5.4. Results .....	105
5.4.1. <i>B. cenocepacia</i> morphology and surface roughness evolution during long-term CF lung infection .....	105
5.4.2. Surface and mechanical properties evolution during long-term CF lung infection .....	107

5.4.3. Growth curves of the <i>B. cenocepacia</i> clonal variants under aerobic or microaerophilic conditions.....	109
5.4.4. Biofilm growth of <i>B. cenocepacia</i> clonal variants.....	110
5.5. Discussion.....	111
6. Final discussion.....	117
7. References.....	123
8. Supplementary Data.....	145
8.1. Supplementary notes on materials and methods.....	145
8.1.1. Data related to chapter II.....	145
8.1.1.1. Genomic DNA sequencing, De novo assembly, and annotation.....	145
8.1.1.1.1. Illumina sequencing by CD Genomics.....	145
8.1.1.1.2. PacBio sequencing by a real-time (SMRT) Pacific Biosciences.....	146
8.1.1.2. Variant calling and SNP/INDEL detections.....	146
8.2. Supplementary Figures.....	147
8.2.1. Data related to chapter II.....	147
8.2.2. Data related to chapter III.....	150
8.2.3. Data related to chapter IV.....	154
8.2.4. Data related to chapter V.....	155
8.3. Supplementary Tables.....	156
8.3.1. Data related to chapter II.....	156
8.3.2. Data related to chapter III.....	193
8.3.3. Data related to chapter IV.....	202

## List of Figures

---

<b>Figure 1.1</b>	Phylogenetic tree based on multiple sequence alignment of 16S rRNA genes obtained from the 24 Bcc species have already deposited in NCBI database and <i>Burkholderia</i> database.	2
<b>Figure 1.2</b>	CF population under surveillance at the major Portuguese CF treatment Centre at Hospital de Santa Maria, in Lisbon, during the last two decades.	6
<b>Figure 1.3</b>	Graphical representation of some virulence factors previously-reported for Bcc bacteria.	9
<b>Figure 1.4</b>	Reported phenotypes found to suffer diversification during Bcc chronic infections.	11
<b>Figure 1.5</b>	Pathways and cellular functions related with the mutated genes previously described to occur during Bcc bacteria evolution inside the CF lung.	17
<b>Figure 1.6</b>	Simplified overview of the LPS biosynthesis based on studies previously published using <i>B. cenocepacia</i> J2315 and K56-2 epidemic strains.	21
<b>Figure 1.7</b>	Atomic force microscopy apparatus and force distance curve.	24
<b>Figure 2.1</b>	(A) A schematic representation of the <i>B. cenocepacia</i> and <i>B. multivorans</i> isolates examined with information on lung function (Forced Expiratory Volume in the first second, FEV1%).	42
<b>Figure 2.2</b>	Genomic islands (A) BcenST281_GIs and (B) BmST836_GIs in the reference strains, <i>B. cenocepacia</i> IST439 and <i>B. multivorans</i> IST419, respectively, examined in the work.	44
<b>Figure 2.3</b>	Phylogenetic tree based on the comparison of the whole-genome of <i>B. cenocepacia</i> and <i>B. multivorans</i> clonal variants and other sequenced strains available in <i>Burkholderia</i> genome database (Table S2.1).	46
<b>Figure 2.4</b>	3D plot of non-synonymous mutations formed among <i>B. cenocepacia</i> (A) and <i>B. multivorans</i> (B) clonal variants during the chronic co-infection.	50
<b>Figure 2.5</b>	Heat-map of non-synonymous mutation frequencies in coding sequences (CDS) under strong parallelism (that acquired $\geq 3$ mutations) among the clonal variants of <i>B. cenocepacia</i> (A) and <i>B. multivorans</i> (B).	52
<b>Figure 2.6</b>	Mutations that were observed in two orthologous genes in the genomes of clonal variants of <i>B. multivorans</i> and of <i>B. cenocepacia</i> (chronological order from left to right) indicated by the amino acid residue change.	54
<b>Figure 3.1</b>	Electrophoretic profile of the LPS extracted from 11 clonal <i>Burkholderia cenocepacia</i> isolates retrieved from a chronically infected CF patient.	69
<b>Figure 3.2</b>	$^1\text{H}$ NMR spectrum of products obtained from IST439 and other three subsequent variants after mild acid hydrolysis.	71
<b>Figure 3.3</b>	LPS moieties structures elucidated by NMR spectroscopy.	71

<b>Figure 3.4</b>	$^1\text{H}$ , $^{13}\text{C}$ -HSQC spectrum of the O-chain moiety from IST439 strain.	73
<b>Figure 3.5</b>	MALDI mass spectrum of the intact LPS from strain IST439 (mass range 1400-4000).	73
<b>Figure 3.6</b>	Genetic organization of gene clusters for core-lipid A and O-antigen biosynthesis in <i>B. cenocepacia</i> IST439 (A) and the reference strains K56-2 and J2315 (B).	75
<b>Figure 3.7</b>	Characterization of the OAg unit after cloning and introduction of <i>B. cenocepacia</i> IST439 <i>wbiI</i> and <i>Bmul_2510</i> genes in all variants lacking OAg production to complement the corresponding mutated genotypes.	77
<b>Figure 3.8</b>	Genetic organization of the O-antigen biosynthetic gene clusters of <i>B. multivorans</i> strains ATCC 17616 (B) and IST419 (A), including genes for lipid A-core biosynthesis and O-antigen biosynthesis within the <i>apaH</i> and <i>ureG</i> flanking genes (in black).	79
<b>Figure 4.1</b>	Schematic representation of the Bcc isolates examined in this study retrieved from different CF patients over 21 years (from 1995 to 2016) of epidemiological survey at the CF Centre of Hospital de Santa Maria.	90
<b>Figure 4.2</b>	Representative SDS-PAGE gels showing the O-antigen banding patterns obtained for the Bcc isolates examined.	93
<b>Figure 5.1</b>	Cell topography and surface roughness.	106
<b>Figure 5.2</b>	Cell morphology.	108
<b>Figure 5.3</b>	Elasticity and adhesion studied by AFM in liquid samples.	109
<b>Figure 5.4</b>	Growth under oxygen limitation.	110
<b>Figure 5.5</b>	Biofilm formation.	111
<b>Figure 6.1</b>	Proposed model for the mechanisms underlying clonal diversification occurring in Bcc bacteria during long-term chronic infection inside the CF lung, based on the experimental evidences gathered during this PhD.	122
<b>Figure S2.1</b>	Antibiotic resistance associated genes – AMR in (A) <i>B. cenocepacia</i> IST439 and (B) <i>B. multivorans</i> IST419.	148
<b>Figure S2.2</b>	Virulence factors associated genes (VF) in (A) <i>B. cenocepacia</i> IST439 and (B) <i>B. multivorans</i> IST419.	149
<b>Figure S3.1</b>	Electrophoretic profiles of the LPS extracted from <i>Burkholderia cenocepacia</i> IST439 isolate and the ( <i>recA</i> lineage IIIA) ET-12 reference strains K56-2 and J2315.	150
<b>Figure S3.2</b>	The MALDI mass spectrum of the intact IST4113 LPS (mass range 1200-420000 Da).	151
<b>Figure S3.3</b>	Comparative genomic of the OAg cluster is visualized by Artemis/ACT among <i>B. cenocepacia</i> IST439 and the reference strains J2315 and k56-2.	152

<b>Figure S3.4</b>	Comparative visualization of the OAg biosynthetic clusters in binary mode between <i>B. cenocepacia</i> IST439 and <i>B. multivorans</i> IST419 (A), between <i>B. cenocepacia</i> IST439 and <i>B. multivorans</i> ATCC 17616 (B), and between <i>B. multivorans</i> isolates IST419 and ATCC 17616 (C) are shown.	153
<b>Figure S4.1</b>	Representative profiles based on Random Amplified Polymorphic DNA (RAPD) analysis of the Bcc isolates examined in this study retrieved from different CF patients during several years of chronic infection.	154
<b>Figure S5.1</b>	The adhesion force maps (A –unattached and B –attached cells) of bacterial surfaces for IST4113 and IST4134, respectively, showed that the Nano-scale surface architectures have less adhesion among the rest of the surfaces.	155



## List of Tables

---

<b>Table 2.1</b>	<i>Burkholderia cepacia</i> complex examined isolates obtained from a cystic fibrosis patient whose genomes were sequenced in this study, ordered based on the isolation date.	36
<b>Table 2.2</b>	Genomic assembly and functional annotation of the Bcc clonal variants.	43
<b>Table 2.3</b>	Pair-wise comparison of the infecting population aligned to the ancestor reference strains <i>B. cenocepacia</i> IST439 and <i>B. multivorans</i> IST419.	47
<b>Table 2.4</b>	List of non-synonymous mutations (highlighted with blue) that became fixed among all clonal variants of <i>B. cenocepacia</i> .	53
<b>Table 3.1</b>	Description of <i>Burkholderia</i> isolates and <i>E. coli</i> strains used in this study.	64
<b>Table 3.2</b>	Chemical shift $\delta$ ( $^1\text{H}/^{13}\text{C}$ ) of the O-chain moiety from <i>B. cenocepacia</i> IST439.	70
<b>Table 4.1</b>	Schematic representation of the sequential isolates from various species chronically infecting the different CF patients examined in this study.	89
<b>Table S2.1</b>	List of the whole genome sequences of Bcc isolates used in the present study for comparative genomic analyses.	156
<b>Table S2.2</b>	Metadata information of the studied clonal variants sequentially ordered based on isolation date; details of genomic assembly and functional annotation are shown as well as <i>in silico</i> MLST profiles.	158
<b>Table S2.3</b>	Published genomic islands present in <i>B. cenocepacia</i> strains J2315 and ST32 and their eventual presence in <i>B. cenocepacia</i> IST439 and <i>B. multivorans</i> IST419.	160
<b>Table S2.4</b>	BcenST218_GIs as Genomic islands of <i>B. cenocepacia</i> IST439.	162
<b>Table S2.5</b>	BmST836_GIs as Genomic islands of <i>B. multivorans</i> IST419.	163
<b>Table S2.6</b>	Antibiotic resistance (AMR) and virulence factors (VF) associated genes in <i>B. cenocepacia</i> IST439 detected by the presented database.	165
<b>Table S2.7</b>	Antibiotic resistance (AMR) and virulence factors (VF) associated genes in <i>B. multivorans</i> IST419 detected by the presented database.	170
<b>Table S2.8</b>	List of mutations found among <i>B. cenocepacia</i> clonal variants.	177
<b>Table S2.9</b>	List of mutations found among <i>B. multivorans</i> clonal variants.	183
<b>Table S3.1</b>	Fasta file of the OAg cluster.	193
<b>Table S3.2</b>	Primer sequences used to amplify the desired regions with detailed description.	200

<b>Table S3.3</b>	Monosaccharide compositional analysis of the LPS/LOS isolated from <i>B. cenocepacia</i> isolates IST4113, IST4129, IST4134 and IST439.	201
<b>Table S4.1</b>	Detailed description of the <i>Burkholderia</i> sequential isolates per patient examined in this study with their isolation dates.	202

## List of Abbreviations

°C	Degrees Celsius	$d_N/d_S$	Non-synonymous substitution to synonymous substitution ratio
$^1\text{H NMR}$	Proton nuclear magnetic resonance	<b>DNA</b>	Deoxyribonucleic Acid
$A_{600\text{nm}}$	Absorbance at 600 nm	<b>DQF-COSY</b>	Double-quantum filtered phase sensitive correlation spectroscopy
<b>aa</b>	Amino acid	$d_S$	Rate of synonymous substitution per synonymous site
<b>AFM</b>	Atomic force microscopy	<b>EDTA</b>	Ethylenediamine tetraacetic acid
<b>AMR</b>	Antimicrobial resistance	<b>ENA</b>	European nucleotide archive
<b>ANI</b>	Average nucleotide identity	<b>EPS</b>	exopolysaccharide
<b>ATP</b>	Adenosine triphosphate	<b>ET-12</b>	Edinburgh-Toronto lineage
<b>Bcc</b>	<i>Burkholderia cepacia</i> complex	<b>FD</b>	Force-distance
<b>cci</b>	Cenocepacia genomic island	<b>FEV1%</b>	Forced expiratory volume in the first second
<b>CD11b</b>	Recombinant human antigen protein	<b>Fw</b>	Forward
<b>CDS</b>	Coding sequences	<b>GI</b>	Genomic island
<b>CF</b>	Cystic fibrosis	<b>GLC-MS</b>	Gas-liquid chromatography mass spectrometry
<b>CFTR</b>	CF-transmembrane conductance regulator	<b>HCl</b>	Hydrochloric acid
<b>CFU</b>	Colony forming unit	<b>HGPA3</b>	Hierarchical genome-assembly process workflow
<b>CHLN</b>	Centr Hospitalar Lisboa Norte	<b>HMBC</b>	Heteronuclear multiple bound correlation
<b>COG</b>	Cluster of orthologous groups	<b>HSM</b>	Hospital de Santa Maria
<b>Core OS</b>	Core oligosaccharide	<b>HSQC</b>	Heteronuclear single quantum coherence
<b>CS</b>	Cepacia syndrome	<b>IL</b>	Interleukins
<b>D<sub>2</sub>O</b>	Deuterated water	<b>IM</b>	Inner membrane
<b>ddH<sub>2</sub>O</b>	Double distilled water	<b>INDEL</b>	Insertion-deletion mutations
$d_N$	Rate of non-synonymous substitution per non-synonymous site		

<b>INFARMED</b>	Portuguese Medicines and Health products authority	<b>pD</b>	Hydrogen potential
<b>KDa</b>	Kilo-Dalton	<b>PG</b>	peptidoglycan
<b>KDO</b>	3-deoxy-D-manno-oct-2-ulosonic acid	<b>PHA</b>	Polyhydroxyalkanoates
<b>KO</b>	D-glycero- $\alpha$ -D-talo-oct-2-ulosonic acid	<b>PHDC</b>	Philadelphia-District of Colombia
<b>LB</b>	Lysogeny broth	<b>RAPD</b>	Random amplified polymorphic DNA
<b>LOS</b>	Lipooligosacchride	<b>RFLP</b>	Restriction fragment length polymorphism
<b>LPS</b>	Lipopolysaccharide	<b>RNA</b>	Ribonucleic acid
<b>lxa</b>	Low-oxygen-activated locus	<b>ROESY</b>	Rotating frame Overhauser enhancement spectroscopy
<b>MALDI-TOF</b>	Matrix-assisted laser ionization and desorption – time of flight mass spectrometer	<b>ROS</b>	Reactive oxygen species
<b>MeOH</b>	Methanol	<b>rpm</b>	Revolutions per minute
<b>MLCT</b>	Metal to ligand charge transfer technology	<b>rRNA</b>	Ribosomal Ribonucleic acid
<b>MLST</b>	Multi-locus sequence typing	<b>Rv</b>	reverse
<b>NaOH</b>	Sodium hydroxide	<b>SDS-PAGE</b>	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
<b>NCBI</b>	National Center for Biotechnology Information	<b>SMRT</b>	Single molecule real-time sequencing
<b>NMR</b>	Nuclear magnetic resonance	<b>SNP</b>	Single nucleotide polymorphism
<b>NOESY</b>	Nuclear Overhauser enhancement spectroscopy	<b>ST</b>	Sequence type
<b>OAg</b>	Lipopolysaccharide O-antigen	<b>T1SS</b>	Type 1 secretion system
<b>OD<sub>640nm</sub></b>	Optical density at 640 nanometer	<b>T2SS</b>	Type 2 secretion system
<b>OM</b>	Outer membrane	<b>T4SS</b>	Type 4 secretion system
<b>ORF</b>	Open reading frame	<b>TAA</b>	Trimeric autotransporter adhesins
<b>PacBio</b>	Pacific bioscience	<b>TLR4</b>	Toll-like receptor type 4
<b>PBP</b>	Penicillin binding proteins	<b>TNF</b>	Tumor necrosis factor
<b>PBS</b>	Phosphate buffer saline	<b>TOCSY</b>	Total correlation spectroscopy
<b>PCR</b>	Polymerase chain reaction	<b>V/V</b>	Volume per volume
		<b>VF</b>	Virulence factors

**wgMLST** Whole genome multi-locus  
sequence typing

**Wt/V** Weight per volume



# **1**                    **General introduction**

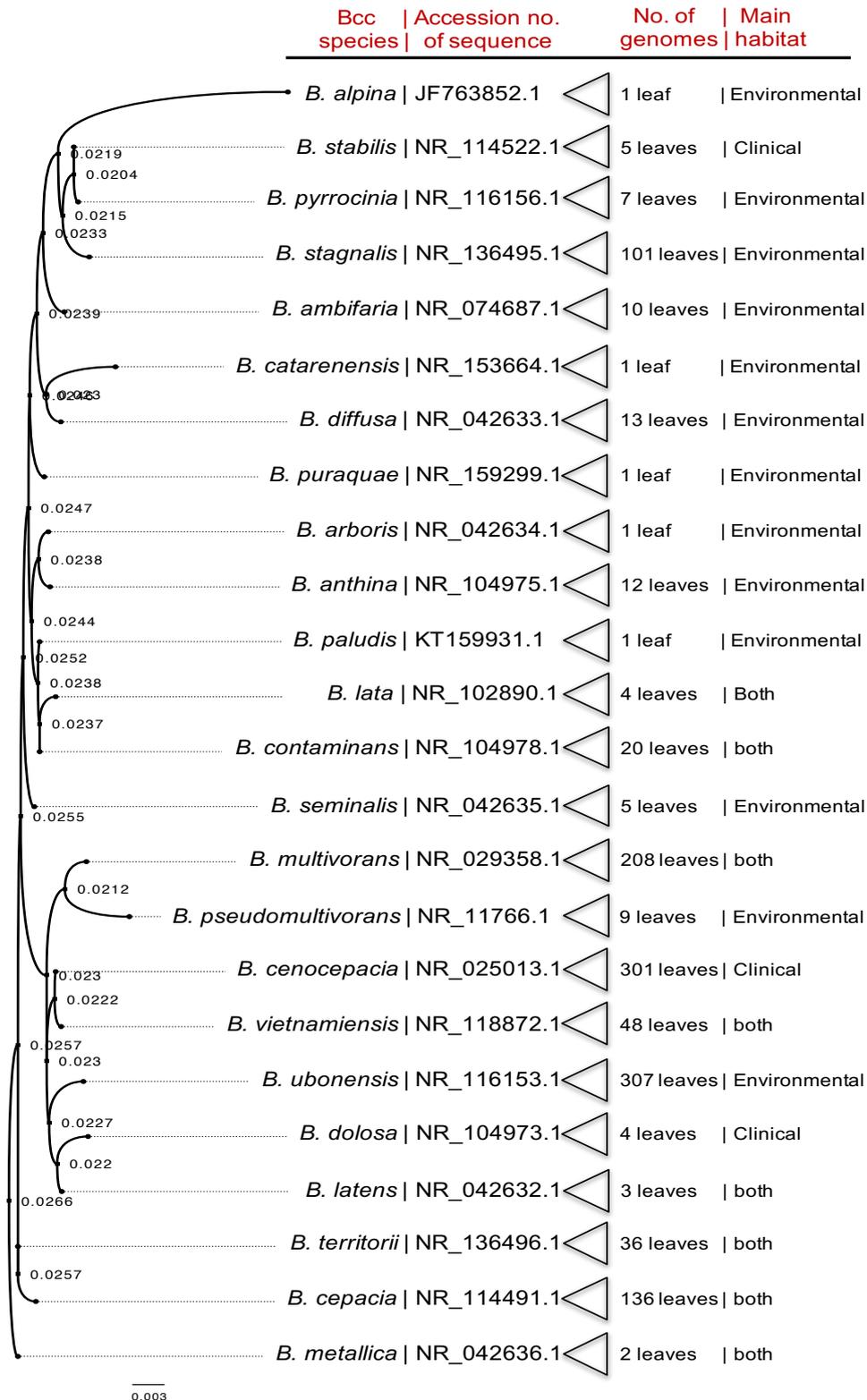
---



### 1.1. *Burkholderia cepacia* complex bacteria (Bcc)

*Burkholderia cepacia* complex (Bcc) bacteria is a group of closely-related Gram-negative, non-fermenting, non-spore-forming bacilli of the  $\beta$ -proteobacteria class (Mahenthiralingam et al., 2005; Compant et al., 2008; Mahenthiralingam et al., 2008; Depoorter et al., 2016). Bcc currently comprises 24 bacterial species (Fig. 1.1) (De Smet et al., 2015; Depoorter et al., 2016; Ong et al., 2016; Bach et al., 2017; Weber and King, 2017; Martina et al., 2018) and exhibits a considerable genotypic and phenotypic variation that confer the capacity of adaptation to challenging environments (Lessie et al., 1996; Parke and Gurian-Sherman, 2001; Coenye and Vandamme, 2003; Loutet and Valvano, 2010; Depoorter et al., 2016). Besides their adaptability to colonize diverse niches in nature, Bcc have emerged as human opportunistic pathogens constituting a serious risk of respiratory infections for susceptible individuals, as immunocompromised individuals (cancer patients submitted to chemotherapy, HIV/AIDS patients, mechanically ventilated infants and the debilitated elderly people) and in patients with cystic fibrosis (CF) or chronic granulomatous diseases (Johnston, 2001; Agodi et al., 2002; Mahenthiralingam et al., 2002; Mahenthiralingam et al., 2005; Greenberg et al., 2009).

Bcc bacteria were originally identified as plant pathogen when Burkholder W. H., in 1950, described *Pseudomonas cepacia* as an etiological agent of soft onion rot (Burkholder, 1950). Bcc species exhibit high degree of 16S rRNA and *recA* genes sequence similarity (98-100% and 94-95%, respectively) as well as moderate and high levels of DNA:DNA hybridization and whole-genome average nucleotide identity (ANI), respectively (Coenye et al., 2001; Vanlaere et al., 2009; Peeters et al., 2013). As a result, Bcc species differentiation is very challenging, however, several efforts over the last decades have been performed for better resolution (Coenye et al., 2001; Vanlaere et al., 2009; Bach et al., 2017). Based on a combination of molecular taxonomic analysis (including 16S rRNA sequencing and DNA:DNA homology values) and phenotypic – physiological and biochemical – characteristics, seven closely-related bacterial species were initially transferred from genus *Pseudomonas* to a new genus *Burkholderia* (Yabuuchi et al., 1992). Later, an advanced-integrated genotypic and phenotypic analyses (polyphasic taxonomy approach) revealed that several isolates classified as a single species belonged to at least five genetically different species then known as genomovars (Vandamme et al., 1997). In the following years, other novel bacterial species were classified in the Bcc bacterial group and the interest in *Burkholderiales* markedly-increased (Vandamme and Dawyndt, 2011). Several molecular approaches (e.g. random amplified polymorphic DNA – RAPD and ribotyping) have also been developed for Bcc species differentiation; however, *recA* gene sequencing, *recA*-restriction fragment length polymorphism (*recA*-RFLP) and the multi-locus sequence typing (MLST) are currently the most widely-used taxonomic tools to identify and genotype Bcc species/strains (Mahenthiralingam et al., 1996; Brisse et al., 2000; Mahenthiralingam et al., 2000a; Vermis et al., 2002; Baldwin et al., 2005; Spilker et al., 2009; Vandamme and Dawyndt, 2011).



**Figure 1.1** | Phylogenetic tree constructed based on multiple sequence alignment of sequence of 16S rRNA genes from the 24 Bcc species that were already deposited in NCBI database [<https://www.ncbi.nlm.nih.gov/>] and *Burkholderia* database [<https://www.burkholderia.com/>]. The number of sequenced-genomes available, for each Bcc species, in the databases was collapsed and introduced as number of leaves. The accession number of the genetic sequences used to prepare the phylogenetic tree is shown as well as the main habitat for each Bcc species.

Based on molecular signatures and phylogenetic analysis of the genus *Burkholderia*, a division of this genus into two major genera (two major phylogenetic clusters) was recently proposed (Sawana et al., 2014). The genus *Burkholderia* was proposed to keep pathogenic bacteria, including Bcc bacteria, and a new genus *Paraburkholderia* was suggested to harbour environmental non-pathogenic species (Sawana et al., 2014). Currently (up to October 2019), the Bcc is a group of 24 closely-related species (De Smet et al., 2015; Depoorter et al., 2016; Ong et al., 2016; Bach et al., 2017; Weber and King, 2017; Martina et al., 2018) [Fig. 1.1 and for a detailed list of *Burkholderia* species, see <http://www.bacterio.net/>]. Interestingly, among these 24 Bcc species, *B. cenocepacia* was found to comprise genetically distinct strains classified into at least four phylogenetic sub-clusters (known as *B. cenocepacia recA* lineages IIIA, IIIB, IIIC, and IIID ) based on the *recA*-RFLP (Vandamme et al., 2003). Among them, *B. cenocepacia recA* lineages IIIA and IIIB contains most of the described clinical isolates (Drevinek and Mahenthiralingam, 2010).

Genomes of Bcc bacteria are complex and relatively large compared with other Gram-negative bacteria (the total genome size range from 6.4 to 10.4 Mbp) (Tatusova et al., 2014; 2015), mainly-organized in three chromosomes and, in a number of strains, with at least one plasmid (Parke and Gurian-Sherman, 2001; Winsor et al., 2008). The biggest chromosome (chromosome I around 3.8 Mbp) is an essential replicon with a higher percentage of protein coding regions while the other replica (chromosome II and III) contain genes that are not involved in core functions (Holden et al., 2009). Recently, the third replicon of Bcc bacteria was proved to confer a competitive advantage concerning Bcc pathogenicity and considered a megaplasmid (0.5 to 1.4 Mbp) (Agnoli et al., 2017). This megaplasmid was proved to be involved in virulence in various virulence models (including rats and zebrafish), in stress tolerance, and in antifungal activity (Agnoli et al., 2012; Agnoli et al., 2014; Agnoli et al., 2017). Bcc genomes were also proved to include a high number of ortholog genes, paralogs, insertion sequences, and mobile elements that confer genomic plasticity and diversity, contributing to bacterial pathogenicity and adaptation (Lessie et al., 1996; Parke and Gurian-Sherman, 2001; Coenye and Vandamme, 2003; Winsor et al., 2008; Holden et al., 2009; Roszniowski et al., 2018). Consequently, Bcc bacteria is known to display a wealth of metabolic diversity and be able to colonize diverse niches in nature. Bcc bacteria are indeed able to persist under nutrient limitation (Ahn et al., 2014) and toxic compounds (Geftic et al., 1979; Kim et al., 2015; Ahn et al., 2016), and resist to antimicrobial peptides (Loutet et al., 2006) and clinically used antibiotics (Aaron et al., 2000; Leitao et al., 2008). Additionally, Bcc bacteria do have biotechnological potential due to their versatile ability to degrade complex aromatic pollutants, to produce antifungal compounds, and fix atmospheric nitrogen (Parke and Gurian-Sherman, 2001; Mahenthiralingam et al., 2005; O'Sullivan and Mahenthiralingam, 2005).

## 1.2. Bcc bacteria as opportunistic human pathogens in cystic fibrosis

### 1.2.1. Bcc bacteria and cystic fibrosis

Cystic fibrosis (CF) is a genetic disease that manifests as clinical syndrome characterized by chronic sino-pulmonary infections, sometimes accompanied with gastrointestinal, nutritional and other deficiencies/pathologies (Ratjen and Doring, 2003; Farrell, 2008). CF is a monogenic recessive-inherited disease affecting Caucasians and caused by mutations in the CF-transmembrane conductance regulator (CFTR) gene that results in the unbalance of the ionic composition and high accumulation of viscous mucous in epithelial cells and to impaired innate immunity of the airways (Lyczak et al., 2002; Ratjen and Doring, 2003; Farrell, 2008). CF patients are more susceptible to chronic pulmonary infections and acute exacerbations that lead to progressive respiratory deterioration and increased morbidity and mortality (Lyczak et al., 2002; Ratjen and Doring, 2003; Hatziagorou et al., 2019). CF lung environment is characterized by high pro-inflammatory cytokine levels, high antibiotic concentrations, high levels of oxidative stress and low oxygen concentration (Moriarty et al., 2007; Palmer et al., 2007; Reid et al., 2007; Williams et al., 2007). The CF lung is colonized by several microbial pathogens and is continuously evolving or involved in chronic pulmonary infections (Zemanick and Hoffman, 2016). In early childhood, *Haemophilus influenza* and *Staphylococcus aureus* are the most predominant bacterial species while *Pseudomonas aeruginosa* becomes the most prevalent pathogen in adult CF patients (Lyczak et al., 2002). However, other bacterial pathogens were found to inhabit the CF lungs, such as *Stenotrophomonas maltophilia*, *Aspergillus fumigatus*, and the Bcc bacteria (Lyczak et al., 2002; Zemanick and Hoffman, 2016; Hatziagorou et al., 2019).

Despite the biotechnological relevance of Bcc bacteria, they have gained notoriety as highly problematic opportunistic pathogens and the underlying cause of serious pulmonary infections in immunocompromised individuals and in patients with chronic granulomatous disease and especially in patients with CF (Bernhardt et al., 2003; Reik et al., 2005; Greenberg et al., 2009; Coutinho et al., 2011b). Bcc bacteria emerged as CF pathogens in 1972 when the first report of Bcc pulmonary colonization in CF patients was published (Ederer and Matsen, 1972). Bcc has been considered as a major threat to the CF community since 1980s, when a minority of the infected CF patients in two North America CF centres exhibited a rapid decline in the clinical condition, resulting in early death (Isles et al., 1984; Govan and Deretic, 1996). In general, respiratory infections caused by Bcc bacteria lead to worse prognosis, reduced life expectancy and, in certain cases, to a lethal uncontrolled clinical deterioration with septicemia and necrotizing pneumonia, known as “cepacia syndrome” (Isles et al., 1984; Govan and Deretic, 1996; Jones et al., 2001; Mahenthalingam et al., 2005). As a result, Bcc infected patients are excluded from lung transplantation (Lyczak et al., 2002). Patient-to-patient transmission of Bcc between individuals with CF has been reported leading to devastating infections

(LiPuma et al., 1990; Govan et al., 1993; Smith et al., 1993; Biddick et al., 2003), however, the acquisition of Bcc from environmental sources has also been observed (LiPuma et al., 2002). Bcc bacteria are inherently resistant to several classes of antimicrobials that hinder their eradication from the CF lung by using the current treatment strategies (Leitao et al., 2008; George et al., 2009). Altogether, Bcc bacteria represent a major concern and their infections are particularly threatening and feared by CF community, clinicians, and caregivers. Moreover, during the chronic infections, only one Bcc strain is usually found, but prolonged co-infection with two or more different strains or species can occur, as well as the replacement of an initial infecting strain with another (Richau et al., 2000; Cunha et al., 2003; Yang et al., 2006; Cunha et al., 2007; Lipuma, 2010).

## 1.2.2. Epidemiology of Bcc respiratory infections in cystic fibrosis patients

### 1.2.2.1. *Epidemiology of Bcc respiratory infections in CF patients and epidemic strains*

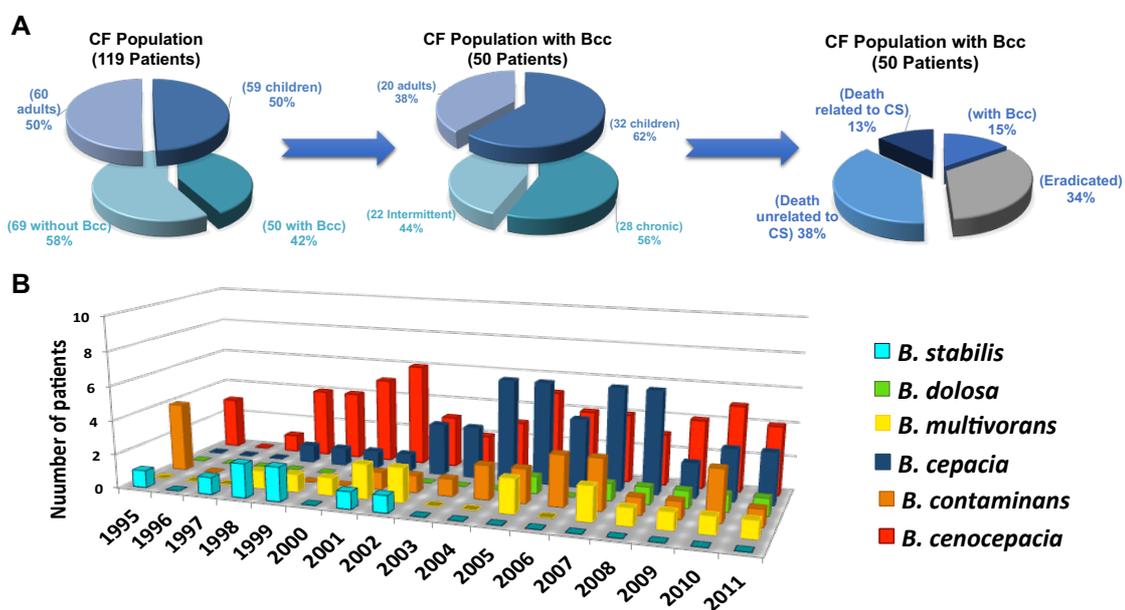
However, although all Bcc species are potential CF pathogens, epidemiological studies showed an uneven geographical and regional distribution among the Bcc species (LiPuma et al., 2001; Speert et al., 2002). *B. cenocepacia* is the most dominant Bcc species with high potential of inter-patient transmission (Speert et al., 2002; Mahenthalingam et al., 2005; Drevinek and Mahenthalingam, 2010) but, in several countries, *B. multivorans* has recently replaced *B. cenocepacia* in this first position (Jones et al., 2004; Lipuma, 2010). Such epidemiological changes among Bcc species may be a result of applying strict infection control measures (e.g. cohort and individual segregation, hand hygiene, and face-masking) (Rowbotham et al., 2019).

The most notorious epidemic *B. cenocepacia* strain, that is considered as an intercontinental clone responsible for several infections and fatalities in CF centres in UK and Canada, is the Edinburgh-Toronto lineage known as the ET-12 clone (Johnson et al., 1994; Holden et al., 2009). Another intercontinental epidemic *B. cenocepacia* strain that was also associated with patient-to-patient transmission (from USA to France, Italy, and UK) is known as Philadelphia-District of Columbia – PHDC (Chen et al., 2001; LiPuma et al., 2002; Coenye et al., 2004). Remarkably, these two intercontinental strains belong to different *B. cenocepacia recA* lineages (IIIA and IIIB, respectively) (Johnson et al., 1994; Chen et al., 2001). Moreover, other epidemic clones were proved to belong to *B. cenocepacia recA* lineage IIIA, a well-studied Bcc lineage: Midwest American clone (Coenye and LiPuma, 2002), CZI Czech epidemic clone (Drevinek and Mahenthalingam, 2010), H111 (Pessi et al., 2013; Carlier et al., 2014) and the epidemic clones ST32 and ST33 (Dedeckova et al., 2013; Miller et al., 2015; Nunvar et al., 2017), whereas other strains belong to *B. cenocepacia recA* lineage IIIB: HI2424 and AU1054 clones (Rozee et al., 1994; Chen et al., 2001; LiPuma et al., 2002; Liu et al., 2003; Coenye et al., 2004). Other Bcc species were also associated with epidemic outbreaks in USA and Europe such as *B. multivorans*, *B. dolosa*, and *B. cepacia* (Biddick et al., 2003). A remarkable case was

registered in the major Portuguese CF Centre at Hospital de Santa Maria (HSM), in Lisbon, where an unusual high representation of *B. cepacia* and, after reclassification of part of the isolates following recent alterations of Bcc taxonomy, of *Burkholderia contaminans* as well was found (Cunha et al., 2007; Coutinho et al., 2011b; Coutinho et al., 2015). Remarkably, an increased incidence of *B. contaminans* as an emerging Bcc species among Argentinian CF patients as well as the Spanish was also recently reported (Martina et al., 2013; Medina-Pascual et al., 2015).

#### 1.2.2.2. Epidemiology of Bcc respiratory infections in the major Portuguese CF centre

The Portuguese CF population under surveillance at the major CF centre at HSM, in Lisbon, included an average 119 CF patients (59 children – up to 18 years old – and 60 adults) [Fig. 1.2; unpublished data and (Coutinho et al., 2011b)]. In 2011, fifty of these patients harbour Bcc bacteria but only 28 were considered to be chronically infected, from whom at least three positive Bcc cultures were obtained during a 6-month period (Cunha et al., 2007; Coutinho et al., 2011b). The epidemiological surveillance of Bcc respiratory infections at this centre was carried out during the past two decades by our research group and a collection of over 800 clinical isolates was sequentially retrieved from persistently infected patients and molecularly identified [(Cunha et al., 2003; Cunha et al., 2007; Correia et al., 2008; Coutinho et al., 2011a; Coutinho et al., 2011b; Moreira et al., 2014; Coutinho et al., 2015) and Chapter III of this thesis].



**Figure 1.2** | CF population under surveillance at the major Portuguese CF treatment Centre at Hospital de Santa Maria, in Lisbon, during the last two decades. A) Characterization of the CF population based on age, presence of Bcc infection, and the clinical outcomes. B) Chronological distribution of different Bcc species based on the last published epidemiological reports (Cunha et al., 2003; Cunha et al., 2007; Coutinho et al., 2011b; Coutinho et al., 2015). CS – cepacia syndrome.

The molecular identification of the above-mentioned isolates revealed that *B. cenocepacia* and *B. cepacia* are the most predominant Bcc species [Fig. I.2; (Coutinho et al., 2011b)]. The unusual increased incidence of *B. cepacia* at the HSM CF Centre was associated to an outbreak between 2003 and 2005 presumably caused by contaminated non-sterile saline solutions for nasal application detected during routine market surveillance by INFARMED – Portuguese Medicines and Health Products Authority (Cunha et al., 2007). Twenty isolates obtained during the aforementioned *B. cepacia* outbreak were reclassified as *B. contaminans* and 21 additional isolates obtained later from a chronically infected patient were also classified as *B. contaminans* (Coutinho et al., 2011b; Coutinho et al., 2015). Such reclassification was performed due to the development in Bcc taxonomy when the species status of several *B. cepacia* isolates with *recA*-RFLP profile K was re-examined since that taxon had been shown to comprise two different species (Vanlaere et al., 2009). Beside the above referred Bcc species, *B. multivorans*, *Burkholderia stabilis*, and *B. dolosa* have also been isolated from the HSM CF population (Cunha et al., 2003; Coutinho et al., 2011b; Moreira et al., 2014). Several cases of prolonged co-infections of different Bcc strains/species were also detected [(Cunha et al., 2003; Coutinho et al., 2011b; Moreira et al., 2014; Coutinho et al., 2015) and chapter III of this thesis].

### 1.2.3. Within-host adaptation of Bcc bacteria during chronic respiratory infections

#### 1.2.3.1. *Bcc bacterial factors involved in host-pathogen interaction and recognition*

Despite the importance of developing therapeutic agents against Bcc infections in CF, our knowledge on Bcc pathogenicity and on how some Bcc species are more successful in the development of respiratory infections are still limited (Sfeir, 2018). Until now, the extent of Bcc bacterial diversification during long-term infection associated with CF patients and the ability of these bacteria to evolve in response to antibiotic pressures, the host immune response system and other host environment imposed stresses is limited compared to other Gram-negative species such as *P. aeruginosa* and *Burkholderia pseudomallei* (Price et al., 2013; Markussen et al., 2014; Marvig et al., 2015; Viberg et al., 2017; Faure et al., 2018; Valentini et al., 2018).

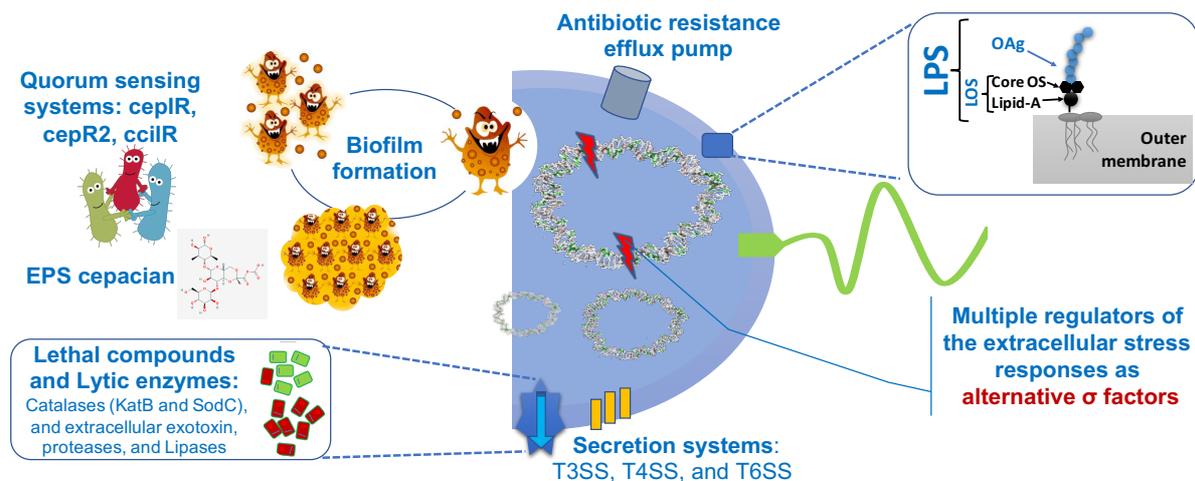
The virulence potential of the Bcc bacteria is believed to be multifactorial (Mahenthiralingam et al., 2005). Several virulence factors have been revealed in Bcc bacteria [reviewed in (Mahenthiralingam et al., 2005; Ferreira et al., 2010; Loutet and Valvano, 2010; Ferreira et al., 2019)], however, the exact roles in the development of CF-infection remain, in general, to be fully understood, especially in the different Bcc species. In this context, the widely-studied Bcc species is *B. cenocepacia* and its epidemic strains that are associated with increased risk of cepacia syndrome' development (Mahenthiralingam et al., 2005). *B. cenocepacia* J2315, the first strain with the complete sequenced-genome (Holden et al., 2009; Agnoli et al., 2012), has been used to identify all the genetic elements associated with *B.*

*cenoepectia* virulence and survival as well as genes involved in regulatory, metabolic, and cell surface biogenesis functions (Hunt et al., 2004; Holden et al., 2009; Agnoli et al., 2012).

During Bcc pulmonary infection, bacterial cells have to adhere to the host mucosal or epithelial cells and this species is eventually followed by subsequent invasion and intracellular survival (Ganesan and Sajjan, 2011; Schwab et al., 2014). For such Bcc cell-CF lung interaction, bacterial surface structures (lipopolysaccharide – LPS, flagella, cable pili, and the 22-KDa adhesin) are considered essential virulence factors promoting adherence and invasion (Sajjan and Forstner, 1993; Tomich et al., 2002; Vinion-Dubiel and Goldberg, 2003; Ganesan and Sajjan, 2011). For instance, studies have shown that Bcc LPS induces a strong immune responses that can contribute to host cell damage (Hutchison et al., 2000) and their chemical structures are unique compared with those from other Gram-negative bacteria (Vinion-Dubiel and Goldberg, 2003). Bcc LPSs contain variable O-antigen (OAg) structures and unusual sugar D-glycero- $\alpha$ -D-talo-oct-2-ulonic acid (KO) in the core oligosaccharide with less phosphates thus reducing the anionic charges of Bcc cell surface and inhibiting the binding of cationic antibiotics (Vinion-Dubiel and Goldberg, 2003). Bcc species also produce metalloproteases, and serine proteases that play a role in the proteolysis of the extracellular matrix impacting the interaction with epithelial cells (Kooi et al., 2006; McClean and Callaghan, 2009; Madeira et al., 2013; Shinoy et al., 2013). All Bcc species produce extracellular lipases that play a role in invasion (Caraher et al., 2007; Mullen et al., 2007; McClean and Callaghan, 2009). Protein secretion in Bcc bacteria was demonstrated to be influenced by several transport and secretion systems as in other Gram-negative bacteria (Spiewak et al., 2019). Type I and type II secretion systems (T1SS and T2SS) from Bcc strains of *B. cenoepectia* ET-12 lineage and *B. veitnamiensis* were found to be responsible for the proteases' secretion, as it is the case of the two zinc metalloproteases ZmpA and ZmpB (Corbett et al., 2003; Kooi et al., 2006) with impact in bacterial pathogenesis (Fehlner-Gardiner et al., 2002; Whitby et al., 2006).

*B. cenoepectia* and *B. multivorans* are known for their ability of survival inside macrophages (Saldias et al., 2009; Saldias and Valvano, 2009; Kotrange et al., 2011; Schmerk and Valvano, 2013) and it was hypothesized that the loss of OAg may benefit the ability of *Burkholderia* to access the intracellular environment, having no similar benefits for *P. aeruginosa* (Saldias et al., 2009). The exopolysaccharide (EPS) biosynthesis, biofilm formation, and resistance to antibiotics and oxidative stress in general (Lefebvre and Valvano, 2001; Mahenthalingam et al., 2005; Ferreira et al., 2019), as well as iron acquisition ability (Miethke and Marahiel, 2007) are also among the more described virulence determinants for Bcc intracellular survival [some of them are illustrated in Fig. 1.3 and (Mahenthalingam et al., 2005; Ferreira et al., 2010; Loutet and Valvano, 2010; Ferreira et al., 2019)]. For instance, cepacian is the most common EPS produced by Bcc from clinical and environmental isolates (Chiarini et al., 2004; Cunha et al., 2004; Zlosnik et al., 2008; Ferreira et al., 2010; Ferreira et al., 2019) that was proved to interfered with the *B. cenoepectia* phagocytosis by neutrophils and altered

the bacterial clearance in mice (Conway et al., 2004). This role of cepacian polysaccharides was proposed to interfere with the innate immune system by neutralizing the reactive oxygen species –ROS and inhibiting neutrophils' functions by modulating the the phagocytosis and scavenging the chemotaxis (Conway et al., 2004; Bylund et al., 2006; Cuzzi et al., 2012), most likely due to masking of the bacterial surface antigen recognized by immune cells (Ferreira et al., 2019). Additionally, global regulators were proved to regulate the ability of intracellular survival of *B. cenocepacia* by the phagolysosomal fusion in macrophages, such as RpoE and RpoN (sigma factor  $\sigma^{54}$ ) (Flanagan and Valvano, 2008; Saldias et al., 2008; Valvano, 2015b; Schaefer et al., 2017).



**Figure 1.3** | Graphical representation of some virulence factors previously-reported for Bcc bacteria. The image of the bacterial cells respecting to the biofilm and respecting to quorum sensing were reproduced from <https://imgbin.com/>.

#### 1.2.3.2. *Bcc phenotypic adaptation during chronic infection*

Different virulence factors have been described in Bcc bacteria (Loutet and Valvano, 2010; Suppiger et al., 2013; Sousa et al., 2017; Butt and Thomas, 2018; Hall and Lee, 2018; Ferreira et al., 2019) but less is known concerning their adaptive traits during chronic infection, compared with *P. aeruginosa* (Valentini et al., 2018). To fulfil this gap, longitudinal studies were performed to compare serial isolates obtained from the same CF patients during chronic infection concerning certain phenotypes.

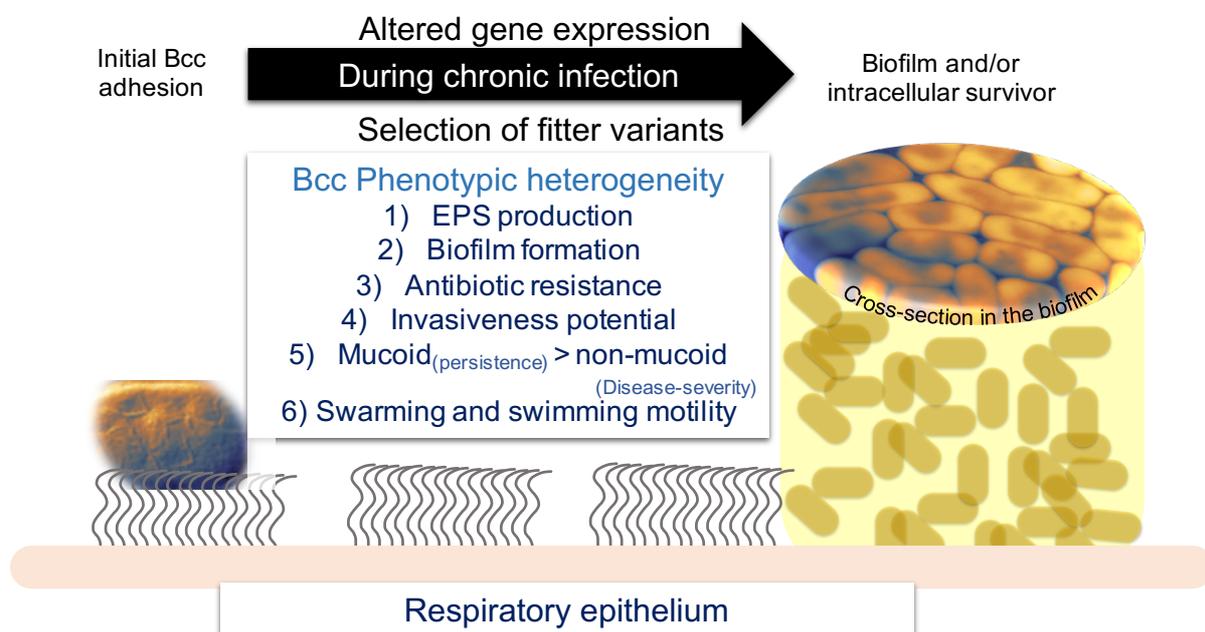
The first retrospective reports using this approach was designed in our research group to demonstrate that although Bcc EPS production is not required for the initiation of biofilm formation, cepacian plays a role in the establishment of thick biofilms (Richau et al., 2000; Cunha et al., 2004). This involved the systematic characterizations of 108 clinical isolates, corresponding 15 strains belonged to *B. cepacia* (*B. contaminans* following re-classification (Coutinho et al., 2015)), *B. multivorans*, *B. cenocepacia*, and *B. stabilis*, obtained during a 7-year period of surveillance from 21 chronically infected CF patients

(Richau et al., 2000; Cunha et al., 2004). Despite the concept that bacteria growing in biofilms display higher resistance to antibiotics and to host phagocyte-killing than planktonically growing cells, the results obtained in this report indicated that the development of thick biofilms involves other strain-dependent factors besides the EPS (Cunha et al., 2004). Altogether with the information on the clinical outcomes of the infected patients, Cunha and her colleagues concluded that the persistence and virulence of respiratory infections caused by Bcc bacteria also depend on other determinants besides the abilities to produce EPS and to form mature biofilms (Cunha et al., 2004). For a better understanding of the prolonged antibiotics-selective pressures that Bcc bacteria have to face during the chronic infections, 94 Bcc isolates obtained from 15 CF patients during 5 years of infections were compared, also in our research group, to evaluate the susceptibility patterns toward 13 antimicrobials (Leitao et al., 2008). These isolates represented 11 different strains of *B. cepacia* (and *B. contaminans* (Coutinho et al., 2015)), *B. multivorans*, *B. cenocepacia* (*recA* lineage IIIA and IIIB), and *B. stabilis* species (Leitao et al., 2008). The results were consistent with the notion that clinical Bcc isolates are resistant to the most clinically relevant antimicrobials and suggested an uneven distribution of resistance rates among the different Bcc species, with *B. cenocepacia recA* lineage IIIA isolates being the most resistant (Leitao et al., 2008). Leitao and colleagues also noticed that the phenotype of tolerance to antimicrobials varied during long-term of infection but the isolation of resistant variants coincided with periods of pulmonary exacerbation and antibiotics therapy (Leitao et al., 2008).

Other retrospective phenotypic longitudinal assessments of the variation of mucoid exopolysaccharide production (Zlosnik et al., 2008) revealed the switch from mucoid to non-mucoid phenotype (Zlosnik et al., 2011) and the loss of the swimming motility was not typical among Bcc bacteria (Zlosnik et al., 2014). The study was performed during chronic infection involving numerous Bcc isolates representing different Bcc species (*B. cenocepacia recA* lineages IIIA and IIIB, *B. multivorans*, *B. cepacia*, *B. stabilis*, *B. vietnamiensis*, *B. dolosa*, *B. ambifaria*, *B. anthina*, and *B. pyrocinia*) and collected from 100 CF patients in the Vancouver area, Canada. It was demonstrated that all nine species of the Bcc can express the mucoid phenotype (Zlosnik et al., 2008) but the *In vitro* incubation of Bcc with ceftazidime and ciprofloxacin, but not meropenem, caused conversion of the mucoid to the nonmucoid phenotype (Zlosnik et al., 2011). This report also suggested an inverse correlation between the quantity of EPS production by Bcc bacteria and the rate of decline in CF lung function (Zlosnik et al., 2011). Conversion of the mucoid-to-non-mucoid phenotype during chronic infection, together with the disproportionately high frequency of non-mucoid isolates among strains of *B. cenocepacia*, the most virulent Bcc species, was proposed to strengthen the concept of the association of the non-mucoid phenotype with the severity of the disease and the mucoid phenotype with persistence (Zlosnik et al., 2008; Zlosnik et al., 2011). This behaviour contrasts with the phenotypic switch from non-mucoid to the mucoid in *P. aeruginosa* chronic infections as a well-established paradigm for infection disease severity and persistence in this species (Valentini et al., 2018). Furthermore, from

another study, non-mucoid late-stage clonal variants showed no exopolysaccharide production, lower motility and chemotaxis, and more biofilm formation, particularly under microaerophilic conditions, than the corresponding mucoid early isolate (Silva et al., 2011; Ferreira et al., 2019). Moreover, upon establishment of chronic infection, subsequent *P. aeruginosa* isolates show a reduction in the swimming ability to be non-motile (Valentini et al., 2018) also contrasting with the observed data in Bcc infections (Zlosnik et al., 2014). This observation was obtained from a retrospective phenotypic longitudinal assessment in 551 clinical Bcc isolates (Zlosnik et al., 2014). Furthermore, while Zlosnik and colleagues observed a statistically significant link between mucoidy and motility, they did not detect any link between the motility phenotype and clinical outcome (Zlosnik et al., 2008; Zlosnik et al., 2011; Zlosnik et al., 2014; Ferreira et al., 2019).

The understanding of the occurrence of these and other adaptive traits have been well-described in *P. aeruginosa* chronic infections (Valentini et al., 2018), but are far less known in Bcc chronic infection (summarized in Fig. 1.4). It is however crucial for a better understanding of Bcc bacterial transformation in the CF lungs during long term infections and infection control. Altogether, these studies also highlight the need for further work to understand the adaptive changes of Bcc bacteria during chronic infection in the CF lung.



**Figure 1.4** | Reported phenotypes found to suffer diversification during Bcc chronic infections (Cunha et al., 2004; Leitao et al., 2008; Zlosnik et al., 2008; Silva et al., 2011; Zlosnik et al., 2011; Zlosnik et al., 2014; Ferreira et al., 2019).

#### 1.2.3.3. Mechanisms underlying genome-wide Bcc adaptation to the CF lung

Inside the CF lung, Bcc bacteria face stressful and changing environmental conditions as a consequence of high pro-inflammatory cytokine levels, high antibiotic concentrations, high levels of oxidative stress, low oxygen concentration and the presence of other co-infecting microbes (Moriarty et al., 2007; Palmer et al., 2007; Reid et al., 2007; Williams et al., 2007; Hogardt and Heesemann, 2010; Döring et al., 2011; Cullen and McClean, 2015). It is likely that the adaptive evolution of each Bcc species strain may vary in different CF patients' lung-environments and consequently, during long-term infection, the original infecting strain(s) undergo(es) extensive genetic and phenotypic diversification inside the CF lung thus enabling the emerging clonal variants more adapted to the CF lung environment and capable of evading the host immune system (Coutinho et al., 2011a; Lieberman et al., 2011; Madeira et al., 2011; Mira et al., 2011; Silva et al., 2011; Madeira et al., 2013; Lieberman et al., 2014; Moreira et al., 2014; Maldonado et al., 2016; Moreira et al., 2016; Nunvar et al., 2016; Silva et al., 2016; Cabral et al., 2017; Hassan et al., 2017; Lee et al., 2017; Moreira et al., 2017; Nunvar et al., 2017; Hassan et al., 2019). The within-patient emergence of multiple clonal variants was first described for the more prevalent CF pathogen *P. aeruginosa* (Markussen et al., 2014; Winstanley et al., 2016) and proposed to provide a pool of mutations affecting virulence and antimicrobial resistance (Lorè et al., 2012; Marvig et al., 2013; Faure et al., 2018). The exploitation of comparative genomic, transcriptomics, proteomics, and metabolomics analyses together with the phenotypic screenings have been essential to get clues on the global strategies used by Bcc bacteria to adapt to the highly stressful environment of the CF lung during chronic infection similar to those strategies described for *P. aeruginosa* (Coutinho et al., 2011a; Lieberman et al., 2011; Madeira et al., 2011; Mira et al., 2011; Silva et al., 2011; Madeira et al., 2013; Lieberman et al., 2014; Markussen et al., 2014; Moreira et al., 2014; Maldonado et al., 2016; Moreira et al., 2016; Nunvar et al., 2016; Silva et al., 2016; Winstanley et al., 2016; Cabral et al., 2017; Hassan et al., 2017; Lee et al., 2017; Moreira et al., 2017; Nunvar et al., 2017; Hassan et al., 2019). Several studies have been detailing alterations occurring in the genome sequences using isolates obtained during long term infections of few Bcc species: *B. dolosa* (Lieberman et al., 2011; Lieberman et al., 2014), *B. multivorans* (Silva et al., 2016) and *B. cenocepacia* (Lee et al., 2017; Nunvar et al., 2017). Additionally, the molecular mechanisms underlying experimental evolution have been performed to elucidate the evolutionary trajectories of *B. cenocepacia recA* lineage IIIA, the highly-virulent lineage among Bcc pathogens (Drevinek and Mahenthiralingam, 2010), and *B. multivorans* (Traverse et al., 2013; Cooper, 2014; Cooper et al., 2014; Dillon et al., 2015) under selective conditions. Altogether, the understanding of the Bcc adaptation (in different species) as a function of the selective pressures of the CF-hostile environments (Fig. 1.5) remains still elusive compared with *P. aeruginosa*.

Retrospective studies involving 112 sequential isolates of *B. dolosa* from 14 CF patients led to the identification of several mutated genes involved in genetic variation of this bacterial pathogen within individual patients, in particular genes required for expression of surface polysaccharides,

lipopolysaccharide O-antigen (OAg), outer membrane components, iron scavenging, and antibiotic resistance (Lieberman et al., 2011) and suggested parallel bacterial evolution within multiple patients (Lieberman et al., 2014). These studies showed that different lineages may coexist for many years within a patient and identified candidate pathogenicity genes (Lieberman et al., 2014). Another comparative genomic analysis was performed using three *B. cenocepacia* clonal variants obtained from the same CF patient during a 10-year period of chronic infection and supported the concept of phenotypic heterogeneity and proposed that several pathways are under strong selective pressures inside the CF lungs (Miller et al., 2015). The proposed pathways and genes under selective pressures were further investigated by another comparative genomic study including 22 isolates of *B. multivorans* recovered over 20 years from CF patients (Silva et al., 2016). Silva and colleagues also found that genes involved in antibiotic resistance, cell wall and membrane composition, lipopolysaccharide O-antigen biosynthesis, metabolism and oxygen-sensing were under convergent evolution inside the CF lungs (Silva et al., 2016). A subsequent comparative genomic analysis including 32 clonal variants of *B. cenocepacia* obtained from 8 CF patients (Nunvar et al., 2017) revealed that, in addition to the aforementioned parallel mutations in gene functions already described for *B. cenocepacia*, *B. dolosa* and *B. multivorans* (Lieberman et al., 2011; Lieberman et al., 2014; Miller et al., 2015; Silva et al., 2016), genes related with transition metal metabolism are hotspots for nucleotide polymorphism (Nunvar et al., 2017). Another comparative study involving 215 genomes from serial *B. cenocepacia* isolates obtained from 16 CF patients during a 20 year-period also support the above mentioned evolutionary trajectories during chronic Bcc infections and the well-established diverse-community model (Lee et al., 2017). This publication also reported the complete loss of chromosome III resulting in genome-size reduction as an adaptive trait of Bcc bacteria (Lee et al., 2017), consistent with a previously reported rare *in-vivo* loss of the same non-essential chromosome among *B. cenocepacia* clonal variants that emerged during chronic infections (Moreira et al., 2017).

As a function of prolonged antibiotic-selective pressures, several studies using sequential *B. cenocepacia* and *B. dolosa* isolates revealed that the most resistant variants were isolated after pulmonary exacerbation and aggressive antibiotic therapy compared with the early clones (Leitao et al., 2008; Coutinho et al., 2011a; Moreira et al., 2014). This concept was further explored, in our research group, by transcriptomic and proteomic analyses using *B. cenocepacia* clinical isolates and revealed that the higher resistance to several antimicrobials, presumably resulting from altered expression of multiple genes implied in antibiotic resistance; genes involved in multidrug efflux, antibiotic degradation/modification, outer membrane permeability, and regulations (Madeira et al., 2011; Mira et al., 2011; Madeira et al., 2013). The observed alterations were also determined in another transcriptomic study involving two *B. cenocepacia* isolates belong to ET-12 lineage (Sass et al., 2011). Based on genomic analysis, a retrospective study involving *B. dolosa* isolates revealed that two non-synonymous polymorphisms (SNPs) in several variants, from different patients, were found in a gene homologous

to *Escherichia coli gyrA* that confers fluoroquinolone resistance (Lieberman et al., 2011). Remarkably, one of these two SNPs was also found in late *B. cenocepacia* variants (Miller et al., 2015). Also, another mutation was observed in late *B. multivorans* clonal variants in a gene encoding a component of an efflux pump (similar to GyrA protein) and confirmed that isolates carrying that mutation had enhanced antibiotic resistance (Silva et al., 2016).

Bcc airway colonization is associated with neutrophil inflammation that is known as a significant cause of tissue damage and lung function deterioration, in which ROS and reactive nitrogen species imposes oxidative and nitrosative stresses to the infecting bacterial species (Hull et al., 1997; Hutchison et al., 2000). As a result of oxidative stresses, it was proved that *B. cenocepacia* bacteria downregulate genes associated with transcription and energy production while upregulate those involved in stress-resistance as it is the case of detoxification processes (Sass et al., 2013). These adaptive strategies were recently suggested by a retrospective comparative genomic study involving *B. cenocepacia* clonal variants where genes functionally related to oxidative stress protection were identified to acquire non-synonymous SNPs resulted in impaired detoxification of hydrogen peroxide (Nunvar et al., 2017). Other adaptive traits used by Bcc bacteria were also proposed: in particular those include bacterial cell translocation across polarized epithelial layers with disruption of tight junctions and invasion of the epithelial cells and macrophages, which presumably leads to the Bcc evasion from the immune responses (Chiu et al., 2001; Duff et al., 2006; Madeira et al., 2011). As a function of CF-immune responses and in agreement of the evasion adaptive-strategy, comparative genomic studies using longitudinal *B. dolosa*, *B. multivorans* and *B. cenocepacia* clonal variants showed that a commonly reported evolutionary trait is the genetic changes occurring in genes present in the OAg biosynthetic cluster (Lieberman et al., 2011; Lieberman et al., 2014; Maldonado et al., 2016; Silva et al., 2016; Hassan et al., 2017; Nunvar et al., 2017). This alteration is believed to facilitate the ability of *B. multivorans* and *B. cenocepacia* intracellular survival inside macrophages, having no similar benefits for *P. aeruginosa* (Saldias et al., 2009; Saldias and Valvano, 2009; Kotrange et al., 2011; Schmerk and Valvano, 2013). Furthermore, *B. cenocepacia* survival inside epithelial cells and macrophages was believed to be linked to impairment of endosome-lysosome fusion and to the expression of type IV secretion system (Lamothe et al., 2007; Sajjan et al., 2008) while *B. multivorans* intracellular survival is apparently based on a different strategy (Schmerk and Valvano, 2013). Given the correlation previously established between *in-vitro* *B. cenocepacia* epithelial cell invasion and *in-vivo* virulence in a mouse model (Cieri et al., 2002), it is expectable that increased epithelial invasiveness can have implications in *in-vivo* virulence. However, the adaptive evolution of Bcc bacteria as a function of the immune system responses and antibiotic therapy during long-term chronic infection requires further understanding. It was also proposed that a better performance of DNA repair system is essential to overcome bacterial cell damage induced by the host-immune responses and the antibiotic therapy (Foti et al., 2012). This concept was supported by a quantitative proteomic analysis that suggested an increase

in nucleotide biosynthetic activity in the late *B. cenocepacia* clonal variant compared with the early isolate (Madeira et al., 2013). Other studies have found that genes associated with purine and pyrimidine biosynthetic pathways, are essential for *B. cenocepacia* virulence in different non-mammalian infection models and crucial for bacterial growth and survival in the human bloodstream (Samant et al., 2008; Schwager et al., 2013).

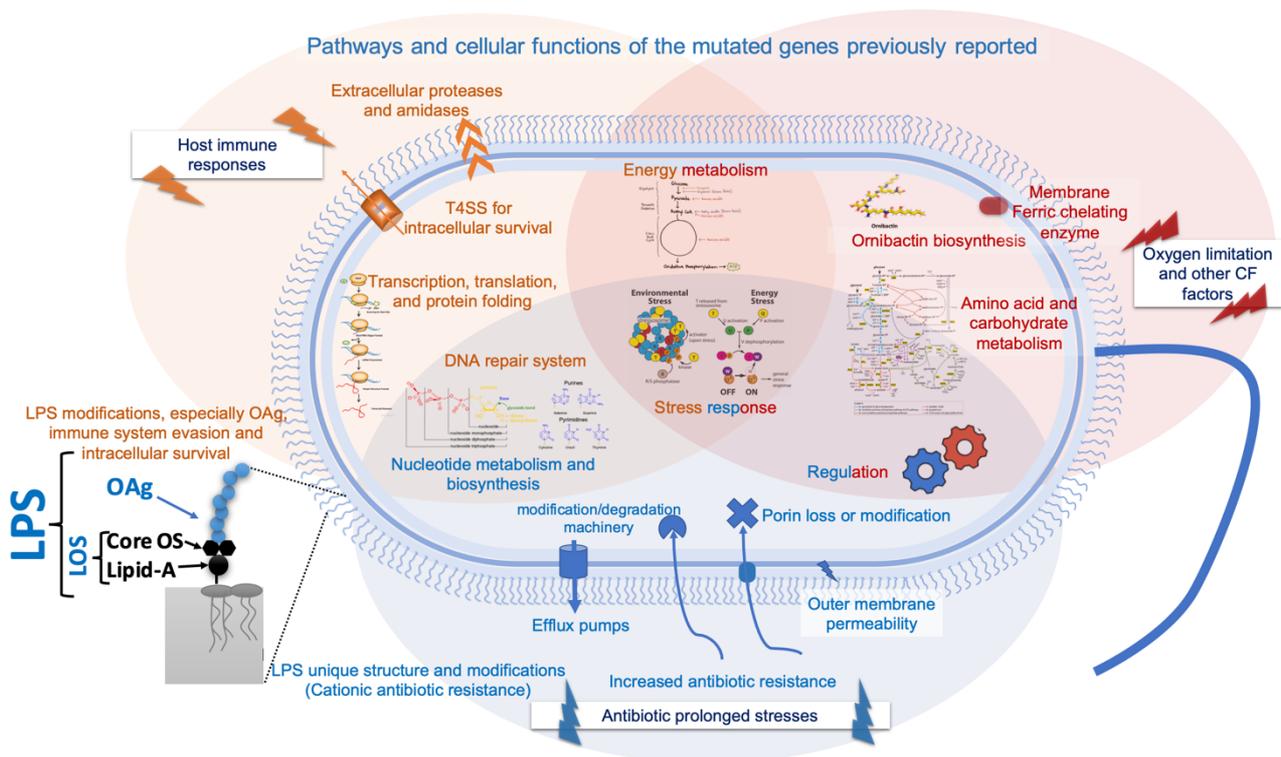
Another characteristic factor of the CF lung-hostile environments is the low oxygen concentration usually present with approximately 7  $\mu$ M dissolved oxygen (Worlitzsch et al., 2002; Yang et al., 2011b). Several studies have shown the capacity of *B. cenocepacia* bacteria to survive and grow in these oxygen-limiting conditions (Pessi et al., 2013; Sass et al., 2013; Schwab et al., 2014). Other comparative studies, at the transcriptomic and proteomic levels using two late *B. cenocepacia* clonal variants in comparison with the early isolate, showed that the expression of genes and proteins associated with metabolic processes, namely amino acid, carbohydrates, and energy production, are altered (Madeira et al., 2011; Mira et al., 2011; Madeira et al., 2013), suggesting that nutritional adaptation might also be due to the microaerobic conditions inside the CF lungs as previously observed in *B. cenocepacia* and in *P. aeruginosa* (Drevinek et al., 2008; Hoboth et al., 2009; Oberhardt et al., 2010; O'Grady and Sokol, 2011). Additional evidence supporting the adaptation of late *B. cenocepacia* clonal variants to low-oxygen tension was obtained by the increased transcript levels of genes coding for CioAB cytochrome subunits (Mira et al., 2011), previously shown to be involved in *P. aeruginosa* aerobic respiration under microaerophilic conditions (Oberhardt et al., 2010). *B. multivorans* and *B. cenocepacia* were also demonstrated to be able to survive under anaerobic conditions using glucose fermentation (Schwab et al., 2014). However, although all Bcc species have genes necessary for lactic and acetic acids fermentation, they lack other essential genes for such survival under the anaerobic conditions as described in *P. aeruginosa*, such as genes involved in arginine fermentation and denitrification, the process implied in ATP generation in similar condition (van Hartingsveldt and Stouthamer, 1973; Vander Wauven et al., 1984; Pessi et al., 2013). This concept was further explored using transcriptomic and proteomic approaches that revealed the upregulation of genes involved in the glyoxylate shunt, polyhydroxyalkanoates (PHA) biosynthesis and of a low-oxygen-activated (*lxa*) locus that was reported in *B. cenocepacia* J2315 genome (Pessi et al., 2013; Sass et al., 2013). This *lxa* locus contains genes associated with stress response, metabolism, transport, electron transfer and regulation, and was proved to be essential for *B. cenocepacia* viability in the absence of oxygen (Sass et al., 2013). Longitudinal genomic analyses also demonstrated that the *fixL* gene from *B. multivorans* and its homolog BDAG\_01161 from *B. dolosa* acquire the highest number of mutations in clinical isolates examined during chronic infection inside the CF lungs (Lieberman et al., 2011; Lieberman et al., 2014; Silva et al., 2016). Orthologues, from other microorganisms, of this *fixL* gene have been implicated in oxygen sensing (Marteyn et al., 2010); due to the low-oxygen pressures present in the CF lungs, it is possible that genetic alterations in this *Burkholderia* regulator are involved in bacterial adaptation to

microaerophilic conditions. Microaerophilic growth conditions lead to profound changes in Bcc physiology, as suggested by the observed increase in biofilm formation, antibiotic resistance and production of virulence factors (Pessi et al., 2013).

Based on DNA microarray analyses, 10% of the 7176 genes of *B. cenocepacia* J2315 genome were altered when *B. cenocepacia* grew in CF sputum, these genes are involved in pathways required for responses to antimicrobial resistance, oxidative stress, and iron metabolism (Drevinek et al., 2008). Recognized virulence factors such as the cable pilus and *Cenocepacia* Pathogenicity Island (*cci*) were unaltered in expression while *B. cenocepacia* show sustained or increased expression of motility-associated genes when grew in sputum, maintaining a potentially invasive phenotype associated with *cepacia* syndrome (Drevinek et al., 2008). Bcc bacteria are proved to degrade mucins, found in high levels in CF mucopurulent secretions, suggesting that mucins and other high molecular weight polymers are *in-vivo* available nutrients during chronic Bcc infections (Schwab et al., 2014). In particular in the Bcc species *B. cenocepacia*, arginine could not be utilized for energy production, unlike other Bcc species and *P. aeruginosa* (Schwab et al., 2014) although in the CF airways the selection and persistence of auxotrophic variants is favoured (Barth and Pitt, 1995). Furthermore, genomic microarray analysis showed a significant upregulation of genes involved in the siderophore pyochelin biosynthesis and membrane ferric reductase, suggesting that mechanisms for iron acquisition are essential for successful Bcc colonization (Drevinek et al., 2008). These mechanisms were also suggested by other transcriptomic and proteomic studies on *B. cenocepacia* clonal variants (Madeira et al., 2011; Mira et al., 2011; Madeira et al., 2013; Sass et al., 2013) or comparative genomic analysis of *B. dolosa* clonal variants (Lieberman et al., 2011; Lieberman et al., 2014), reinforcing the idea that iron acquisition has a key role in Bcc adaptive response to CF lung environment.

Comprehensive comparative genomic sequence analyses performed in isolates retrieved during chronic infection on the most feared *B. cenocepacia* (Nunvar et al., 2017) and *B. multivorans* (Silva et al., 2016) and on the far less distributed *B. dolosa* (Lieberman et al., 2011) revealed that the genes affected during *B. cenocepacia* and *B. multivorans* adaptive evolution are different from *B. dolosa*. This fact indicates that evolution in *B. dolosa* during chronic CF infection is driven by different selective forces presumably linked to host immune responses (Lieberman et al., 2014; Nunvar et al., 2017). Specifically, while the most mutated genes reported to undergo adaptive within-patient evolution in *B. cenocepacia* were associated with oxidative stress response and transition metal metabolism, the same was not observed either in *B. dolosa* or in *P. aeruginosa* (Lieberman et al., 2014; Nunvar et al., 2017). These genes encode proteins required for protection against the ROS produced by leukocytes thus suggesting the involvement of the host immune system in driving *B. cenocepacia* evolution during chronic CF infection (Nunvar et al., 2017). Given that persistent inflammation and neutrophil infiltration often accompany chronic lung infections, the authors hypothesized that under increased

stress encountered in CF macrophages, the global stress response might be activated by the evolved bacterial population and thus modulate the course of infection (Chua et al., 2016; 2017; Nunvar et al., 2017). Since the interactions between pathogens and the immune system are highly complex, affecting pathogen adaptation to different host immune system stimuli, further investigation into the roles of macrophages and defence mechanisms in chronic infection outcome is needed (Chua et al., 2016; 2017). Since *B. cenocepacia* is the most-widely studied Bcc species, a working model summarizing the most relevant findings proved by previous reports in the literature is proposed in Fig. 1.5. Although these studies are considered significant input into the current knowledge about the underlying global mechanisms of *B. cenocepacia* adaptive strategies, much remains to be elucidated. Also, studies focusing on the mechanisms of adaptation of other Bcc species are still lacking, though a number of studies on this topic were published in the last few years. Finally, the complementation of these results with other Omics approaches, in particular metabolomic analyses, is expected to increase our knowledge into the adaptive mechanisms used by Bcc bacteria during long-term chronic infection inside the CF lung (Shommu et al., 2015; Moreira et al., 2016).



**Figure 1.5** | Pathways and cellular functions related with the mutated genes previously described to occur during Bcc bacteria evolution inside the CF lung. This model is based on the current knowledge gathered from phenotypic, genomic, transcriptomic, proteomic, and metabolomic analyses of sequential isolates retrieved from chronically infected CF patients. The reported mechanisms are introduced as a function of the well-characteristic -CF lung- environmental factors: Host immune responses (Orange – up left), Antibiotic prolonged stresses (Blue – bottom center) and Oxygen limitation and other CF factors (Red – up right).

## 1.2.4. Bcc cell surface in host-pathogen interaction and adaptation to the CF lung

### 1.2.4.1. *The cell surface of Bcc bacteria*

During chronic infection, Bcc bacterial need to sense their proximity and location within the host environment and consequently Bcc have developed various means to promote their survival in host environments. For such host-pathogen interaction, bacterial cell envelope and its surface structures are considered essential virulence factors promoting bacterial adhesion and invasion (Sajjan and Forstner, 1993; Tomich et al., 2002; Vinion-Dubiel and Goldberg, 2003; Ganesan and Sajjan, 2011). Like other Gram-negative bacteria, Bcc cell envelope is a complex multi-layered structure that acts as a protective barrier from the unpredictable and often hostile environment (Silhavy et al., 2010; Cunneen et al., 2011; Gislason et al., 2017). The lipopolysaccharide (LPS) is a tripartite structure allocated in the outer leaflet of the outer membrane of the cell envelope, considered a central unique compartment in Gram-negative bacteria that is believed to play a key role in pathogenesis (Whitfield and Trent, 2014; Fathy Mohamed et al., 2019; Oppy et al., 2019). The LPS-Lipid A was known as an endotoxin responsible for the endotoxic shock/fever associated with septicemia caused by Gram-negative bacteria (Raetz and Whitfield, 2002; De Soyza et al., 2008). Together with the lipid A, the core oligosaccharide (core OS) mainly form the LPS associated with rough colony-morphotype and often called lipooligosaccharides (LOS) while when to both structures the OAg is linked, the complete capped LPS is formed providing smooth colony-morphotype (Raetz and Whitfield, 2002; Whitfield and Trent, 2014). The lipid A-core OS are highly conserved among the Gram-negative bacteria (Raetz and Whitfield, 2002; De Soyza et al., 2008; Whitfield and Trent, 2014). As above-mentioned for other Gram-negative bacteria, the LPS from Bcc induces a strong immune response that can contribute to host cell damage (Hutchison et al., 2000; Vinion-Dubiel and Goldberg, 2003). However, Bcc bacteria LPS differs from those from other Gram-negative bacteria LPSs due to the presence of 4-amino-4-deoxyarabinose (Ara4N) residues, either in the core or in lipid A, and to the unique disaccharide structure of the core OS – that is described below (Vinion-Dubiel and Goldberg, 2003; De Soyza et al., 2008; Olagnon et al.). The unusual structure of Bcc LPS decreases the anionic charge of the Bcc cell surface, thus inhibiting the binding and subsequent effects of cationic antibiotics (Vinion-Dubiel and Goldberg, 2003; De Soyza et al., 2008). Also, the addition of the positively charged groups in Gram-negative bacteria, including Bcc, as it is the case of phosphoethanolamine (Lee et al., 2004; Kim et al., 2006), 4-L-amino-4-deoxy-arabinose (L-Ara4N) (Trent et al., 2001; Gislason et al., 2017; Olagnon et al.), galactosamine (Wang et al., 2009), and glucosamine (Marr et al., 2008), reduces the net negative charges of the bacterial surfaces promoting resistance to antimicrobial peptides as polymyxin B. The degree of acylation and phosphorylation of the lipid A is also known to be vital for the human Toll-like receptor 4 - Myeloid differentiation factor 2 (TLR4-MD2) recognition and outer membrane (OM) stability and reduce the diffusion of hydrophobic substances (Needham and Trent, 2013).

There are only a few studies discussing the adaptation of different Bcc species focusing on the LPS and *B. cenocepacia* is the only well-studied species (Lorenzo et al., 2013; Schmerk and Valvano, 2013; Valvano, 2015a; Maldonado et al., 2016). As mentioned in previous sections, Bcc bacteria modify and remodel the LPS macromolecules in response to the environmental hostile conditions in order to survive and develop infection inside the host (Raetz and Whitfield, 2002; De Soyza et al., 2008; Valvano, 2015a). Bcc LPS is believed to be the major constituent to activate the host innate system; the lipid A is detected by Toll-like receptor 4 (TLR4) forming a complex with the amyloid differentiation factor 2 (both were found on the cell surface of macrophages and monocytes) (Di Lorenzo et al., 2015). Several studies have demonstrated that Bcc LPS when interact with neutrophils, the expression of CD11b on neutrophil' surface increases, stimulating neutrophil respiratory burst response (Hughes et al., 1997). In addition, macrophages and human blood cells are also stimulated by Bcc LPS, producing pro-inflammatory cytokines such as TNF- $\alpha$ , IL-6 and IL-8 (Hutchison et al., 2000; Shimomura et al., 2001). Additionally, the OAg expression in Bcc bacteria was illustrated to reduce phagocytosis by macrophage (Saldias et al., 2009; Saldias and Valvano, 2009).

#### 1.2.4.2. LPS biosynthesis in Bcc

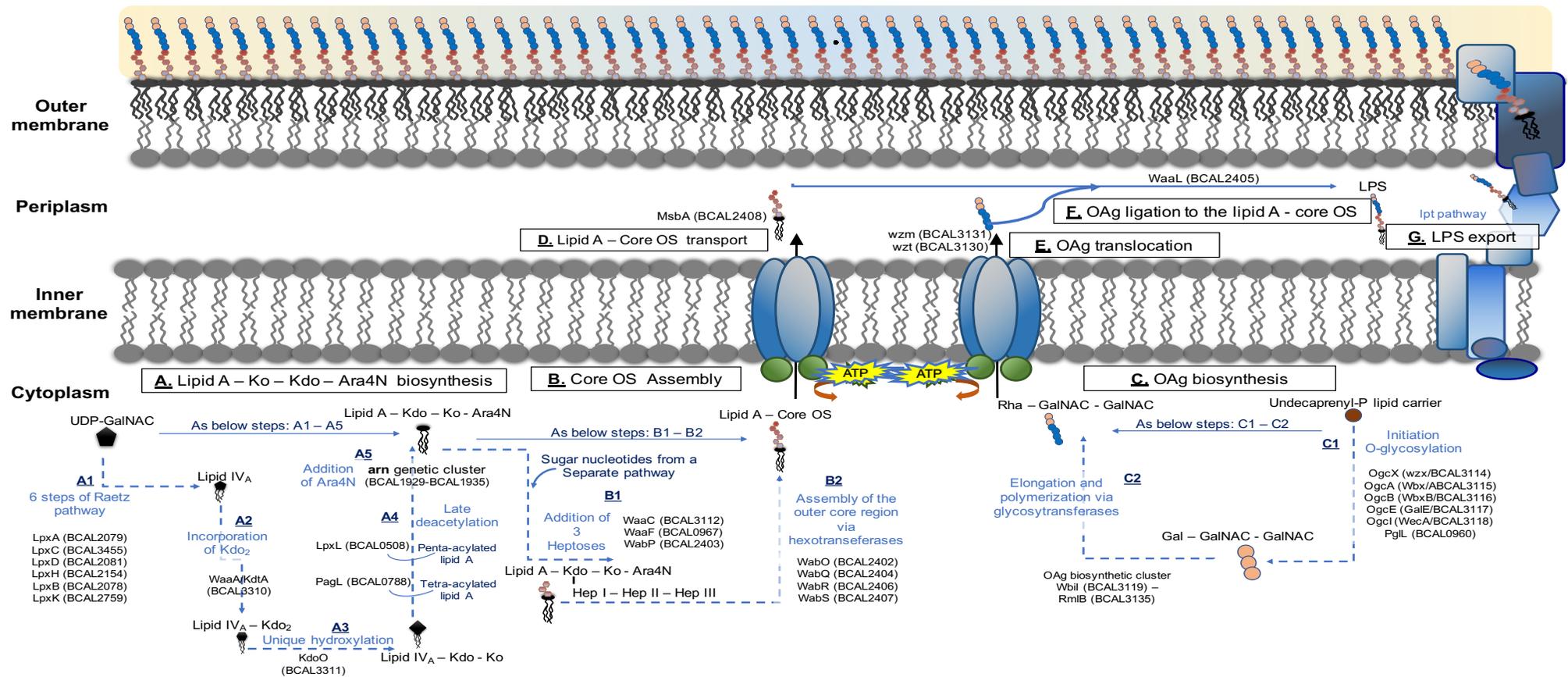
The *B. cenocepacia* lipid A consists of two linked glucosamine residues that contain phosphodiester-linked to L-Ara4N residues (Silipo et al., 2005). Unlike other bacteria, L-Ara4N residues have been proved to be essential for *B. cenocepacia* viability and LPS translocation into the outer membrane (Ortega et al., 2007; Hamad et al., 2012; Fathy Mohamed et al., 2017). Structurally, core OS made of a disaccharide, 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) and D-glycero-D-talo-oct-2-ulosonic acid (Ko), and anchored to the lipid A to form distinct Bcc Lipid A-core OS, instead of the prototypical Kdo-Kdo disaccharide core OS present in other Gram-negative (Silipo et al., 2005). Such Ko monosaccharide has only been found, as component of bacterial LPS in a small number of bacteria: *Acinetobacter* (Vinogradov et al., 1997; Vinogradov et al., 1998; Vinogradov et al., 2002a), *Yersinia pestis* (Vinogradov et al., 2002b), and *B. caryophylli* (Molinaro et al., 2002). The rest of the core OS is attached to the Kdo residue of the lipid A-Kdo-Ko-L-Ara4N molecule (De Soyza et al., 2008). In fact, the Ko residue is also glycosylated by L-Ara4N (Silipo et al., 2005; De Soyza et al., 2008). Several *B. cenocepacia* isolates also contain an OAg ligated to lipid A-core OS, mainly composed of a repeating trisaccharide unit made of two residues of N-acetyl galactosamine and one rhamnose residue (Ortega et al., 2005; Ortega et al., 2009). While lipid A-core is universally present in all Bcc bacteria LPS, not all bacteria LPS exhibits the OAg (Valvano, 2015a).

In general, the genes involved in LPS production by *B. cenocepacia* J2315 are in Chromosome 1, and organized in three loci (BCAL1929 to BCAL1935; BCAL2402 to BCAL 2408; BCAL3110 to BCAL 3125) (Holden et al., 2009), together with additional genes encoding sugar modification enzymes

[Fig. 1.6 and (Loutet et al., 2006; Fathy Mohamed et al., 2017; Fathy Mohamed et al., 2019; Oppy et al., 2019)]. In brief, the biosynthesis of the lipid A-core OS is initiated at the cytoplasmic leaflet of the inner membrane by the conserved pathway “Lpx” [sequential conversion of the precursor UDP-*N*-acetyl-glucosamine into lipid A-Kdo<sub>2</sub> or A-Kdo-Ko] [Fig. 1.6 and (Raetz and Whitfield, 2002; Vinion-Dubiel and Goldberg, 2003; De Soyza et al., 2008; Whitfield and Trent, 2014; Fathy Mohamed et al., 2017)]. Core OS is assembled by sequential glycosyltransferase reactions, to be subsequently received by the lipid A-Kdo-Ko molecule, and converted into lipid A-core OS (Valvano, 2015a; Fathy Mohamed et al., 2017). This lipid A-core OS is transported across the inner membrane to the periplasmic side via the ABC transporter MsbA (Valvano, 2015a; Fathy Mohamed et al., 2017). When expressed, OAg starts to be assembled at the cytoplasmic side of the inner membrane and then translocated to the periplasm where it is linked to the lipid A-core OS, catalysed by WaaL ligase (Fig. 1.6) (Valvano, 2015a). The mature LPS is exported to the outer side of the outer membrane through the LPS transport system Lpt, which form a multiprotein complex that links to the inner membrane and outer membrane consuming energy provided by ATP hydrolysis (Hamad et al., 2012; Okuda et al., 2016). The LPS biosynthesis has been reviewed in detail several times (Raetz and Whitfield, 2002; De Soyza et al., 2008; Whitfield and Trent, 2014; Valvano, 2015a; Maldonado et al., 2016) and Fig. 1.6 summarize all the steps involved.

#### *1.2.4.3. LPS modification in Bcc during chronic respiratory infection in CF and pathogenicity role*

Most of the alterations in the LPS registered molecules is found in the OAg polysaccharide and the lipid-A components (Pier, 2007; Maldonado et al., 2016; Faure et al., 2018). The lipid A of *B. cenocepacia*, is a penta-acylated tetrasaccharide backbone consisting of two linked glucosamine residues that contain phosphodiester-linked L-Ara4N residues (Silipo et al., 2005). Unlike those from *E. coli* and *S. enterica*, the lipid A of *B. cenocepacia* consists of a mixture of tetra and penta-acylated forms containing only one secondary acyl chain (Silipo et al., 2005; Silipo et al., 2007; Fathy Mohamed et al., 2017). Also, in *Burkholderia*, the distal Kdo residue is converted to Ko by the dioxygenase KdoO (Chung et al., 2014) (BCAL3311) (Fathy Mohamed et al., 2017). The classical lipid A biosynthesis in the most of Gram-negative bacteria requires a nine-enzymatic step machinery (Whitfield and Trent, 2014) but the *B. cenocepacia* genome has homologues of genes encoding eight of these enzymes (Fathy Mohamed et al., 2017). Several studies were done in order to identify the exact role of these 8 genes and the underlying pathway for the lipid A biosynthesis (Anderson et al., 1985; Anderson and Raetz, 1987; Silipo et al., 2005; Whitfield and Trent, 2014; Fathy Mohamed et al., 2017). LPS and its lipid A is an important constituent of *B. cenocepacia*, however for different Bcc species the contribution of the lipid A biosynthetic machinery and modification have not been systematically investigated (Maldonado et al., 2016; Fathy Mohamed et al., 2017).



**Figure 1.6** | Simplified overview of the LPS biosynthesis based on studies previously published using *B. cenocepacia* J2315 and K56-2 epidemic strains (Ortega et al., 2005; Loutet et al., 2006; Ortega et al., 2007; Ortega et al., 2009; Hamad et al., 2012; Lee et al., 2013; Valvano, 2015b; Fathy Mohamed et al., 2017; Fathy Mohamed et al., 2019; Olgagnon et al., 2019; Oppy et al., 2019). Lipid A - Kdo - Ko - Ara4N is synthesized on the cytoplasmic surface of the inner membrane (A). The rest of the core OS is assembled to the lipid A - Kdo - Ko - Ara4N (B) and MsbA proteins flips the whole complex to the periplasmic side of the inner membrane (D). Independently, the OAg biosynthesis is initiated by cytoplasmic membrane associated enzymes recently described as O-glycosylation genetic cluster and then chain elongation is assembled by another genetic OAg-biosynthetic locus (C). The OAg of *B. cenocepacia* J2315 was proved to be translocated by ABC transporter dependent pathway, wzm and wzt (E). Once on the periplasmic side, the OAg is linked to the lipid A-core OS by the WaaL ligase (F). Finally, the complete LPS is transported to the outer leaflet of the outer membrane by the Ipt pathway which is known as LPS transport system.

The chemical structures and biosynthetic pathways for the OAg of the different Bcc species are not well studied. However, there are efforts to understand the OAg alterations exhibited by *B. cenocepacia* clinical isolates [(Ortega et al., 2005; Ortega et al., 2009; Fathy Mohamed et al., 2019) and chapters III of this thesis] and their impact on pathogenesis (Ortega et al., 2005; Ortega et al., 2009; Saldias et al., 2009; Kotrange et al., 2011; Fathy Mohamed et al., 2019; Oppy et al., 2019). The OAg is synthesized by cytoplasmic membrane associated enzyme complexes and requires the C55-undecaprenyl phosphate (Und-P), which serves as an acceptor for the OAg units (Valvano et al., 2011; Fathy Mohamed et al., 2019). The OAg assembly occurs by the action of diverse glycosyltransferases and most of these enzyme are encoded by genes of the *wb\** locus (Valvano et al., 2011; Fathy Mohamed et al., 2019). The *B. cenocepacia* K56-2 OAg was found to be composed of a trisaccharide repeating units made of one rhamnose and two N-acetylgalactosamine residues, while *B. cenocepacia* J2315 (the corresponding epidemic clone) was observed to lack the OAg (Ortega et al., 2005). This OAg loss in this J2315 epidemic clone was related to an insertion element in the OAg biosynthetic cluster (Ortega et al., 2005). This loss is considered as an advantage and this idea was strengthened by a study proving that the lack of OAg in *B. cenocepacia* lead to an increased internalization into macrophages upon phagocytosis (Saldias et al., 2009) and led to an impaired biofilm formation and siderophore activity (Oppy et al., 2019). Consistent with the higher invasiveness described for this epidemic strain, J2315 (Ortega et al., 2005; Silipo et al., 2007; Saldias et al., 2009; Saldias and Valvano, 2009). The OAg loss could therefore facilitate access of Bcc bacteria to macrophages, in agreement with a study showing that in human lungs, Bcc bacteria, but not *P. aeruginosa*, were found mainly inside macrophages (Schwab et al., 2014). In *P. aeruginosa*, the modification of the LPS structure, including the loss of OAg expression during chronic pulmonary infection do not confer phagocytic resistance *in-vitro* (Demirdjian et al., 2017), contrasting with study done using OAg deficient *B. cenocepacia* that was reported to be more susceptible to phagocytic internalization (Saldias et al., 2009). The OAg is also believed to be essential in several pathogens for motility (swarming and twitching) (Toguchi et al., 2000; Berry et al., 2009; Bowden et al., 2013), protection against oxidative stress (Berry et al., 2009) and evasion from host immune defences (Murray et al., 2003; Saldias et al., 2009; Kotrange et al., 2011; Kintz et al., 2017). LPS OAg is believed to be an immunodominant molecule that can modulate host-pathogen interaction (Ranf, 2016) and proposed to be under selective pressure in Gram-negative bacteria (King et al., 2009; Maldonado et al., 2016; Kintz et al., 2017).

Longitudinal studies of *P. aeruginosa* isolates from CF patients with chronic lung infections revealed the conversion from smooth (in the early isolates) to rough LPS at the late-stage of infections (King et al., 2009; Maldonado et al., 2016; Demirdjian et al., 2017). In Bcc bacteria, a few recent reports also described the variation of the presence and/or modification of the OAg during persistent infection of the CF lung and on how this alteration may affect Bcc pathogenicity [(Lieberman et al., 2011; Maldonado et al., 2016; Silva et al., 2016) and chapters III and IV of this thesis]. An extensive study on

an epidemic outbreak of *Burkholderia dolosa* for 16 years, involving 14 CF patients and 112 isolates, has shown that several late isolates produce an LPS exhibiting the OAg that was absent in the LPS of the ancestral strain; this fact was related with the appearance of two different mutations in the glycosyltransferase encoding gene *wbaD* (Lieberman et al., 2011). The genome analysis of clones and metagenomes of evolved *B. cenocepacia* biofilms, after more than one thousand generations, unveiled a mutation in the gene *manC* – a gene involved in mannose metabolism located within the LPS biosynthesis gene cluster- resulting in the loss of the LPS OAg (Traverse et al., 2013). Remarkably, mutations in the same *manC* gene were previously found to disrupt OAg biosynthesis in late *B. cenocepacia* (Saldias et al., 2009) or early *B. dolosa* (Lieberman et al., 2011) isolates retrieved during chronic infection. A comparative genomic analysis focused on sequential *B. multivorans* isolates obtained over 20 years of CF chronic infection showed that late isolates accumulate three different mutations in a locus homologous to the *wbi* gene cluster, involved in LPS OAg biosynthesis, leading to OAg loss (Silva et al., 2016). Other recent comparative genomic analyses using *B. cenocepacia* clonal variants have proved the above evolutionary trajectory, in which genes involving in the OAg biosynthetic cluster acquired non-synonymous SNPs during chronic CF-pulmonary infection resulting OAg loss (Lee et al., 2017; Nunvar et al., 2017).

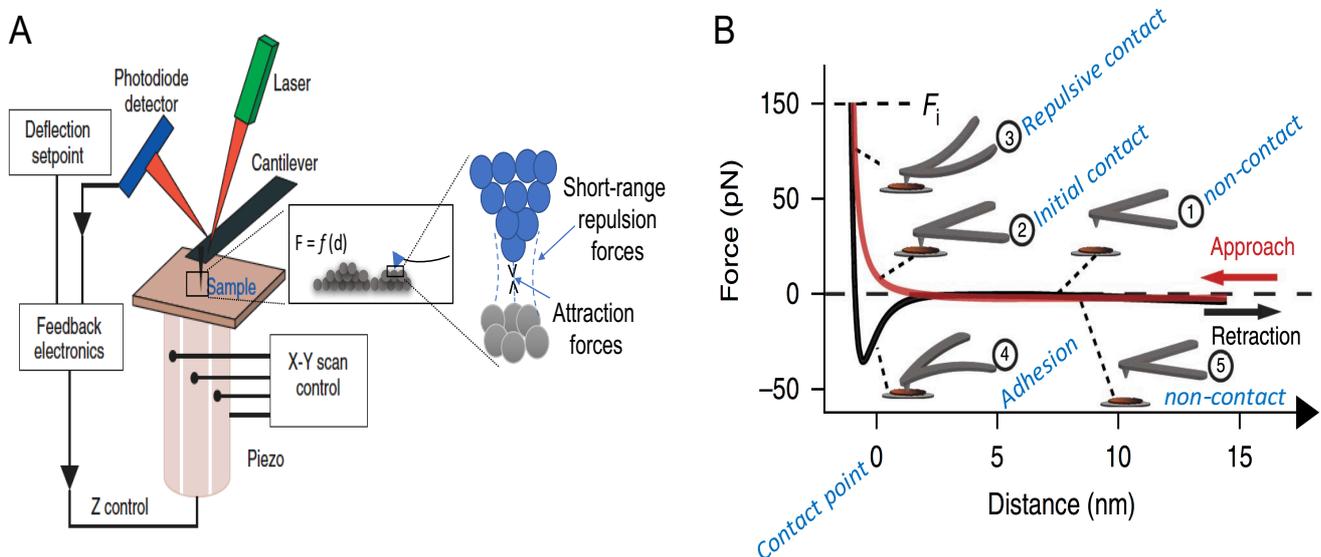
### 1.3. Atomic Force Microscopy (AFM) as an advanced technique to study bacteria at the cell surface level

#### 1.3.1. Fundamentals of AFM

Emerging microscopic techniques useful for the study of cell biophysics are essential to the better understanding of the infection processes and pathogenicity mechanisms (Stewart, 2014). Among of them is Atomic force microscopy (AFM) that can be used to study cell mechanics which affects cellular functions including shape, motility, differentiation, division, and adhesion to an extracellular matrix (Atabek and Camesano, 2007; Dufrene, 2014; Moeendarbary and Harris, 2014; Gavara, 2017). In addition, AFM can contribute to the understanding of the molecular interactions driving microbial adhesion in a single-cell basis (Dufrene, 2004; 2014; Moeendarbary and Harris, 2014; Gavara, 2017).

In 1986, AFM was introduced as a technique capable of investigating the surfaces of insulators at the atomic scale, providing three-dimensional images (Binnig et al., 1986). AFM becomes commercially available in 1989 and since then it became one of the widely-used microscopic method in biological research (Allison et al., 2010). AFM allows the measurement of attractive and repulsive forces acting between the sample's surface and a sharp tip. This sharp tip is mounted on a very flexible cantilever and scans the surfaces through the last atoms or the probing-molecule in the tip that feels the forces exerted by the surrounding molecules (fig. 1.7 – panel A) (Binnig et al., 1986; Allison et al.,

2010). The size of the tip determines the lateral resolution of the AFM and is designed to have a radius of less than 10 nm (Angeloni, 2016; Gavara, 2017), while the cantilever has tens to hundreds of micrometers in length and width, constituting the spring element that bends, towards or away from the sample, when attractive or repulsive forces are present. The cantilever chip is firmly attached to a piezoelectric positioner to allow ultraprecise positioning in the vertical direction (z-axis) of the cantilever that is perpendicular to the surface of the sample (fig. 1.7 – panel A). In most of the AFMs, the bending of the cantilever (typically referred to as deflection) is detected by optical means; a laser light is reflected from the cantilever and is detected by a quadrant photodiode. While the cantilever is un-deflected (usually when resting far away from the sample), the photodiode is manually positioned in such a way that half of the laser spot reaches the top quadrants, and the other half the bottom quadrants. When properly adjusted, the difference between the output of the top and the bottom quadrants is zero (Gavara, 2017).



**Figure 1.7** | A) Atomic force microscopy apparatus (adapted from (Allison et al., 2010) and B) force-distance (FD) curve (adapted from (Pfreundschuh et al., 2014)). AFM monitors the interaction forces between the surface and a sharp tip at the end of a microcantilever, giving rise FD curves. Approach (red) and retraction (black) FD curves. Zero distance indicates the contact point of the tip and the sample. The analysis of the FD curves provides information such as the sample height, deformation, elasticity (Young's modulus) and adhesion.

AFM can be functioning to obtain high-resolution topographical images, in contact (static) or non-contact (dynamic) (Angeloni, 2016; Gavara, 2017). In contact mode, the tip is brought into direct contact with the sample and the atoms at the surface push the tip back, since the repulsive forces dominate in this scenario. A complementary approach is the dynamic one, later called “tapping mode”, in which the cantilever is forced to oscillate near its resonance frequency and is kept at pre-determined amplitude. The AFM system also continuously monitors the amplitude of the cantilever and then adjusts in real-time the vertical position of the cantilever with respect to the sample, to keep the amplitude

constant (Hansma et al., 1994). The dynamic mode has evolved dramatically when larger amplitudes (>10 nm) started to be applied to stiffer cantilevers in air, creating what is today termed the tapping or intermittent contact modes that present several practical advantages (Zhong et al., 1993). By increasing the oscillatory amplitude, the adsorption of the tip by the surface is dramatically decreased since the cantilever receives enough restoring energy to detach. Consequently, imaging of biological samples became possible even when adhesion forces are very high. Also, tapping mode usually allows much higher resolution than contact mode when compared in the same condition (Binnig et al., 1986). While obtaining high-resolution topography of cellular surfaces, the key advantage of AFM for cell mechanics is the possibility to perform force measurements at desired cellular locations using the tip of the cantilever as indenter. Precise force measurements are possible because the cantilever behaves as a hookean spring onto the sample, while tracking how the sample deforms in response to said force (Gavara, 2017). During the measurements, the vertical displacement of the cantilever and its deflection are recorded simultaneously, and later converted to force-versus-displacement curves, briefly called force-distance (FD) curves, Fig. 1.7 – panel B.

### 1.3.2. Experimental considerations for the AFM imaging of the bacterial cells

There are different artifacts that should be considered before and over the course of an AFM experiment (Gavara, 2017). In brief, one of them has been a source of heated debates: the cantilever tip shape. There is now unanimity to consider that sharp conical (or pyramidal) tips should be used for cellular imaging, while the large radius (>1  $\mu\text{m}$ ) colloidal probes should be used for mechanical characterization (Dimitriadis et al., 2002). Several studies have proved that the sharp tips can be reliably used to study the cell stiffness similarly to the colloidal probes (Chiou et al., 2013; Vargas-Pinto et al., 2013), while mapping the topography without having to exchange tips halfway through the experiment. However, it is worth mentioning some less prevalent tip shapes as the tip-less cantilevers that have been usually used to probe loosely attached spherical objects, such as non-adherent cells or isolated cell nuclei to be then used as biological probes to measure cell-cell interactions (Moreno-Cencerrado et al., 2017). On the other hand, the cantilever stiffness should match those of the probed sample due the optimal force measurements that should display a large amount of cantilever deformation and a marked degree of sample indentation (Torre et al., 2011; Gavara, 2017). Consequently, the cantilever chips, commercially available, frequently contain an array of 4-6 cantilevers altogether lined-up closely in a chip, each with a different stiffness, to ease the switch between them to optimize the measurements (Torre et al., 2011). However, the slightly stiffer cantilever is favored because it may hardly remain bound to very adhesive samples – biological samples - for the entire of the withdraw curve as much as the soft one (Gavara, 2017) and because of its higher resonance frequency, that has proven crucial in the newest generations of AFMs (Smolyakov et al., 2016). Unlike tip shape and cantilever stiffness, other parameters (as indentation depth, location, and probing frequency) are readily changed during the

course of an experiment (Gavara, 2017). Due to the nature of the biological cells, such parameters should be optimized.

Prior to AFM measurements and during sample preparation, the problem of cell immobilization is a significant barrier to the application of AFM in microbiology research (Dufrêne, 2011). It is crucial to firmly attach the microbial cells to a solid substrate and use suitable and nondestructive methods to prevent displacement of the cell by the scanning tip. Drying bacteria cells attachment onto the solid surface, e.g. glass slides or mica discs, before imaging is a popular choice for cell immobilization (Fernandes et al., 2009; Sahu et al., 2009). Unfortunately, drying cells can result in cell dehydration and a flattened or collapsed appearance in the resulting AFM cell images (Doktycz et al., 2003). Moreover, cells immobilized this way may not be viable and are frequently not stable when imaged in liquid (Doktycz et al., 2003). Imaging in liquid is a requirement for live-cell imaging but brings additional challenges (Dufrêne, 2011). Given that Gram-negative bacteria surfaces composed of different compartments (proteins, saccharides and appendages), the ionic content and pH of the imaging buffer have to be considered (Muller et al., 1997; Yao et al., 2002). Imaging living microbes in liquid requires careful consideration of the immobilization technique used to ensure that the physiology of the bacteria is not compromised by the immobilization and imaging conditions, therefore, immobilization strategies have to be optimized for an individual bacterial strain. All of the successful approaches generally involve physical trapping or chemical attachment (Dufrêne, 2011). In the physical trapping, also known as mechanical trapping, a spherical microbial cells is simply trapped by a porous polymer membrane and can be used in AFM imaging in air or liquids, however, it is not the best method to preserve cell viability and mimic physiological conditions when measuring the biomechanical properties of the cell surface (Ahimou et al., 2003). Moreover, this strategy is not applicable for mounting the rod-shaped bacteria because of the spherical-like holes of the porous polymer membrane. On the contrary, chemical attachment is based on the bacterial attachment to a chemically modified-surfaces, such as gelatin-coated surfaces (Doktycz et al., 2003; Allison et al., 2011). Apparently, in this technique, not all the commercially-available gelatin can be effectively used (Allison et al., 2011) so the customized surfaces have to be tested using other cationic polymers, as poly-*L*-lysine (Dufrêne, 2011).

### 1.3.3. Mechanical and adhesion properties of Gram-negative bacteria studied by AFM

The characterization of the mechanical properties related to cell envelope constituents (LPS, peptidoglycan (PG), and membrane proteins), may provide some answers to important questions, as it is the case of the acquisition of resistance to current antibiotics and other environmental stresses in bacteria (Stewart, 2014). Bacterial cell mechanics is mostly described by the Young's Modulus and the bending rigidity (Tuson et al., 2012; Amir et al., 2014). The Young's Modulus (or the tensile elasticity) is characterized by the relationship between the applied stress on the cell (force per unit area) and the

resulting strain (fractional change in length). If a physical load is applied to the cell in the linear region, the cell will deform, and following the removal of the load the cell will return to its pre-load state (Tuson et al., 2012). The bending rigidity is the resistance of the cell to bending under a load and represents the output of the Young's modulus and the second moment of inertia (Amir et al., 2014). The bending rigidity can provide insight into the orientation of structural elements within cells, for example of the biomolecular elements that play a mechanical role — such as peptide bonds within the peptidoglycan — that are oriented perpendicularly to the long axis of bacterial cells (Wang et al., 2010). The bending rigidity can also be used to estimate the Young's modulus through its inherent relation to bending rigidity (Dufrene, 2014).

Microbial adhesion is believed to be mediated by a multitude of molecular interactions that could either be specific (i.e., molecular recognition between receptor and ligands, as a ligand-binding phenomenon) or non-specific (i.e., hydrogen bonding, hydrophobic, van de Waals, electrostatic, and macromolecular forces) (Busscher et al., 2008). In general, forces driving microbial adhesion measured by AFM, may help to study the binding mechanisms (specificity, binding strength, and mechanics) of adhesins as well as to understand how bacterial pili respond to mechanical cues and to determine cell-cell interactions (Dufrene, 2015).

As aforementioned, among cell envelope constituents that affect cell mechanical properties and can be studied by AFM, are the LPS, the PG, and membrane proteins (Stewart, 2014). The LPS are negatively charged molecules that electrostatically repel other LPS molecules, leading to the physical separation between the molecules (Delcour, 2009). The LPS binds tightly to divalent cations ( $\text{Ca}^{+2}$  or  $\text{Mg}^{+2}$ ) to overcome the electrostatic repulsion and increase stability (Delcour, 2009). *In vitro* studies using *E. coli* proved that the addition of  $\text{Ca}^{2+}$  to a “deeprough” LPS monolayer containing only lipid A and 2 Kdo sugars produces a cross-linked elastic gel (Herrmann et al., 2015). Furthermore, increasing the polysaccharide length of the *P. aeruginosa* LPS enabled the formation of a similar gel in the absence of  $\text{Ca}^{2+}$ , possibly due to increased hydrogen bonding, which resulted in additional lateral compression of the LPS monolayer (Ivanov et al., 2011; Herrmann et al., 2015). These *in vitro* studies support the idea of LPS stabilization achieved by divalent cations and/or by the length of the polysaccharide chain in different species of the Gram-negative bacteria (Ivanov et al., 2011; Herrmann et al., 2015). Additional *in vivo* studies illuminated the magnitude of the mechanical interaction between divalent cations and LPS, the mechanical response of the LPS to changing environmental conditions (Gaboriaud et al., 2006), and the role of LPS in outer membrane homeostasis (Smit et al., 1975; Kamio and Nikaido, 1976). Such roles of the LPS by inhibiting the binding and consequent effects on cell resistance to cationic antibiotics was also illustrated as a result of the unusual structure of the Bcc LPS that have a lower anionic charge, as discussed before (Vinion-Dubiel and Goldberg, 2003; De Soyza et al., 2008).

The PG of Gram-negative bacteria is a cross-linked polymeric meshwork that encapsulates bacterial cells and a cellular compartment examined during current AFM studies (Vollmer and Bertsche, 2008; Turner et al., 2014; Turner et al., 2016). The relative abundance of peptide cross-links in PG is an indicator of the cell stiffness examined by AFM (Desmarais et al., 2013). The PG is considered as a porous layer that contains different membrane proteins (Turner et al., 2013; Dover et al., 2015). This organization enables the peptide bonds to swell and shrink with changes in turgor pressure, while the more rigid glycan chains remain relatively unchanged (van den Bogaart et al., 2007). Recent genomic studies strengthened the concept that cell wall PG remodelling relates with the alteration of bacterial mechanical properties (Auer et al., 2016; Trivedi et al., 2018). For example, *Escherichia coli* mutants, deleted for genes encoding proteins associated with cell-wall synthesis, exhibit different stiffness defects (Auer et al., 2016) and the accumulation of the peptidoglycan *D*-Alanine residues is tightly regulated in *P. aeruginosa* since their accumulation reduces PG cross-linking and cell stiffness (Trivedi et al., 2018). Higher rigidity and increased elasticity was also recently reported to be associated with a lower outer membrane permeability which may lead to the reduction of antibiotic diffusion into the cells (Uzoechi and Abu-Lail, 2019a).

In Gram negative bacteria, the surface trimeric autotransporter proteins called adhesins (TAAs) were found in Bcc to mediate adhesion, aggregation and colonization of the respiratory epithelium, and AFM unraveled the binding mechanisms of TAAs (El-Kirat-Chatel et al., 2013; Dufrene, 2015). Also, AFM was able to examine the LPS roles – in particular involving the OAg component – in bacterial adhesion in *P. aeruginosa* and *E. coli*, although with controversial results (Abu-Lail et al., 2007; Atabek and Camesano, 2007; Atabek et al., 2008; Strauss et al., 2009). AFM was used to study Gram-negative bacteria pili type IV that was proved to be able to mediate the attachment of *P. aeruginosa* and *Neisseria gonorrhoeae* to the host cells (Beaussart et al., 2014). This is in line with an earlier single-molecule study using AFM on the Gram-negative pili type I that illustrated that type I pili mediates *E. coli*-host cell interaction (Miller et al., 2006).

The use of AFM to study cell-cell adhesion in detail, including aggregation, mixed biofilms, and the microbe-host interactions was reviewed (Dufrene, 2015). AFM was used to understand the molecular bases of polymicrobial interactions, for example to reveal strong adhesion forces between *Lactobacilli* and virulent *S. aureus* strain (Younes et al., 2012) as well as the strong physical interaction due to the interaction between chitin-binding proteins of *P. aeruginosa* and the hyphae of *Candida albicans* (Ovchinnikova et al., 2013). Adhesion forces between *P. aeruginosa* and host epithelial cells were also measured using AFM, and it was found to involve a complex mixture of short-range cohesive interactions between the bacterial outer membrane and the host membrane and long-range constant force interactions reflecting the extension of type IV pili and the host cell membrane proteins, tethers (Beaussart et al., 2014). In a conclusion, recent studies clearly show that AFM has emerged as an

essential tool for understanding the nanomechanics of living systems, in particular of Gram-negative bacterial cells (Dufre ne, 2011; Dufrene, 2014; Auer and Weibel, 2017) although, more studies are still required.

#### 1.4. Thesis outline and research objectives

Inside the CF lung, Bcc bacteria face stressful and changing environmental conditions as a consequence of high pro-inflammatory cytokine levels, high antibiotic concentrations, high levels of oxidative stress and low oxygen concentration and the presence of other co-infecting microbes (Moriarty et al., 2007; Palmer et al., 2007; Reid et al., 2007; Williams et al., 2007; Hogardt and Heesemann, 2010; D oring et al., 2011; Cullen and McClean, 2015). It might be likely that the adaptive evolution of each Bcc species and strain may vary in different CF patients' lung-environments and consequently, during long-term infection, the original infecting strain(s) undergo(es) extensive genetic and phenotypic diversification inside the CF lung thus enabling the emerging clonal variants to adapt to the CF lung environment and evade the host immune system (Coutinho et al., 2011a; Lieberman et al., 2011; Madeira et al., 2011; Mira et al., 2011; Silva et al., 2011; Madeira et al., 2013; Lieberman et al., 2014; Moreira et al., 2014; Maldonado et al., 2016; Silva et al., 2016; Cabral et al., 2017; Lee et al., 2017; Moreira et al., 2017; Nunvar et al., 2017). However, our knowledge on Bcc pathogenicity and on how some Bcc species or strains develop lethal respiratory infections while others do not are still limited (Sfeir, 2018). Until now the extent of Bcc bacterial diversification associated to long-term infection in CF patients and the ability of these bacteria to evolve in response to drug pressures, immune responses and other lung environment stresses is limited compared to *P. aeruginosa* and *Burkholderia pseudomallei* (Price et al., 2013; Markussen et al., 2014; Marvig et al., 2015; Viberg et al., 2017; Faure et al., 2018; Valentini et al., 2018).

Among of the above-mentioned efforts to understand Bcc pathogenesis and evolution in the CF lungs, are those of our research group led by Isabel S a-Correia in the Institute of Bioengineering and Biosciences (iBB). A two decade-collaboration between our research team and the major Portuguese CF treatment centre at Hospital de Santa Maria, in Lisbon, led to first studies describing the Portuguese situation concerning the molecular epidemiology of the respiratory infections caused by Bcc bacteria among the CF population under surveillance (Richau et al., 2000; Cunha et al., 2003). Based on the collection of isolates gathered over years, retrospective studies allowed the study of the molecular mechanisms underlying Bcc diversification of genotypes and phenotypes during adaptation to the CF lung environment of chronic infected patients (Richau et al., 2000; Cunha et al., 2003; Cunha et al., 2004; Leitao et al., 2008; Coutinho et al., 2011a; Coutinho et al., 2011b; Madeira et al., 2011; Mira et al., 2011; Madeira et al., 2013; Moreira et al., 2016; Moreira et al., 2017). Such retrospective studies suggested the occurrence of clonal expansion and emergence of phenotypic variations among Bcc

bacteria, especially among 11 *B. cenocepacia* longitudinal clones obtained from Patient J during 4.4 years of infection until death caused by cepacia syndrome (Coutinho et al., 2011a; Moreira et al., 2017). Three of these 11 isolates were extensively studied based on transcriptomic, quantitative proteomic and metabolomics analyses, in particular the first and the last isolates retrieved from the patient and one intermediary isolate with a remarkably higher resistance to several classes of antibiotics (Madeira et al., 2011; Mira et al., 2011; Madeira et al., 2013; Moreira et al., 2016). This thesis work was performed to enrich the above referred knowledge and to elucidate the adaptive strategies used by Bcc bacteria during long-term colonization in the CF lungs. It is organized into six main chapters. Chapter 1 provides a review on Bcc bacteria, their taxonomy and epidemiology in the context of CF lung infections and on their adaptive strategies to the CF lungs. Recent studies focused on the outermost cellular compartment of Gram-negative bacteria, the lipopolysaccharide (LPS), and its modification during the chronic respiratory infection are also discussed as well as studies on bacterial cell mechanics by the atomic force microscopy as an advanced technique to study the corresponding alterations during bacterial adaptation.

Although the above referred comprehensive comparative genomic analyses have been essential to establish the field, there is no available information on the comparison of the evolutionary patterns occurring in strains of different species when exposed to the same CF lung environment and host immune responses during co-infection. This was the objective of the second thesis chapter (and first experimented chapter) that was focused on the two most prevalent and feared Bcc species among the CF community worldwide: *B. cenocepacia* and *B. multivorans* (Jones et al., 2004; Mahenthiralingam et al., 2005; Drevinek and Mahenthiralingam, 2010; Lipuma, 2010). This work was designed to compare the genome sequences of *B. cenocepacia* and *B. multivorans* clonal variants co-inhabiting the same host-selective environment to try to identify species-specific and common evolutionary patterns. For this, a retrospective study on twenty clonal variants derived from two ancestor strains (9 *B. multivorans* isolates and 11 *B. cenocepacia* isolates) was performed. They were sequentially retrieved from the same CF patient (patient J) over a period of 4.4 years, from the onset of infection with *B. multivorans* followed by co-infection with *B. cenocepacia* until the patient's death from cepacia syndrome (Correia et al., 2008; Coutinho et al., 2011a). This work involved an international collaboration with Prof. Vaughn Cooper laboratory, USA, that provided the genome sequences and contributed to data analysis.

Recent studies at the genome sequence [(Spencer et al., 2003; Lieberman et al., 2011; Traverse et al., 2013) and results presented in chapter 1], transcriptome (Mira et al., 2011) and proteome (Madeira et al., 2011; Madeira et al., 2013) levels indicate that the LPS and particularly the OAg undergo alterations during chronic infection, which could be attributed to bacterial adaptation to biofilm lifestyle (Lieberman et al., 2011; Traverse et al., 2013), immune evasion and adaptation to antibiotics pressure

(reviewed by (Maldonado et al., 2016)). However, knowledge on the phenotypic alterations occurring at the level of Bcc OAg expression and their underlying mechanisms during the chronic respiratory infection, is limited (Maldonado et al., 2016). In chapter 3, the chemical structure of the LPS molecule of 4 *B. cenocepacia* isolates was analyzed, in an international collaboration with Prof. Antonio Molinaro laboratory, Italy, as well as the genetic organization of the OAg biosynthetic cluster in the above-mentioned 11 *B. cenocepacia* serial isolates, in an international collaboration with Prof. Miguel Valvano laboratory, UK. The obtained results reveal that the early *B. cenocepacia* IST439 isolate encodes a functional genetic cluster responsible for OAg biosynthesis, while all subsequent *B. cenocepacia* isolates lost the ability to produce an OAg molecule. This output was in agreement with longitudinal studies of *P. aeruginosa* isolates from CF patients with chronic lung infections that revealed the conversion from smooth (in the early isolates) to rough LPS with no OAg side chain at the late-stage of infections rendering the bacteria non-typable and less immunogenic (King et al., 2009; Maldonado et al., 2016; Demirdjian et al., 2017). Part of the work described in this chapter and published in (Hassan et al., 2017) was performed by Rita Maldonado during the start of her PhD in our laboratory and in the laboratory of Miguel Valvano. The international scientific collaboration, described in this chapter, was carried out in the laboratories of Miguel Valvano and Antonio Molinaro in the context of the EU COST Action BM1003.

Although the Bcc comprises 24 bacterial species (De Smet et al., 2015; Depoorter et al., 2016; Ong et al., 2016; Bach et al., 2017; Weber and King, 2017; Martina et al., 2018) and several were already demonstrated before to be involved in CF respiratory infections, only *B. cenocepacia* (Ortega et al., 2005; Saldias et al., 2009; Hassan et al., 2017; Lee et al., 2017; Nunvar et al., 2017), *B. multivorans* (Silva et al., 2016), and *B. dolosa* (Lieberman et al., 2011; Lieberman et al., 2014) strains were examined concerning the variation of the OAg presence over chronic infection and on how this alteration may affect Bcc pathogenicity. The studies described in chapter 4 were performed to understand whether OAg loss in different Bcc species during respiratory infection in CF patients can be considered a general phenomenon that affects immune evasion favoring chronic infection. For this, a systematic retrospective and longitudinal screening was performed based on a collection of isolates, recovered from 1995 to 2016 from 19 CF patients under surveillance at the major Portuguese CF treatment center at Hospital de Santa Maria (HSM), in Lisbon, over the duration of chronic infection (ranging from 1.2 to 15.2 years). The 357 sequential isolates examined were molecularly identified at the species level and genotyped during this study or in previous studies and belong to *B. cenocepacia*, *B. multivorans* or to the more rare species in CF population worldwide, *B. dolosa*, *B. stabilis*, *B. cepacia* and *B. contaminans* (Cunha et al., 2003; Cunha et al., 2007; Coutinho et al., 2011b; Moreira et al., 2014; Coutinho et al., 2015).

Bacterial cell envelope plays a central role in cell physiology and the alteration of surface properties can implicate the variation of phenotypes that play a crucial role in the pathogenesis of infectious diseases, such as antibiotic resistance and biofilm formation (Maldonado et al., 2016; Hill et al., 2017; Trivedi et al., 2018). However, very few bacterial species have been on the focus of studies related to cell surface physical properties (Auer et al., 2016; Auer and Weibel, 2017; Trivedi et al., 2018) and information on the diversification and adaptive evolution at the level of Bcc bacteria cell wall mechanical properties during CF chronic lung infections is missing. In this context, over the last years atomic force microscopy (AFM) emerged as an essential tool for understanding the nanomechanics of live systems (Touhami et al., 2003; Muller et al., 2009; Costa et al., 2014). Hence, the objective of chapter 5 was to obtain the lacking knowledge by studying cell surface morphology and mechanical properties of 3 *B. cenocepacia* clonal variants (the widely-studied variants among those 11 isolates obtained from patient J) isolated from the lungs of a CF patient during long term infection using AFM.

At last, chapter 6 presents an integrated final discussion of the significant findings of this thesis work and propose some working hypotheses that can be exploited in the future.

The contribution of the PhD candidate for the experimental and bioinformatics work presented in each chapter in this thesis is as follows. The candidate wrote the entire chapter 1. In chapter 2, within an international collaboration, the candidate performed the assembly (except the PacBio sequencing), the functional annotations and the comparative genomic analysis of *B. cenocepacia* and *B. multivorans* isolates and confirmed the results obtained for *B. cenocepacia* variants obtained by another colleague. In chapter 3, the candidate was involved in the cloning and expression of OAg biosynthetic genes and performed the electrophoretic profiles of the LPS and the complementation assays. The candidate was also involved in the *in-silico* analysis of Bcc OAg biosynthetic clusters and the corresponding comparative genomic. For chapter 4, most of the experimental work including the species/lineage identification and genotyping of Bcc isolates was done by the candidate as well as all the electrophoretic profiles of the LPS. In chapter 5, within a national collaboration with the group of Prof. Mário Rodrigues, Faculdade de Ciências da Universidade de Lisboa, the candidate prepared the AFM samples, examined the growth behaviour of the cells, performed the biofilm formation assays, performed AFM measurements, together with other colleagues, and analysed the roughness measurements. The candidate was also actively involved in data analysis and interpretation of all the work included in this thesis, as well as in the writing and preparation of tables, figures and the *in-silico* data of the accepted manuscripts and of the chapter that is still unpublished.

## 2 Comparative evolutionary patterns of *Burkholderia cenocepacia* and *Burkholderia multivorans* during chronic co-infection of a cystic fibrosis patient lung

---

**This chapter contains results included in a manuscript in preparation:**

A. Amir Hassan<sup>1,2</sup>, Sandra C. dos Santos<sup>2</sup>, Robert Sebra<sup>3</sup>, Marcus Dillon<sup>4</sup>, Vaughn S. Cooper<sup>4</sup>, Isabel Sá-Correia<sup>1,2</sup>, **Comparative evolutionary patterns of *Burkholderia cenocepacia* and *Burkholderia multivorans* during chronic co-infection of a cystic fibrosis patient lung**

<sup>1</sup> iBB - Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisbon, Portugal

<sup>2</sup> Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisbon, Portugal

<sup>3</sup> Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

<sup>4</sup> Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, Bridgeside Point II, 450 Technology Drive, Pittsburgh, PA 15219, USA



## 2.1. Abstract

During chronic respiratory infection of cystic fibrosis (CF) patients, extensive genetic and phenotypic diversification occurs enabling the emerging bacterial population variants to adapt and evade the host immune system and become established in the patients' airways. The present study was designed to gain insights into the molecular mechanisms underlying genomic diversification occurring in the two most prevalent and feared *Burkholderia cepacia* complex (Bcc) species and to identify common and relevant pathoadaptive mechanisms. For this, a comparative genome analysis of two *B. cenocepacia* and *B. multivorans* strains co-inhabiting the same host-selective lung environment, for at least 3 years, was performed. The genome sequences of 9 *B. multivorans* and 11 *B. cenocepacia* clonal variants sequentially retrieved from the same CF patient over a period of 4.4 years, from the onset of infection with *B. multivorans*, followed by co-infection with *B. cenocepacia* until the patient's death with cepacia syndrome, were obtained and compared. During the course of co-infection, both species acquired several mutations with accumulation rates of 2.08 (*B. cenocepacia*) and 2.27 (*B. multivorans*) SNPs/year. The high number and distribution of non-synonymous mutations that we observed to occur in several genes, during the co-infection involved periods of diversification, suggested the existence of positive selection as an important driving force underlying the adaptation of Bcc to the CF patient lung. Most of the mutated genes are associated with oxidative stress response, transition metal metabolism and defense mechanisms against antibiotics, consistent with the idea that positive selection might be driven by the action of the host immune system, antibiotic therapy and low oxygen and iron concentrations. Two orthologous genes were found to be under strong selection and accumulate mutations related to clade emergence in *B. cenocepacia* and *B. multivorans*. One gene encodes a nucleotide sugar dehydratase involved in lipopolysaccharide O-antigen (OAg) biosynthesis (*wbiI*). The other gene encodes a putative two-component regulatory sensor kinase protein required to sense and adapt to oxidative- and heavy metal- inducing stresses. This study provides relevant information to understand the common and the specific evolutionary patterns occurring in *B. cenocepacia* and *B. multivorans* under positive selection in the same CF lung environment and host immune system during a co-infection that ultimately led to the cepacia syndrome.

## 2.2. Introduction

Chronic pulmonary infections are considered the leading cause for morbidity and premature death of patients at risk, in particular cystic fibrosis (CF) patients and immunocompromised individuals (Lipuma, 2010; Cullen and McClean, 2015). Several opportunistic pathogens may populate the CF airways and the heterogeneous microbial populations usually present are very hard to eradicate (Lyczak et al., 2002; Lipuma, 2010). Among the bacterial pathogens, species of the *Burkholderia cepacia* complex (Bcc) are particularly feared by CF patients (Mahenthiralingam et al., 2005). Currently, the Bcc includes 24 closely related bacterial species (De Smet et al., 2015; Depoorter et al., 2016; Ong et al., 2016; Bach et al., 2017; Weber and King, 2017; Martina et al., 2018). Pulmonary infection with Bcc can involve a single Bcc species or co-infection with more than one Bcc species and, besides interpatient transmission, these bacteria are very difficult to eradicate and can cause a marked decline of lung functions and decreased life expectancy. In certain cases, they can lead to a lethal uncontrolled clinical deterioration with septicemia and necrotizing pneumonia (the “cepacia syndrome”) (Mahenthiralingam et al., 2005).

During chronic infection by Bcc, the original infecting strain(s) establishes a population that evolves and diversifies into genetically and phenotypically distinct lineages. These lineages are thought to acquire multiple adaptations to the CF lung environment (Coutinho et al., 2011a; Lieberman et al., 2011; Madeira et al., 2011; Mira et al., 2011; Silva et al., 2011; Madeira et al., 2013; Lieberman et al., 2014; Moreira et al., 2014; Maldonado et al., 2016; Silva et al., 2016; Cabral et al., 2017; Hassan et al., 2017; Lee et al., 2017; Moreira et al., 2017; Nunvar et al., 2017; Hassan et al., 2019). Over the last decade, several Bcc virulence factors and adaptive traits have been identified in Bcc species (Mahenthiralingam et al., 2005; Drevinek and Mahenthiralingam, 2010; Loutet and Valvano, 2010; Zlosnik and Speert, 2010; Madeira et al., 2011; Mira et al., 2011; Zlosnik et al., 2011; Madeira et al., 2013; Zlosnik et al., 2014; Sousa et al., 2017; Hassan et al., 2019). A number of studies have used genome-wide methods to identify genetic changes during long term infection by *B. dolosa* (Lieberman et al., 2011; Lieberman et al., 2014), *B. multivorans* (Silva et al., 2016) and *B. cenocepacia* (Lee et al., 2017; Nunvar et al., 2017) within individual hosts. For example, retrospective studies involving 112 sequential isolates of *B. dolosa* from 14 CF patients, identified several mutated genes involved in genetic variation of this bacterial pathogen within individual patients, in particular genes required for expression of surface polysaccharides, lipopolysaccharide O-antigen (OAg) biosynthesis, outer membrane components, iron scavenging, and antibiotic resistance (Lieberman et al., 2011) and suggested parallel bacterial evolution within multiple patients showing that different lineages may coexist for many years within a patient and identified candidate pathogenicity genes (Lieberman et al., 2014). Another comparative genomic study including 22 isolates of *B. multivorans* recovered over 20 years from CF patients also found parallel adaptive variations resulting from host related selection

pressures in genes functionally also involved in antibiotic resistance, cell wall and membrane composition, lipopolysaccharide O-antigen biosynthesis, metabolism and oxygen-sensing (Silva et al., 2016). A subsequent comparative genomic analysis including 32 clonal variants of *B. cenocepacia* obtained from 8 CF patients (Nunvar et al., 2017) revealed that, in addition to the aforementioned parallel mutations in gene functions already described for *B. dolosa* and *B. multivorans* (Lieberman et al., 2011; Lieberman et al., 2014; Silva et al., 2016), genes related with transition metal metabolism are hotspots for nucleotide polymorphism (Nunvar et al., 2017). Another comparative study involving 215 genomes from serial *B. cenocepacia* isolates obtained from 16 CF patients during a 20 year-period also support the above mentioned evolutionary trajectories during chronic Bcc infections and the well-established diverse-community model (Lee et al., 2017). This publication also reported the complete loss of chromosome III resulting in genome-size reduction as an adaptive trait of Bcc bacteria (Lee et al., 2017), consistent with a previously reported rare *in-vivo* loss of the same non-essential chromosome among *B. cenocepacia* clonal variants emerging during chronic infection (Moreira et al., 2017).

Although the above genomic analyses have established the field of *Burkholderia* evolution during chronic infection of the CF airway, we lack a comparison of the evolutionary patterns when strains of different species co-infect the same CF patient. This was the objective of the present study focused on the most prevalent and feared Bcc species among the CF community worldwide: *B. cenocepacia* and *B. multivorans* (Jones et al., 2004; Mahenthiralingam et al., 2005; Drevinek and Mahenthiralingam, 2010; Lipuma, 2010). *B. cenocepacia* is the dominant species with high potential for inter-patient transmission (Mahenthiralingam et al., 2005; Drevinek and Mahenthiralingam, 2010) but, in several countries, *B. multivorans* has recently replaced *B. cenocepacia* in this first position (Jones et al., 2004; Lipuma, 2010). Given that the CF environment is characterized by high pro-inflammatory cytokine levels, high antibiotic concentrations, high levels of oxidative stress and low oxygen concentration (Moriarty et al., 2007; Palmer et al., 2007; Reid et al., 2007; Williams et al., 2007), it is likely that the adaptive evolution of each Bcc species strain may vary in different CF patients' lung-environments. This work was designed to compare the genome sequences of *B. cenocepacia* and *B. multivorans* clonal variants co-inhabiting the same host-selective environment to try to identify species-specific and common evolutionary patterns. For this, a retrospective study on twenty clonal variants derived from two ancestor strains (9 *B. multivorans* isolates and 11 *B. cenocepacia* isolates) was performed. They were sequentially retrieved from the same CF patient over a period of 4.4 years, from the onset of infection with *B. multivorans* followed by co-infection with *B. cenocepacia* until the patient's death from cepacia syndrome (Correia et al., 2008; Coutinho et al., 2011a; Hassan et al., 2019). These isolates were obtained in the major Portuguese CF treatment Center at Hospital de Santa Maria during consultation routines and pulmonary exacerbations that compelled the patient to hospitalization and intravenous therapy with gentamicin and ceftazidime (Correia et al., 2008; Coutinho et al., 2011a). The design and successful execution of the present study provided relevant and useful information to

contribute to a better understanding of the common and the specific evolutionary patterns occurring in *B. cenocepacia* and *B. multivorans* under positive selection in the same CF lung environment and host immune system during a co-infection that ultimately led to the cepacia syndrome.

## 2.3. Materials and Methods

### 2.3.1. Bcc bacterial isolates and growth conditions

Eleven *B. cenocepacia* (*recA* lineage IIIA) sequential isolates and nine *B. multivorans* sequential isolates obtained from the same cystic fibrosis patient (patient J) (Cunha et al., 2003; Coutinho et al., 2011a; Hassan et al., 2019) were examined in this study. These isolates were collected during hospital routines from the sputum of a chronically infected CF patient who was under surveillance at Hospital de Santa Maria (HSM), Centro Hospitalar Lisboa Norte (CHLN) EPE, Lisbon, Portugal, from February 1998 to July 2002 (Cunha et al., 2003; Correia et al., 2008; Coutinho et al., 2011a; Hassan et al., 2019) (Table 2.1). These isolates were obtained from the onset of the Bcc infection until the patient's death with cepacia syndrome after 4.4 years of Bcc infection and were selected at random among the colonies isolated in selective *Burkholderia cepacia* Selectatab medium at the Hospital, at a specific date of isolation. Bacterial cultures are stored at -80°C in 1:1 (v/v) glycerol. Bacteria were grown in Lysogeny broth (LB; Conda, Pronadisa) at 37°C with shaking at 250 rpm or in LB agar plates.

**Table 2.1** | *Burkholderia cepacia* complex examined isolates obtained from a cystic fibrosis patient whose genomes were sequenced in this study, ordered based on the isolation date.

Bcc isolate	Isolation date	Bcc species
IST419	26 Feb 1998	<i>B. multivorans</i>
IST424	4 Jun 1998	<i>B. multivorans</i>
IST439	30 Jan 1999	<i>B. cenocepacia recA</i> lineage IIIA
IST453	19 Jul 1999	<i>B. multivorans</i>
IST455A/IST455B	1 Feb 2000	<i>B. multivorans</i>
IST461	4 Apr 2000	<i>B. multivorans</i>
IST495A/IST495B	29 May 2001	<i>B. multivorans</i>
IST4103	24 Jul 2001	<i>B. cenocepacia recA</i> lineage IIIA
IST4110	25 Sep 2001	<i>B. cenocepacia recA</i> lineage IIIA
IST4112	11 Oct 2001	<i>B. cenocepacia recA</i> lineage IIIA
IST4113	6 Nov 2001	<i>B. cenocepacia recA</i> lineage IIIA
IST4119	22 Jan 2002	<i>B. multivorans</i>
IST4116A/IST4116B	11 Feb 2002	<i>B. cenocepacia recA</i> lineage IIIA
IST4131	26 Feb 2002	<i>B. cenocepacia recA</i> lineage IIIA
IST4129	26 Mar 2002	<i>B. cenocepacia recA</i> lineage IIIA
IST4130	14 May 2002	<i>B. cenocepacia recA</i> lineage IIIA
IST4134	2 Jul 2002	<i>B. cenocepacia recA</i> lineage IIIA

### 2.3.2. Genomic DNA sequencing, assembly and annotation

For genomic DNA extraction, bacterial cultures were prepared by suspending isolated colonies from LB agar plates in 3 mL LB broth, followed by overnight growth at 37°C with shaking at 250 rpm. Genomic DNA was extracted and purified using a DNeasy Blood and Tissue kit (Qiagen, Germany) according to manufacturer instructions. DNA concentration and purity were assessed using a Nanodrop ND-1000 spectrophotometer.

Genomic DNA samples of all the studied clonal variants were processed according to Illumina's instructions for generating paired-end libraries. In brief, *Burkholderia cenocepacia* IST439, IST4113, IST4129 and IST4134 were sequenced using a whole-genome shotgun sequencing strategy and Illumina Genome Analyser sequencing technology at CD Genomics (New York, USA), generating short 100-bp paired-end reads with high coverage (~300x), details of this protocol attached as supplementary materials. All of the *B. multivorans* clonal variants were sequenced on Illumina NextSeq 500 platform at the University of Pittsburgh, PA, United States; isolates of *B. cenocepacia* IST4103, IST4110, IST4112, IST4116A, IST4116B, IST4131 and IST4130 were sequenced using the 151-bp paired-end Illumina HiSeq platform at the University of New Hampshire Hubbard Center for Genomic Studies, after library preparation using a modified Illumina Nextera protocol designed for inexpensive library preparation of microbial genomes (Baym et al., 2015). Raw fastQ reads were analyzed using fastQC, which revealed that all isolates were sequenced at sufficient depth to accurately identify single nucleotide polymorphisms and indel mutations. Raw fastq paired-end files were processed for removal of Illumina adapters, trimming, and quality-based filtering using Trimmomatic v0.32 (Bolger et al., 2014).

*B. multivorans* IST419 and *B. cenocepacia* IST439, used as ancestor reference strains, were sequenced for a second time to generate a high-quality contiguous assembly. *B. multivorans* IST419 was likewise processed using Illumine NextSeq 500 platform producing a second sets of pair-ended reads, while *B. cenocepacia* IST439 was later-sequenced to generate a complete assembly by using a combination of single molecule, real-time (SMRT) Pacific Biosciences – PacBio reads and Illumina 100-bp paired-end reads, details of this PacBio protocol described in supplementary materials. We used the hierarchical genome-assembly process workflow (HGAP3) to generate a completed assembly of *B. cenocepacia* IST439 and polished our assembly using the Quiver algorithm (Chin et al., 2013).

The trimmed reads of all the studied clonal variants (except *B. cenocepacia* IST439) were *de novo* assembled using VelvetOptimiser v2.2.4 (Zerbino and Birney, 2008; Zerbino, 2010) and SPAdes v3.11.1-linux (Bankevich et al., 2012) with automated optimization of the assembly parameters. The quality of the assemblies obtained was estimated by Quast v2.3 (Gurevich et al., 2013) and the best

assembled outputs were automatically improved using Pilon v1.22 (Walker et al., 2014). The three trimmed data sets (two paired-end libraries and one mate-pair library) were used for scaffolding by SSPACE-standard v3.0 (Boetzer et al., 2011) followed by an automated improvement using Pilon v1.22 to gain the final draft genome sequences.

With draft genome sequences, contigs/scaffolds were reordered using MAUVE Contig Mover (Darling et al., 2010). All assembled-drafted genomes were reordered versus *B. cenocepacia* J2315 and *B. multivorans* ATCC\_17616. We then annotated the completed assembly of *B. cenocepacia* IST439 and the remaining draft genome sequences with the prokaryotic genome annotation tool prokka (v1.11) with a local, customized *Burkholderia* database of closely related Bcc strains (*B. cenocepacia* H111, *B. cenocepacia* J2315, *B. cenocepacia* ST32, *B. cenocepacia* VC1254, *B. dolosa* AU0158, *B. multivorans* ATCC\_BAA-247, and *B. multivorans* ATCC 17616) (Seemann, 2014). BLAST Ring Image Generator (BRIG) was used to generate visual genome comparisons with the reference genomes of *B. cenocepacia* H111, *B. cenocepacia* J2315, *B. cenocepacia* ST32, *B. multivorans* ATCC\_BAA-247, and *B. multivorans* ATCC 17616 (Alikhan et al., 2011). For enriched-functional annotation, eggNOG-mapper v4.5.1 was additionally applied for both strains (*B. cenocepacia* IST439 and *B. multivorans* IST419) that were used as references for the variant calling (Huerta-Cepas et al., 2017). This COG mapping was used to count the CDS per COG category for each species.

### 2.3.3. Variant calling and SNP/INDEL detections

Trimmed paired-end reads were mapped against the ancestor reference complete genome of *B. cenocepacia* IST439 (for all *B. cenocepacia* clonal variants) and/or against the ancestor reference draft genome sequence of *B. multivorans* IST419 for the corresponding clonal variants using BWA-MEM packages of Burrows-Wheeler Aligner (BWA v.0.7.10) (Li and Durbin, 2010) and NovoAlign (www.novocraft.com). Variants [Single nucleotide polymorphisms (SNPs) and insertion-deletion mutations (INDELs)] were called as described previously by using two independent standard variant calling pipelines; GATK and SAMtools/BCFtools toolbox (Li et al., 2009; Van der Auwera et al., 2013; Dillon et al., 2015; Dillon et al., 2017). Detailed pipelines are attached as supplementary materials.

To perform a quality control of the results obtained by the above reported method, a parallel comparison was performed using Breseq (Barrick et al., 2014), and similar predictions were in general observed in the case of base-substitutions, with some discrepancies in indels. All putative SNPs/INDELs were then manually inspected and evaluated using the Integrative Genomics Viewer – IGV (Robinson et al., 2011; Thorvaldsdottir et al., 2013), and discarded if the BWA and Novoalign-produced alignments did not provide enough confidence (poor coverage, as described in the supplementary materials). Thus, we are confident that nearly all base-substitution and indels identified in this study were genuine events that arose during the *in-vivo* evolutionary process. Finally, Functional

annotation of the called variants was performed by using SnpEff v3.1 (Cingolani et al., 2012) with manual BLAST verification against the NCBI Microbes genome database.

#### 2.3.4. Assessment of population structure and phylogeny

*In silico* whole genome multi-locus sequence typing (wgMLST) (Jolley and Maiden, 2010) was performed to confirm the isolate sequence type (ST) and the clonality of all the isolates examined.

To obtain a comprehensive phylogeny of the studied clonal variants, the SNP calls and their matrices were automatically processed by using a set of algorithms previously described (Kaas et al., 2014) via the online web server CSI phylogeny v1.4 with default settings; min depth at SNP position – 10x, relative depth at SNP position – 10 %, minimum distance between SNPs – 10 bp, min SNP quality – 30, and min read mapping quality – 25 (Kaas et al., 2014). Phylogenetic trees were then constructed using 70 and 110 variant positions among *B. cenocepacia* and *B. multivorans* clonal variants, respectively, by using FastTree (Price et al., 2010).

Seven *B. cenocepacia* *recA* lineage IIIA, two *B. multivorans*, and one *B. dolosa* deposited-reference genomes in *Burkholderia* genome database (<https://www.burkholderia.com/>) were used to describe the phylogenetic relationship of the studied clinical isolates among the Bcc bacteria (Table S2.1).

#### 2.3.5. Examining the co-infecting *B. cenocepacia* and *B. multivorans* evolving populations for Selection

The rates of non-synonymous ( $d_N$ ) to synonymous ( $d_S$ ) substitutions ( $d_N/d_S$  ratio) within the two evolving populations associated to the *B. cenocepacia* and *B. multivorans* isolates examined were calculated using SNAP v2.1.1 (Synonymous Non-Synonymous Analysis Program – <https://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html>). The relative abundance of non-synonymous to synonymous polymorphisms ( $p_n/p_s$  ratio) was also calculated. In brief, the concatenated variant positions created to construct the phylogeny trees (in the previous step, Fig. 2.1 – panels B and C) were used to estimate the number of synonymous and nonsynonymous nucleotide substitutions per site. This estimation was computationally performed by SNAP based on the method described by Nei and Gojobori (1986), incorporating a statistical method developed by Ota and Nei (1994) (Nei and Gojobori, 1986; Ota and Nei, 1994).

#### 2.3.6. Genomic features and structural analysis of the ancestor *B. cenocepacia* and *B. multivorans* strains genomes

*B. cenocepacia* IST439 and *B. multivorans* IST419 were considered the ancestor reference strains and used to identify subsequent evolved mutations. Reported genomic islands (GIs: BcenGIs and GiST32-s) (Holden et al., 2009; Graindorge et al., 2012; Nunvar et al., 2017) were searched for to identify their presence in *B. cenocepacia* IST439 and *B. multivorans* IST419 genomes using basic local alignment search tool – BLAST (Altschul et al., 1990; Zhang et al., 2000).

The presence of Putative GIs was also predicted using IslandViewer 4 (Bertelli et al., 2017), unique GIs were manually confirmed and curated in the genome sequences of *B. cenocepacia* IST439 and *B. multivorans* IST419. For this confirmation, it was considered that they are unique continuous DNA regions absent in the available genome sequences of the published epidemic clones and longer than 10 kbp, flanked, at both sides, by homologous regions (especially those inserted immediately downstream of tRNA), and have lower (or higher) GC content compared with the rest of the genome (Guo et al., 2017; Nunvar et al., 2017). The absence of DNA regions homologous to these unique GIs in the genomes of epidemic clones – *B. cenocepacia* ET12 and ST32 (Holden et al., 2009; Nunvar et al., 2017) and *B. multivorans* ATCC\_17616 and ATCC\_BAA-247 (Komatsu et al., 2003; Johnson et al., 2015; Johnson et al., 2016) – was confirmed and visualized by Progressive MAUVE whole genome pairwise alignment and ACT/Artemis (Carver et al., 2005; Darling et al., 2010). Large DNA regions that respected those criteria but were not predicted by IslandViewer, were also manually identified and considered genuine-unique GIs.

The assembled genomes of IST419 and IST439 were also examined to identify acquired antimicrobial resistance –AMR – genes through different databases ResFinder 3.2 (Zankari, 2014) and CARD (Jia et al., 2017) and virulence factors via VFAnalyzer from VF-database (Liu et al., 2019). All the pathoadaptive associated genes obtained *in silico* were confirmed by using ABRicate version 0.5 (<https://github.com/tseemann/abricate>). Visual presentation of all the GIs, AMR genes, and virulence factors associated genes, together with visual genome comparison was performed using Blast Ring Image Generator – BRIG (Alikhan et al., 2011).

### 2.3.7. Ethics statement

Studies involving the clinical isolates compared in this study were approved by Centro Hospitalar Lisboa Norte (CHLN)´ ethics committee and the anonymity of the patient is preserved. Informed consent for the use of these isolates in research was obtained from the patient and/or the legal guardians. The patient’s isolates and clinical data was collected as part of the epidemiological survey of Bcc bacteria involved in pulmonary infections among the CF patients receiving care followed at Hospital de Santa Maria (Cunha et al., 2003; Correia et al., 2008).

### 2.3.8. Data and nucleotide sequence accession numbers

DNA sequence reads for all isolates obtained by Illumina sequencing are available at the EMBL's European Nucleotide Archive (ENA) under accession number PRJEB200521 and PRJEB35836.

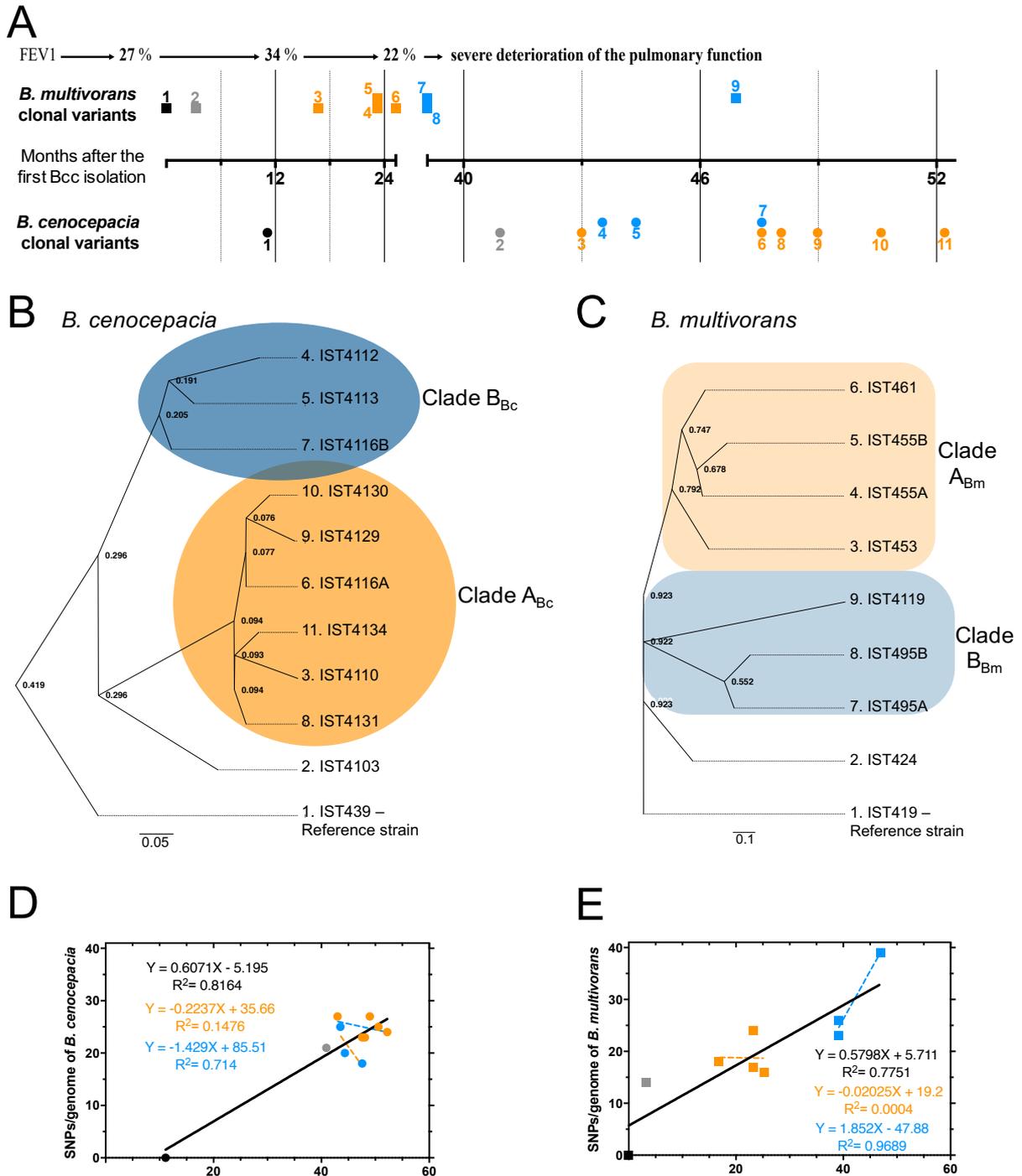
## 2.4. Results

### 2.4.1. Genomic analysis of *B. cenocepacia* and *B. multivorans* ancestor strains

The genomes of the eleven *B. cenocepacia* clonal variants and of the nine *B. multivorans* clonal variants were examined to identify the within-host evolutionary traits for these Bcc species retrieved from the same CF patient's lungs during 4.4 years of chronic infection (Fig. 2.1 – panel A). These isolates were sequentially obtained from the onset of the Bcc infection with the *B. multivorans* strain IST419 followed by a co-infection with *B. cenocepacia* (*B. cenocepacia* IST439 was the first isolate retrieved), for most of the infection period (at least 3 years of co-infection duration) until the patient's death with cepacia syndrome (Correia et al., 2008; Coutinho et al., 2011a; Hassan et al., 2019). The genome of two colonies of different morphotypes that were isolated at the same isolation date for each species (Fig. 2.1 (Coutinho et al., 2011a; Hassan et al., 2019) were also compared.

The comparative genomic analysis of the clonal isolates of *B. cenocepacia* and *B. multivorans* under study with the complete available genome sequences of *B. cenocepacia* J2315 (Holden et al., 2009) and *B. multivorans* ATCC 17616 (Komatsu et al., 2003) and ATCC BAA-247 (Johnson et al., 2015; Johnson et al., 2016), respectively, showed a thousands of false mutation calls (data not shown). For this reason, the first isolates obtained from patient J, *B. cenocepacia* IST439 and *B. multivorans* IST419, were used as the ancestor reference strains for comparative genomic analysis.

The PacBio sequencing of *B. cenocepacia* IST439 was performed and polished using Illumina short reads for high-resolution mutation-calling. The complete genome sequence of *B. cenocepacia* IST439 was assembled and 3 scaffolds obtained associated to the 3 expected replicons (Holden et al., 2009) with total genome size of 7.63 Mb (Table S2.2). Plasmid pBCJ2315, present in the genome of *B. cenocepacia* J2315 strain (Holden et al., 2009), was found to be absent from *B. cenocepacia* IST439 genome by mapping the IST439 Illumina reads against the *B. cenocepacia* J2315 genome. This non-essential replicon was absent from all the studied *B. cenocepacia* clonal isolates.



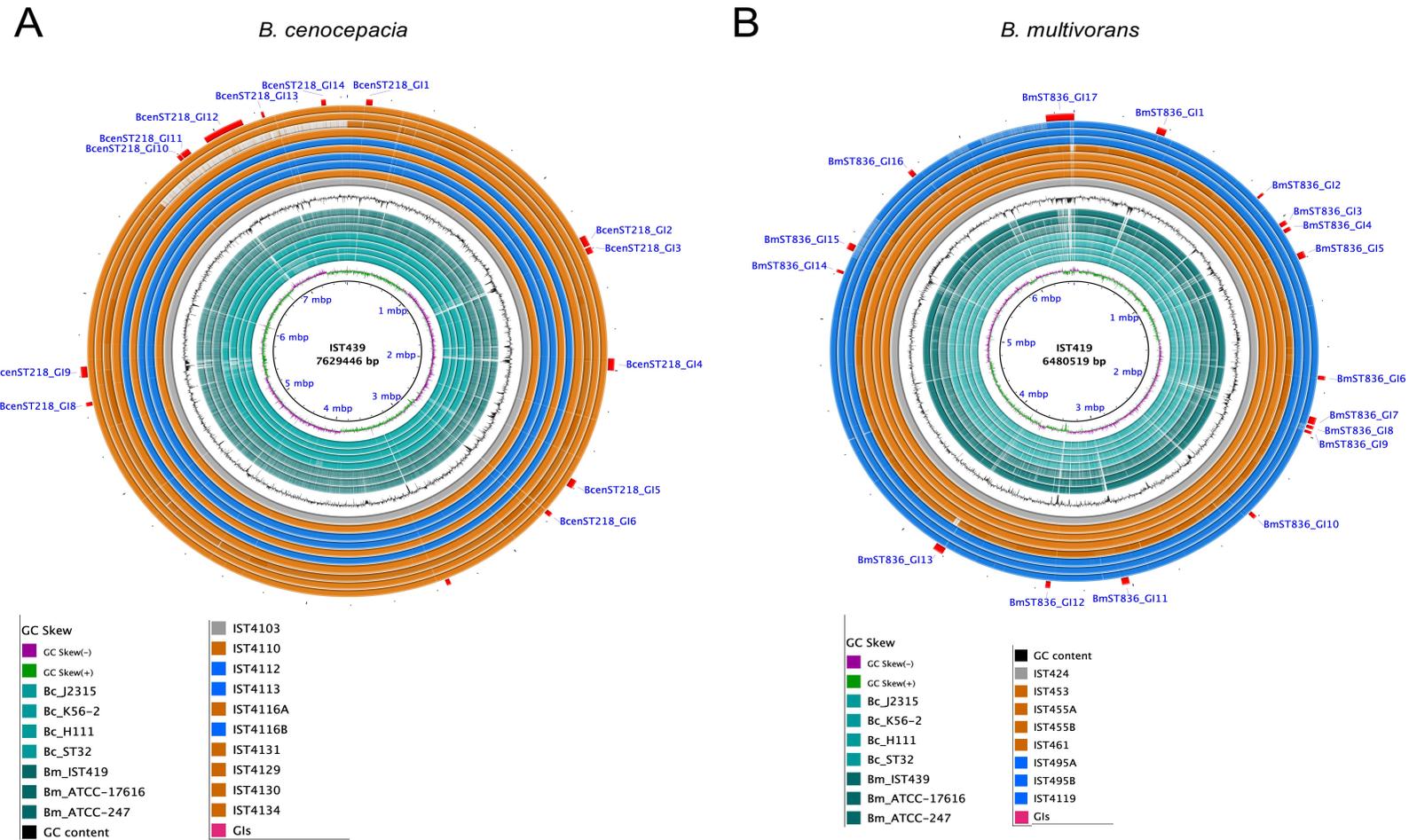
**Figure 2.1** | (A) A schematic representation of the *B. cenocepacia* (●) and *B. multivorans* (■) isolates examined with information on lung function (Forced Expiratory Volume in the first second, FEV1%). Time points (time of isolation in months after the first Bcc strain was retrieved) are presented and showing the numerical order (also presented in the phylogenetic trees [B and C] of *B. cenocepacia* and *B. multivorans*, respectively) and the two colours correlate to the two clades detected in *B. cenocepacia* and *B. multivorans* clonal variants. (D and E) are SNP accumulation during chronic infections and the relation between time elapsed from the first Bcc culture positivity to the point of bacterial isolation. A linear fit with a slope was plotted. Bc and Bm indicates *B. cenocepacia* and *B. multivorans*, respectively.

The average genome size calculated for *B. cenocepacia* and *B. multivorans* clonal isolates is 7.52 Mb and 6.41Mb, respectively (Table 2.2 and detailed metadata in Table S2.2), compared with *B. cenocepacia* J2315 (~8 Mb) and *B. multivorans* ATCC\_17616 (~7 Mb) (Komatsu et al., 2003; Holden et al., 2009). The average number of annotated genes was 6795 for *B. cenocepacia* isolates and 5821 for *B. multivorans* isolates (Table 2.2).

**Table 2.2** | Genomic assembly and functional annotation of the Bcc clonal variants. For further details of the metadata information (Table S2.2).

	<i>B. cenocepacia</i> clonal variants				<i>B. multivorans</i> clonal variants			
	Minimum	Average	Median	Maximum	Minimum	Average	Median	Maximum
<b>No. of Contigs/Scaffolds</b>	3	49.82	54	70	25	55.22	54	85
<b>Size of Largest Contigs</b>	717001	1165238.27	916088	3562400	326140	1590349.67	586375	6456554
<b>Genome Size (Bp)</b>	6495866	7438152.91	7520933	7629436	5815930	6412654	6454872	6588994
<b>Gc Contents (%)</b>	67.41	67.48	67.5	67.51	66.99	67.12	67.15	67.16
<b>N50</b>	228859	527977.55	271233	3006387	131227	1328427.11	332277	6456554
<b>No. of Cds</b>	5870	6690.18	6761	6848	5182	5737.56	5768	5947
<b>No. of Genes</b>	5972	6794.91	6866	6960	5260	5821	5846	6047
<b>No. of Misc_Rna</b>	26	30.91	31	33	19	22.33	23	25
<b>No. of Tmrna</b>	1	1	1	1	1	1	1	1
<b>No. of Trna</b>	68	72.82	70.5	85	51	60.11	56	78

The genome sequences of the ancestor reference strains, *B. cenocepacia* IST439 and *B. multivorans* IST419, were *in silico* mass-screened to search for the presence of the previously reported Genomic Islands (GIs), BcenGIs and GiST32-s (Holden et al., 2009; Graindorge et al., 2012; Nunvar et al., 2017). Most of these metabolic GIs were absent from both genomes but resistance GIs (in chromosome II) and pathogenic GI (Prophage BcepMu in chromosome III) were found (Table S2.3). These shared GIs are believed to be essential for persistence as it is the case of *cenocepacia* island cci that contains genes associated with arsenic resistance, antibiotic resistance, inorganic-ion and sulfate transporter, and stress response (Baldwin et al., 2004; Holden et al., 2009) and the low-oxygen-activated *lxa* locus with genes related to the anoxic persistence of *B. cenocepacia* (Sass et al., 2013). Another shared pathogenic GI, BcepMu prophage, is known to contain genes involved in replication, regulation and pathogenesis (Summer et al., 2004) that were found to be overexpressed in response to high doses of exogenous reactive oxygen species (ROS) applied during experimental treatment of *B. cenocepacia* J2315 biofilm (Peeters et al., 2010). Moreover, 14 putative unique GIs were predicted in the ancestor reference *B. cenocepacia* IST439 (MLST profile ST218) and termed BcenST218\_GI1 to BcenST218\_GI14 (Fig. 2.2 – panel A and Table S2.4). Also, 17 unique GIs were putatively detected in the ancestor reference *B. multivorans* IST419 (MLST profile ST836) and termed BmST836\_GI1 to BmST836\_GI17 (Fig. 2.2



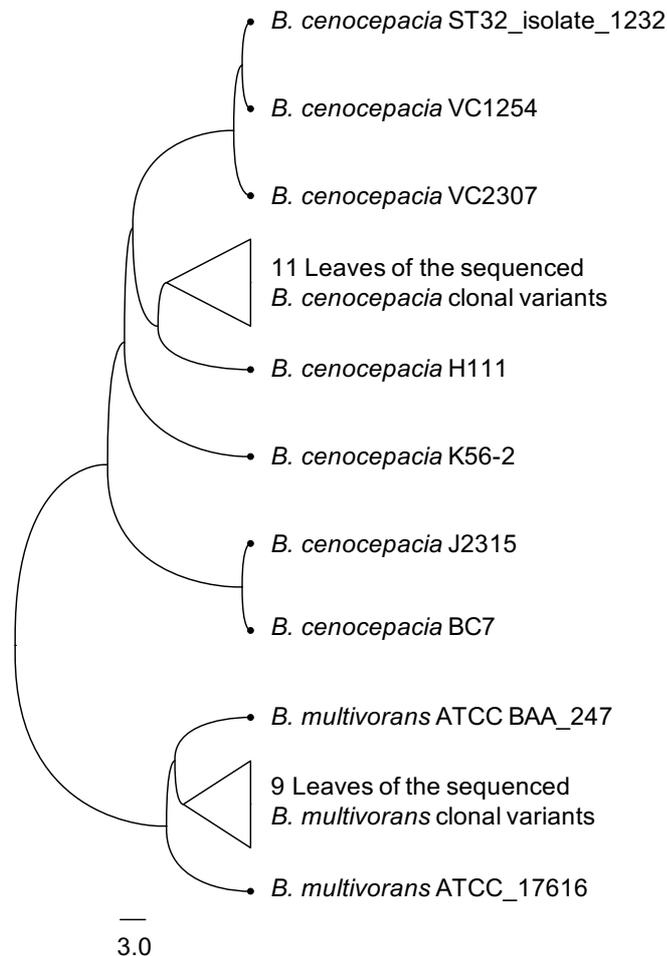
**Figure 2.2** | Genomic islands (A) BcenST281\_GIs and (B) BmST836\_GIs in the reference strains, *B. cenocepacia* IST439 and *B. multivorans* IST419, respectively, examined in the work. The inner circles denote genomes of the reference isolates and the corresponding clonal variants, ordered as indicated in the legend; Bc – *B. cenocepacia* isolates and Bm – *B. multivorans* strains. Blue and orange colors denote different clades as indicated in Fig. 1. All clonal variants were chronologically ordered based on isolation date. Visualizations were carried out by BRIG (Alikhan et al., 2011). For more details, see tables S2.4 and S2.5

– panel B and Table S2.5). All these unique GIs are presumably advantageous extra DNA sequences that include multiple phage-related genes and/or other genes associated with metabolism, replication, regulation and pathogenesis, and hypothetical membrane proteins (Tables S2.4 and S2.5). These unique GIs, together with the shared GIs (*cci*, *lxa* and *BcepMu*), corroborate the idea of the functional importance of these regions as pathoadaptive targets since some of the genes involved accumulated advantageous *in-vivo* mutations (as found below) and are believed to be a factor leading to genetic diversification (Ochman et al., 2000). This *in-silico* analysis of the ancestor reference genomes of *B. cenocepacia* IST439 and *B. multivorans* IST419 also identified genetic information present in these essential replicons related with virulence and antibiotic resistance (Figs. S2.1 and S2.2 and Tables S2.6 and S2.7).

#### 2.4.2. Comparative analysis of the genome sequences of *B. cenocepacia* and *B. multivorans* sequential clonal isolates

Based on the available genome sequences, an *in silico* multilocus sequence typing (MLST) analysis was performed and the eleven *B. cenocepacia* isolates and the nine *B. multivorans* isolates examined were confirmed as clonal variants sharing the same MLST profiles, specifically, MLST-sequence types ST-218 and ST-836, respectively (Table S2.2). Two phylogenetic trees were prepared using *B. cenocepacia* IST439 or *B. multivorans* IST419 as ancestor reference strains for the corresponding clonal variants (Fig. 2.1 – panels B and C) and 70 or 110 variant positions among *B. cenocepacia* and *B. multivorans* clonal variants, respectively. Concerning the spatiotemporal relationships of the variants, the phylogenetic trees obtained suggest that the two Bcc populations originated from a single colonization event followed by subsequent diversification leading to two different sub-populations (clades  $A_{Bc}$  and  $A_{Bm}$ ;  $B_{Bc}$  and  $B_{Bm}$ ) for each Bcc species (Fig. 2.1 – panels B and C – orange and blue for clade  $A_{Bc}$ ,  $A_{Bm}$  and  $B_{Bc}$ ,  $B_{Bm}$  – respectively). The two *B. cenocepacia* clades (clades  $A_{Bc}$  and  $B_{Bc}$ ) co-emerged during a common infection period while *B. multivorans* clade  $B_{Bm}$  was apparently developed subsequently to clade  $A_{Bm}$  (Fig. 2.1 – panels A and B). This difference is further supported by the distribution in the clades of the two isolates of *B. cenocepacia* and *B. multivorans* obtained at the same isolation date (Fig 2.1 A and B). The branching of *B. cenocepacia* phylogenetic tree did not always follow the chronology of isolation of the clonal variants. For instance, there are very similar isolates that are not chronologically sequential, as it is the case of *B. cenocepacia* IST4110 and IST4134 (Fig. 2.1 – panels A and B). This indicates long-standing co-existence of the two *B. cenocepacia* subpopulations which evolved from the original colonizing strain. Interestingly, when the genetic relationship of the Bcc clonal variants of both Bcc species were investigated with other Bcc strains with complete genomes available in *Burkholderia* genome database (<https://www.burkholderia.com/>), the *B. cenocepacia* clone here examined (ST-218) was found to be closely related to the German *B. cenocepacia* H111 isolate (ST1506) obtained from sputum of a CF patient (Carlier et al., 2014) and

further away from the worldwide ET-12 epidemic clones J2315, BC7, and K56-2 (Holden et al., 2009; Varga et al., 2013) (Fig. 2.3 and Table S2.1). The *B. multivorans* clonal variants (ST-836) examined were found to be closer to a Belgian clone *B. multivorans* ATCC\_BAA-247 (ST650) retrieved from sputum of a CF patient (Johnson et al., 2015; Johnson et al., 2016) than to the soil/American isolate *B. multivorans* ATCC\_17616 (ST21) (Komatsu et al., 2003) (Fig. 2.3 and Table S2.1).



**Figure 2.3** | Phylogenetic tree based on the comparison of the whole-genome of *B. cenocepacia* and *B. multivorans* clonal variants and other sequenced strains available in *Burkholderia* genome database (Table S2.1).

#### 2.4.3. Evolutionary dynamics of *B. cenocepacia* and *B. multivorans* co-infecting populations inside the CF lung

A total of 63 single-nucleotide polymorphisms (SNPs) and 9 insertions-deletions (INDELs) and of 97 SNPs and 16 INDELs were accumulated over the course of infection among the *B. cenocepacia* and *B. multivorans* clonal variants examined, respectively (Tables S2.8 and S2.9). SNP accumulation rates (calculated as linear regression slope, Fig. 2.1 – panels D and E) indicate that the mutation rates for *B. cenocepacia* and *B. multivorans* populations were 2.08 and 2.27 SNPs/year, respectively. As

aforementioned, each *B. cenocepacia* and *B. multivorans* population includes two clades that emerged from a single colonization event, followed by subsequent diversification (Fig. 2.1 – panels B and C). Most of the mutations, especially those observed along with the clades' emergence and shared among multiple isolates, are nonsynonymous (Tables S2.8 and S2.9) and observed in coding sequences (CDSs). For instance, 22 nonsynonymous and 1 synonymous mutations were associated with two clades A<sub>Bc</sub> and B<sub>Bc</sub> that emerged within the *B. cenocepacia* population and 35 nonsynonymous and 7 synonymous were related to the two clades A<sub>Bm</sub> and B<sub>Bm</sub> that emerged in the *B. multivorans* population during long term infection. Such enrichment of nonsynonymous mutations is greater than the reported substitution rate (27.8%) under neutral evolution given the codon usage and GC content of *B. cenocepacia* bacteria (Dillon et al., 2015). These results suggest that the clonal variants of both species are under a positive selection inside the CF lung.

**Table 2.3** | Pair-wise comparison of the infecting population aligned to the ancestor reference strains *B. cenocepacia* IST439 and *B. multivorans* IST419.

	$d_S$ Rate of synonymous substitution per synonymous site	$d_N$ Rate of non-synonymous substitution per non-synonymous site	$d_N/d_S$ Comparison ratio of the two rates
<b><i>B. cenocepacia</i></b>			
entire population	0.4289	0.3808	0.888
Clade A	0.3423	0.2475	0.723
Clade B	0.6373	0.242	0.380
<b><i>B. multivorans</i></b>			
entire population	0.6171	0.5237	0.849
Clade A	0.2743	0.2215	0.808
Clade B	0.6166	0.6096	0.988

The observed enrichment of nonsynonymous mutations and the proposed positive selection inside the CF lung were statistically validated, as proposed before (Hedge and Wilson, 2016), to understand and detect the host-CF lung-selection in the Bcc populations especially the periods of diversification. The rate of synonymous nucleotide substitution per synonymous site ( $d_s$ ) and the rate of non-synonymous substitution per non-synonymous site ( $d_n$ ) were estimated. The averages of all rates per synonymous and nonsynonymous sites are presented in Table 2.3. Over a period of 4.4 years of Bcc infection in this case-study, *B. cenocepacia* and *B. multivorans* populations last (from the first isolated clone to the last variant of the same species) 3.4 and 4 years in the CF lung, respectively (Fig. 2.1 –

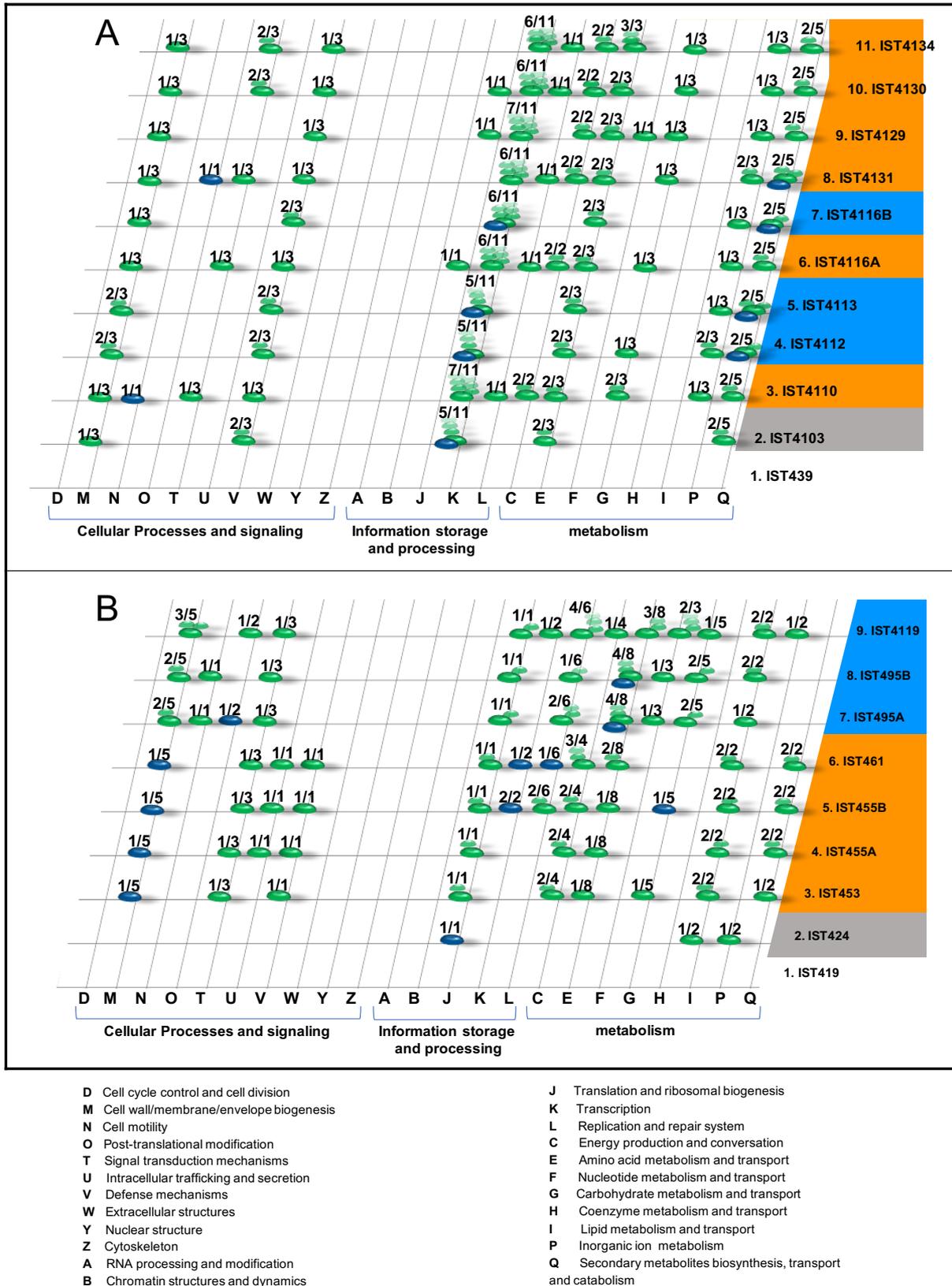
panel A and Table 2.1). However, although the comparison ratio of the two rates ( $d_N/d_S$ ) is considered to be time-dependent for closely related bacterial-genomes (Rocha et al., 2006), the evaluation of the  $d_N/d_S$  values was not helpful in determining the existence of positive selection, which was not surprising considering that these determinations can be biased in similar populations characterized by small number of isolates with uneven chronological distribution (Kryazhimskiy and Plotkin, 2008). Nevertheless, the high number and distribution of non-synonymous mutations that we observed to occur in several genes (Fig. 2.4) suggest the existence of positive selection as an important driving force underlying the adaptation of Bcc to the CF patient lung. Similar positive selection has been reported for other Bcc strains, belonging to *B. cenocepacia*, *B. multivorans* and *B. dolosa* species, inhabiting other patients and considered the result of the action of the CF-lung hostile-environment (Lieberman et al., 2011; Lieberman et al., 2014; Silva et al., 2016; Lee et al., 2017; Nunvar et al., 2017).

Consistent with the increased selection observed for *B. cenocepacia* clonal variants of Clade B<sub>Bc</sub>, only a few genes accumulated mutations that were not found in the other clade A<sub>Bc</sub> (Table S2.8). For instance, among clade B<sub>Bc</sub>, the gene homologous to the IST439\_04198 locus, a *penA* homolog from *B. cenocepacia* J2315 that encodes a putative beta-lactamase, was found to acquire Asp262Gly mutation in the variants IST4112, IST4113 and IST4116B. This mutation might be a result of the drastic intravenous therapy with gentamicin and ceftazidime and in agreement with the previously reported high-resistance profile to beta-lactam antibiotic ceftazidime, a beta-lactam antibiotic (Coutinho et al., 2011a). Also, among clade B<sub>Bc</sub> variants, one indel was observed (Pro599\_Ala601deletion) in genes homologous to IST439\_04244 from the ancestor strain. This IST439\_04244 gene is homologous to BCAM2211 from *B. cenocepacia* J2315 that is an orthologue of *acoR* from *Bacillus subtilis*. This *acoR* gene encodes a sigma-54 dependent transcriptional regulator that is involved in the regulation of the acetoin catabolic pathway (Huang et al., 1999; Ali et al., 2001). The acetoin catabolic pathway is believed to be a defense mechanism to avoid the lethal acidification produced by CF pathogens during fermentation and anaerobic respiration as a result of the decreased oxygen availability inside the CF lung (Whiteson et al., 2014). Moreover, other two mutations (Ser382Leu and Val449Leu) were observed in genes homologous to IST439\_00883 and IST439\_00884 loci, respectively. These two loci are homologous to *orbA* and *pvdA*, respectively, that encode ornibactin siderophore biosynthesis and transport proteins (Agnoli et al., 2006). The *pvdA* deletion mutant were found to be less virulent than the parent strain in chronic and acute models of respiratory infection (Sokol et al., 1999) and the *B. cenocepacia* K56-2 *orbA* mutant (K56*orbA*::tp, *orbA* allelic exchange mutant) has also an attenuated virulence in the rat chronic respiratory infection model (Sokol et al., 2000). The reported function of the four genes identified, among *B. cenocepacia* clonal variants of Clade B<sub>Bc</sub>, is consistent with the idea that positive selection might be driven by different environmental factors of the CF lung, in particular high antibiotic concentrations, low oxygen concentrations and low iron concentrations.

#### 2.4.4. *B. cenocepacia* and *B. multivorans* genes under convergent evolution in the CF lung

Synonymous and non-synonymous mutations accumulated in the clonal variants that evolved within the CF lung during chronic infection appeared in 45 and 61 protein-coding *sequences* (CDSs) of *B. cenocepacia* and *B. multivorans* genomes, respectively. These CDSs belong to different clusters of orthologous groups – COG (Fig 2.3 and Tables S2.8 and S2.9). The genes of well-characterized functions that accumulated non-synonymous mutations (41 and 48 *B. cenocepacia* and *B. multivorans* CDSs, respectively) are schematized in Fig. 2.4. Most of the mutated genes in the *B. cenocepacia* were found in “metabolism” [41.4%] (COGs E, F, G, P, Q and C) and “cellular processes and signaling” [26.8%] (COGs M, T, N, O and V) categories, as registered for *B. multivorans*: (metabolism [54.2%] (COGs E, F, G, P, I, Q and C) and cellular processes and signaling [27.1%] (COGs M, T, N, O, U and V)). Among the five COGs representative of “information storage and processing” functions (COGs A, B, K, J and L – Fig. 2.4), only COGs K, J, and L include the mutated genes in *B. cenocepacia* and *B. multivorans* populations (31.8% and 18.7%, respectively).

Considering that the most obvious genes under host-selection are those with non-synonymous mutations that become fixed over the entire period of chronic infection, one mutation was found in all the *B. cenocepacia* clonal variants in a gene linked to defense mechanisms against antibiotic stresses (IST439\_01475, encoding a Putative ABC transporter ATP-binding membrane protein; homologous to BCAL1039 from *B. cenocepacia* J2315) (Table 2.4). Three other mutations were registered in genes, involved in transcription regulation, encoding the ribose operon repressor RbsR and two sigma factor cytoplasmic proteins homologous to *B. cenocepacia* J2315 RpoC and RpoD proteins. Three mutations were also detected in genes associated with metabolism (BCAL1610 and *glnQ* homologs – encoding periplasmic cysteine-binding protein and glutamine ABC transporter ATP-binding protein – and *pcaC* homolog encoding 4-carboxymuconolactone decarboxylase, involved in the metabolism of amino acids and secondary metabolites, respectively). On the other hand, no single non-synonymous mutation was identified among all the *B. multivorans* clonal variants that became fixed during the entire period of chronic co-infection (Table S2.9).



**Figure 2.4** | 3D plot of non-synonymous mutations formed among *B. cenocepacia* (A) and *B. multivorans* (B) clonal variants during the chronic co-infection. All mutated genes of well-characterized functions are depicted as green oval shapes (with SNPs) and/or blue shapes (with INDELS). The number of mutated genes per each COG category is illustrated above oval shapes as a ratio (number of genes per isolate/number of all mutated genes in the entire population for this particular COG). Accumulation of small green shapes indicates the number of SNP

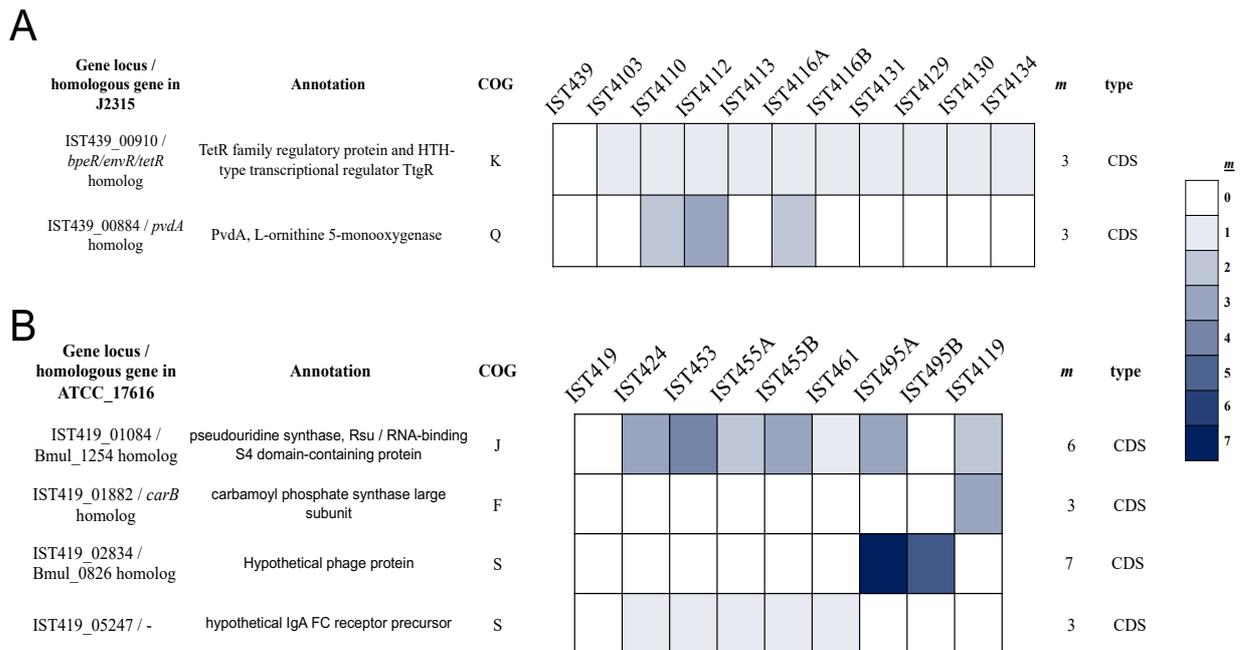
mutations that was observed per isolate. For detailed list of mutations, see Tables S2.8 and S2.9 for *B. cenocepacia* and *B. multivorans*, respectively. Abbreviation letters representing the COGs are detailed in the bottom panel.

Among the mutated genes, two polymorphic genes (that acquired multiple independent mutations) from *B. cenocepacia* genome and four polymorphic genes from *B. multivorans* genome accumulated three or more mutations during chronic co-infection (Fig. 2.5) providing evidence for convergent evolution and selection (Didelot et al., 2016). Regardless the Bcc species, “information storage and processing” (COGs K and J) and “metabolism” (COGs Q and F) are the functional groups in which all the polymorphic genes were sorted (Fig. 2.5). Notably, among *B. cenocepacia* clonal variants, one polymorphic gene, IST439\_00910, homologous to *B. cenocepacia* J2315 *bpeR*, evolved by acquiring three different mutations during co-infection; one mutation (Gly196Asp) became fixed in clade A<sub>Bc</sub> variants while mutation Cys63Arg emerged in clade B<sub>Bc</sub> and one INDEL (Lys113fs [-57bp]) was only found in the second sequential clonal variant isolated, IST4103, the isolate that does not belong to any of the two defined clades (Fig. 2.1 – panel B). The *bpeR* gene encodes a TetR family regulatory protein which is a conserved transcription regulator across the *Burkholderia* genus (Ramos et al., 2005). This gene was involved in the transcriptional control of multidrug efflux pumps, antibiotic biosynthesis, response to osmotic stress and toxic chemicals, and in pathogenicity in other Gram-negative pathogens (Ramos et al., 2005). The IST439\_00910/*bpeR* is also an ortholog to *B. dolosa* AU0158 BDAG\_00732 that was found before to acquire several mutations in the promoter region of 112 CF-clinical isolates (Lieberman et al., 2014). Among the four polymorphic genes in *B. multivorans* clonal variants, one polymorphic gene (IST419\_01084) accumulated six different mutations and one of these variant positions (Val285Ala) became fixed in the first 6 clonal variants (Fig. 2.5 – panel B). This IST419\_01084 gene encodes a putative RNA pseudouridine synthase that structurally-includes a small RNA-binding protein domain (S4 domain) which delivers nucleotide-modifying enzymes to RNA, regulating translation through structure specific RNA binding (Aravind and Koonin, 1999; Del Campo et al., 2001).

#### 2.4.5. Orthologous Genes, under selection in the co-evolved *B. cenocepacia* and *B. multivorans* populations

Two orthologous genes were found to be under selection and accumulate mutations related to clade emergence in *B. cenocepacia* and *B. multivorans* (Fig. 2.6). Glu487Lys and Leu493Pro mutations were observed in IST439\_01755 gene in all *B. cenocepacia* clonal variants forming clades A<sub>Bc</sub> and B<sub>Bc</sub>, respectively, while Met69fs (+64bp) insertion were observed in IST419\_02260 gene in all *B. multivorans* clonal variants forming clade A<sub>Bm</sub>. The IST439\_01755 gene from *B. cenocepacia* IST439 is an ortholog of *B. multivorans* IST419 IST419\_02260 gene and is homologous to BCAL3119/*wbiI* from *B. cenocepacia* J2315 and Bmul\_2504 from *B. multivorans* ATCC\_17616, respectively. The *wbiI*

gene encodes a nucleotide sugar dehydratase involved in lipopolysaccharide O-antigen (OAg) biosynthesis (Ortega et al., 2005). Interestingly, regardless the Bcc species, all the isolates with a mutated *wbiI* gene expressed a lipopolysaccharide that lacks the OAg (Hassan et al., 2019). Remarkably, this *wbiI* gene was reported to be under strong selection in *B. cenocepacia*, *B. multivorans* and *B. dolosa* during chronic lung infection (Lieberman et al., 2011; Lieberman et al., 2014; Silva et al., 2016; Hassan et al., 2017).



**Figure 2.5** | Heat-map of non-synonymous mutation frequencies in coding sequences (CDS) under strong parallelism (that acquired  $\geq 3$  mutations) among the clonal variants of *B. cenocepacia* (A) and *B. multivorans* (B). Homologous genes in *B. cenocepacia* J2315 and *B. multivorans* ATCC\_17616 are mentioned. Total numbers of mutations acquired per gene (m) and the type of mutation (coding sequence [CDS]) are presented on the right.

The other gene is *B. cenocepacia* IST439 gene IST439\_06597, an ortholog of *B. multivorans* IST419 IST419\_05375 gene that is homologous to *B. cenocepacia* J2315 BCAS0619 and *B. multivorans* ATCC\_17616 Bmul\_05359. This gene is located in the third chromosome and encodes a two-component regulatory sensor kinase protein. The mutation Val176Met is found within *B. cenocepacia* variants that forms clade A<sub>Bc</sub>, with the exception of the IST4129 isolate that lacks chromosome III, while IST419\_05375 gene has a Leu181Phe mutation in two *B. multivorans* variants (IST495A and IST495B).

**Table 2.4** | List of non-synonymous mutations (highlighted with blue) that became fixed among all clonal variants of *B. cenocepacia*. The only presented mutations are those were acquired by CDS associated with well-characterized functions as shown in Table S2.8.

Gene locus / homologous gene in J2315	Annotation	Mutation al effect	COG	IST439	IST4103	IST4110	IST4112	IST4113	IST4116A	IST4116B	IST4131	IST4129	IST4130	IST4134
IST439_01475 / BCAL1039 homolog	Putative (ABC transporter) multidrug export ATP-binding/permease membrane protein	Gln347*	V											
IST439_00921 / <i>rbsR</i> homolog	Ribose operon repressor	Val287Asp	K											
IST439_02295 / <i>rpoC</i> homolog	DNA-directed RNA polymerase subunit beta	Tyr669Cys	K											
IST439_05755 / <i>rpoD/sigE</i> homolog	RpoD/SigE, RNA polymerase sigma factor RpoD	Arg167Thr	K											
IST439_00975 / BCAL1610 homolog	FliY_1, Cystine-binding periplasmic protein	Asp13Gly	E											
IST439_01981 / <i>glnQ</i> homolog	Glutamine transport ATP-binding protein GlnQ/ABC transporter-like protein	Gln136*	E											
IST439_4949 / <i>pcaC</i> homolog	PcaC, 4-Carboxymuconolactone decarboxylase family protein	Ser52Arg	Q											

<i>B. multivorans</i> clonal variants										<i>B. cenocepacia</i> clonal variants																		
Amino acid residue change	IST419	IST424	IST453	IST455A	IST455B	IST461	IST495A	IST495B	IST4119	Gene locus in IST419	Homologous gene in ATCC_17616	Homologous gene in J2315	Gene locus in IST439	Annotation	COG	IST439	IST4103	IST4110	IST4112	IST4113	IST4116A	IST4116B	IST4131	IST4129	IST4130	IST4134	Amino acid residue change	
Leu181Phe										IST419_05375	Bmul_5359	BCAS0619	IST439_06597	Two-component regulatory system, sensor kinase protein, FgrL/YycG	T													Val176Met
Met69fs (+64bp)										IST419_02260	Bmul_2504	BCAL3119	IST439_01755	UDP-N-acetyl- $\alpha$ -D-glucosamine C6 dehydratase / nucleotide sugar epimerase/dehydratase, WbiI	M													Glu487Lys
																												Leu493Pro
<b>Clades</b>			A <sub>Bm</sub>	A <sub>Bm</sub>	A <sub>Bm</sub>	A <sub>Bm</sub>			B <sub>Bm</sub> B <sub>Bm</sub> B <sub>bm</sub>																			

**Figure 2.6** | Mutations that were observed in two orthologous genes in the genomes of clonal variants of *B. multivorans* and of *B. cenocepacia* (chronological order from left to right) indicated by the amino acid residue change.

## 2.5. Discussion

The present study was designed to gain insights into the molecular mechanisms underlying genomic diversification occurring in two *B. cenocepacia* and *B. multivorans* strains that co-infected, for a period of at least 3 years, the same cystic fibrosis (CF) patient and to identify common and relevant pathoadaptive mechanisms. Both species were found to have acquired several mutations with accumulation rates of 2.08 (*B. cenocepacia*) and 2.27 (*B. multivorans*) SNPs/year during the course of chronic co-infection in the CF-lung hostile environment. These mutation rates are comparable with those reported for *B. dolosa* [2.1 SNPs/year] (Lieberman et al., 2011; Lieberman et al., 2014) and *B. multivorans* [2.4 SNPs/year] (Silva et al., 2016) during chronic CF infection and slightly higher than the mutation rate reported for the Czech epidemic *B. cenocepacia* ST32 strains [average rate 1.66 SNPs/year] (Nunvar et al., 2017). Such mutation rates are also comparable to those occurring in the major CF pathogen *P. aeruginosa* [2.7 SNPs/year] during chronic infection in the CF lung (Markussen et al., 2014; Marvig et al., 2015) but lower than the mutation rates reported for the non-Bcc member and etiological agent of melioidosis *B. pseudomallei*, obtained from chronically infected CF patients [3.6 SNPs/year] (Viberg et al., 2017). The aforementioned mutation rates are believed to be a common feature of the within-CF lung-host evolution associated with fast diversification of the infecting bacterial strains (Didelot et al., 2016) that diverge into sublineages, each one with their own functional and genomic signatures and rates of adaptation to different CF-lung environments (Markussen et al., 2014; Didelot et al., 2016).

Besides point mutations, the acquisition of genetic sequences from unrelated organisms through horizontal gene transfer is known to drive genome diversification (Ochman et al., 2000) as well as large genetic losses (Rau et al., 2012). Supporting this concept, the genomic structure of the two ancestor strains, *B. cenocepacia* IST439 and *B. multivorans* IST419, compared to the Bcc reference strains *B. cenocepacia* J2315 and *B. multivorans* ATCC\_17616, miss several genomic islands previously reported (Holden et al., 2009; Graindorge et al., 2012; Nunvar et al., 2017) due to large deletions or substitutions but other genomic islands are present, such as the *cenocepacia* island *cci* and the low-oxygen-activated *lxa* locus (Baldwin et al., 2004; Holden et al., 2009; Sass et al., 2013). The complete loss of the non-essential replicon chromosome III/megaplasmid was found to occur during chronic co-infection but only for the clonal variant *B. cenocepacia* IST4129. Such loss of chromosome III/megaplasmid during chronic infection was previously experimentally detected in our laboratory as a rare event (for only one *B. cenocepacia* isolate among 39 clinical *B. cenocepacia* isolates obtained from 3 different CF patients) (Moreira et al., 2017). The deletion of genetic regions in pathogenic bacteria is well documented as an example of reductive genome evolution due to positive selection (Rau et al., 2012; Merhej et al., 2013; Bentley and Parkhill, 2015).

In the co-infecting *B. cenocepacia* and *B. multivorans* populations examined, several genes were found to acquire different mutations either related or unrelated with the emergence of different clades. Remarkably, most of these genes have not been previously reported as being involved in parallel evolution in *B. dolosa*, *B. cenocepacia*, and *B. multivorans* (Lieberman et al., 2011; Lieberman et al., 2014; Silva et al., 2016; Nunvar et al., 2017). However, most of the mutated genes found in this work are involved in cell envelope/wall/membrane biogenesis, and in regulatory and metabolic processes, as reported before for Bcc strains as being under selective pressure in the CF lungs (Lieberman et al., 2011; Lieberman et al., 2014; Silva et al., 2016; Nunvar et al., 2017). Mutated genes with fixed non-synonymous variant positions over the entire period of chronic infection were found in the present work among the *B. cenocepacia* clonal variants examined but, no single fixed non-synonymous variant position was found when the genomes of the *B. multivorans* clonal variants were examined. The differences observed for the two co-infecting Bcc species could be genetically-grounded predetermining the way how the initial infecting strains establish their independent adaptive evolution mechanisms, as suggested by the number of the registered non-synonymous mutations compared with synonymous nucleotide substitution. Similar positive selection has been reported for other Bcc strains, belonging to *B. cenocepacia*, *B. multivorans* and *B. dolosa* species, inhabiting other patients and considered the result of the action of the CF-lung hostile-environment (Lieberman et al., 2011; Lieberman et al., 2014; Silva et al., 2016; Lee et al., 2017; Nunvar et al., 2017). In summary, although coexisting in the same hostile-lung environment for several years, the registered evolutionary patterns for *B. cenocepacia* and *B. multivorans* appear to be species specific. Although we cannot discard the possibility that the spatiotemporal characteristics of the isolated clonal variants examined in this work are not representative of the population heterogeneity present at each isolation time in the CF airways, the distinct evolutionary patterns here attributed to *B. cenocepacia* and *B. multivorans* when co-infecting the same CF environment for several years, appear to be very consistent and in line with the literature.

The comparative genomic analysis performed during the present work indicates that among the mutated genes in *B. cenocepacia* and *B. multivorans* during long-term co-infection of the CF patient are genes encoding oxidative stress regulatory proteins and heavy metal-sensing proteins (Tables S2.8 and S2.9). Such regulatory proteins (e.g. OxyR and FixK) are involved in bacterial protection against reactive oxygen species (ROS) produced by phagocytic leukocytes (Reyrat et al., 1993; Saito et al., 2003; Bobik et al., 2006; Panmanee et al., 2008). Consistent with this conclusion, the host immune system was proposed to play a fundamental role in driving *B. cenocepacia* and *B. multivorans* evolution during chronic CF infection and genes linked to the oxidative stress responses (homologous to *B. cenocepacia* J2315 *katG*, *yedY*, *moeA1*, *fixL* and *ompR*) reported to suffer mutation (Silva et al., 2016; Nunvar et al., 2017). This selection driven by the host immune system was not however reported within-patient evolution either in *B. dolosa* or *P. aeruginosa* (Lieberman et al., 2011; Lieberman et al., 2014;

Winstanley et al., 2016). Increased resistance to antibiotics is also a well-described evolutionary trait associated with many CF pathogens and genes related with antibiotic resistance have been among the genes mostly reported to be under selective pressure (Jeannot et al., 2008; Lieberman et al., 2011; Lieberman et al., 2014; Silva et al., 2016; Winstanley et al., 2016; Nunvar et al., 2017). In agreement with these reports, some of the mutated genes found in the present study in both Bcc species are involved in defense mechanisms and their regulation (as those encoding PenA, EmrA, ABC-2 type transporter, and DNA gyrase proteins) (Tables S2.8 and S2.9). Another commonly reported evolutionary trait is the genetic changes occurring in genes present in the OAg biosynthetic cluster (Lieberman et al., 2011; Lieberman et al., 2014; Maldonado et al., 2016; Silva et al., 2016; Hassan et al., 2017; Nunvar et al., 2017). Longitudinal comparative genomic studies of Bcc isolates from CF patients with chronic lung infections revealed the conversion from smooth (in the early isolates) to rough LPS with no OAg at the late-stages of infections by *B. cenocepacia* and *B. multivorans* (Silva et al., 2016; Hassan et al., 2017). However, another extensive study focused on *B. dolosa* during chronic lung infections has shown that late isolates produce an LPS exhibiting the OAg that was absent from the LPS of the initial infecting isolate (Lieberman et al., 2011). Comprehensive comparative genomic studies performed during chronic infection on the most feared *B. cenocepacia* (Nunvar et al., 2017) and *B. multivorans* (Silva et al., 2016) and on the far less distributed *B. dolosa* (Lieberman et al., 2011); revealed that the genes affected during adaptive evolution in *B. cenocepacia* or *B. multivorans* are different from *B. dolosa*. Such phenotypic switch of the OAg presence/absence was recently examined among several Bcc species and the most prevalent and feared species, *B. cenocepacia* and *B. multivorans*, showed a marked tendency to lose the OAg along chronic infection while in the more distantly related *B. dolosa* species, the OAg-chain was absent from the isolates obtained since the beginning of the 5.5-year infection until the patient died (Hassan et al., 2019). This indicates that evolution in *B. cenocepacia*, *B. multivorans* and *B. dolosa* during chronic CF infection is driven by different selective forces presumably linked to host immune responses (Lieberman et al., 2014; Nunvar et al., 2017). Specifically, while the most mutated genes reported to undergo adaptive within-patient evolution in *B. cenocepacia* were associated with oxidative stress response and transition metal metabolism, the same was not observed either in *B. dolosa* or *P. aeruginosa* (Lieberman et al., 2014; Winstanley et al., 2016; Nunvar et al., 2017). It is also known that the OAg absence leads to increased Bcc survival in eukaryotic cells as amoebae, epithelial cells and human macrophages (Saldias et al., 2009; Saldias and Valvano, 2009; Kotrange et al., 2011; Maldonado et al., 2016), to increased internalization of *B. cenocepacia* into macrophages upon phagocytosis (Saldias et al., 2009; Kotrange et al., 2011) and to facilitate *B. multivorans* growth inside macrophages (Schmerk and Valvano, 2013). Consistent with those reports, it was found that different mutations were accumulated and related to the clade emergence in the two co-infecting populations in genes, IST439\_01755 and IST419\_02260 (homolog and ortholog to *wbiI*, respectively – Fig. 2.6), that encodes a nucleotide sugar dehydratase that was found to be under parallel evolution in *B. cenocepacia* or in *B. multivorans* (Silva et al., 2016; Hassan et al., 2017). These results are consistent with the idea

that OAg loss could promote Bcc persistence through intracellular survival inside the macrophages. Remarkably, the comparison of the ability to subvert the host's immune function, assessed by internalization assays using human dendritic cells, of three of the *B. cenocepacia* isolates examined in the present work showed that the late variants, IST4113 and IST4134, were significantly more internalized exhibiting increased survival within dendritic cells than the early isolate IST439 (Cabral et al., 2017) the sole with an LPS with OAg (Hassan et al., 2017). Altogether, mutated genes' functions involved in the observed adaptive evolution of the two strains co-inhabiting the same host-selective environment are consistent with the idea that positive selection in both *B. cenocepacia* and *B. multivorans* strains might be driven by the action of the host immune system.

Since the CF patient from whom the clonal isolates examined here were retrieved died from the cepacia syndrome, the results obtained might be useful to unveil the molecular mechanisms underlying the development of this feared outcome which remain poorly characterized. Former studies have suggested that *B. cenocepacia* genes involved in oxidative stress and transition metals metabolism might be the targets of convergent evolution underlying higher affinity to intra-macrophage persistence and development of the fatal cepacia syndrome (Lee et al., 2017; Nunvar et al., 2017). The present study provides evidences supporting this concept since among the few mutated regulatory proteins identified are the histidine kinase protein (IST439\_06759, *hisS* homolog and IST419\_05375) that enable the bacterial cells to sense and adapt to oxidative- and heavy metal- inducing stresses (Wolanin et al., 2002; Skerker et al., 2005). This histidine kinase encoding gene acquired two mutations (Val176Met and Leu181Phe) in the two co-infecting *B. cenocepacia* and *B. multivorans* populations, respectively (Fig. 2.6). This global regulator encoded by IST439\_06597 or by IST419\_05375 had not been reported before to undergo adaptive within-patient evolution either in *B. dolosa*, *B. multivorans*, or in *P. aeruginosa* (Lieberman et al., 2011; Lieberman et al., 2014; Silva et al., 2016; Winstanley et al., 2016). Signal transduction histidine kinase-related proteins are involved in signaling cascades contributing to several simultaneous coordinated responses (Wolanin et al., 2002; Skerker et al., 2005) such as in osmoregulation (Tomomori et al., 1999) and chemotaxis (Bilwes et al., 1999) in other bacterial pathogens. Putative orthologous of these particular global regulators, encoded by IST439\_06597 and IST419\_05375, were found to interact with several other regulatory proteins in other bacterial pathogens (Skerker et al., 2005; Yamamoto et al., 2005; Mechaly et al., 2014; Babu et al., 2018). For instance, it was found to interact with the two-component regulatory protein, IrlR – homologous to BCAM0443, that is known to be a heavy metal response regulator and cross-talk with putative cyclic-di-GMP signaling protein, homolog to BCAL1635, that is known to influence vital processes in persistent infection such as in biofilm formation (Fazli et al., 2014). Such interactions have been experimentally proved to influence cell envelope biogenesis in other pathogens (Skerker et al., 2005; Babu et al., 2018). Therefore, it is likely that global regulators histidine kinases, encoded by IST439\_06597 and IST419\_05375 might cross-talk with heavy metal response regulators and

regulatory proteins required for protection against reactive oxygen species (ROS) produced by leukocytes, reinforcing the idea of the involvement of the host immune system in driving *B. cenocepacia* and *B. multivorans* evolution during chronic CF co-infection, as suggested before (Nunvar et al., 2017). Given that persistent inflammation and neutrophil infiltration often accompany chronic lung infections, it has been hypothesized that under increased stress encountered in CF macrophages, the global stress response might be activated by the evolved bacterial population and thus modulate the course of infection (Chua et al., 2016; 2017; Nunvar et al., 2017). Altogether, it is hypothesized that the observed regulatory system that suffered mutation and is under selective pressure during chronic co-infection inside the CF lung is required for bacteria adaptation to oxidative and heavy-metal stresses eventually underlying a yet undisclosed regulatory mechanism related with cepacia syndrome development (Lee et al., 2017; Nunvar et al., 2017)].

In summary, the comparative genomic analysis performed in this work identified genes which are under strong positive selection in the most prevalent and most feared Bcc species that co-infected the lung of a CF patient who died from cepacia syndrome. Among them, polymorphic genes and others involved mainly in regulatory and metabolic functions were sorted out. Since cystic fibrosis is a genetic disorder associated with inflammation, sub-optimal antioxidant protection and the continuous use of antimicrobial therapy, all resulting in marked oxidative stress (Galli et al., 2012), it is likely that the evolved strains might have activated several physiological processes via the global stress responses which in turn can modulate the course of infection. Global stress responses were found to be under parallel evolution in experimentally-evolved *P. aeruginosa* biofilms (Winstanley et al., 2016) and were recently observed during chronic lung infections, irrespective to the presence/absence of other members of the CF microbiome (Vandeplassche et al., 2019). More recently, in a comparative transcriptomic analysis of two different *Vibrio* species in the oyster host model, species-specific mechanisms involving global stress responses (cross-talk with genes conferring resistance to ROS and heavy metals) were found to converge and compromise host immune responses, allowing evasion from the host immune system (Rubio et al., 2019). Given that both co-infecting Bcc species are known for their ability of intracellular survival inside macrophages (Saldias et al., 2009; Saldias and Valvano, 2009; Kotrange et al., 2011; Schmerk and Valvano, 2013), the observed mutated genes associated with cell envelope biogenesis may suggest that this cellular compartment, including the lipopolysaccharide OAg, could promote Bcc persistence through intracellular survival and eventually promote cepacia syndrome development. The present study provides relevant and useful information that may contribute to a better understanding of the common and the specific evolutionary patterns occurring in *B. cenocepacia* and *B. multivorans* under positive selection in the same CF lung environment and host immune system during co-infection and to the understanding of the mechanisms underlying the development of the cepacia syndrome.



### 3      **Structure of O-antigen and hybrid biosynthetic locus in *Burkholderia cenocepacia* clonal variants recovered from a cystic fibrosis patient**

---

#### **This chapter contains results published in:**

A. Amir Hassan<sup>1§</sup>, Rita F. Maldonado<sup>1§</sup>, Sandra C. dos Santos<sup>1§</sup>, Flaviana Di Lorenzo<sup>2§</sup>, Alba Silipo<sup>2</sup>, Carla P. Coutinho<sup>1</sup>, Vaughn S. Cooper<sup>4</sup>, Antonio Molinaro<sup>2</sup>, Miguel A. Valvano<sup>3</sup> and Isabel Sá-Correia<sup>1</sup>, **Structure of O-antigen and hybrid biosynthetic locus in *Burkholderia cenocepacia* clonal variants recovered from a cystic fibrosis patient. (2017)**, *Frontiers in Microbiology* 8, 1027. doi: 10.3389/fmicb.2017.01027.

<sup>1</sup> iBB-Institute for Bioengineering and Biosciences, Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal

<sup>2</sup> Department of Chemical Sciences, University of Napoli Federico II Complesso Universitario Monte Santangelo, Napoli, Italy

<sup>3</sup> The Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast, Belfast, UK

<sup>4</sup> Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, Pittsburgh, USA

§ Authors contributed equally



### 3.1. Abstract

*Burkholderia cenocepacia* is an opportunistic pathogen associated with chronic lung infections and increased risk of death in patients with cystic fibrosis. In this work, we investigated the lipopolysaccharide (LPS) of clinical variants of *B. cenocepacia* that were collected from a cystic fibrosis patient over a period of 3.5 years, from the onset of infection until death by necrotizing pneumonia (cepacia syndrome). We report the chemical structure of the LPS molecule of various sequential isolates and the identification of a novel hybrid O-antigen (OAg) biosynthetic cluster. The OAg repeating unit of the LPS from IST439, the initial isolate, is a  $[\rightarrow 2)\text{-}\beta\text{-D-Ribf-(1}\rightarrow 4)\text{-}\alpha\text{-D-GalpNAc-(1}\rightarrow ]$  disaccharide, which was not previously described in *B. cenocepacia*. The IST439 OAg biosynthetic gene cluster contains 7 of 23 genes that are closely homologous to genes found in *B. multivorans*, another member of the *Burkholderia cepacia* complex. None of the subsequent isolates expressed OAg. Genomic sequencing of these isolates enabled the identification of mutations within the OAg cluster, but none of these mutations could be associated with the loss of OAg. This study provides support to the notion that OAg LPS modifications are an important factor in the adaptation of *B. cenocepacia* to chronic infection and that the heterogeneity of OAgs relates to variation within the OAg gene cluster, indicating that the gene cluster might have been assembled through multiple horizontal transmission events.

### 3.2. Introduction

*Burkholderia cenocepacia* is a Gram-negative opportunistic human pathogen of the *Burkholderia cepacia* complex (Bcc), relevant in immunocompromised individuals and cystic fibrosis (CF) patients (Mahenthiralingam et al., 2002; Mahenthiralingam et al., 2005). *B. cenocepacia* lung infections in CF patients are associated with poor prognosis and increased risk of death (Mahenthiralingam et al., 2002; Mahenthiralingam et al., 2005; Drevinek and Mahenthiralingam, 2010). In comparison to *Pseudomonas aeruginosa*, the major CF pathogen, less is known about the molecular mechanisms involved in the adaptation of *B. cenocepacia* to the CF lung (Coutinho et al., 2011b; Maldonado et al., 2016).

The lipopolysaccharide (LPS) O-antigen (OAg) biosynthetic cluster is under strong selective pressure during chronic infection (reviewed in (Maldonado et al., 2016)). LPS is a major component of the Gram-negative bacterial outer membrane, which participates in host-bacterium interactions, such as adhesion, immune evasion, persistence, and antimicrobial resistance (Raetz and Whitfield, 2002; De Soyza et al., 2008; Valvano et al., 2011; Maldonado et al., 2016). LPS consists of a core oligosaccharide (core) that is covalently linked to a lipophilic glycan termed lipid A (Whitfield and Trent, 2014). In many bacteria, the LPS is extended by an OAg polysaccharide that is linked to the core. The lipid A is made of a  $\beta$ -(1 $\rightarrow$ 6)-glucosamine disaccharide acylated with primary fatty acids at positions 2 and 3 of both glucosamine residues, which are in turn phosphorylated at the 1- and 4'-positions. Secondary acyl chains can further substitute primary fatty acids at their hydroxyl positions (Whitfield and Trent, 2014). The core, subdivided into inner and outer core, comprises conserved monosaccharide residues, such as heptoses and 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo), which are typically unique to the LPS molecule (Whitfield and Trent, 2014). In *Burkholderia* species, one of the Kdo residues is modified to D-glycero-D-talo-oct-2-ulosonic acid (Ko) (Silipo et al., 2005). The OAg extends away from the outer membrane surface becoming exposed to the extracellular milieu; it is composed of linear or branched homo- or heteropolysaccharides of variable length, with subunits consisting of up to eight different sugars (Valvano et al., 2011).

Recent studies at genome (Spencer et al., 2003; Lieberman et al., 2011; Traverse et al., 2013), transcriptome (Mira et al., 2011), and proteome (Madeira et al., 2011; Madeira et al., 2013) levels indicate that the LPS and particularly the OAg undergo alterations during chronic infection, which could be attributed to bacterial adaptation to biofilm lifestyle (Lieberman et al., 2011; Traverse et al., 2013), immune evasion and selective pressure from antibiotics (reviewed in (Maldonado et al., 2016)). The airways of CF patients are colonized by polymicrobial communities that show variability in composition and diversity (Coburn et al., 2015). The CF microbiota is also exposed to a fluctuating environment and multiple selective pressures arising from nearly constant antibiotic treatment, oxygen limitation, and the persistent host inflammatory response (Döring et al., 2011; Cullen and McClean,

2015). Therefore, infecting bacteria give rise to genetically heterogeneous populations in which different phenotypes adaptive for growth and survival are selected (Harrison, 2007; Yang et al., 2011a). Understanding the adaptive and evolutionary mechanisms within chronic infections of the CF airway may help improve the management of these infections.

We investigated a collection of serial clonal isolates of *B. cenocepacia* obtained from a CF patient over a period of 3.5-years since the onset of infection until the patient's death with the cepacia syndrome (Cunha et al., 2003; Coutinho et al., 2011a). These isolates belong to the epidemic ET-12 lineage (represented by the prototypic strains J2315 and K56-2 (Mahenthalingam et al., 2000b; Holden et al., 2009). This patient was also co-infected with *Burkholderia multivorans*, another member of the Bcc, which appeared before the isolation of the early *B. cenocepacia* strain (isolate IST439); co-infection continued until near the patient's death (Cunha et al., 2003). The *B. cenocepacia* isolates from this patient were characterized by phenotypic (Coutinho et al., 2011a), transcriptomic (Mira et al., 2011), proteomic (Madeira et al., 2011; Madeira et al., 2013) and metabolic profiling (Moreira et al., 2016), as well as by comparative genomics (still unpublished). Isolates of particular interest are *B. cenocepacia* IST4113 (a highly antibiotic resistant variant retrieved after an exacerbation treated with intravenous therapy), IST4134 (obtained just before the patient's death), and IST4129 (a variant that exhibits attenuated virulence related with the loss of the third replicon (Moreira et al., 2017). The virulence potential of these isolates in non-mammalian host models (Moreira et al., 2017) and their ability to modulate dendritic cell function were also compared (Cabral et al., 2017).

In this study, we analyzed the chemical structure of the LPS molecule and the genetic organization of the predicted OAg biosynthetic cluster in these serial isolates. Our results reveal that the early *B. cenocepacia* IST439 isolate encodes a functional genetic cluster responsible for OAg biosynthesis, with a hybrid composition including genes highly homologous to *B. multivorans* genes. Further, this isolate produces a structurally different OAg from that previously reported in the ET12 lineage strains, while all subsequent *B. cenocepacia* isolates lost the ability to produce an OAg molecule.

### 3.3. Materials and Methods

#### 3.3.1. Bacterial strains and growth conditions

The isolates investigated in this study are indicated in Table 1 (Cunha et al., 2003; Coutinho et al., 2011a). *B. cenocepacia* and *B. multivorans* isolates were recovered from the sputum of a CF patient at the major Portuguese CF Center in Hospital de Santa Maria (HSM), from Centro Hospitalar Lisboa Norte EPE, Lisbon, Portugal, between 1998 and 2002. These isolates belong to the same clonal complex and ET-12 lineage (Moreira et al., 2017). Studies involving these isolates were approved by the ethics committee of the Hospital, and the anonymity of the patient was preserved. The genome of the soil

strain *B. multivorans* ATCC 17616 (NCBI nucleotide accession number: NC\_010804.1) (Vandamme et al., 1997) was used for comparisons. Bacterial cultures were stored at -80°C in 1:1 (v/v) glycerol. Bacteria were grown in LB Lennox (LB; Conda, Pronadisa) at 37°C with shaking at 250 rpm or in LB agar plates. *Escherichia coli* strains (Table 3.1) were grown in the same conditions. When needed, growth media were supplemented with antibiotics at the following concentrations: for *B. cenocepacia*, trimethoprim (TMP) at 100 µg/ml, and for *E. coli*, TMP at 50 µg/ml and kanamycin (KAN) at 40 µg/ml.

**Table 3.1** | Description of *Burkholderia* isolates and *E. coli* strains used in this study

Bacterial Isolate	Species	Description
IST419	<i>B. multivorans</i>	
IST439	<i>B. cenocepacia</i> IIIA	
IST4103	<i>B. cenocepacia</i> IIIA	
IST4110	<i>B. cenocepacia</i> IIIA	
IST4112	<i>B. cenocepacia</i> IIIA	Clinical isolates used in chapter II (Table 2.1) that were obtained from a chronically infected patient followed at the Cystic Fibrosis Center of Hospital Santa Maria, Lisbon, Portugal
IST4113	<i>B. cenocepacia</i> IIIA	
IST4116A	<i>B. cenocepacia</i> IIIA	
IST4116B	<i>B. cenocepacia</i> IIIA	
IST4131	<i>B. cenocepacia</i> IIIA	
IST4129	<i>B. cenocepacia</i> IIIA	
IST4130	<i>B. cenocepacia</i> IIIA	
IST4134	<i>B. cenocepacia</i> IIIA	
ATCC 17616	<i>B. multivorans</i>	Soil isolate (Berkeley, California, USA) (Vandamme et al., 1997; Nishiyama et al., 2010)
DH5- $\alpha$ (pRK2013)	<i>E. coli</i>	Helper strain for triparental conjugation (Figurski and Helinski, 1979)
ER2925 (dam-dcm-)	<i>E. coli</i>	Host for plasmids used to transform resilient <i>Burkholderia</i> strains (Craig et al., 1989)

### 3.3.2. LPS extraction, purification, and compositional analyses

Single-colony purified cells of the early isolate (IST439) and three late-stage clonal variants (IST4113, IST4129, and IST4134) were used for LPS purification and structural analysis. Bacteria were first grown overnight in LB broth until mid-exponential phase at 37°C with shaking (250 rpm). Cultures were diluted to an OD<sub>640nm</sub> of 0.2, and 100 µl of the cellular suspensions were plated onto LB agar plates and incubated for 24 h at 37°C. Bacteria were scraped from the agar surface, collected, autoclaved and lyophilized. LPS from bacterial dried cells was extracted by the hot phenol/water method (Westphal and Jann, 1965). The nature of the extracted material was checked by SDS-PAGE after gel silver

staining (Kittelberger and Hilbink, 1993). To remove contaminants the extracts were treated with RNase (Roth, Germany), DNase (Roth, Germany) and Proteinase K (Roth, Germany) at 37 °C and 56 °C, followed by dialysis against distilled water. The LPS was further purified by ultracentrifugation (4°C, 30,000 rpm, 24 h) and gel-filtration chromatography. The monosaccharide content of the sample was determined by analysis of the acetylated O-methyl glycoside derivatives after treatment with HCl/MeOH (1.25 M, 85 °C, 24 h) plus acetylation with acetic anhydride in pyridine (85 °C, 30 min) using gas-liquid chromatography mass spectrometry (GLC-MS). The sugar linkages were determined as described (Ciucanu and Kerek, 1984). The total fatty acid content was determined on intact LPS by treatment with 4 M HCl (100 °C, 4 h), followed by 5 M NaOH (100 °C, 30 min). After extraction in chloroform, fatty acids were methylated with diazomethane and analyzed by GLC-MS. The ester bound fatty acids were released by base-catalyzed hydrolysis with aqueous NaOH 0.5 M, MeOH (1:1, v/v, 85 °C, 2 h), and then the product was acidified, extracted in chloroform, methylated with diazomethane, and analyzed by GLC-MS. The absolute configuration of fatty acids was determined as previously described (Rietschel, 1976). Authentic 3-hydroxy fatty acids were used to assign the (R) configuration to all LPS/LOS acyl chains.

### 3.3.3. NMR spectroscopy

Prior to NMR spectroscopy an aliquot of the purified LPS (20 mg) was hydrolyzed with acetate buffer (100 °C, 2 h). After centrifugation, the supernatant, containing the saccharide fraction, was collected, lyophilized, and purified by size exclusion chromatography. 1D and 2D <sup>1</sup>H NMR spectra were recorded on a Bruker 600 DRX equipped with a cryo probe. The solvent employed was D<sub>2</sub>O and the temperature was 298 K and pD was 7. Spectra calibration was performed with internal acetone ( $\delta_{\text{H}}$  2.225 ppm,  $\delta_{\text{C}}$  31.45 ppm). The double-quantum filtered phase sensitive correlation spectroscopy (DQF-COSY) experiment was executed by using data sets of 4096 x 256 points. Total correlation spectroscopy (TOCSY) experiments were carried out with spinlock times of 100 ms, using data sets ( $t_1 \times t_2$ ) of 4096 x 256 points. Rotating frame Overhauser enhancement spectroscopy (ROESY) and Nuclear Overhauser enhancement spectroscopy (NOESY) experiments were performed by using data sets ( $t_1 \times t_2$ ) of 4096 x 256 points and by using mixing times between 100 and 400 ms. In all homonuclear experiments the data matrix was zero-filled in both dimensions to give a matrix of 4 K x 2 K points and was resolution enhanced in both dimensions by a cosine-bell function before Fourier Transformation. The determination of coupling constants was obtained by 2D phase sensitive DQF-COSY (Piantini et al., 1982; Rance et al., 1983). Heteronuclear single quantum coherence (<sup>1</sup>H, <sup>13</sup>C HSQC) and heteronuclear multiple bound correlation (<sup>1</sup>H, <sup>13</sup>C HMBC) experiments were recorded in <sup>1</sup>H-detection mode by single-quantum coherence with protein decoupling in the <sup>13</sup>C domain using data sets of 2048 x 256 points. <sup>1</sup>H, <sup>13</sup>C HSQC was performed using sensitivity improvement and in the phase-sensitive mode using Echo/Antiecho gradient selection, with multiplicity editing during selection step

(States et al., 1982). The  $^1\text{H}$ ,  $^{13}\text{C}$  HMBC experiment was optimized on long-range coupling constants with low-pass  $J$  filter to suppress one-bound connectivity, using gradient pulses for selection. A delay of 60 ms was employed for the evolution of long-range correlations. It was used a long-range coupling constant value of 6 Hz. The data matrix in both heteronuclear experiments was extended to  $2048 \times 1024$  points using forward linear prediction (Stern et al., 2002).

#### 3.3.4. MALDI mass spectrometry

MALDI-TOF mass spectra of the intact LPS were recorded in reflectron mode and negative ion polarity on a Perseptive (Framingham, MA, USA) Voyager STR equipped with delayed extraction technology. Ions formed by a pulsed UV laser beam (nitrogen laser,  $\lambda = 337$  nm) were accelerated by 24 kV. LPS/LOS preparation was executed as described before (Sturiale et al., 2005; Lombardi et al., 2013).

#### 3.3.5. Small-scale LPS extraction for SDS-polyacrylamide gel electrophoresis analysis

LPS was extracted as previously described (Marolda et al., 1990) with small modifications. Briefly, *Burkholderia* isolates were harvested from overnight liquid cultures by centrifugation for 1 min after  $\text{OD}_{640\text{nm}}$  adjustment to 2.0 in 1 ml of PBS, suspended in 150  $\mu\text{l}$  of lysis buffer containing 2% SDS, 4% 2- $\beta$ -mercaptoethanol, and 500 mM Tris-HCl (pH 6.8), and boiled for 10 min. Proteinase K (20 mg/ml) was added, and the sample was incubated at 60°C for 2 h. Finally, samples were mixed with the tracking dye solution (125 mM Tris-HCl [pH 6.8], 2% SDS, 20% [v/v] glycerol, 0.002% bromophenol blue, and 10% mercaptoethanol) and boiled for 5 min before the gels were loaded. LPS samples were resolved by electrophoresis (at 150V for about 1h40m) in 14% polyacrylamide gels with a Tricine-SDS system followed by silver staining as previous described to visualize the banding patterns of the O-antigen (OAg) (Marolda et al., 2006).

#### 3.3.6. Genomic DNA sequencing, assembly, and annotation

Bacterial cultures were prepared by suspending isolated colonies from LB agar plates in 3 mL LB broth, followed by overnight growth at 37°C with shaking at 250 rpm. Genomic DNA was extracted and purified using a DNeasy Blood and Tissue Kit (Qiagen, Germany) according to manufacturer instructions. DNA concentration and purity were assessed using a Nanodrop ND-1000 spectrophotometer.

*B. multivorans* IST419 was sequenced on an Illumina NextSeq500 at the University of Pittsburgh, PA, USA. Raw fastq paired-end files were processed for removal of Illumina adapters, trimming, and quality-based filtering using Trimmomatic (Bolger et al., 2014). *De novo* assembly was performed

using Velvet with automated optimization of assembly parameters (Zerbino and Birney, 2008), using the two sets of pair-ended reads, and annotated using the prokaryotic genome annotation tool Prokka (Seemann, 2014). For the identification of mutational events in the OAg genetic clusters, read data sets were mapped against the relevant reference genomes (*B. cenocepacia* or *B. multivorans*) using BWA-MEM of BWA (Li and Durbin, 2010) and Bowtie 2 (Langmead and Salzberg, 2012), and visualized in IGV (Robinson et al., 2011).

DNA sequence reads for all isolates are available in the EMBL's European Nucleotide Archive (ENA) under accession number PRJEB20052 (<http://www.ebi.ac.uk/ena/data/view/PRJEB20052>). The sequence of the OAg cluster of *B. cenocepacia* IST439 is given as supplemental material (Table S3.1).

### 3.3.7. *In silico* characterization of O-antigen biosynthetic gene clusters

Genome sequences of 11 *B. cenocepacia* sequential isolates and the sequence of *B. multivorans* IST419 obtained from a CF patient (Coutinho et al., 2011a) were used to investigate the OAg biosynthetic cluster (the detailed analysis of the remaining genomic differences will be reported elsewhere). The gene composition of the OAg biosynthetic gene cluster was determined by examining the region flanked by genes *apaH* and *ureG* (Ortega et al., 2005). These flanking genes were detected by BLASTN (Altschul et al., 1990) using as queries the gene sequences from *B. cenocepacia* J2315 against each of the corresponding nucleotide databases created from the available sequences under study. The assembled scaffolds of all studied isolates were cropped to obtain only the OAg cluster (*apaH-ureG*) for further examination. Computer-assisted analysis for these clusters was performed in Artemis (Carver et al., 2008) and by BLASTP (Altschul et al., 1990) to reveal the open reading frames. Then, the clusters were visualized and compared with the OAg cluster of *B. cenocepacia* K56-2 (NCBI nucleotide accession number: NZ\_ALJA00000000.2) by Artemis and ACT (Carver et al., 2008), and gene function was assigned based on the analysis of predicted ORF homologies.

### 3.3.8. Molecular cloning of *wbiI* and *bmul\_2510* genes in *B. cenocepacia* clonal variants lacking O-antigen

Genes *wbiI* and *bmul\_2510* were PCR amplified from *B. cenocepacia* IST439 chromosomal DNA using a Hot Start High Fidelity polymerase (Qiagen) and cloned into pSCrhaB2 (Cardona and Valvano, 2005; Cardona et al., 2006). Primers were designed based on the genome sequence of IST439 and included specific restriction enzyme sites and a sequence encoding the FLAG epitope tag (Table S3.2). Primers WbiI-flag-NdeI and WbiI\_439\_XbaI were used to clone *wbiI* fused to an N-terminal FLAG tag; primers wbiI\_439\_NdeI and WbiI-flag-XbaI were used to clone *wbiI* fused to a C-terminal FLAG tag; primers Bmul\_2510-439-NdeI and Bmul\_2510-flag-XbaI were used to clone *bmul\_2510* fused to a C-terminal FLAG tag; primers P1, P2, P3 and P4 were used to clone both *wbiI* and *bmul\_2510* with

C-terminal FLAGs using the Gibson Assembly strategy (New England BioLabs). The amplification conditions were 5 min at 95°C, 30 cycles of 95°C for 30 s, 60°C for 1 min, and 72°C for 2 min, and a final extension of 10 min at 72°C. The resulting products were digested with *NdeI* and *XbaI*, ligated to pSCrhaB2 and introduced into *E. coli* ER2925 (New England BioLabs) (Craig et al., 1989) by transformation. Transformants carrying recombinant plasmids with the DNA insert were screened by colony PCR with primers 824 and pSC rev, which anneal to vector sequences flanking the cloning sites. All constructs were confirmed by DNA sequencing.

### 3.3.9. Plasmid conjugation into *B. cenocepacia*

Plasmids were mobilized by triparental mating (Craig et al., 1989) into all *B. cenocepacia* clonal variants lacking the OAg that possess single nucleotide polymorphisms (SNP) in *Bmul\_2510* and/or *wbiI* using *E. coli* DH5- $\alpha$  carrying the pRK2013 helper plasmid (Figurski and Helinski, 1979). Exconjugants were selected on LB agar plates supplemented with 100  $\mu$ g/ml of TMP and 200  $\mu$ g/ml of ampicillin (AMP).

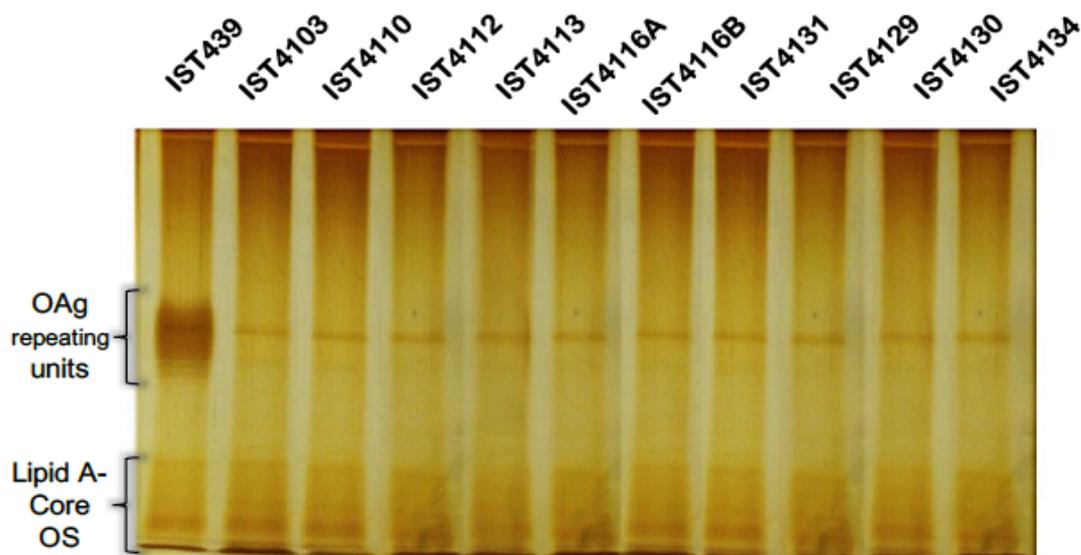
### 3.3.10. Western blot

Protein expression in *B. cenocepacia* clonal variants expressing *wbiI* or *bmul\_2510* from IST439 was confirmed by western blot. Briefly, electrophoresed whole cell lysate samples were transferred to nitrocellulose membranes. The membranes were incubated with a 1:10,000 dilution of anti-FLAG mouse monoclonal antibody followed by incubation with a 1:5,000 dilution of a goat anti-mouse IgG monoclonal antibody conjugated to Alexa Fluor 680 (Life Technologies). Images were acquired using an Odyssey infrared imaging system (LI-COR Biosciences).

## 3.4. Results

### 3.4.1. Only the first of the 11 sequential clonal variants of *B. cenocepacia* produce OAg

Chronically infecting bacteria in CF lungs often display changes in their LPS OAg (Maldonado et al., 2016). Therefore, we investigated the electrophoretic profiles of LPS extracted from the series of 11 sequential clonal variants of *B. cenocepacia*. SDS-PAGE followed by silver staining revealed that OAg was only present in the early isolate (IST439), while all other variants produce only lipid A-core devoid of OAg (Fig. 3.1). The banding pattern of the OAg in IST439 was distinct from that in the strain K56-2 (Fig. S3.1). Therefore, although both IST439 and K56-2 belong to the same *B. cenocepacia* clonal group (<http://pubmlst.org/bcc>) they produce seemingly different OAg molecules.



**Figure 3.1** | Electrophoretic profile of the LPS extracted from 11 clonal *Burkholderia cenocepacia* (*recA* lineage IIIA) isolates retrieved from a chronically infected CF patient (in chrononological order). LPS loading was standardized based on culture optical density. Samples were run in 14% polyacrylamide gels (at 150V for about 1h40m) in a Tricine-SDS system and developed by silver staining. The band that appears present in all the gels in the zone of the O-antigen repeating units corresponds to residual proteinase K.

### 3.4.2. Chemical structure of the LPS of *B. cenocepacia* isolates with or without O-antigen

We investigated the chemical structure of the LPS produced by *B. cenocepacia* IST439 and also the structure of 3 randomly selected subsequent isolates that lacked OAg. Monosaccharide compositional analysis of the LPS from variants IST4113, IST4129 and IST4134 (Table 3.1) gave identical results (Table S3.3), which also agreed with previously reported data for *B. cenocepacia* J2315 (Silipo et al., 2007). In contrast, the IST439 LPS had D-ribose and D-galactosamine residues as the main constituents. Results of linkage analyses of LPS from variants IST4113, IST4129 and IST4134 was also identical to that of J2315 (Silipo et al., 2007), but data obtained from IST439 LPS indicated also the presence of 2-substituted ribofuranose and 4-substituted galactosamine. The fatty acids content of the LPS in the four IST strains also matched the archetypal *Burkholderia* lipid A components, namely (*R*)-3-hydroxyhexadecanoic acid (C16:0 (3-OH)) in amide linkage, and (*R*)-3-hydroxytetradecanoic (C14:0 (3-OH)) and tetradecanoic acid (C14:0) in ester linkage (Silipo et al., 2007).

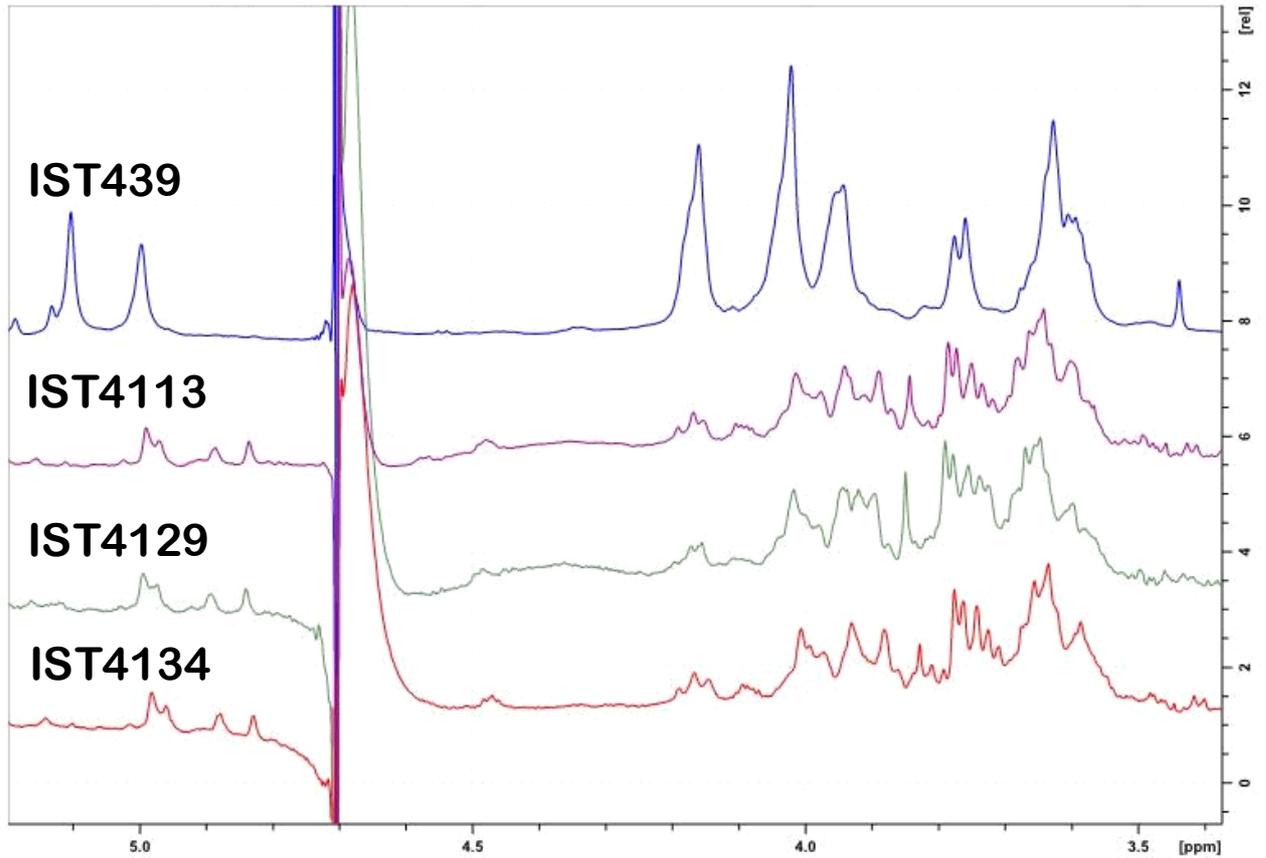
Mild acidic conditions were used to characterize the monosaccharide sequence of the carbohydrate fraction of the purified LPS samples. This process cleaves the labile glycosidic linkage between the saccharide and the lipid A moieties. The carbohydrate fractions were purified by gel-permeation

**Table 3.2** | Chemical shift  $\delta$  ( $^1\text{H}/^{13}\text{C}$ ) of the O-chain moiety from *B. cenocepacia* IST439

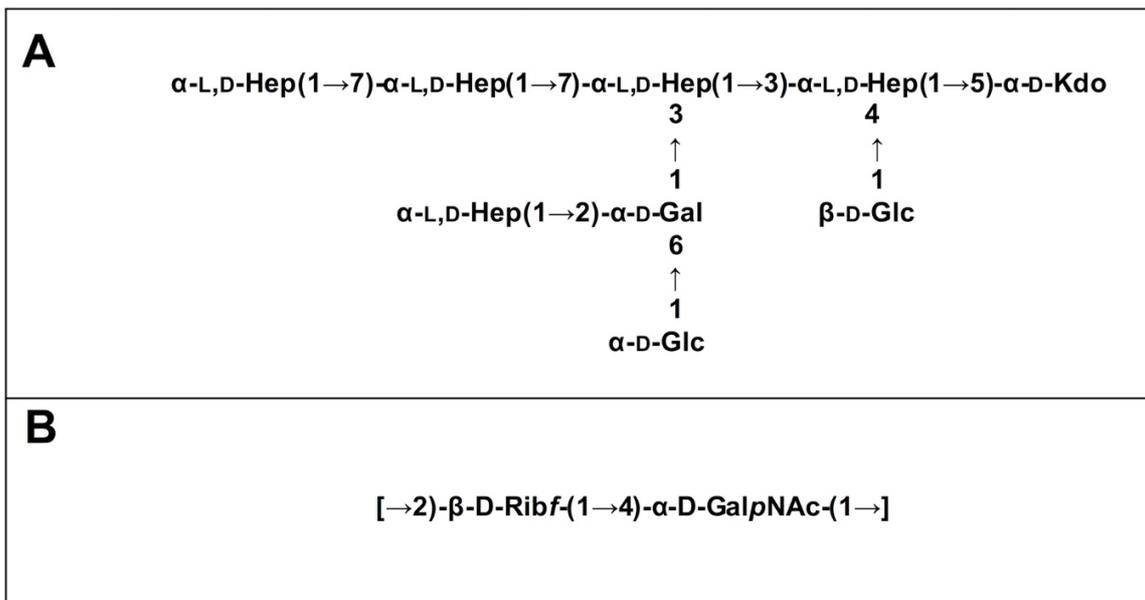
Chemical shift $\delta$ ( $^1\text{H}/^{13}\text{C}$ )						
Unit	1	2	3	4	5	6
A	5.09	4.15	4.03	3.95	3.77-3.59	---
2- $\beta$ -Rib	106.6	78.7	67.2	82.3	62.5	----
B	4.99	4.02	3.93	3.61	3.63	4.03
4- $\alpha$ -GalNAc	95.5	50.0	76.2	78.9	60.5	70.8

chromatography and analyzed by 1D and 2D NMR spectroscopy. The overlapped  $^1\text{H}$ -NMR spectra relative to all the isolates carbohydrate fractions were reported in Fig. 3.2. It was immediately evident that the  $^1\text{H}$  NMR spectra of IST4113, IST4129 and IST4134 were identical, highlighting the presence of eight anomeric signals relative to the sugar units composing the core moieties. In contrast, the IST439  $^1\text{H}$  NMR spectrum showed two main signals in the anomeric region (Fig. 3.2, Table 3.2), attributed to the OAg domain. The complete LPS core oligosaccharide structures of isolates IST4113, IST4129, IST4134 and IST439 was assigned by the combination of data obtained from chemical analyses and 2D NMR spectroscopy using DQF-COSY, TOCSY, NOESY, ROESY,  $^1\text{H},^{13}\text{C}$ -HSQC and  $^1\text{H},^{13}\text{C}$ -HMBC. In all of these cases, the core was structurally identical to that previously described for *B. cenocepacia* J2315 (Silipo et al., 2007) except that the IST samples lacked the rhamnose-*N*-acetyl quinovosamine (Rha-QuiNAc) disaccharide linked to the outer core. In J2315, this disaccharide corresponds to a truncated component of the OAg (Ortega et al., 2009). Fig. 3.3 – panel A shows the core structure, elucidated by NMR spectroscopy, of isolates IST4113, IST4129, IST4134.

As stated above, NMR spectroscopy of the carbohydrate fraction from isolate IST439 identified two spin systems in the  $^1\text{H}$ -NMR spectrum, belonging to the OAg, designed as **A** (H-1 at  $\delta = 5.09$  ppm) and **B** ( $\delta = 4.99$  ppm) (Fig. 3.2). Spin system **A** was identified as ribofuranose due to the correlations observed in the DQF-COSY and TOCSY spectra from the H-1 **A** signal up to the diastereotopic methylene signal at position H-5. The furanose form was established due to the presence of the typical low field shifted ring carbon atom signals (Table 3.2 and Fig. 3.4). The anomeric configuration was attributed on the basis of the low  $^3J_{\text{H-1,H-2}}$  value attained from the DQF-COSY spectrum, indicative of a  $\beta$ -configuration in aldofuranose rings (less than 2 Hz), and by the chemical shift of C-4 (around 83.0 ppm in case of  $\beta$ -configuration) (Ahrazem et al., 2002; Ahrazem et al., 2007). The downfield displacement of C-2 signal (Fig. 3.4) indicated glycosylation at this position. Spin system **B** (H-1 at  $\delta = 4.99$  ppm) was identified as an  $\alpha$ -galacto- configured monosaccharide, as attested by the low  $^3J_{\text{H3,H4}}$  and  $^3J_{\text{H4,H5}}$  values (3 Hz and 1 Hz, respectively), and the TOCSY correlations between the anomeric



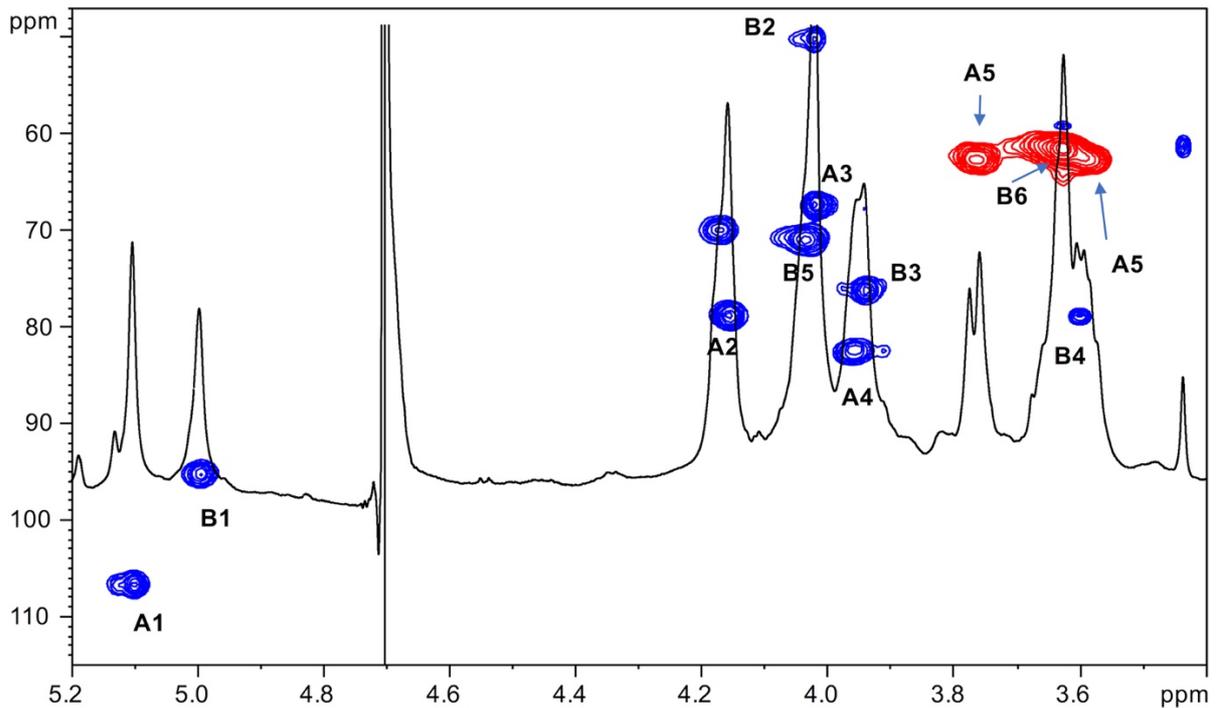
**Figure 3.2** |  $^1\text{H}$  NMR spectrum of products obtained from IST439 and other three subsequent variants after mild acid hydrolysis.



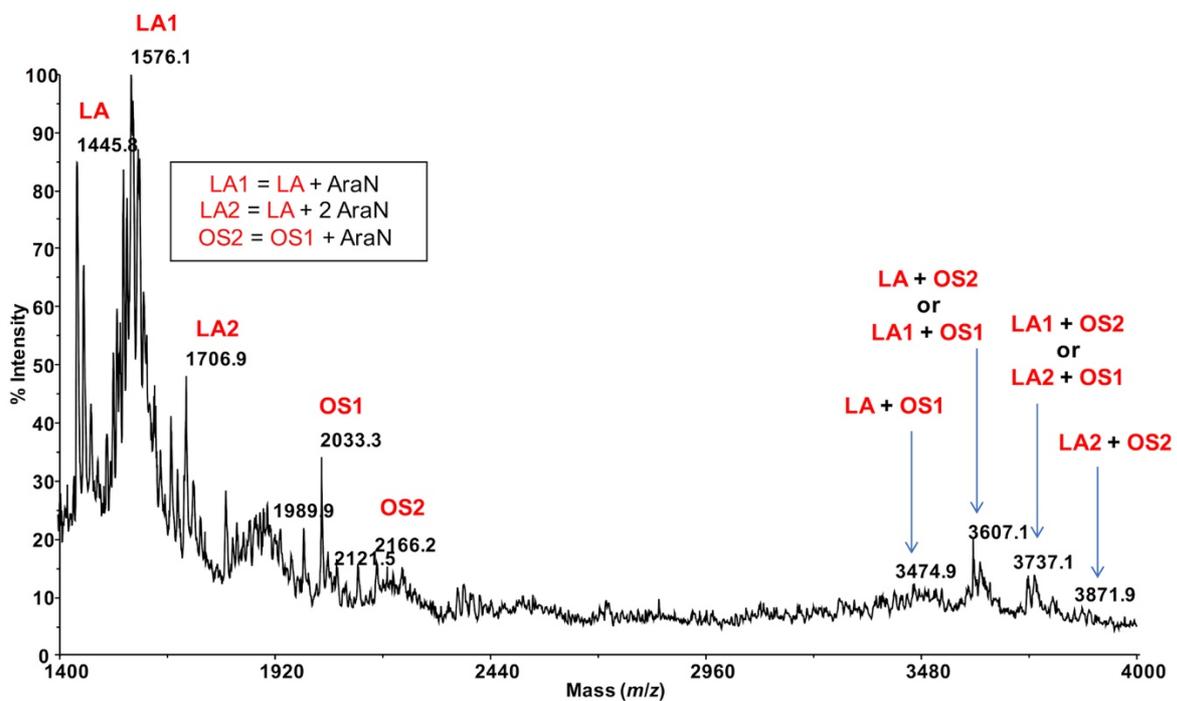
**Figure 3.3** | LPS moieties structures elucidated by NMR spectroscopy. Core structure of isolates IST439, IST4113, IST4129, and IST4134 (A) and OAg repeating disaccharide of isolate IST439 (B).

signal and the other ring protons up to the H-4 resonance. Moreover, the anomeric proton and carbon chemical shifts ( $\delta = 4.99$  ppm and 95.5 ppm), the *intra*-residue NOE correlation of H-1 **B** with H-2 **B**, and the  $^3J_{H1,H2}$  were all in agreement with an  $\alpha$ -anomeric configuration and a  $^4C_1$  ring conformation. The detection of a signal for C-2 at 50.0 ppm (Fig. 3.4) indicated that C-2 was a nitrogen-bearing carbon atom. Moreover, the downfield shift of proton resonance of H-2 **B** was indicative of *N*-acetylation at this position, as also corroborated by the NOE contact of H-2 **B** with the methyl protons of the acetyl group resonating at  $\delta = 1.92$  ppm. Therefore, residue **B** was identified as a  $\alpha$ -GalNAc. The downfield shift of its C-4 ( $\delta = 78.9$  ppm) was indicative of glycosylation at this position. ROESY and NOESY spectra (not shown) allowed to detect the dipolar correlations necessary to assign the primary sequence of the IST439 LPS OAg repeating unit. The anomeric proton of GalNAc **B** gave a NOE correlation with H-2 signal of ribose **A**, whose anomeric signal showed a NOE correlation with the H-4 of residue **B**. These results, confirmed by the long scalar range connectivity derived from the  $^1H, ^{13}C$ -HMBC spectrum (not shown), demonstrated that the IST439 OAg structure consisted of a repeating disaccharide of  $\beta$ -D-ribofuranose ( $\beta$ -D-Ribf) and  $\alpha$ -D-*N*-acetylgalactosamine ( $\alpha$ -D-GalpNAc; Fig. 3.3 – panel B). This structure is similar to that previously identified in *B. cepacia* (formerly *Pseudomonas cepacia*) serogroups O3 and O5 LPS OAg structures (Cox and Wilkinson, 1989).

Negative-ion MALDI mass spectrometry (MS) of the intact IST439 LPS showed ion peaks corresponding to lipid A and core species, which arose from the cleavage of the labile glycosidic linkage between Kdo and lipid A. Ion peaks corresponding to lipid A species were assigned as previously described (Silipo et al., 2007). The ion peak **OS1** ( $m/z = 2033.3$ ) agreed with an oligosaccharide composed of five heptoses, three hexoses, one Kdo, one *D-glycero-D-talo*-oct-2-ulosonic acid (Ko) and one 4-L-amino-4-deoxyarabinose (Ara4N), while ion peak **OS2** ( $m/z = 2166.2$ ) corresponded to the same oligosaccharide carrying an additional Ara4N residue (Fig. 3.5). Thus, the ion peaks derived from the IST439 core agreed with the structure previously elucidated (Silipo et al., 2007). The mass range 3000-4000 Da (Fig. 3.5), containing peaks relative to either the oligosaccharide and the lipid A species, also confirmed previous studies (Silipo et al., 2007). Likewise, negative-ion MALDI mass spectra executed on the intact LPS isolated from the IST4113, IST4129 and IST4134 variants were identical each other and to the previously described mass spectra performed on intact *B. cenocepacia* ET-12 strain J2315 LPS (Ortega et al., 2009). The MALDI mass spectrum of the intact IST4113 LPS, presented in Supplemental Fig. S3.2, showed at mass range  $m/z$  1300-2500 Da, OS species composed of four to six heptoses and three to five hexoses, one Kdo, one Ko, and with or without Ara4N. In particular, the ion peak termed **OS** ( $m/z = 2032.5$ , Supplemental Fig. S3.2) was indicative of the core previously elucidated (Silipo et al., 2007). The LPS molecular ions in the mass range 3000-4200 Da (Supplemental Fig. S3.2) also confirmed previously reported data and agreed with chemical and NMR spectroscopy analyses.



**Figure 3.4** |  $^1\text{H}$ ,  $^{13}\text{C}$ -HSQC spectrum of the O-chain moiety from IST439 strain. The heteronuclear correlations are indicated.



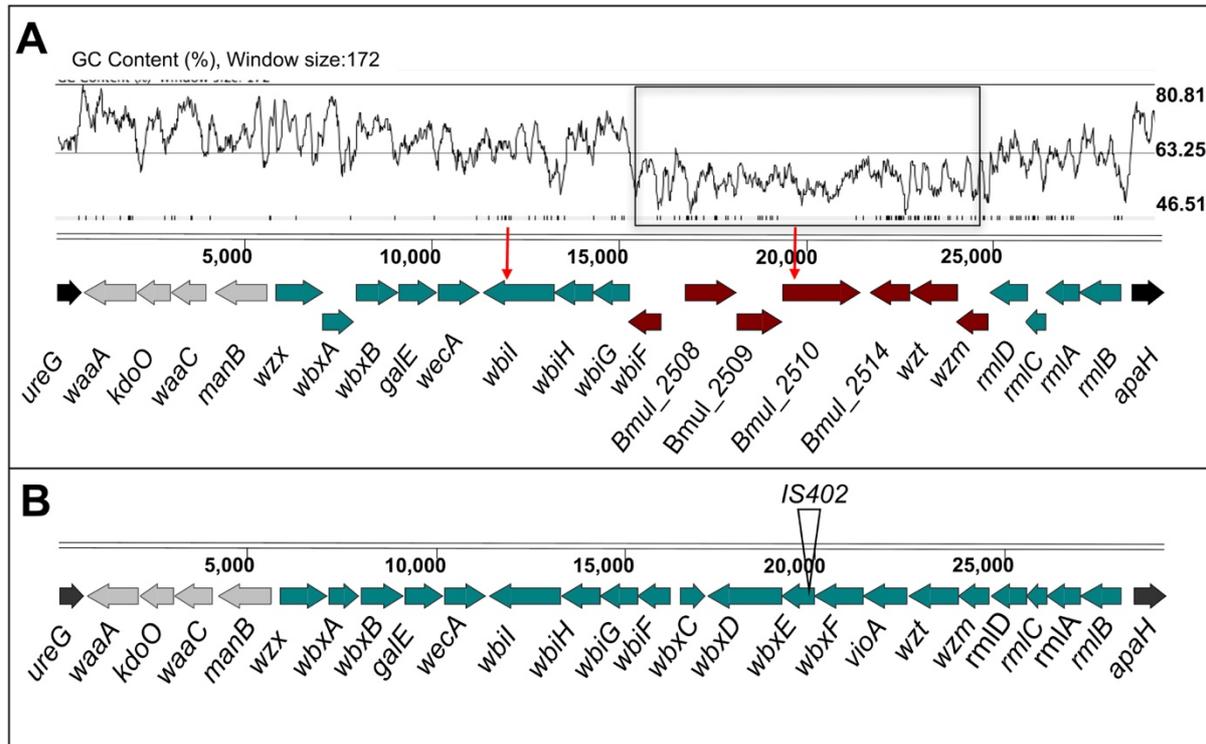
**Figure 3.5** | MALDI mass spectrum of the intact LPS from strain IST439 (mass range 1400-4000). Lipid A-core oligosaccharide molecular ions and their ion fragments, attributable to the core OS and to the reported lipid A structure(s), are indicated.

### 3.4.3. Genomic sequence of the *B. cenocepacia* clonal isolates reveals of a novel hybrid O-antigen biosynthetic cluster

The 11 *B. cenocepacia* sequential isolates studied here belong to the *recA*-lineage IIIA/ET-12, as the prototypic epidemic strains K56-2 and J2315 (Coutinho et al., 2011a; Mira et al., 2011). The genome sequences of the 11 isolates were used to investigate the genetic organization of the OAg biosynthetic gene cluster in comparison to that of *B. cenocepacia* K56-2 (Ortega et al., 2005) (Fig. 3.6). The *cenocepacia* K56-2 OAg biosynthetic cluster is on chromosome 1 and spans approximately 29 kb comprising 24 genes (Fig. 3.6) flanked by *ureG* and *apaH* (Ortega et al., 2005). The *B. cenocepacia* IST439 OAg biosynthetic cluster is also located in chromosome 1, flanked by *ureG* and *apaH*, and spans a region of 29.5 kb. The cluster consists of 23 open reading frames (ORFs) (Fig. 3.6), most of them functionally assigned based on the bioinformatic analysis of predicted polypeptides. Like *B. cenocepacia* K56-2, the IST439 OAg cluster does not have a transposon insertion element inserted within the cluster (Fig. S3.3) (Ortega et al., 2005). Sixteen of the 23 ORFs in IST439 are homologs to *B. cenocepacia* K56-2/J2315 genes, including *waaA*, *wbxY* (*kdoO*), *waaC*, *manB*, *wzx*, *wbxA*, *wxB*, *galE*, *wecA*, *wbiI*, *wbiH*, *wbiG*, *rmlD*, *rmlC*, *rmlA* and *rmlB*. However, the remaining 7 genes (*wbiF*, *bmul\_2508*, *bmul\_2509*, *bmul\_2510*, *bmul\_2514*, *wzt* and *wzm*) are closer in terms of sequence similarity to ORFs in *B. multivorans* ATCC 17616 (environmental isolate with a complete genome deposited at NCBI and Bcc databases (Nishiyama et al., 2010)), ranging from 79 to 96 % of amino acid identity, versus no homology (at an expect value cut-off of 1E-4) with any known *B. cenocepacia* strains for *bmul\_2508*, *bmul\_2509*, *bmul\_2510*, and amino acid identity of 31 to 74 % with the closely related ET-12 *B. cenocepacia* J2315 for the remaining 4 genes. This observation suggested a hybrid origin of this OAg cluster in IST439 (Figs. S3.2 and S3.3 – panel B). Analysis of the GC content across the IST439 OAg cluster shows a sharp drop in the region containing the seven genes related to *B. multivorans* (average content 56.1% in *B. cenocepacia* IST439 and 55.3% in *B. multivorans* ATCC 17616; Figs. 3.6 – panel A and 3.8 – panel B, respectively) in comparison to the rest of the cluster (average content 63.1%, and 66.6% when subtracting the seven genes) and the IST439 complete genome (average content 67.4%). This variation in GC content is usually found in genomic islands of prokaryotic genomes and is a hallmark of regions undergoing frequent recombination, as is the case of the LPS synthetic locus (Zhang et al., 2014; Maldonado et al., 2016).

The *waaA*, *kdoO*, *waaC* and *manB* genes (represented in gray in Fig. 3.6) encode proteins involved in lipid A-core biosynthesis (Ortega et al., 2005). In particular, the predicted amino acid sequence of WaaA and WaaC are highly identical to 3-deoxy-D-manno-octulosonic acid and heptosyltransferase I, respectively, and they are therefore involved in inner core LPS synthesis. *wbxY* (BCAL3311), previously annotated as a gene encoding a conserved protein of unknown function (Ortega et al., 2005), encodes a Kdo dioxygenase that is responsible for the conversion of the distal Kdo residue of the

*Burkholderia* lipid A-core into Ko (Chung and Raetz, 2011; Chung et al., 2014; Fathy Mohamed et al., 2017), and it was renamed as *kdoO*. Another conserved functional feature in this OAg cluster is the predicted transmembrane protein encoded by *wbiH*, which includes functional features of the *wecA*

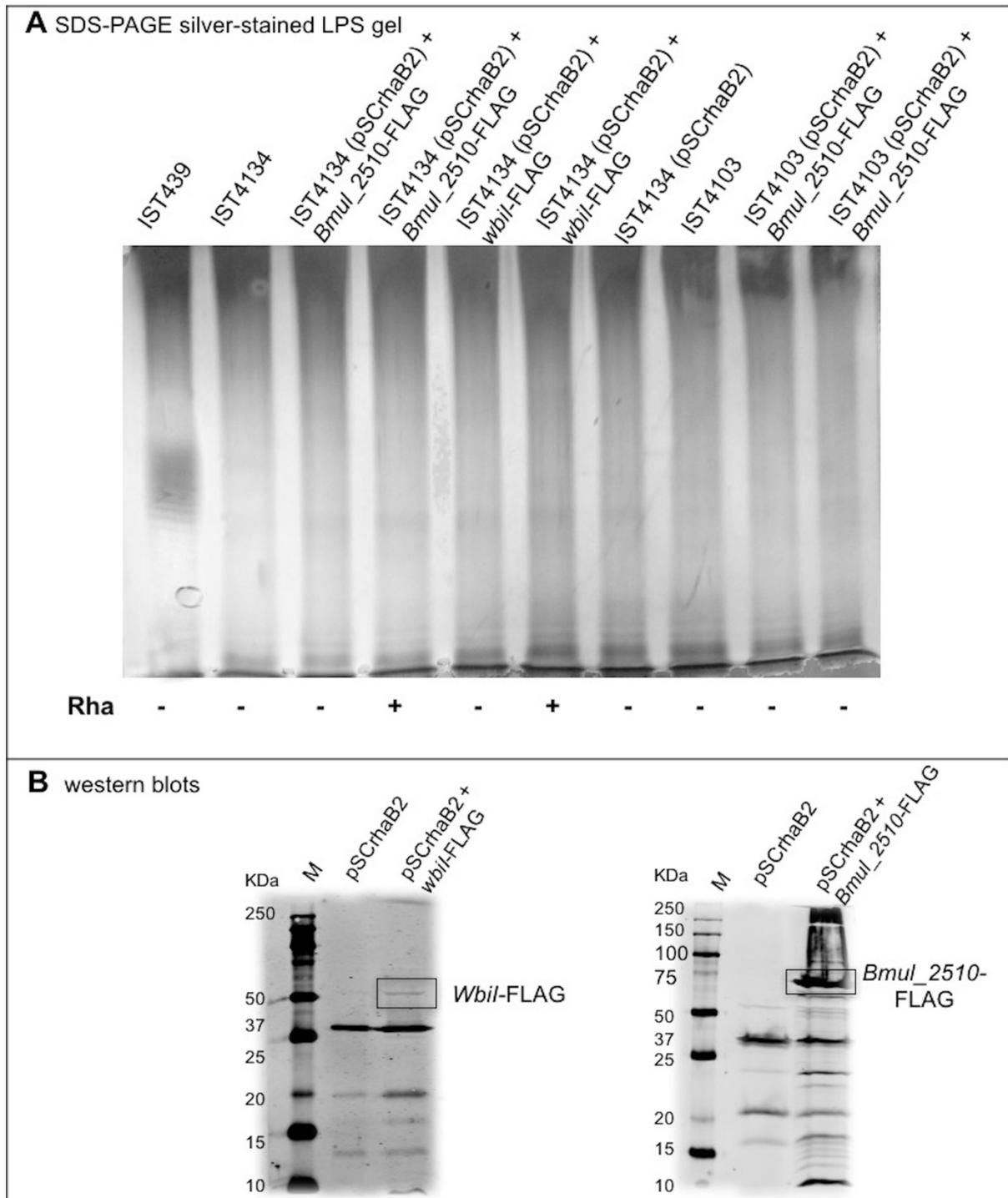


**Figure 3.6** | Genetic organization of gene clusters for core-lipid A and O-antigen biosynthesis in *B. cenocepacia* IST439 (A) and the reference strains K56-2 and J2315 (B). The flanking genes *ureG* and *apaH* are indicated in black and the four genes represented in gray encode proteins putatively involved in lipid A-core biosynthesis. The genes in dark-red correspond to genes present in IST439 without a counterpart in the reference strain J2315 (Figure S3.3), but with a degree of homology to *B. multivorans* ATCC 17616 (see also Figure S3.4 – panel B). The red-vertical arrows above the genes represent the two non-synonymous point mutations in genes *bmul\_2510* and *wbiI* of the 10 sequential isolates (comparative genomic analysis to be described elsewhere). A GC content plot is also represented for IST439 (drawn using Artemis (Carver et al., 2008)) above the display line of the sequence, where the genes annotated as *B. multivorans* are highlighted in a black rectangle. *waaA*, 3-deoxy-D-manno-octulosonic acid transferase; *kdoO*, Kdo dioxygenase; *waaC*, heptosyltransferase I; *manB*, phosphomannomutase; *wzx*, O-antigen exporter; *wbxA*, glycosyltransferase; *wbxB*, glycosyltransferase; *galE*, UDP-glucose epimerase; *wecA*, UDP-N-acetylglucosamine 1-P transferase; *wbiI*, nucleotide sugar epimerase-dehydratase; *wbiH*, UDP-N-acetylglucosamine 1-P transferase; *wbiG*, nucleotide sugar epimerase-dehydratase; *wbiF*, glycosyltransferase; *bmul\_2508*, conserved hypothetical protein; *Bmul\_2509*, group 1 glycosyl transferase; *Bmul\_2510*, conserved hypothetical protein; *Bmul\_2514*, type 11 methyltransferase; *wzt*, ABC transporter ATP-binding protein; *wzm*, ABC transporter membrane permease; *rmlDCAB*, dTDP-rhamnose biosynthesis; *wbxC*, acetyltransferase; *wbxD*, glycosyltransferase; *wbxE*, glycosyltransferase; *vioA*, nucleotide sugar aminotransferase.

family (Valvano, 2003; Ortega et al., 2005). The *wecA* family proteins and other proteins, encoded by *galE*, *wbiI*, and *wbiG*, are probably involved in the transfer of *N*-acetylhexosamines to the undecaprenol-phosphate intermediate to initiate the assembly of the OAg subunits (Raetz and Whitfield, 2002; Valvano, 2003; 2015a), while genes *wbxA*, *wxB*, *wbiF* and *bmul\_2509* encode glycosyltransferases likely involved in OAg elongation. Genes *bmul\_2508* and *bmul\_2510* encode proteins of unknown function, while *bmul\_2514* encodes a type-11 methyltransferase, possibly involved in the termination of OAg assembly. Genes *wzm* and *wzt* encode a two-component ABC transporter involved in OAg export across the cytoplasmic membrane (ABC-transporter-dependent pathway) after OAg polymerization on the cytoplasmic side (Raetz and Whitfield, 2002; Valvano, 2003; 2015a). There is an alternative pathway of OAg export, the *wzy*-dependent pathway based on the concerted action of proteins Wzx and Wzy (Raetz and Whitfield, 2002; Valvano, 2003; 2015a), but the OAg biosynthetic cluster of IST439 does not encode a homolog of *wzy*. This gene is also lacking in *B. cenocepacia* K56-2 (Ortega et al., 2005). Together, these observations indicate that the IST439 OAg, like in strain K56-2, is exported by an ABC transporter dependent pathway.

#### 3.4.4. Mutations in *wbiI* and *Bmul\_2510* are not involved in loss of OAg

Whole-genome comparison of the 10 *B. cenocepacia* clonal variants against IST439 identified only three mutational events spanning the entire OAg cluster, corresponding to three non-synonymous SNPs in coding regions *IST439\_01746* (homolog to *B. cenocepacia* *bmul\_2510*) and *wbiI* (Fig. 3.6 – panel A, red arrows). The point mutation in *bmul\_2510* is conserved in all sequential isolates and results in the replacement of a threonine for a proline residue (T116P) in the corresponding Bmul\_2510 polypeptide. In contrast, *wbiI* has two different point mutations resulting in L493P replacement in three isolates (IST4112, IST4113, and IST4116B) and E489K in other six isolates (IST4110, IST4116A, IST4131, IST4129, IST4130, and IST4134). No other mutations were identified within the OAg biosynthetic locus. Further, no mutations were found in the OAg ligase *waaL* and in the *wabO-dnaE* cluster that contains the remainder of the core-oligosaccharide biosynthesis genes (Ortega et al., 2009). However, eleven conserved SNPs were found outside the OAg cluster as well as one mutational event – as a deletion - in the late variants, compared with IST439, but none of them appear to be directly related to the loss of the OAg. Since the *bmul\_2510* mutation is in all isolates after IST439, and *wbiI* mutations are present in 9 isolates (except for IST4103), we performed complementation experiments in IST4103 with the *bmul\_2510* homolog gene cloned from IST439 as an attempt to reconstitute OAg biosynthesis (Cardona and Valvano, 2005). The *bmul\_2510* coding sequence including a FLAG tag epitope was placed under the control of a rhamnose-inducible promoter and the complemented strain was grown at various rhamnose concentrations. However, we could not detect OAg production in IST4103 (Fig. 3.7). We also attempted complementation experiments for *bmul\_2510* and *wbiI* singly or in combination using strain IST4134, but none of these experiments resulted in restoration of OAg



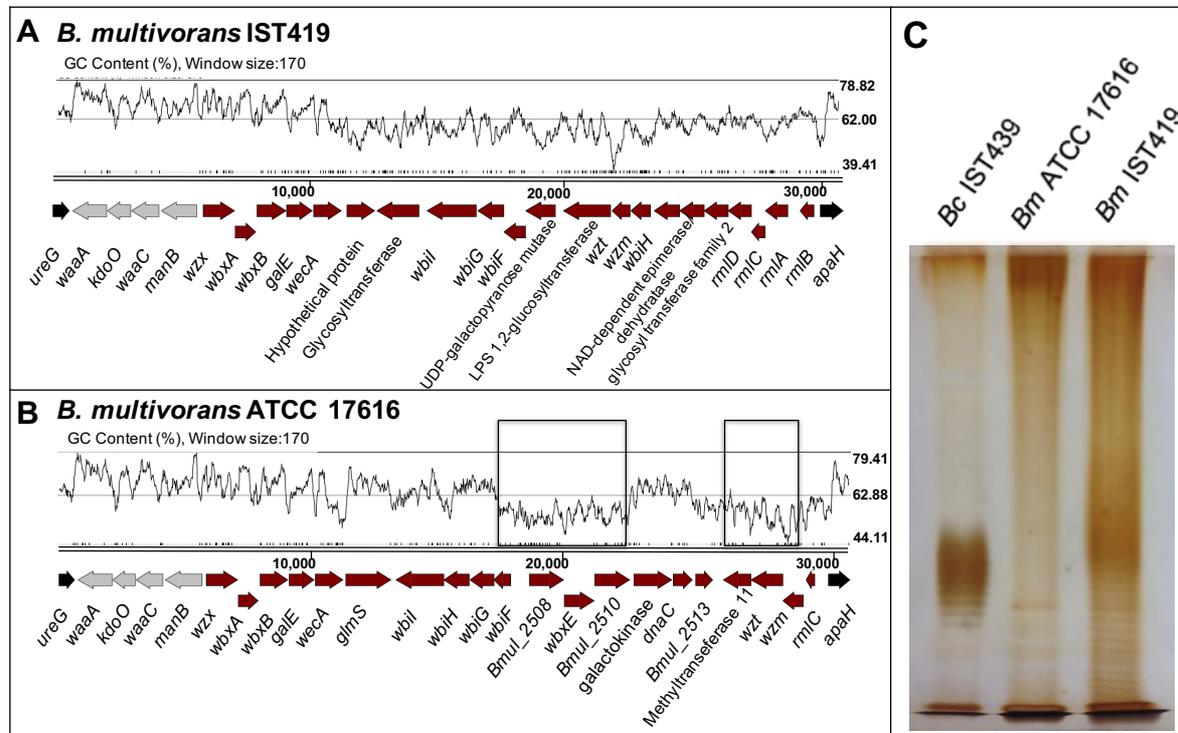
**Figure 3.7** | Characterization of the OAg unit after cloning and introduction of *B. cenocepacia* IST439 *wbiI* and *Bmul\_2510* genes in all variants lacking OAg production to complement the corresponding mutated genotypes. (A) LPS electrophoretic profiles for IST439, IST4134, and IST4103 showing the presence/absence of OAg units, and OAg absence in all transconjugants selected; (+) denotes 1% (w/v) rhamnose and (-) denotes no rhamnose added to the medium. (B) Western blots showing bands at around 60 KDa (left) and 75 KDa (right), which indicate *WbiI* and *Bmul\_2510* polypeptides, respectively. Lanes shown (left to right): protein ladder (M), cloning vector pSCrhaB2, cloning vector + *WbiI*-FLAG or *Bmul\_2510*-FLAG.

synthesis (Fig. 3.7). Failure to complement OAg synthesis was not due to lack of protein expression since *bmul\_2510*- and *wbiI*-encoded polypeptides containing FLAG tags were detectable by western blot analysis (Fig. 3.7 – panel B). Additional bands reactive against the FLAG monoclonal antibodies correspond to cross-reactive proteins, as they also appear in the control lanes of lysates from bacteria containing empty vector. Therefore, from these results we conclude that the gene polymorphisms in the OAg biosynthesis cluster cannot explain the loss of OAg production in the sequential isolates obtained after IST439.

#### 3.4.5. Analysis of the OAg cluster of a co-infecting *B. multivorans* isolate rules out that the hybrid cluster in IST439 arose from gene transfer in the lung

Because of the hybrid nature of the IST439 OAg biosynthetic cluster and the isolation of co-infecting *B. multivorans* in the same patient, we investigated the possibility that this cluster could have arisen from a gene transfer event in patient. *B. multivorans* IST419 was collected approximately 1 year before *B. cenocepacia* IST439 in patient J (Table 3.1) (Cunha et al., 2003). The combination of two sequencing rounds of *B. multivorans* IST419 allowed us to assemble 155 contigs (N50=171331) with a predicted genome size of ~6.5 Mb and 5818 genes (Zerbino and Birney, 2008; Seemann, 2014). Using this annotated assembly, the corresponding OAg genetic cluster was identified and compared to that of *B. multivorans* ATCC 17616 (Nishiyama et al., 2010), *B. multivorans* D2095 (NCBI nucleotide accession number: JFHP00000000.1) (Silva et al., 2016), and *B. cenocepacia* IST439. The OAg biosynthetic cluster in *B. multivorans* IST419, also located in chromosome 1, includes 25 ORFs spanning a 30.8 kb segment flanked by *ureG* and *apaH* (Fig. 3.8 – panel A). Functional assignments could be made to most of the genes based on bioinformatics. This OAg cluster has more ORFs than IST439 but shows 3 conserved features: first, the region involved in lipid A-core biosynthesis (*waaA*, *kdoO*, *waaC*, and *manB*), second, the presence of *wbiH* and *wecA* encoding the putative initiating enzyme for OAg biosynthesis, and third, the presence of conserved genes *rmlBACD* encoding enzymes for dTDP-rhamnose synthesis (Valvano, 2003; Ortega et al., 2005; Valvano, 2015a). A detailed analysis of this cluster, together with the alignment of filtered fastq files of *B. multivorans* IST419 against the same region of *B. multivorans* ATCC 17616 (Fig. 3.8 – panel B), shows that *B. multivorans*-like genes in the IST439 OAg cluster are not present in *B. multivorans* IST419 (Fig. S3.4 – panel A). Further, a similar analysis using a second clinical *B. multivorans* isolate as reference (*B. multivorans* D2095 (Silva et al., 2016)), also revealed the complete absence of homologs to those specific *B. multivorans* ATCC 17616 genes. The electrophoretic analysis of LPS samples obtained from *B. cenocepacia* IST439, *B. multivorans* IST419, and *B. multivorans* ATCC 17616 showed that the first two express an OAg possessing repetitive units with different banding patterns, but *B. multivorans* ATCC 17616 does not produce an OAg unit (Fig. 3.8 – panel C). Therefore, we conclude that *in vivo* gene transfer over the

course of the infection of the patient can't explain the hybrid gene organization of the OAg cluster in IST439 and related serial isolates.



**Figure 3.8** | Genetic organization of the O-antigen biosynthetic gene clusters of *B. multivorans* strains ATCC 17616 (B) and IST419 (A), including genes for lipid A-core biosynthesis and O-antigen biosynthesis within the *apaH* and *ureG* flanking genes (in black). Conserved genes among *Burkholderia* species involved in lipid A-core biosynthesis are indicated in gray. GC content plots are represented for both clusters (drawn using Artemis (Carver et al., 2008)), where the *B. multivorans* ATCC 17616 with homology in the corresponding cluster of *B. cenocepacia* IST439 are highlighted in a black rectangle. A silver nitrate-stained SDS-PAGE gel (C) shows the banding pattern of LPS samples extracted from *B. cenocepacia* IST439 (Bc IST439), *B. multivorans* IST419 (Bm IST419), and *B. multivorans* ATCC 17616 (Bm ATCC17616).

To shed more light on the origin of the hybrid OAg biosynthetic cluster in *B. cenocepacia* IST439, we determined the genetic organization of the corresponding region in the environmental isolate *B. multivorans* ATCC 17616 (Fig. 3.8 – panel B). The gene cluster spans a region of 31.4 kb in chromosome 1 containing 24 predicted ORFs. In terms of functional conservation, the *B. multivorans* ATCC 17616 region is similarly organized to that of *B. cenocepacia* IST439 (Fig. S3.4 – panel B) although, unlike *B. cenocepacia* IST439, the seven genes annotated in *B. multivorans* ATCC 17616 are not consecutively organized in the OAg cluster (indicated by black boxes in Figs. 3.6 – panel A and 3.8 – panel B).

### 3.5. Discussion

Phenotypic diversity and genotypic flexibility of *B. cenocepacia* during long-term infection of cystic fibrosis lungs was recently well established using genomic and phenotypic analyses of serially collected clinical isolates (Lee et al., 2017). In this study, we characterized a novel hybrid genetic LPS OAg locus in the early *B. cenocepacia* isolate of a series of sequential isolates obtained over a 3.5-year period from a patient with cystic fibrosis. We also demonstrate that all of the 10-subsequent clonal *B. cenocepacia* variants collected from the same patient until shortly before death do not produce OAg (Fig. 3.1). This observation provides further evidence for the OAg loss during *B. cenocepacia* adaption over chronic lung infection (Maldonado et al., 2016).

We elucidated the chemical structure of the lipid A-core of four of the 11 isolates, as well as the structure of the OAg in the early isolate (IST439). The lipid A-core of these strains had the same structure as that described for other members of the *Burkholderia* genus (De Soyza et al., 2004; De Soyza et al., 2008). Indeed, the lipid A moiety shows the typical *Burkholderia* glucosamine disaccharide backbone with the [P→4-β-D-GlcpN-(1→6)-α-D-GlcpN1→P] sequence that comprises two Ara4N residues linked to phosphate groups. The inner core contains the characteristic Ara4N-Ko-Kdo trisaccharide, previously elucidated in other *Burkholderia* species (Isshiki et al., 1998; Gronow et al., 2003), and is extended by a heptose-rich sequence. The OAg repeating unit of the LPS in IST439 was a disaccharide [→2)-β-D-Ribf-(1→4)-α-D-GalpNAc-(1→], which was not previously described in *B. cenocepacia*, differs from the OAg structure of the prototypic strain K56-2 (Ortega et al., 2005; Ortega et al., 2009), and is similar to that occurring in the OAg of *B. cepacia* serotypes O3 and O5 (Cox and Wilkinson, 1989). In addition to the presence of a novel OAg, the absence of the Rha-QuiNAc disaccharide, linked to the α-Hep-(1→2)-α-Gal disaccharide of the outer core in all the isolates investigated, represented a further novelty of these strains.

In keeping with the OAg structure, a new OAg biosynthetic locus was also discovered in this study. This genetic locus exhibits a hybrid origin, comprising genes with homology to *B. multivorans* ATCC 17616 and low to no amino acid similarity with other sequenced *B. cenocepacia* strains. Comparisons with the OAg locus of a co-infecting *B. multivorans* isolate (IST419), which was present at the time of acquisition of *B. cenocepacia* IST439 (Fig. 3.8 – panel A) (Cunha et al., 2003), showed that the genetic organization of the OAg cluster in *B. multivorans* IST419 is different from that of *B. cenocepacia* IST439, with no homology with the *B. multivorans*-like genes shared by IST439 and environmental strain *B. multivorans* ATCC 17616 genomes. Together, our findings demonstrate that no transfer of genetic material occurred between the 2 co-infecting species in the lung of this CF patient during the infection within the OAg biosynthetic cluster.

The origin of the unique OAg locus of *B. cenocepacia* IST439 is unknown, but it was found that the *B. multivorans*-like region differs significantly from the GC content of the entire cluster and core genome (Fig. 3.6 – panel A), which is a common indicator for the presence of a putative genetic island and former mobile element (Hacker and Carniel, 2001; Zhang et al., 2014). Studies mostly based on *Salmonella enterica* and *E. coli* (D'Souza et al., 2005; Hu et al., 2010; Reeves et al., 2013) have previously suggested that OAg biosynthetic operons in bacterial pathogens were acquired by horizontal genetic transfer from species with a low GC content. Moreover, horizontal exchange of O-specific antigen biosynthetic genes among phylogenetically distinct *P. aeruginosa* strains was observed and serotype switching was found to be the result of horizontal transfer and genetic recombination of LPS biosynthetic genes originating from an multidrug-resistant taxonomic outlier *P. aeruginosa* strain (Thrane et al., 2015). The high variation within the cluster also indicates that these OAg clusters might have been assembled by multiple transmission events over time (Lerouge and Vanderleyden, 2002; Thrane et al., 2015; Maldonado et al., 2016). In the case of *B. cenocepacia* IST439, the core genome and the OAg cluster (minus the 7-gene region) have a similar GC content (66-67%), whereas the region containing the 7 genes annotated as *B. multivorans* ATCC 17616, with a 56.1% GC content, is an obvious candidate for a mobile element that was acquired by horizontal genetic transfer and may be adaptive. However, the corresponding 7 genes in *B. multivorans* ATCC 17616 are not encoded consecutively, being interrupted by a group of 3 high-GC content genes (Fig. 3.8 – panel B), and cannot be matched to the same region in *B. cenocepacia* IST439 without assuming the occurrence of multiple mutational events. Based on these observations and the variability observed within the OAg cluster of the *B. multivorans* sequences examined, it can be suggested that this putative mobile element identified in the OAg biosynthetic cluster of *B. cenocepacia* IST439 has a foreign non-Bcc origin, which at some point might have also been acquired by an ancestor of *B. multivorans* ATCC 17616.

We could not establish a direct link between OAg loss in all 10 late-stage variants with the conserved mutation in the gene homolog to *Bmul\_2510* or with the mutations in *wbiI* since none of our complementation efforts using genetic constructions with *B. cenocepacia* IST439 *Bmul\_2510* and/or *wbiI* in different backgrounds could reconstitute OAg biosynthesis. Although it is possible that lack of complementation of OAg synthesis is associated with insufficient expression of the cloned gene or poor translocation of the protein to the membrane, it cannot be excluded that these mutations require the complementation of another genetic alteration in a distinct part of the genome with some role in OAg biosynthesis that remains to be identified. Alterations in the LPS molecule during chronic CF infections are thought to contribute to adhesion, evasion of immune defenses and overall adaptation to the infection niche (Maldonado et al., 2016). In *P. aeruginosa*, it is well established that adaptation to the lung during chronic infection includes the loss of OAg and/or lipid A modifications that allow the bacterium to avoid host immune responses (Lyczak et al., 2002; Cigana et al., 2009; Dettman et al., 2013; Maldonado et al., 2016). LPS virulence resides both in the endotoxin activities of the lipid A and

in the ability of the core oligosaccharide and OAg to provide the bacterium with resistance to host defenses. The ability of 3 of the 11 *B. cenocepacia* clonal variants to subvert host defenses was recently assessed using dendritic cells (Cabral et al., 2017), revealing that the late variants IST4113 and IST4134 were significantly more internalized than IST439, the only isolate that expresses the OAg unit, in line with previous studies by Saldias *et al.* (Saldias et al., 2009). Moreover, the late-stage isolates also exhibited improved survival within dendritic cells than the early isolate IST439, corroborating the idea that loss of the OAg may participate in providing an adaptive advantage to chronically infecting *B. cenocepacia*. Collectively, our results lend support to the notion that LPS OAg modifications are an important factor in the adaptation of *B. cenocepacia* to chronic infection and that that OAg heterogeneity relates to variation within the OAg gene cluster.

## 4 *Burkholderia cepacia* complex species differ in the frequency of variation of the lipopolysaccharide O-antigen expression during cystic fibrosis chronic respiratory infection

---

**This chapter contains results published in:**

A. Amir Hassan, Carla P. Coutinho, and Isabel Sá-Correia, ***Burkholderia cepacia* complex species differ in the frequency of variation of the lipopolysaccharide O-antigen expression during cystic fibrosis chronic respiratory infection. (2019), *Frontiers in Cellular and Infection Microbiology* 9(273). doi: 10.3389/fcimb.2019.00273.**

iBB - Institute for Bioengineering and Biosciences, Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal



## 4.1. Abstract

*Burkholderia cepacia* complex (Bcc) bacteria can adapt to the lung environment of cystic fibrosis (CF) patients resulting in the emergence of a very difficult to eradicate heterogeneous population leading to chronic infections associated with rapid lung function loss and increased mortality. Among the important phenotypic modifications is the variation of the lipopolysaccharide (LPS) structure at level of the O-antigen (OAg) presence, influencing adherence, colonization and the ability to evade the host defense mechanisms. The present study was performed to understand whether the loss of OAg expression during CF infection can be considered a general phenomenon in different Bcc species favoring its chronicity. In fact, it is still not clear why different Bcc species/strains differ in their ability to persist in the CF lung and pathogenic potential. The systematic two-decade-retrospective-longitudinal-screening conducted covered 357 isolates retrieved from 19 chronically infected patients receiving care at a central hospital in Lisbon. The study involved 21 Bcc strains of six/seven Bcc species/lineages, frequently or rarely isolated from CF patients worldwide. Different strains/clonal variants obtained during infection gave rise to characteristic OAg-banding patterns. The two most prevalent and feared species, *B. cenocepacia* and *B. multivorans*, showed a tendency to lose the OAg along chronic infection. *B. cenocepacia* *recA* lineage IIIA strains known to lead to particularly destructive infections exhibit the most frequent OAg loss, compared with lineage IIIB. The switch frequency increased with the duration of infection and the level of lung function deterioration. For the first time, it is shown that the rarely found *B. cepacia* and *B. contaminans*, whose representation in the cohort of patients examined is abnormally high, keep the OAg even during 10- or 15-year infections. Data from co-infections with different Bcc species reinforced these conclusions. Concerning the two other rarely found species examined, *B. stabilis* exhibited a stable OAg expression phenotype over the infection period while for the single clone of the more distantly related *B. dolosa* species, the OAg-chain was absent from the beginning of the 5.5-year infection until the patient dead. This work reinforces the relevance attributed to the OAg-expression switch suggesting marked differences in the various Bcc species.

## 4.2. Introduction

Chronic respiratory infections involving opportunistic pathogens remain the leading cause of premature death of cystic fibrosis (CF) patients (Cullen and McClean, 2015), in particular the long-term infections caused by *Pseudomonas aeruginosa* and bacteria of the *Burkholderia cepacia* complex (Bcc). The Bcc is a group of more than 20 closely-related species (Depoorter et al., 2016) particularly feared by CF patients due to their very difficult eradication, inter-patient transmission, and related increased morbidity and decreased life expectancy (Mahenthiralingam et al., 2005; Drevinek and Mahenthiralingam, 2010; Kenna et al., 2017). The majority of CF patients who acquire Bcc develops a chronic infection while a minority may clear the infection spontaneously (Mahenthiralingam et al., 2005). Pulmonary colonization with Bcc can cause an accelerated decline in lung function and the “cepacia syndrome” that is characterized by an uncontrolled deterioration with septicaemia and necrotizing pneumonia that usually results in early death (Mahenthiralingam et al., 2005). Inside the CF lung, Bcc bacteria face stressful and changing environmental conditions as a consequence of host immune defenses, inflammatory responses, antimicrobials and reactive oxygen species (ROS), fluctuating levels of nutrients, oxygen deprivation, high osmolarity, low pH, biofilm growth, and the presence of other co-infecting microbes (Hogardt and Heesemann, 2010; Döring et al., 2011; Cullen and McClean, 2015). Consequently, during long-term infection, the initial Bcc strain-population accumulates genetic alterations leading to genotypic and phenotypic diversification forming a heterogeneous bacterial community very difficult to eradicate therapeutically (Lieberman et al., 2011; Madeira et al., 2011; Silva et al., 2011; Madeira et al., 2013; Lieberman et al., 2014; Moreira et al., 2014; Silva et al., 2016; Nunvar et al., 2017). Such within-patient emergence of multiple clonal variants was first described for the more prevalent CF pathogen *Pseudomonas aeruginosa* (Markussen et al., 2014; Winstanley et al., 2016) and proposed to provide a pool of mutations affecting virulence and antimicrobial resistance (Lorè et al., 2012; Marvig et al., 2013; Faure et al., 2018). Several studies have shown that, during chronic infection, alterations in the lipopolysaccharide (LPS) molecule, a complex glycolipid covering the cell surface and shielding Gram-negative bacteria from adverse host environments, may occur (Pier, 2007; Maldonado et al., 2016; Faure et al., 2018). The LPS is one of the most important components of pathogenic Gram-negative bacteria cellular envelope being considered a major virulence factor and contributing to adhesion, immune defenses evasion, host colonization and adaptation to the infection niche (Vinion-Dubiel and Goldberg, 2003; Pier, 2007; Maldonado et al., 2016). The LPS structure is composed by three constituents: a highly acylated lipid A, also known as endotoxin, covalently linked to the central core oligosaccharide, and the O-antigen (OAg), composed by repeating polysaccharide units of variable length (Vinion-Dubiel and Goldberg, 2003). Most of the heterogeneity in LPS molecules is found in the OAg polysaccharide and lipid-A components (Pier, 2007; Maldonado et al., 2016; Faure et al., 2018). The OAg component of the LPS is essential in several pathogens for motility (swarming and twitching) (Toguchi et al., 2000; Berry et

al., 2009; Bowden et al., 2013), protection against oxidative stress (Berry et al., 2009) and evasion from host immune defenses (Murray et al., 2003; Saldias et al., 2009; Kotrange et al., 2011; Kintz et al., 2017). LPS OAg is believed to be an immunodominant molecule that can modulate host-pathogen interaction (Ranf, 2016) and proposed to be under selective pressure in Gram-negative bacteria (King et al., 2009; Maldonado et al., 2016; Kintz et al., 2017). Longitudinal studies of *P. aeruginosa* isolates from CF patients with chronic lung infections revealed the conversion from smooth (in the early isolates) to rough LPS with short or no OAg side chain at the late-stage of infections rendering the bacteria non-typable and less immunogenic (King et al., 2009; Maldonado et al., 2016; Demirdjian et al., 2017). In Bcc bacteria, a few recent reports also described the variation of the presence and/or modification of the OAg during persistent infection of the CF lung and on how this alteration may affect Bcc pathogenicity (Lieberman et al., 2011; Maldonado et al., 2016; Silva et al., 2016; Hassan et al., 2017). An extensive study on an epidemic outbreak of *Burkholderia dolosa* during 16 years, involving 14 CF patients and 112 isolates, has shown that several late isolates produce an LPS exhibiting the OAg that was absent in the LPS of the ancestral strain; this fact was related with the appearance of two different mutations in the glycosyltransferase encoding gene *wbaD* (Lieberman et al., 2011). The genome analysis of clones and metagenomes of evolved *B. cenocepacia* biofilms, after more than one thousand generations, unveiled a mutation in the gene *manC* – a gene involved in mannose metabolism located within the LPS biosynthesis gene cluster- resulting in the loss of the LPS OAg (Traverse et al., 2013). Remarkably, mutations in the same *manC* gene were previously found to disrupt OAg biosynthesis in late *B. cenocepacia* (Saldias et al., 2009) or early *B. dolosa* (Lieberman et al., 2011) isolates retrieved during chronic infection. A comparative genomic analysis focused on sequential *B. multivorans* isolates obtained over 20 years of CF chronic infection showed that late isolates accumulate three different mutations in a locus homologous to the *wbi* gene cluster, involved in LPS OAg biosynthesis, leading to OAg loss (Silva et al., 2016). Another recent study involving eleven serial clonal variants of *B. cenocepacia* *recA* lineage IIIA obtained from a CF patient over a 3.5 year period, since the onset of infection until death with the cepacia syndrome, revealed that the OAg was only present in the early isolate but not in any of the late variants (Hassan et al., 2017). Genomic sequencing of these serial isolates enabled the identification of mutations, also within the OAg cluster, but none of these mutations could definitively be associated to OAg loss (Hassan, et al. 2017).

Although the Bcc comprises 24 bacterial species (De Smet et al., 2015; Depoorter et al., 2016; Ong et al., 2016; Bach et al., 2017; Weber and King, 2017; Martina et al., 2018) and several were found to be involved in CF respiratory infections, only *B. cenocepacia* (Hassan et al., 2017), *B. multivorans* (Silva et al., 2016), and *B. dolosa* (Lieberman et al., 2011) were examined concerning the variation of the OAg presence over chronic infection. *B. cenocepacia* is recognized as the dominant species with high potential for inter-patient transmission (Drevinek and Mahenthalingam, 2010), and *B. multivorans* has recently replaced *B. cenocepacia* in this first position in several countries (Lipuma,

2010). However, it is recognized that the less represented species may also be associated with poor clinical outcome, but it is not clear why different Bcc species/strains differ in their ability to persist in the CF lung and in their pathogenic potential (Woods et al., 2004; Kalish et al., 2006; Cunha et al., 2007; Moehring et al., 2014; Coutinho et al., 2015; Nunvar et al., 2016; Marquez et al., 2017; Roux et al., 2017). It also remains to be elucidated whether LPS OAg loss confers an advantage to the different Bcc species in particular and to Gram-negative bacteria, in general, over the course of infection.

The present study was performed to understand whether OAg loss in different Bcc species during respiratory infection in CF patients can be considered a general phenomenon that affects immune evasion favoring chronic infection. For this, a systematic retrospective and longitudinal screening was performed based on a collection of isolates, recovered from 1995 to 2016 from 19 CF patients under surveillance at the major Portuguese CF treatment center at Hospital de Santa Maria (HSM), in Lisbon, over the duration of chronic infection (ranging from 1.2 to 15.2 years). The 357 sequential isolates examined were molecularly identified at the species level and genotyped during this study or in previous studies and belong to *B. cenocepacia*, *B. multivorans* or to the more rare species in CF population worldwide, *B. dolosa*, *B. stabilis*, *B. cepacia* and *B. contaminans* (Cunha et al., 2003; Cunha et al., 2007; Coutinho et al., 2011b; Moreira et al., 2014; Coutinho et al., 2015). The abnormally high representation of the species *B. cepacia* and *B. contaminans* in our collection is considered related with a contamination of saline solutions for nasal application (Cunha et al., 2007; Coutinho et al., 2015). Information on the health condition of several patients at the time of isolation based on the forced expiratory values (FEV), an indicator of pulmonary function and, consequently, of oxygen availability in the patient's lung, is also available (Correia et al., 2008; Moreira et al., 2014). This systematic retrospective study allowed to answer the question of whether the described variation of the LPS OAg expression for a limited number of the Bcc species/strains and patients is extensive to the more rarely found Bcc species and can be considered a general phenomenon in the Bcc during long-term respiratory infections in CF patients.

### 4.3. Materials and methods

#### 4.3.1. Bacterial isolates and growth conditions

The 357 *Burkholderia cepacia* complex (Bcc) isolates examined in this study were retrieved from the respiratory secretions of 19 chronically infected cystic fibrosis (CF) patients under surveillance at the major Portuguese CF treatment center at Hospital de Santa Maria (HSM), Centro Hospitalar Lisboa Norte EPE, Lisbon, from 1995 to 2016, during hospital routines. The isolates examined in this retrospective study were selected at random among the colonies isolated in selective *Burkholderia cepacia* Selectatab medium at the Hospital, at a specific date of isolation. Over the years, a significant

part of the isolates were molecularly confirmed to belong to six Bcc species: *Burkholderia cenocepacia*, *Burkholderia multivorans*, *Burkholderia cepacia*, *Burkholderia contaminans*, *Burkholderia dolosa*, and *Burkholderia stabilis* and genotyped (Cunha et al., 2003; Cunha et al., 2007; Coutinho et al., 2011b; Coutinho et al., 2015) (Table 4.1 and Table S4.1). Bacterial cultures are stored at -80°C in 1:1 (v/v) glycerol. During this study, bacteria were grown in Lysogeny broth agar (LB agar; Conda, Pronadisa) at 37°C with shaking at 250 rpm or in LB agar plates.

#### 4.3.2. Species identification and genotyping of Bcc isolates

The isolates examined in this study that had not been previously identified at the species level and genotyped (by ribotyping or *recA* gene PCR Restriction fragment length polymorphism - *recA* PCR RFLP –and species-specific *recA* direct PCR) were identified and genotyped in this study based on *recA* PCR RFLP analysis and Random Amplified Polymorphic DNA (RAPD) (Mahenthalingam et al., 1996; Mahenthalingam et al., 2000a). Genomic DNA was extracted from cells of those isolates grown overnight in LB broth with orbital agitation at 37°C, using the Puregene DNA isolation kit (cell and tissue kit, Gentra Systems, Qiagen, Germany). The concentration of the resulting genomic DNA solutions was estimated using a ND-1000 spectrophotometer (NanoDrop Technologies). Species identification was based on the *recA* PCR RFLP analysis that was performed by amplification of the entire *recA* gene using specific primers for Bcc bacteria (BCR1-<sup>5'</sup>TGACCGCCGAGAAGAGCAA<sup>3'</sup>; BCR2-<sup>5'</sup>CTCTTCTTCGTCCATCGCCTC<sup>3'</sup>) (Mahenthalingam et al., 2000a). Bcc *recA* amplicons were digested with *Hae*III (Amersham Biosciences) and the restriction fragments separated by electrophoresis in 2% (w/v) agarose gels. The RFLP patterns obtained were compared with those reported in the literature (Mahenthalingam et al., 2000a; Vermis et al., 2002; Cunha et al., 2003; Coutinho et al., 2015). Strain assignment was performed by RAPD (Mahenthalingam et al., 1996). Genotyping of the isolates tested was performed using primer 270 (<sup>5'</sup>TGCGCGCGGG<sup>3'</sup>), followed by confirmation of the strain types established by this primer with primers 208 (<sup>5'</sup>ACGGCCGACC<sup>3'</sup>) and 272 (<sup>5'</sup>AGCGGGCCAA<sup>3'</sup>). RAPD profiles were visually compared by GelJ software (Heras et al., 2015). Similarity coefficients were calculated using the RAPD profiles obtained with each primer for the complete set of isolates examined using Pearson coefficient with Arithmetic Mean (UPGMA) linkage method.

#### 4.3.3. LPS extraction and SDS-PAGE analysis

LPS was extracted from cell envelopes obtained from bacterial cells as described before (Marolda et al., 1990) with small modifications (Hassan et al., 2017). Briefly, Bcc cells were harvested from overnight liquid cultures by centrifugation for 1 min after OD<sub>640nm</sub> adjustment to 2.0 in 1 mL of phosphate buffer saline – PBS (at pH 7.4: 1.8 mM/L KH<sub>2</sub>PO<sub>4</sub>, 137 mM/L NaCl, 2.7 mM/L KCl and 10

mM/L Na<sub>2</sub>HPO<sub>4</sub>). These cells were suspended in 150µL of lysis buffer, containing 2% SDS, 4% 2-β-mercaptoethanol, and 500 mM Tris-HCl (pH 6.8), and boiled for 10 min. Proteinase K (20 mg/mL) was added, and the samples were incubated at 60°C for 2 hours. Finally, samples were mixed with a tracking dye solution (125 mM Tris-HCl [pH 6.8], 2% SDS, 20% [v/v] glycerol, 0.002% bromophenol blue, and 10% mercaptoethanol) and boiled for 5 min before loading the gels. LPS samples were resolved by electrophoresis – at 150 V for about 1h 40m - in 14% polyacrylamide gels with a Tricine-SDS system followed by silver staining to visualize the O-antigen (OAg) banding patterns (Marolda et al., 2006).

#### 4.3.4. Ethics

The present retrospective work involves clinical Bcc isolates obtained as a part of the hospital routine and its use in clinical research was approved by the hospital ethics committee. Consent was obtained from the patient or legal representative for the use of these isolates in research. Patients' anonymity is preserved.

### 4.4. Results

#### 4.4.1. During chronic infections, *B. cenocepacia recA* lineage IIIA and *B. multivorans* isolates were more prone to lose the OAg present in early isolates, compared with *B. cenocepacia recA* lineage IIIB.

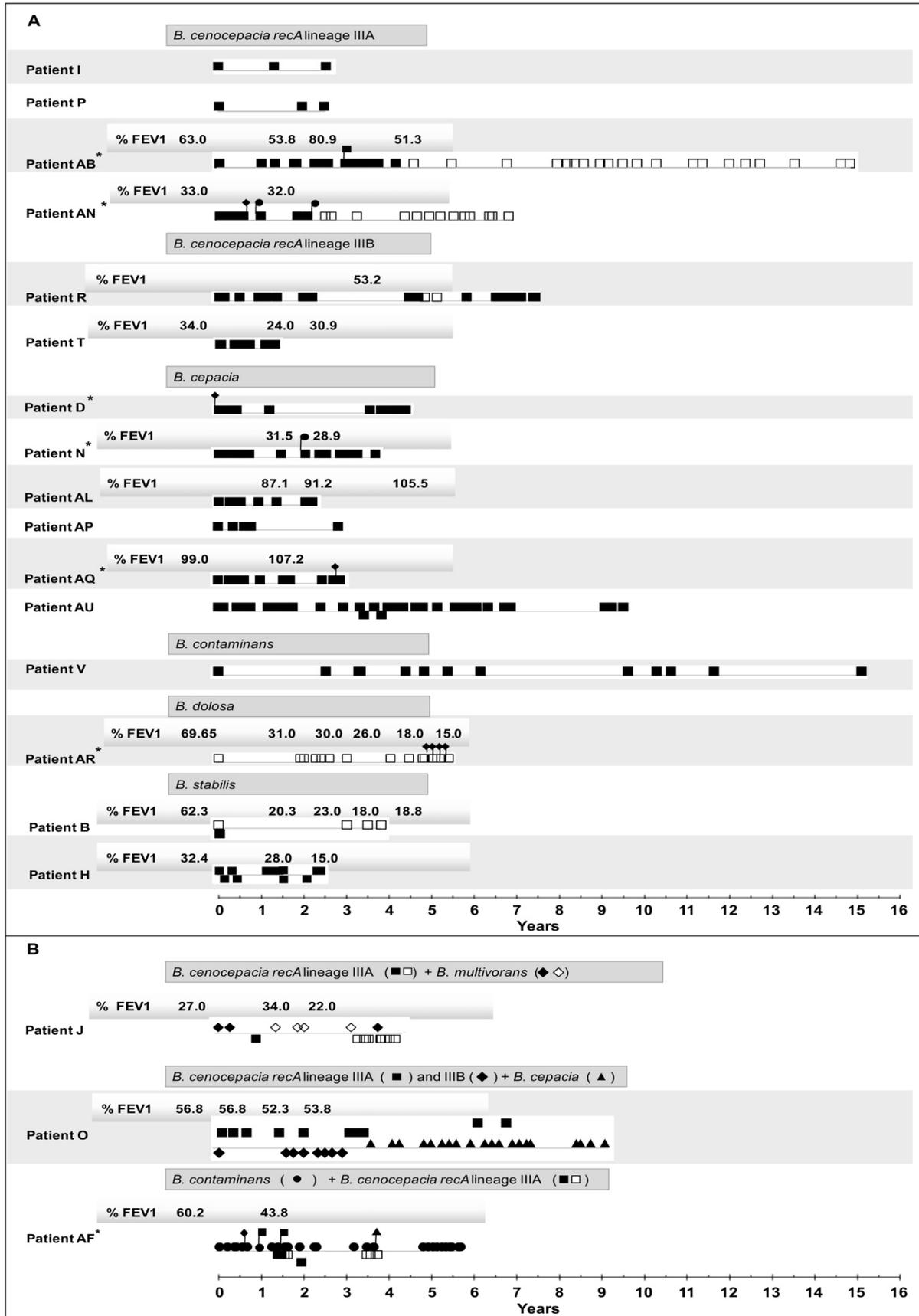
The 357 *Burkholderia cepacia* complex (Bcc) isolates tested in this study were examined with the goal of systematically compare the O-antigen (OAg) banding patterns when this LPS component is present and to demonstrate the eventual absence of the OAg in late isolates (Fig. 4.1 and Table 4.1). It was found that the different strains tested and the corresponding clonal variants obtained during long-term-infection, as demonstrated by Multilocus sequence typing (MLST), gave rise to characteristic banding patterns of the OAg whenever this component is present (Fig. 4.2 – panels A and B).

Concerning the 153 isolates of *B. cenocepacia recA* lineages IIIA (112 isolates from 7 patients) and IIIB (41 isolates from 3 patients) (Table 4.1) tested for the presence of OAg polysaccharide during lung infection (timelines shown in Fig. 4.1), it was found that strains from *B. cenocepacia* lineage IIIA were more prone to lose the OAg than lineage IIIB strains.

**Table 4.1** | Schematic representation of the sequential isolates from various species chronically infecting the different CF patients examined in this study. Information on the number of sequential clonal isolates in which the OAg is present (●) or absent (○) and the corresponding OAg profiles (Figure 2), when present, the ripopatterns of the isolates (Cunha et al., 2003; Cunha et al., 2007; Coutinho et al., 2015) and their RAPD profiles (Fig. S4.1) and the duration of the chronic infection and clinical outcome of the infected CF patients (Cunha et al., 2003; Cunha et al., 2007; Coutinho et al., 2011b; Moreira et al., 2014) is provided.

CF patients	Number of sequential isolates with (●) or without (○) O-antigen	O-antigen profiles (in Figure 2)	Isolate ribo-patterns (R) / RAPD profiles	Patient clinical outcome	Duration of the chronic infection with the Bcc strain (years)
<b><i>Burkholderia cenocepacia IIIA</i></b>					
I	● 3	A1	R08 / RAPD01	infection eradicated	2.6
J	● 1 ○ 10	A2 & A3	R11 / RAPD02	Dead *	3.4
O	● 10	A4	R11 / RAPD03	infection eradicated before death	3
P	● 3	A5	R13 / RAPD04	infection eradicated before death	2.5
AB	● 17 ○ 29	A6 & A7	R07 / RAPD05	still infected	14.7
AF	● 2 ○ 5 ● 1	A8 & A10 A9	ND / RAPD06 RAPD07	Dead	2.5 0.1
AN	● 10 ○ 21	A11 & A12	R21 / RAPD08	Dead	6.8
<b><i>Burkholderia cenocepacia IIIB</i></b>					
O	● 8 ● 2	B1 B2	R14 / RAPD09 RAPD10	infection eradicated before death	3.4 0.7
R	● 12 ○ 2 ● 11	B3 & B4	R15 / RAPD10	infection eradicated	7.5
T	● 6	B5	R16 / RAPD11	infection eradicated before death	1.2
<b><i>Burkholderia cepacia</i></b>					
D	● 11	C1	ND / RAPD12	Dead	4.4
N	● 28	C2	R12	Dead	3.8
O	● 24	C2	R12	infection eradicated before death	5.6
AL	● 10	C1	R19 / RAPD12	infection eradicated	2.2
AP	● 5	C1	R19	infection eradicated	2.8
AQ	● 10	C5	R24	infection eradicated	2.9
AU	● 32 ● 2	C3 C4	ND / RAPD12 RAPD13	still infected	9.6 0.4
<b><i>Burkholderia contaminans</i></b>					
V	● 13	D1	R17	Alive	15.2
AF	● 29	D2	R02	Dead	5.8
<b><i>Burkholderia multivorans</i></b>					
J	● 2 ○ 6 ● 1	E1 & E2	R9	Dead *	4
<b><i>Burkholderia dolosa</i></b>					
AR	○ 14	F	ND	Dead	5.5
<b><i>Burkholderia stabilis</i></b>					
B	○ 4 ● 1	G1 G2	R01 / RAPD15 RAPD14	Dead	3.8 0.1
H	● 8 ● 4	G3 G4	R01 / RAPD14 RAPD15	Dead *	2.5 2

R – Isolate ribopattern; RAPD – Random amplified polymorphism DNA profile (Fig. S4.1); ND – not determined; \* – cepacia syndrome



**Figure 4.1** | Schematic representation of the Bcc isolates examined in this study retrieved from different CF patients over 21 years (from 1995 to 2016) of epidemiological survey at the CF Centre of

Hospital de Santa Maria. Time zero marks the date of isolation of the first Bcc isolate from a specific patient with essentially a single Bcc species/strain (panel A) or co-infected with different Bcc species (panel B). The presence or absence of the LPS O-antigen is represented by closed or open symbols, respectively. The same symbol in a different line is used to indicate a different strain of the same species co-infecting the patient, as detailed in Table 4.1 and Table S4.1, as in the case for patients AU, B, H, O, and AF. Whenever available, the Forced Expiratory Values (%) in the first second (FEV1), before and during infection, are provided (Correia et al., 2008; Coutinho et al., 2011a; Moreira et al., 2014). These values are an indication of the pulmonary function in the infected patients.

\* - Sporadic isolation of strains from other Bcc species (in Table S4.1): *B. cenocepacia* III B -  $\blacklozenge$ , *B. cepacia* -  $\blacksquare$ , *B. multivorans* -  $\bullet$ , and *B. dolosa* -  $\blacktriangle$ .

There were two cases of extended chronic infection of patients AB and AN by *B. cenocepacia* lineage IIIA, with a duration of 14.7 and 6.8 years, respectively, (Fig. 4.1 – panel A and Table 4.1) where the late isolates retrieved after 4.5 or 2.5 years, respectively, were found to lack the OAg that was present in the early isolates (Fig. 4.1 and Table S4.1). The loss of the OAg in late isolates was also previously reported for another *B. cenocepacia* lineage IIIA strain infecting patient J for a shorter period of 3.4 years (Hassan et al., 2017) and (Fig. 4.1 and Table 4.1). Patient J was co-infected with a *B. multivorans* strain and the LPS of the majority of late isolates of this *B. multivorans* strain also lack the OAg present in the early isolates, as shown in the electrophoretic profiles of the extracted LPS (Fig. 4.2, columns E1 and E2). However, one *B. multivorans* clonal isolate retrieved at the last isolation date exhibited again the OAg, consistent with the pattern observed before by others (Silva et al., 2016). Late isolates of *B. cenocepacia* and *B. multivorans* were retrieved from patient J during a period of oxygen depletion in the CF lung that accompanied disease progression and the deterioration of lung function. In fact, according to the hospital records, the forced expiratory values in the first second (FEV1) available are low during co-infection (22-34 %) and no further values of FEV1 were registered during the later stages of the patient's life due to the highly severe deterioration of pulmonary function (Fig. 4.1 – panel B; (Correia et al., 2008).

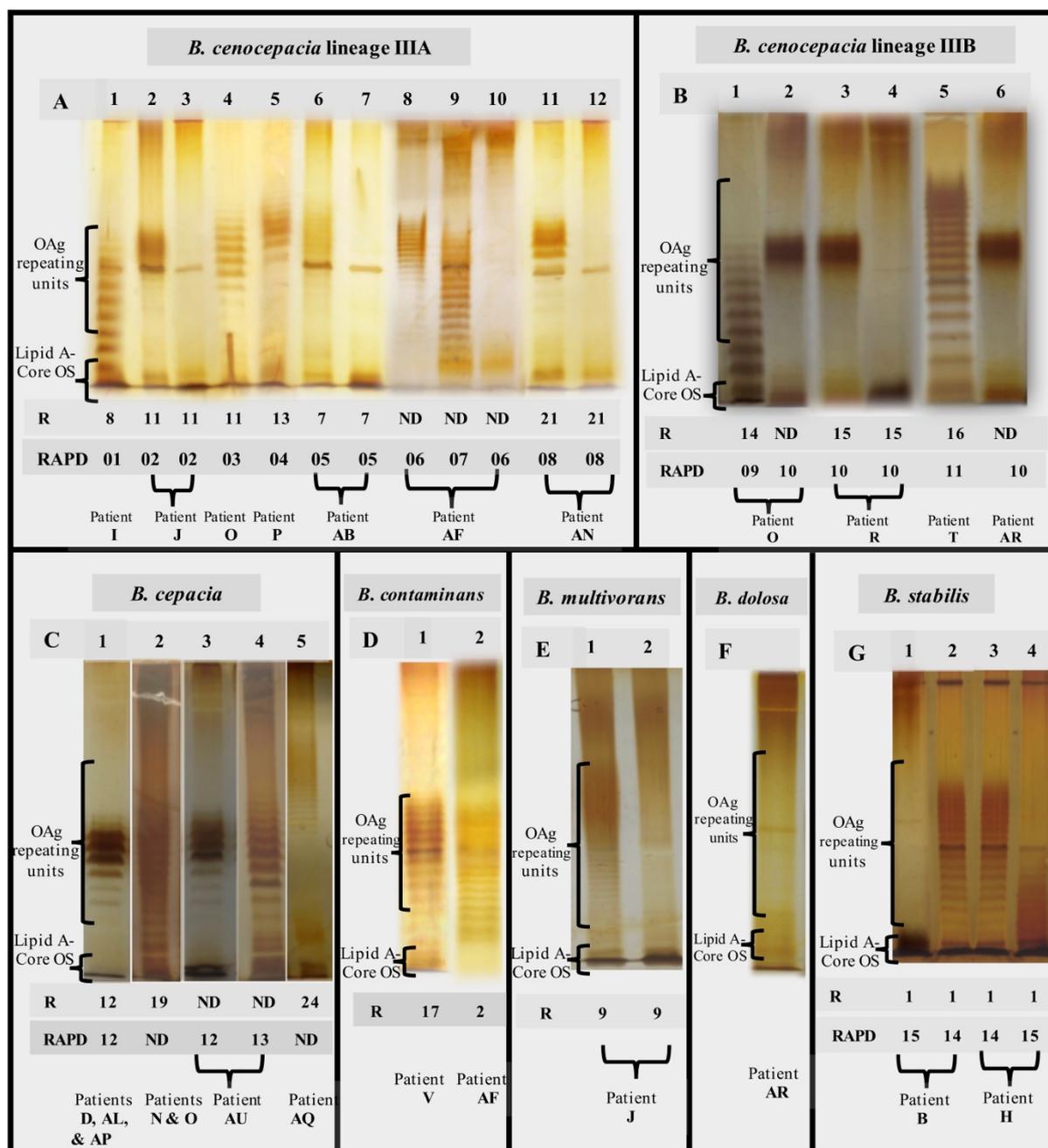
Differently, the *B. cenocepacia* lineage IIIA strain that infected patient O for 3 year-infection period kept the LPS OAg and the same was found for the co-infecting Bcc strains of *B. cenocepacia* lineage III B and *B. cepacia* chronically infecting patient O for 3.4 and 5.6 years, respectively. Remarkably, the FEV1 values available were above 50% during the co-infection period and, after 9 years, the *Burkholderia* infection was eradicated (Correia et al., 2008); Fig. 4.1, Table S4.1). When examining the 11 and 10 *B. cenocepacia* IIIA isolates obtained from patients' J and O, respectively, we were surprised by the fact that the electrophoretic profiles of the extracted LPS obtained by SDS-PAGE analysis and silver-staining produced two different profiles (Fig. 4.2, columns A2 and A4, respectively) even though they were considered to share the same ribopattern 11 [table S4.1; (Coutinho et al., 2011b)]. However,

the RAPD analysis of the tested isolates revealed that they are clonal isolates but from two different ancestral strains, RAPD02 for patient J and RAPD03 for patient O (Table 4.1), in line with the different MLST sequence types, ST-218 and ST-280 at the MLST database [<http://pubmlst.org/bcc/>], obtained for the isolates infecting patient J and O, respectively.

For the *B. cenocepacia* IIIA strain co-infecting patient AF for 2.5 years, the loss of the OAg was also observed when the FEV1 values were close to 40%. However, as demonstrated by the different RAPD profiles obtained (Table 4.1), another *B. cenocepacia* IIIA strain was obtained at an isolation date after the loss of the OAg by the original infecting strain; this new strain exhibit an OAg with a distinct banding pattern from the early isolates of the series (Fig. 4.2 – panel A, column A9), Remarkably, all the *B. contaminans* clonal isolates co-infecting this patient kept the OAg when, for identical isolation times, *B. cenocepacia* IIIA clonal isolates lack the OAg (Fig. 4.1 – panel B).

Only for *B. cenocepacia* IIIA isolates chronically infecting patients I and P over a period of time below 3 years, the loss of OAg along infection was not observed (Fig. 4.1 – panel A). Interestingly, after this shorter period of chronic infection, the *Burkholderia* strains were eradicated from both patients.

Concerning the sequential *B. cenocepacia* lineage IIIB isolates retrieved from patient T over 1.2 years of infection before eradication, all the sequential isolates exhibit the OAg even though the FEV1 value was in the low range of 25-30%. In the case of patient R, among the 25 sequential isolates of the *B. cenocepacia* lineage IIIB strain obtained during a very extended infection period (7.5 years) two isolates retrieved after the fifth year of infection were found to lack the OAg (Figs. 4.1 and 4.2, columns B3 and B4, and Table 4.1). Interestingly, the OAg banding pattern obtained for isolates with ribopattern 15 infecting patient R was identical to the OAg profile extracted from two isolates retrieved from patient O (Fig. 4.2, columns B3 and B2) and identical to those from four *B. cenocepacia recA* lineage IIIB isolates also present in patient AR who was mainly infected with *B. dolosa* (Fig. 4.2, columns B3 and B6). All these isolates producing the same banding pattern were confirmed to share the same RAPD profile (Table 4.1 and Fig. S4.1). Consistent with these results the MLST database showed that the *B. cenocepacia* lineage IIIB isolates retrieved from patients R and AR also have the same MLST sequence type (ST-43) [<http://pubmlst.org/bcc/>].



**Figure 4.2** | Representative SDS-PAGE gels showing the O-antigen banding patterns obtained (A1-A12, B1-B6, C1-C5, D1, D2, E1, E2, F, G1-G4) for the Bcc isolates examined. Genotyping data for the different strains tested are also shown: the ribopatterns were obtained before (Cunha et al., 2003; Cunha et al., 2007; Coutinho et al., 2011b) and the RAPD profiles were obtained during this study (Fig. S4.1).

R – Isolate ribopattern; RAPD – Isolate profile based on Random amplified polymorphism DNA profiles (Figure S1); ND – not determined

**The isolates tested are (Table S4.1):** A1 – IST416, A2 – IST439, A3 – IST4103, A4 – IST462, A5 – IST432, A6 – IST4121, A7 – IST4893, A8 – IST4240a, A9 – IST4272, A10 – IST4253a, A11 – IST4197, A12 – IST4386, B1 – IST435, B2 – IST4177, B3 – IST438, B4 – IST4155, B5 – IST466, B6 – IST4481 IIIB, C1 – IST4152, C2 – IST4128, C3 – IST4283, C4 – IST4546, C5 – IST4198, D1 – IST481, D2 – IST4193, E1 – IST419, E2 – IST453, F – IST4208, G1 – IST402, G2 – IST409, G3 – IST413, G4 – IST412.

#### 4.4.2. The OAg is stably expressed in *B. cepacia*, *B. contaminans* and *B. stabilis*

Regarding the usually less frequent species worldwide that were examined in this study, the *B. cepacia* isolates (a total of 87) retrieved from five patients (N, AL, AP, AQ and AU) infected with this species for a period of time ranging from 2.2 to 9.6 years, all exhibit the OAg even when the FEV1 value was as low as 30% (patient N). Patients AL, AP and AQ, eradicated the infection after the screening period. Interestingly, the OAg banding pattern obtained for isolates with ribopattern 19 that infected patients AL and AP was identical to the OAg profile extracted from isolates retrieved from patients AU and D (Fig. 4.2, columns C1 and C3), the sole exception were the two isolates retrieved from patient AU during the third year of infection with a different RAPD profile (Figs. 4.1 and 4.2, columns C3 and C4, and Table 4.1). All these isolates producing the same banding pattern were confirmed to share the same RAPD profile (Table 4.1 and Fig. S4.1).

Concerning the 13 *B. contaminans* clonal isolates retrieved from patient V infected for a very long period of 15.2 years (Cunha et al., 2003; Coutinho et al., 2011b), all of them were found to express the OAg (Fig. 4.2, column D1).

Regarding the 17 *B. stabilis* isolates retrieved from patients B and H for a period of 3.8 and 2.5 years of chronic infection, respectively (Fig. 4.1 and Table 4.1), the two strains isolated from each patient exhibit the same RAPD profiles (Fig. S4.1, RAPD14 and RAPD15, Table 4.1). Remarkably, a stable phenotype concerning the presence or absence of the OAg over the infection period was found. The four isolates with RAPD15 profile retrieved from patient B exhibit an LPS that lack the OAg, independently of being early or late isolates, while the 4 clonal isolates of the strain with the same RAPD profile obtained from patient H, all have the OAg (Table 4.1 and Fig. 4.2, columns G1 and G4). Interestingly, patient H died from the cepacia syndrome while patient B death was not related with the cepacia syndrome. Isolates with RAPD14 profile do have an LPS that stably maintained the OAg. When present, the banding patterns of the OAg were different for the two strains but identical for the clonal variants of each strain obtained from both patients (Fig. 4.2, columns G2-G4).

#### 4.4.3. The LPS from early and late *B. dolosa* isolates lack the OAg

The analysis of 14 *B. dolosa* isolates obtained from a single patient (AR) chronically infected with this species for 5.5 years showed that this strain did not have an LPS with the OAg chain from the beginning of the infection until dead (Fig. 4.2 – panel F). The very peculiar *B. dolosa* LPS chemical structure of the first isolate IST4208 obtained from patient AR (Table S4.1) was examined before and a novel complete structure of the lipooligosaccharide (LOS) lacking the OAg component was revealed (Lorenzo et al., 2013). This LOS was found to show a strong proinflammatory activity (Lorenzo et al., 2013).

## 4.5. Discussion

Different virulence factors have been described in *Burkholderia cepacia* complex (Bcc) bacteria (Loutet and Valvano, 2010; Sousa et al., 2017) but less is known concerning their adaptive traits during chronic infection, compared with *P. aeruginosa*, the most prevalent and studied CF pathogen (Valentini et al., 2018). Before the systematic retrospective-longitudinal-screening, carried out in the present study, a retrospective phenotypic longitudinal assessment of the variation of mucoid exopolysaccharide production (Zlosnik et al., 2008), the switch from mucoid to non-mucoid (Zlosnik et al., 2011) and the swimming motility (Zlosnik et al., 2014) was performed during chronic infections involving numerous Bcc isolates representing different Bcc species and collected from 100 CF patients in the Vancouver area. Conversion of the mucoid-to-non-mucoid phenotype during chronic infection strengthen the concept of the association of the non-mucoid phenotype with the severity of the disease and the mucoid phenotype with persistence (Zlosnik et al., 2008; Zlosnik et al., 2011). This behavior is in contrast with the phenotypic switch from non-mucoid to mucoid in *P. aeruginosa* chronic infections as a well-established paradigm for infection disease severity and persistence in this species (Valentini et al., 2018). Moreover, upon establishment of chronic infection, subsequent *P. aeruginosa* isolates show a reduction in the swimming ability to be non-motile (Valentini et al., 2018), contrasting with the observed data in Bcc infections (Zlosnik et al., 2014). The understanding of the occurrence of these and other adaptive traits that have been well-described in *P. aeruginosa* chronic infections (Valentini et al., 2018), but are far less known in Bcc chronic infection, is instrumental for a better understanding of bacterial transformations in the CF lungs during long term infections and infection control.

The LPS forms the outer leaflet of the outer membrane and has roles in antigenicity, the inflammatory response and exclusion of external molecules (King et al., 2009). Although the lipid A can vary, the variability in the O-antigen (OAg) is greater influencing adherence, colonisation and the ability to evade the host's defense mechanisms (Kintz and Goldberg, 2008; King et al., 2009; Saldias et al., 2009; Kotrange et al., 2011; Maldonado et al., 2016; Ranf, 2016; Kintz et al., 2017). The ability of Bcc species to evade the host immune response often leads to chronic infections that are associated with significant loss of lung function due to the hyperactive inflammatory response that ultimately can lead to fatalities. The systematic, comprehensive and retrospective longitudinal screening conducted in this work, covering multiple sequential isolates from different Bcc strains of species, frequently or rarely isolated from CF patients, to search for the eventual lack of OAg expression in the late-stage of infections, reinforce the relevance that has been attributed to this switch. This was particularly frequent and clear for *B. cenocepacia recA* lineage IIIA and *B. multivorans*, also occurring at lower frequency in *B. cenocepacia recA* lineage IIIB. However, in the case of *B. multivorans*, there is a late isolate obtained within the series of sequential cloned isolates lacking the OAg that was found to express it. This behavior was previously observed for *B. multivorans* by others (Silva et al., 2016).

For the first time, it was shown that *B. cepacia* and *B. contaminans* do not lose the OAg even during long term infections that last for periods of time as extended as 9.6 and 15.2 years, respectively. Among the less frequent species, *B. contaminans* is currently considered an emerging CF pathogen (Vanlaere et al., 2009; Martina et al., 2013; Coutinho et al., 2015; Medina-Pascual et al., 2015; Nunvar et al., 2016; Power et al., 2016). Concerning the other two rarely found species examined, *B. stabilis* exhibits in two infected patients a stable phenotype concerning the presence or absence of the OAg over the infection period, consistent with the description of *B. stabilis* sp. Nov, as having a relatively stable genome (Vandamme et al., 2000). For the single patient chronically infected for 5.5 years with the more distantly related Bcc species *B. dolosa*, the OAg chain was absent from the beginning of the infection until the patient death. The novel LPS structure of this *B. dolosa* strain was examined and its strong pro-inflammatory activity reported before (Lorenzo et al., 2013).

The conclusions from this work were in general supported by the results gathered and suggest that the observed switch frequency of the OAg expression appears to be species dependent and also dependent on the duration of the infection and the level of deterioration of the lung function, increasing with both. This observation is consistent with the fact that cystic fibrosis is a genetic disorder associated with inflammation, sub-optimal antioxidant protection and the continuous use of antimicrobial therapy, all resulting in marked oxidative stress (Galli et al., 2012) that causes an increase in *mutation* rate. Results obtained from co-infected patients with different Bcc species reinforce those conclusions as it is for example the case for patient AF who was infected with *B. contaminans* isolates that kept the OAg during 5.8 years of chronic infection while late clonal isolates of the co-infecting *B. cenocepacia* IIIA strain during a much shorter period of 2.5 years lost the OAg.

It is interesting to note that the two Bcc species that showed a higher tendency to lose the OAg expression along chronic lung infections in CF (*B. cenocepacia* and *B. multivorans*) are also the two most prevalent worldwide and the most feared among the CF community (Jones et al., 2004; Mahenthalingam et al., 2005). Remarkably, *B. cenocepacia* and *B. multivorans* cell lysates are more potent stimulators of pro-inflammatory cytokines compared with those from other Bcc species (De Soyza et al., 2004). However, even though the prevalence of *B. multivorans* has recently replaced *B. cenocepacia* in several countries (Baldwin et al., 2008; Lipuma, 2010; Medina-Pascual et al., 2012; Peeters et al., 2017), it is clear that *B. cenocepacia* *recA* lineage IIIA remains the most dominant species (Manno et al., 2004; Nunvar et al., 2017; Scoffone et al., 2017; Teri et al., 2018). Remarkably, the particularly destructive nature of *B. cenocepacia* *recA* group IIIA infections in CF patients when compared with other Bcc bacteria has been reported (Manno et al., 2004; Zlosnik et al., 2015).

Remarkably, the single clone of the more distantly related *B. dolosa* Bcc species tested did not exhibit the OAg chain from the beginning to the end of the 5.5 year-infection as reported before

(Lorenzo et al., 2013). Comprehensive comparative genomic studies performed during chronic infection on the most feared *B. cenocepacia* (Nunvar et al., 2017) and *B. multivorans* (Silva et al., 2016) and on the far less distributed *B. dolosa*; (Lieberman et al., 2011) revealed that the genes affected during adaptive evolution in *B. cenocepacia* or in *B. multivorans* are different from *B. dolosa*. This indicates that evolution in *B. dolosa* during chronic CF infection is driven by different selective forces presumably linked to host immune responses (Lieberman et al., 2014; Nunvar et al., 2017). Specifically, while the most mutated genes reported to undergo adaptive within-patient evolution in *B. cenocepacia* were associated with oxidative stress response and transition metal metabolism, the same was not observed either in *B. dolosa* or in *Pseudomonas aeruginosa* (Lieberman et al., 2014; Nunvar et al., 2017). These genes encode proteins required for protection against the reactive oxygen species (ROS) produced by leukocytes thus suggesting the involvement of the host immune system in driving *B. cenocepacia* evolution during chronic CF infection (Nunvar et al., 2017). Given that persistent inflammation and neutrophil infiltration often accompany chronic lung infections, the authors hypothesized that under increased stress encountered in CF macrophages, the global stress response might be activated by the evolved bacterial population and thus modulate the course of infection (Chua et al., 2016; 2017; Nunvar et al., 2017). Since the interactions between pathogens and the immune system are highly complex, affecting pathogen adaptation to different host immune system stimuli, further investigation into the roles of macrophages and defense mechanisms in chronic infection outcome is needed (Chua et al., 2016; 2017).

The presence/absence of the OAg at the Bcc bacteria surface is important for cell evasion from the host immune response and pathogenicity. The OAg is considered highly immunogenic and induces the production of antibodies that may activate the complement pathway, either through the classic pathway or an alternate pathway, which leads to cellular death or phagocytosis (Reyes et al., 2012). Certain modifications in the oligosaccharide chain of the OAg may alter the interaction of the complement pathway. Several OAg of pathogens are similar to host molecules and this facilitate invasion through mimicking in the host (Reyes et al., 2012). An OAg deficient *B. cenocepacia* strain was found to be more susceptible to phagocytic internalization (Saldias et al., 2009) while, the loss of OAg expression in *P. aeruginosa* during chronic pulmonary infection did not confer phagocytic resistance *in vitro* (Demirdjian et al., 2017). Given that *Burkholderia* is a non-obligate intracellular pathogen while *P. aeruginosa* is considered an extracellular pathogen, it was hypothesized that the loss of OAg may benefit the ability of *Burkholderia* to access an intracellular environment, having no similar benefits for *P. aeruginosa* (Saldias et al., 2009). Recent studies suggest that there is a high selective pressure on the OAg biosynthetic locus leading to alterations both at the structural, genetic and regulatory levels strongly suggestive of an adaptive mechanism that potentially contribute to evade the host immune defenses in several Gram-negative pathogens (Murray et al., 2003; Yang et al., 2011a; Maldonado et al., 2016; McCarthy et al., 2017; Faure et al., 2018; Neiger et al., 2019). Moreover, the OAg absence

was shown to increase Bcc survival in eukaryotic cells as amoebae, epithelial cells and human macrophages (Saldias et al., 2009; Saldias and Valvano, 2009; Kotrange et al., 2011; Maldonado et al., 2016), leading to increased internalization of *B. cenocepacia* into macrophages upon phagocytosis (Saldias et al., 2009; Kotrange et al., 2011) and to facilitate *B. multivorans* growth inside macrophage (Schmerk and Valvano, 2013). These results may suggest that OAg loss could promote Bcc persistence through intracellular survival. Remarkably, the comparison of the ability to subvert the host's immune function, assessed by internalization assays using human dendritic cells, of three of the *B. cenocepacia* isolates examined in this work and retrieved from CF patient J showed that the late variants, IST4113 and IST4134, were significantly more internalized exhibiting increased survival within dendritic cells than the early isolate IST439 (Cabral et al., 2017). This early isolate was the only one that expresses the OAg unit (Hassan et al., 2017) supporting the idea that the loss of the OAg is advantageous for Bcc persistence. In addition, the absence of the OAg in these same isolates was suggested to reduce *B. cenocepacia* virulence potential using the *Galleria mellonella* infection model (Moreira et al., 2017). Such decrease in the virulence potential of early *B. cenocepacia* clonal variants exhibiting the OAg compared with late clonal variants lacking the OAg was also observed for two other chronically infected patients (AB and AN) with *B. cenocepacia* (Moreira et al., 2017). Also, a *Salmonella typhimurium* mutant missing the entire OAg was found to be avirulent in *G. mellonella* infection model while the shortening of the OAg chain length reduced the pathogenic potential by one-half compared to wild-type strain (Bender et al., 2013) suggesting that the OAg length is also a key-determinants of virulence in *G. mellonella*. The OAg truncation has also been implicated in increased neutrophil-mediated killing, complement-mediated susceptibility and phagocytosis of *Salmonella* by macrophages (Murray et al., 2003).

It is essential to obtain comprehensive knowledge about the pathogenesis of the various Bcc species involved in acute or chronic infections in different patients in order to fight the associated infections. This work reinforces the relevance that has been attributed to the switch involved in the lack of expression of the OAg during chronic infection in the most feared pathogenic Bcc species *B. cenocepacia* (in particular the *recA* lineage IIIA) and *B. multivorans*. It also suggests that the evolution of different Bcc species in chronic CF infection also at the level of OAg expression may be driven by different selective forces presumably associated to host immune responses.

## 5 Variation of *Burkholderia cenocepacia* cell wall morphology and mechanical properties during cystic fibrosis lung infection, assessed by atomic force microscopy

---

**This chapter contains results published in:**

A. Amir Hassan<sup>1,2</sup>, Miguel V Vitorino<sup>3,4</sup>, Tiago Robalo<sup>3</sup>, Mário S Rodrigues<sup>3,4</sup> and Isabel Sá-Correia<sup>1,2</sup>, **Variation of *Burkholderia cenocepacia* cell wall morphology and mechanical properties during cystic fibrosis lung infection, assessed by atomic force microscopy. (2019), *Scientific Reports* 9(1), 16118. doi: 10.1038/s41598-019-52604-9.**

<sup>1</sup> iBB – Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, 1049-001, Portugal

<sup>2</sup> Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, 1049-001, Portugal

<sup>3</sup> BioISI – Biosystems and Integrative Sciences Institute, Faculdade de Ciências, Universidade de Lisboa, 1749-016, Lisboa, Portugal

<sup>4</sup> Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, 1749-016, Lisboa, Portugal



## 5.1. Abstract

The influence that *Burkholderia cenocepacia* adaptive evolution during long-term infection in cystic fibrosis (CF) patients has on cell wall morphology and mechanical properties is poorly understood despite their crucial role in cell physiology, persistent infection and pathogenesis. Cell wall morphology and physical properties of three *B. cenocepacia* isolates collected from a CF patient over a period of 3.5 years were compared using atomic force microscopy (AFM). These serial clonal variants include the first isolate retrieved from the patient and two late isolates obtained after three years of infection and before the patient's death with cepacia syndrome. A consistent and progressive decrease of cell height and a cell shape evolution during infection, from the typical rods to morphology closer to cocci, were observed. The images of cells grown in biofilms showed an identical cell size reduction pattern. Additionally, the apparent elasticity modulus significantly decreases from the early isolate to the last clonal variant retrieved from the patient but the intermediary highly antibiotic resistant clonal isolate showed the highest elasticity values. Concerning the adhesion of bacteria surface to the AFM tip, the first isolate was found to adhere better than the late isolates whose lipopolysaccharide (LPS) structure loss the O-antigen (OAg) during CF infection. The OAg is known to influence Gram-negative bacteria adhesion and be an important factor in *B. cenocepacia* adaptation to chronic infection. Results reinforce the concept of the occurrence of phenotypic heterogeneity and adaptive evolution, also at the level of cell size, form, envelope topography and physical properties during long-term infection.

## 5.2. Introduction

The Gram-negative opportunistic bacterial pathogens *Pseudomonas aeruginosa* and *Burkholderia cepacia* complex (Bcc) exhibit extensive genetic and phenotypic heterogeneity during persistent infection and evolution in the lungs of cystic fibrosis (CF) patients over the years (Govan et al., 2007; Lieberman et al., 2014; Winstanley et al., 2016; Lee et al., 2017). The molecular mechanisms underlying adaptation to the lung and genotypic and phenotypic diversification have been intensively studied in the more prevalent CF pathogen *P. aeruginosa* (Govan et al., 2007; Winstanley et al., 2016). However, Bcc lung infections in CF are highly feared because they are associated with poor prognosis and increased risk of death due to rapid lung function deterioration and, in certain cases to a necrotizing pneumonia, bacteraemia, and sepsis, (the cepacia syndrome) (Jones et al., 2004; Mahenthiralingam et al., 2005; Drevinek and Mahenthiralingam, 2010; Coutinho et al., 2011b).

During long-term lung infection in CF patients, *P. aeruginosa* and Bcc bacteria face multiple selective pressures in the highly challenging, fluctuating, and stressful environment of the patients' airways, in particular due to antimicrobial therapy, the action of the host immune system and of other members of the microbiome and the decrease of oxygen availability as the result of lung function deterioration (Döring et al., 2011; Cullen and McClean, 2015). Under those stresses, several genetic changes accumulate in the initial infecting bacterial strain leading to phenotype and genotype heterogeneity. CF bacterial pathogens phenotypic diversification can be recognized in terms of colony morphology diversity (Deretic et al., 1994; Drenkard and Ausubel, 2002; Zlosnik and Speert, 2010; Coutinho et al., 2011a; Silva et al., 2011; Zlosnik et al., 2011; Moreira et al., 2014) and variation of clinically relevant phenotypes such as antibiotic resistance (Leitao et al., 2008; Coutinho et al., 2011a; Moreira et al., 2014; Lopez-Causape et al., 2015; Rhodes and Schweizer, 2016), ability to form biofilms (Drenkard and Ausubel, 2002; Lee et al., 2005; Sriramulu et al., 2005; Fazli et al., 2014; Valentini et al., 2018), virulence potential (Zlosnik et al., 2011; Lorè et al., 2012; Madeira et al., 2013; Moreira et al., 2017), among many others (Mowat et al., 2011; Silva et al., 2011; Sousa et al., 2011; Warren et al., 2011; Moreira et al., 2014; Maldonado et al., 2016; Hill et al., 2017). Remarkably, such phenotypic heterogeneity within human hosts has important clinical implications. For example, antimicrobial susceptibility diversity within the bacterial population isolated from an individual sputum sample may affect the treatment of life-threatening infections given that the results from antimicrobial testing carried out on single isolates randomly collected can be a poor predictor of the clinical outcome of antibiotic therapy (Coutinho et al., 2011b; Lopez-Causape et al., 2015; Rhodes and Schweizer, 2016).

Bacterial cell envelope plays a central role in cell physiology and the alteration of surface properties can implicate the variation of phenotypes that play a crucial role in the pathogenesis of infectious diseases, such as antibiotic resistance and biofilm formation (Maldonado et al., 2016; Hill et al., 2017;

Trivedi et al., 2018). However, very few bacterial species have been on the focus of studies related to cell surface physical properties (Auer et al., 2016; Auer and Weibel, 2017; Trivedi et al., 2018) and information on the diversification and adaptive evolution at the level of Bcc bacteria cell wall mechanical properties during CF chronic lung infections is missing. In this context, over the last years atomic force microscopy (AFM) emerged as an essential tool for understanding the nanomechanics of live systems (Touhami et al., 2003; Muller et al., 2009; Costa et al., 2014). Hence, the objective of the present study was to obtain this knowledge by studying cell surface morphology and mapping the mechanical properties of *Burkholderia cenocepacia* clonal variants isolated from the lungs of a CF patient during long term infection using AFM. The *B. cenocepacia* isolates examined are from a collection of 11 serial clonal variants obtained from the same CF patient over a period of 3.5 years, from the onset of infection until the patient's death (Cunha et al., 2003; Coutinho et al., 2011a). The clonal variants tested were: IST439, the first isolate retrieved; IST4113, obtained three years later after an exacerbation with the patient hospitalization and treatment with intravenous therapy with gentamicin and ceftazidime and found to be highly resistant to different classes of antimicrobials; and IST4134, obtained 3 months later, just before the patient's death with cepacia syndrome (Cunha et al., 2003; Coutinho et al., 2011a; Mira et al., 2011). These isolates were picked at random from selective agar plates obtained in the major Portuguese CF Center at Hospital de Santa Maria during consultation routines. The clinical isolates examined are of high interest in the context of this study because they were previously characterized by phenotypic (Coutinho et al., 2011a), transcriptomic (Mira et al., 2011), proteomic (Madeira et al., 2011; Madeira et al., 2013) and metabolic profiling (Moreira et al., 2016). Results on the comparison of the virulence potential of these isolates using non-mammalian infection models and of their ability to modulate dendritic cell function are also available (Cabral et al., 2017; Moreira et al., 2017). The two late variants were found to have lost the ability to produce the OAg molecule of the lipopolysaccharide (Hassan et al., 2017) present in the early isolates and to be more internalized by dendritic cells and show improved survival within dendritic cells when compared to the initial isolate (Cabral et al., 2017). Inflammatory cytokines were highly expressed in all the sequential clonal isolates but this pro-inflammatory trait was more pronounced in dendritic cells infected with the late variants compared with the isolate retrieved at the first stages of infection (Cabral et al., 2017).

Results of the present study, in which AFM cell wall morphology and mechanical properties of these three sequential *B. cenocepacia* clonal variants were studied, reinforce the concept of the occurrence of phenotypic variation and adaptive evolution also at the level of cell size, form, envelope topography and physical properties during long-term infection.

## 5.3. Methods

### 5.3.1. Bacterial Strains and growth conditions.

The three *Burkholderia cenocepacia* clonal variants examined in this study (IST439, IST4113 and IST4134) were recovered, as part of the hospital routine, from the sputum of a CF patient under surveillance at the major Portuguese CF Center in the Hospital de Santa Maria, Centro Hospitalar Lisboa Norte (CHLN) EPE, from 1999 to 2002 (Cunha et al., 2003; Coutinho et al., 2011a; Moreira et al., 2017). Studies involving these isolates were approved by CHLN ' ethics committee and the anonymity of the patient was preserved. Informed consent was also obtained from all participants and/or their legal guardians. All the methods were performed in accordance with the relevant guidelines and regulations. Bacterial cultures are stored at -80°C in 1:1 (v/v) glycerol. Bacterial growth was carried out in Lysogeny Broth, Lennox (LB; Conda, Pronadisa), at 37°C and 250 rpm, or in LB agar plates obtained by supplementation of LB with 2% agar (Iberagar, Portugal). LB medium at 37°C was also used in biofilm experiments.

### 5.3.2. Preparation of the AFM samples

Bacterial cells used for AFM analysis were deposited onto the gelatin coated mica for the observations and measurements done in liquid environment and onto freshly cleaved mica surfaces for the observations done in air. *B. cenocepacia* isolates were cultured overnight in LB medium, at 37°C with shaking at 250 rpm, and then sub-cultured until mid-exponential phase. Bacterial planktonic cells in suspension were collected by centrifugation and washed three times with phosphate buffer saline (PBS). For observations and measurements in liquid environment, the gelatin-coated mica was prepared and the bacterial immobilization was done as described before (Liu and Camesano, 2008; Allison et al., 2011; Uzoechi and Abu-Lail, 2019a). Two types of gelatin with different concentrations (0.25%, 0.5% and 1% (w/v)), porcine gelatin Sigma G-6144 and G-2625 and bovine gelatin Sigma G-9382, were tested (Allison et al., 2011). Briefly, a gelatin solution was prepared by dissolving 0.25g, 0.5g or 1g gelatin in 100 ml of deionized water at 90°C and cooled to 60-70 °C prior to vertically dipping several discs of the freshly cleaved mica into the solution. Following optimization, gelatin G-6144 was found to allow the best immobilization effectiveness and used thereafter. The gelatin-G-6144-coated mica surfaces were supported on edge on a paper towel and then air dried overnight. 20-40 µl of the bacterial suspension in PBS ( $10^8$  CFU/ml) was applied onto a gelatin-coated mica surface after being sonicated in ultrasonic bath (40 kHz, 19W – Branson, Model 200, NL) for ~5 min (Liu and Camesano, 2008). The sample was allowed to rest for 10-20 min before it was rinsed in PBS and imaged in the liquid cell of the AFM.

For the images/observations taken in air, the deposition of bacteria on the mica surfaces was carried out as described before (Osiro et al., 2012; Saar-Dover et al., 2012) and used for other Gram-negative bacteria (Eaton et al., 2008; Turner et al., 2013; Yang et al., 2018), with few modifications. In brief and during the optimization of the protocol, PBS and deionized water were tested to prepare the bacterial suspension aliquots prior to deposition on mica. Given that the samples prepared with PBS formed aggregates at the freshly cleaved-mica surface, whereas those prepared with deionized water did not, 10  $\mu\text{l}$  of those aliquots in ddH<sub>2</sub>O, for observation done in air, with a final bacterial concentration of  $10^8$  CFU/ml, were immobilized onto freshly cleaved mica surfaces and left to rest for 15-20 min. The mica surfaces were rinsed twice with deionized water to detach the weakly adherent and the non-adherent cells and allowed to dry before AFM analysis for another 15-20 min. The biofilm samples were prepared for AFM observation as described before (Qin et al., 2009) with few modifications. The microtiter plates were incubated without shaking at 37°C for 12 hours. The unattached planktonic bacteria were washed twice with sterile saline solution by pipetting in and out. The remaining biofilms were fixed by 10 % formalin in PBS for 10 min. Next, the plates were inverted to remove all the fixation solution and the fixed biofilms were washed twice with sterile saline solution by pipetting in and out. Finally, the biofilm was resuspended in ddH<sub>2</sub>O, and 10  $\mu\text{l}$  of the suspended biofilm was deposited into the freshly cleaved micas and air-dried for AFM imaging.

### 5.3.3. AFM observations and measurements

Cell samples were analysed using a PicoSPM LE system of Molecular Imaging in a liquid cell containing PBS and in air at room temperature. Bruker MLCT-F microlevers with nominal cantilever stiffness of 0.6 N/m and nominal tip radius of 20 nm were used for all experiments. Images obtained in air were taken in contact mode whereas in liquid environment both contact and tapping modes were used depending on which proved best. To measure cell surface roughness  $2.5 \times 2.5 \mu\text{m}^2$  (approximately) images with  $512 \times 512$  pixels were obtained. Twelve representative bacteria of each isolate were selected for roughness measurements.

For nanomechanics and adhesion measurements in liquid environment force spectroscopy maps, consisting of  $32 \times 32$  approach/retract force-distance (FD) curves, were obtained over an area slightly larger than the cell footprint. The maximum cantilever deflection was set constant in all experiments, yielding a maximum applied force of roughly 15 nN. The tip-sample approach speed was also set constant, to 0.4  $\mu\text{m/s}$ . To reduce bias due to different cantilevers being used on different populations, each cantilever was used to measure 2-3 bacteria of each isolate population and the order in which the different isolates was measured was randomized (Wagner et al., 2011; Schillers et al., 2017). In total, about 40 bacteria of each population were measured in liquid environment and 14 cantilevers were used. Nanomechanical analysis was performed using a custom-made software. For each grid, we selected

only curves obtained at the top of the bacteria (above 85% of the total height of the bacteria), as shown in the inset of Fig. 5.3 – panels A-C. The contact part of the approach curves was analysed according to Sneddon contact model (Touhami et al., 2003) that establishes a relationship between load and indentation and from which it is possible to extract the reduced Young's modulus Fig. 5.3 – panel D. We have estimated the adhesion force, Fig. 5.3 – panel E, from the minimum of the retract part of the curves as illustrated in Fig. 5.3 – panel F. For each cell, all curves obtained at the top of the cell were analyzed and the median was kept, then for each isolate sample we rejected values further away from 3 standard deviations. To determine the apparent Young's modulus of the cell we assumed a non-deformable tip and a Poisson's ratio of 0.5 (Touhami et al., 2003; Butt et al., 2005). We assumed nominal values for the tip radius and used Sader method to calibrate the cantilever spring constant (Sader et al., 2016).

#### 5.3.4. Growth curves

The growth curves of the three clonal variants examined under aerobic and microaerophilic conditions were monitored by measuring culture optical density at 640 nm ( $OD_{640}$ ). Cells were grown in LB medium at 37 °C in shaking flasks (100 ml with 30 ml of liquid medium) in an orbital shaker at 250 rpm (for aerobic growth) or standing in a microaerophilic atmosphere, containing 5-8% oxygen and 12-15% carbon dioxide, generated in sealed jars using the GENbox microaerator (bioMérieux, Marcy L'Etoile, France). Results are from three independent growth experiments.

#### 5.3.5. Biofilm formation assays

Biofilm formation assays were based on a described methodology (Cunha et al., 2004; Coutinho et al., 2011a). Overnight liquid cultures of each CF isolate were transferred to LB medium and grown at 30°C with orbital agitation until the mid-exponential phase was reached. The cultures were subsequently diluted to a standardized culture  $OD_{640}$  of 0.5, and 20  $\mu$ l of this cell suspension was used to inoculate the wells of a 96-well polystyrene microtiter plate (Greiner Bio-One) containing 180  $\mu$ l of LB medium. Wells containing sterile growth medium were used as negative controls. Plates were incubated at 37°C from 4 to 48 h without agitation. For biofilm quantification, the culture media and unattached bacterial cells were removed from the wells by careful rinsing with water (three times, 200  $\mu$ l for each rinse). Adherent bacteria were stained with 200  $\mu$ l of a 1% crystal violet solution for 15 min at room temperature (50 ml of the solution was prepared by adding 1% [wt/vol] crystal violet in 10 ml of 95% ethanol to 40 ml of water containing 0.4 g of ammonium oxalate). After three gentle rinses with 200  $\mu$ l of water each time, the dye associated with the attached cells was solubilized in 200  $\mu$ l of 95% ethanol and the biofilm was quantified by measuring the absorbance of the solution at 600 nm ( $A_{600nm}$ ) in a microplate reader.

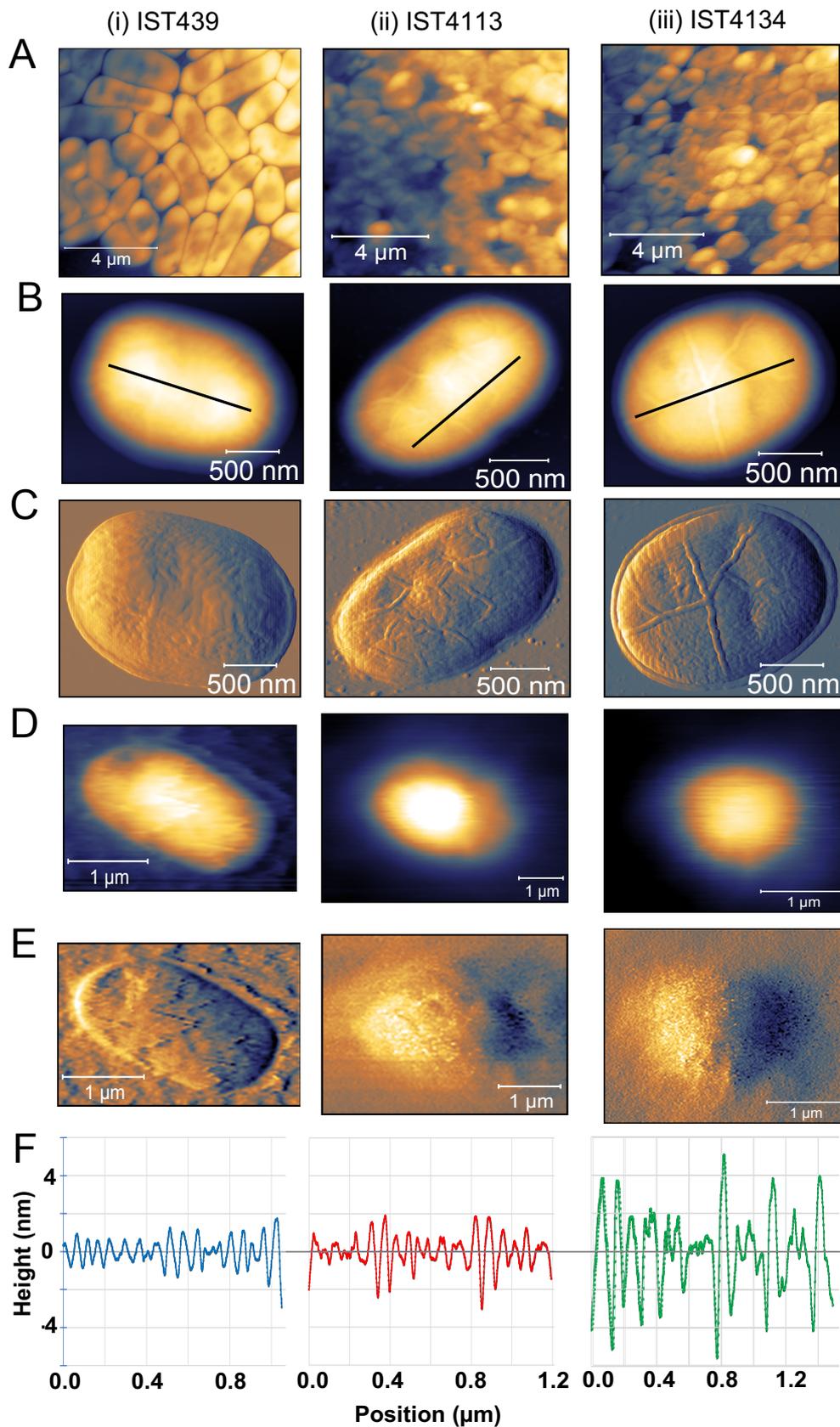
### 5.3.6. Statistics

A non-parametric Mann-Whitney u-test was used to determine statistical significance of the observed variations (GraphPad Prism 7; GraphPad Software, CA).  $P \leq 0.05$  was considered statistically significant.

## 5.4. Results

### 5.4.1. *B. cenocepacia* morphology and surface roughness evolution during long-term CF lung infection

The cell morphology and topography of the three *B. cenocepacia* clonal variants were examined using AFM in both planktonic and biofilm forms. The images for individual cells were obtained in air and in liquid environment. Concerning cell topography in air (Fig. 5.1), the easiest way to clearly visualize individual cells, the three clonal variants examined in their planktonic form exhibit the porous network architecture of the cell wall previously reported by others (Turner et al., 2013; Turner et al., 2016). However, late variants IST4113 and IST4134 show features, characterized by string-like formations on the surface of the bacteria, not found in the early isolate IST439. Specifically, the last isolate retrieved from the patient, IST4134, displays longer and well-organized string-like structures that span the entire length of the cell while IST4113 exhibits shorter and less organized structures (Fig. 1(b,c)). However, the same structures could not be observed in the biofilm images Fig. 5.1 – panel A. The surface roughness of the cells examined in air in their planktonic form was assessed by defining longitudinal cross sections along the cell surface, as shown in Fig. 5.1 – panels B and F. The first isolate, IST439, was found to be smoother, with average roughness (root mean square of the cross sections) of  $0.9 \pm 0.1$  nm. The late variants showed an average roughness of  $1.0 \pm 0.1$  nm and  $2.0 \pm 0.1$  nm respectively. Alternatively, we have also measured the roughness by taking the root mean square over an area defined on the top of the cell after subtracting the cell envelope. The resulting values for the first, second and third isolate were respectively 1.7 nm, 2.0 nm and 2.5 nm with 0.5 nm standard deviation. We find the first method less prone to errors because it is easier to separate the cell roughness from the cell contour/envelope - nonetheless, both methodologies indicate the same trend. The surface roughness from the images obtained in liquid was impossible to visualize due to the poor resolution of the images seemingly caused by some mobility of either the cell or the cell surface. We have used both contact and tapping mode but the results were similar.



**Figure 5.1 | Cell topography and surface roughness.** AFM images of the studied clonal variants (i - IST439, ii - IST4113 and iii - IST4134) show (A) topography of biofilm; (B) topography of planktonic and (C) corresponding deflection images in air; (D) topography of planktonic and

(E) corresponding deflection in liquid samples; (F) representative roughness profiles (marked by the black lines) for the same variants. The late variants exhibit distinct string-like surface features, whereas both the images and roughness profile of the IST439 variant show a smoother surface.

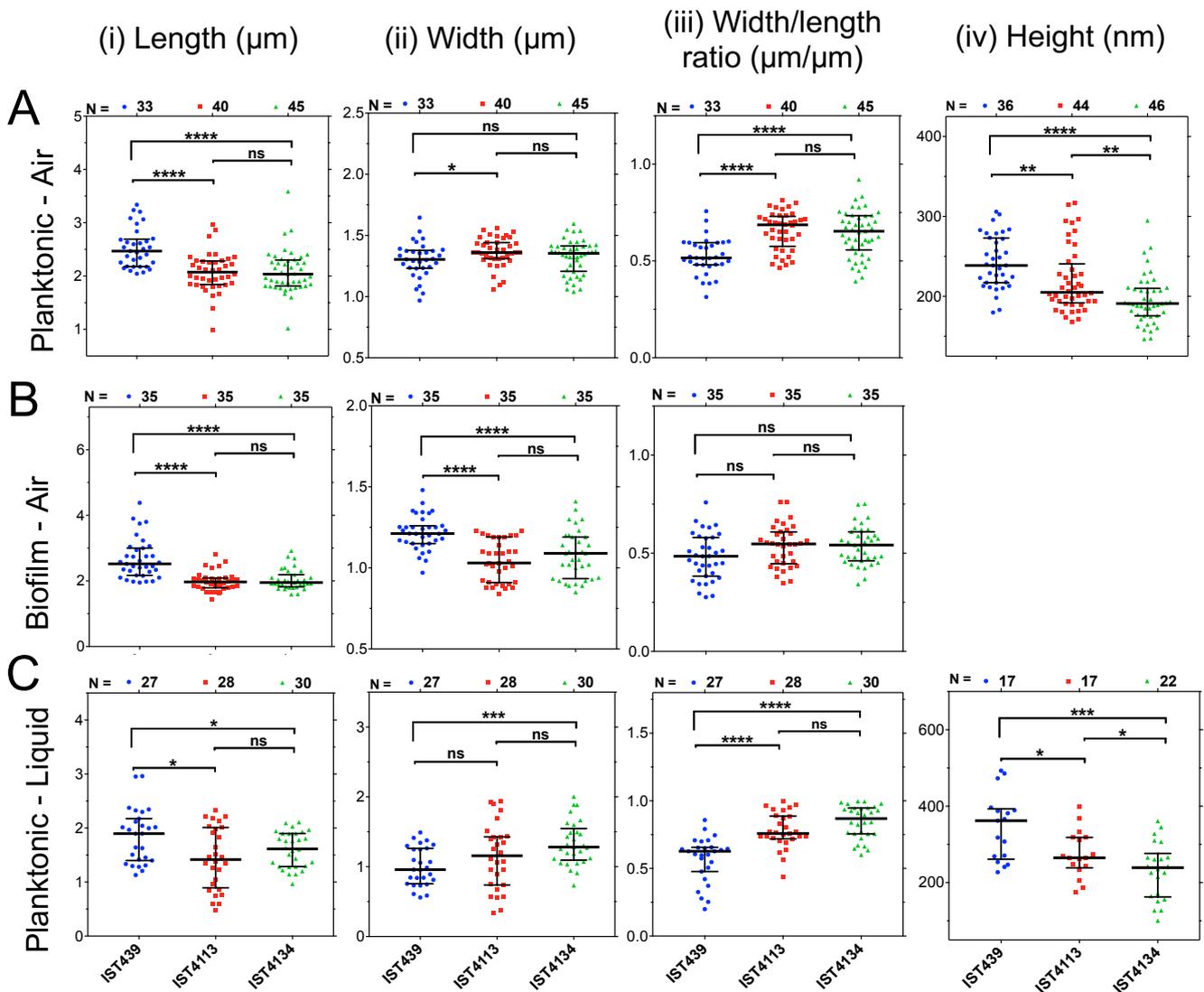
Concerning cell dimensions measurements, the late variants, IST4113 and IST4134, were found to be significantly smaller in length than the early isolate IST439, but no significant differences were found between the lengths of late variants, as shown on Fig. 5.2 – column (i). Similar significant differences of the cell length were observed both in air and liquid, however the absolute values registered in liquid environment were smaller than those in air (Fig. 5.2 – column (i)). Additionally, either in air or in liquid environment, the height of the cells also decreases from the first to the last clonal isolate by about 30% (Fig. 5.2 – column (iv)). Overall, the observed trend points towards an evolution of the cell shape from rod to a more rounded shape, with the ratio width/length increasing from around 0.52/0.62 (air/liquid) for the early isolate to about 0.68/0.75 (air/liquid) and 0.65/0.86 (air/liquid) for the second and third isolates, respectively (Fig. 5.2 – column (iii)). The AFM images of the biofilm (Fig. 5.1 – panel (A)) show cell dimensions consistent with the reduction of the cell size and increase of the ratio width/length (Fig. 5.2 – column (iii, panel B)) observed for the individual cells. However, due to the fact that cells are tightly packed it is more difficult to accurately determine cell dimensions in biofilm.

#### 5.4.2. Surface and mechanical properties evolution during long-term CF lung infection

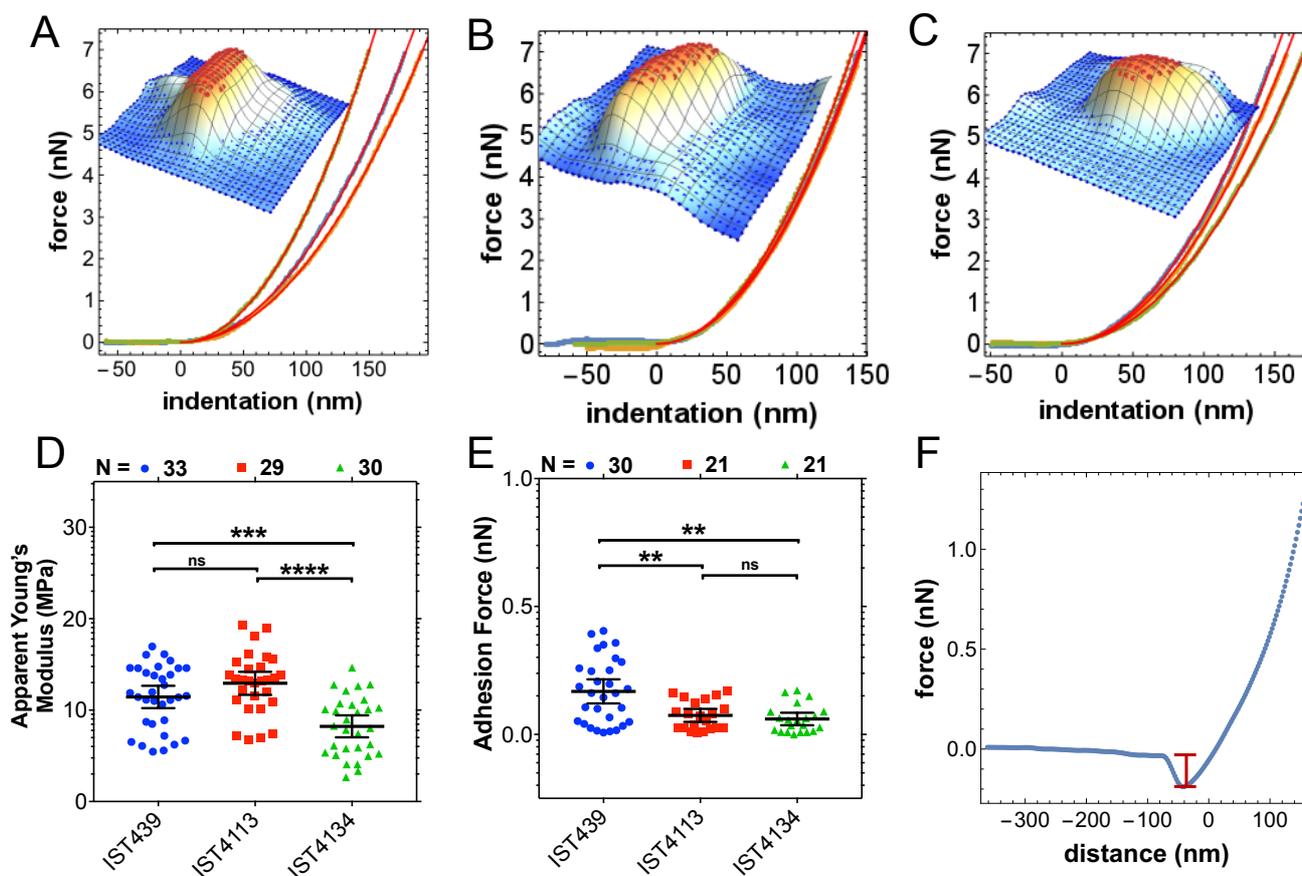
Cell elasticity was examined at selected points (inset of Fig. 5.3 – panels A-C) along the cell surface and the resulting force-distance curves were compared with the Sneddon contact mechanics model (Fig. 5.3 – panels A-C), as described in the Methods section. All these measurements were performed in liquid environment because the measurement of the mechanical properties of dried cells is questionable (Eaton et al., 2008; Cerf et al., 2009; Deng et al., 2011; Francius et al., 2011; Elbourne et al., 2019). In fact, when measurements were made in air, the apparent Young's modulus of the cell surface was about 40 times larger.

Concerning the elasticity, and in particular the apparent Young's modulus of the cell surface measured in liquid environment, the values show a significant decrease from the early isolate to the late variant (Fig. 5.3 – panel D), but the highly antibiotic resistant clonal variant IST4113 (Coutinho et al., 2011a) exhibited the highest values. Concerning the adhesion of the bacteria surface to the Si<sub>3</sub>N<sub>4</sub> AFM tip (Fig. 5.3 – panel E), the first isolate was found to adhere better than the late isolates whose lipopolysaccharide (LPS) structure loss, during CF infection, the O-antigen (OAg) present in the early isolate (Hassan et al., 2017). No significant differences were found between the adhesion of the two

late isolates missing the OAg, Fig. 5.3 – panel D). Additionally, we have spatially characterized the adhesion force for the IST4113 and IST4134, Fig. S5.1, to study the impact of the above mentioned string-like structures on the adhesion. These measurements showed that the string-like formations of these variants have less adhesion to the AFM tip than the rest of the bacterial surfaces, however, it is not easy to validate their exact impact in the complex bacterial adhesion.



**Figure 5.2 | Cell morphology.** Distribution of the cell dimensions (i - length, ii - width, iii - width/length ratio and iv - height) measured with AFM for N individual cells of the studied clonal variants (blue circle - IST439, red square - IST4113 and green triangle - IST4134), for (A) planktonic and (B) biofilm growth conditions both measured in air and for the (C) planktonic form measured in liquid environment. An increase of the width/length ratio and a decrease of cell height, as well as the evolution of the cell shape during long-term infection from a rod-like to a more cocci-like morphology can be observed. The results of the Mann-Whitney u-test (\*  $P < 0.05$ , \*\*  $P \leq 0.005$ , \*\*\*  $P \leq 0.0005$ , ns not significant) are indicated.

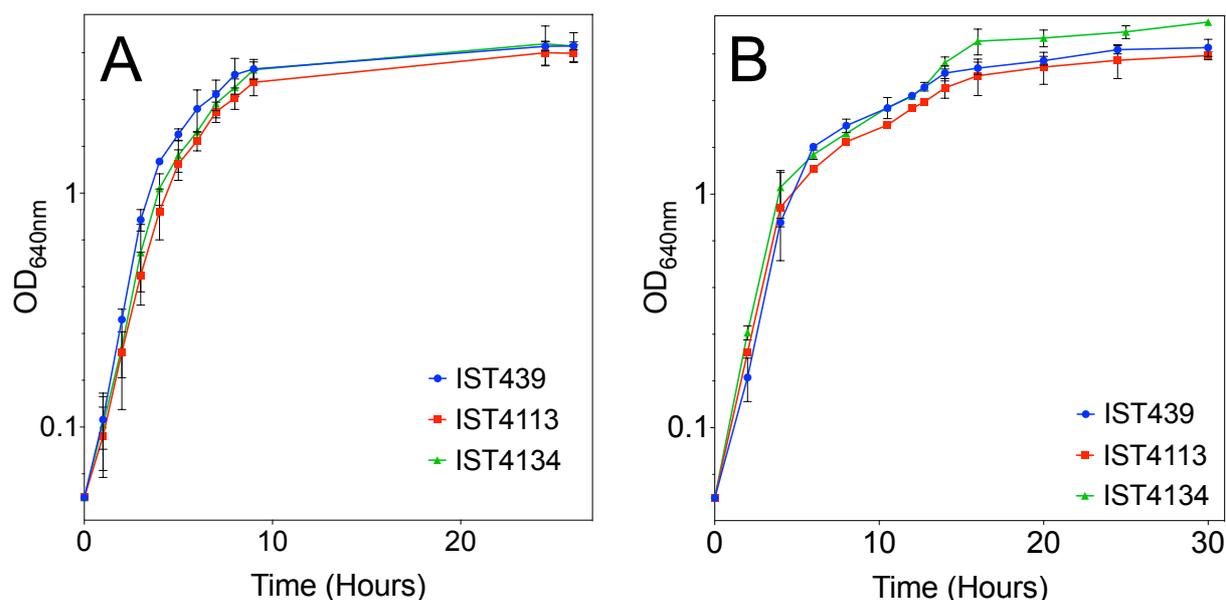


**Figure 5.3 | Elasticity and adhesion studied by AFM in liquid samples.** (A-C) Example of indentation curves and respective fits using Sneddon model for the first, second and third clonal variants with insets showing the respective 3D maps and the selection of points at the top of the cell; (D and E) distribution of young modulus and adhesion forces respectively, for each isolate; (F) a retract curve obtained for the first isolate illustrating the adhesion measurement. The results of the Mann-Whitney u-test (\*  $P < 0.05$ , \*\*\*  $P \leq 0.001$ , \*\*\*\*  $P \leq 0.0005$ , ns not significant) are indicated.

#### 5.4.3. Growth curves of the *B. cenocepacia* clonal variants under aerobic or microaerophilic conditions

The growth curves of the three clonal variants examined were compared in the same Lysogeny broth (LB) medium under aerobic and microaerophilic conditions, at 37°C (Fig. 5.4). The general conclusion is that under aerobic conditions, the growth performance of the first isolate is slightly better (higher specific growth rate and higher final biomass concentration attained) than the late isolates, with the highly antibiotic resistant intermediary isolate (IST4113) exhibiting the slowest and less efficient growth. This behaviour contrasts with the growth performance observed under microaerophilic

conditions which are conditions closer to those expected to occur in the CF patient lung, especially at late stages of disease progression and very low values of Forced expiratory volume in one second (FEV1) (Correia et al., 2008). Although the reported differences are small, the early isolate consistently exhibited, under oxygen limitation, the lowest specific growth rate while the last isolate showed the more rapid growth and efficient biomass production suggesting that the late isolates are better adapted to the CF lung. Moreover, under microaerophilic conditions, the growth curves of the 3 isolates exhibit a pattern consistent with diauxic growth in the complex LB medium, a behaviour observed before (unpublished data).

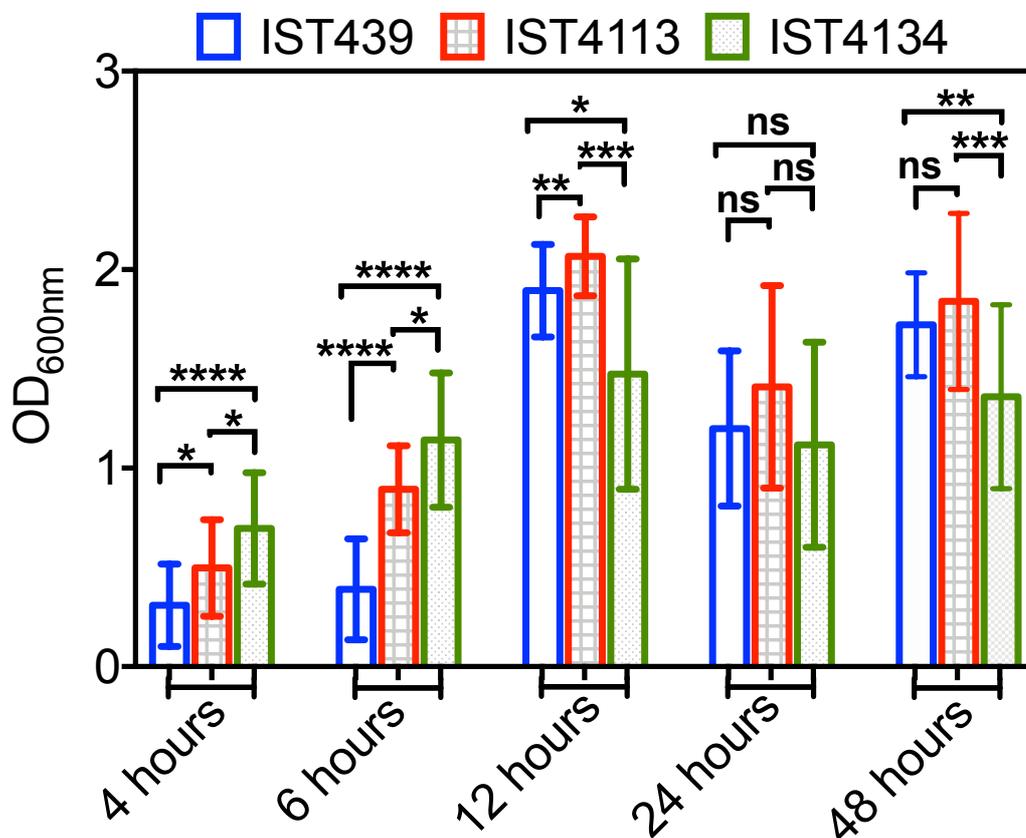


**Figure 5.4 | Growth under oxygen limitation.** Evolution of the optical density values during the incubation time in the same medium (LB) for the three clonal variants examined under (A) aerobic condition and (B), microaerophilic conditions. Experimental values and error bars represent the mean and the estimated standard deviation, respectively, for three independent growth experiments. These results indicate that the late clonal variant grows better under microaerophilic conditions whereas the early IST439 grows better under aerobic conditions.

#### 5.4.4. Biofilm growth of *B. cenocepacia* clonal variants

The quantification of biofilms in terms of biomass using crystal violet staining after 4 and 6 hours of incubation is consistent with the specific growth rates and final biomass attained by the three clonal variants, grown under microaerophilic conditions in the planktonic lifestyle (Figs. 5.4 and 5.5). In fact, the late isolates produce more rapidly immature biofilms of higher biomass, formed after 4-6 hours. However, the relative biomass of the mature biofilms formed after 24-48h of incubation is consistent

with the level of exopolysaccharide produced by the three variants, as reported before for the same isolates/growth medium (Coutinho et al., 2011a), being maximal for the intermediary isolate.



**Figure 5.5 | Biofilm formation.** Quantification of the biofilm formed at different incubation times, based on the OD<sub>600nm</sub> values of crystal violet-stained biomass. Bar height and error bars represent the mean and standard deviation obtained from three independent growth experiments with 7 measurements each. The Mann-Whitney u-test (\*  $P < 0.05$ , \*\*  $P \leq 0.001$ , \*\*\*  $P \leq 0.0005$ , \*\*\*\*  $P \leq 0.0005$ , ns not significant) values are indicated.

## 5.5. Discussion

During long term infection, the genetic adaptation of bacteria of the *Burkholderia cenocepacia* complex (Bcc) to the challenges of multiple selective pressures occurring in the cystic fibrosis (CF) airways is known to take place (Lieberman et al., 2011; Lieberman et al., 2014; Silva et al., 2016; Lee et al., 2017; Nunvar et al., 2017). The present study provides the first insights into the adaptive evolution of these bacteria at the level of cell size, form, envelope topography and physical properties during long-term infection. However, it cannot be guaranteed that the properties reported here are identical to those exhibited by bacteria grown in the lungs and measured in their native environment. Given the highly

relevant information gathered over the years on the three sequential *B. cenocepacia* variants examined here, it is possible to speculate on how the variation in surface properties can impact phenotypes of clinical relevance in the pathogenesis of infectious diseases, such as adhesion, resistance to antibiotics, biofilm formation and growth efficiency in the CF environment.

Former studies have shown that in *B. cenocepacia* the O-antigen (OAg) of the lipopolysaccharide (LPS), occurring in the outermost layer of the cell, is lost or modified during persistent infection of the lungs (Maldonado et al., 2016; Hassan et al., 2017). The loss or modification of the OAg appears to play an important role during the infection process, in particular in the colonization step (adherence) and ability to overcome host defence mechanisms (King et al., 2009; Cunneen et al., 2011; Maldonado et al., 2016). Recently, we have shown that the tendency of the most prevalent and feared species *B. cenocepacia* and *B. multivorans* to lose the OAg during chronic infection is higher than the one of the rarely found *B. cepacia* and *B. contaminants* that keep the OAg even during decades of infection (Hassan et al., 2019). Moreover, *B. cenocepacia* *recA* lineage IIIA strains, as it is the case of the isolates examined in the present study known to lead to particularly destructive infections, exhibit the most frequent OAg loss, compared with lineage IIIB (Hassan et al., 2019). Concerning the clonal isolates tested in the present study, the OAg is only present in the early isolate IST439 (Hassan et al., 2017). Mutations within the OAg cluster of the serial isolates examined were identified but none of them could definitely be associated to OAg loss (Hassan et al., 2017). In the present study, we found that the ability of the earlier isolate to adhere to the AFM Si<sub>3</sub>N<sub>4</sub> tip, when assessed in liquid environment, was significantly higher than the estimated ability of the late variants, lacking the OAg whose adhesion values were similar. These results are consistent with the concept that the variability in the LPS OAg affects bacterial adherence and colonisation and the ability to evade the host's defence mechanisms being an important factor in *B. cenocepacia* adaptation to chronic infection (Saldias et al., 2009; Maldonado et al., 2016; Ranf, 2016; Hassan et al., 2017; Hassan et al., 2019).

In agreement with the characterization of Gram-negative sacculus (Gan et al., 2008; Turner et al., 2014), images in air have clearly revealed tube-like features not regularly spaced mostly in the plane of the sacculus and roughly perpendicular to the long axis of the cell, for the two late isolates, but not for the early isolate. These formations are short and organized randomly in IST4113 cell surface but well developed in IST4134. These features do not seem to be associated to the division septum and its equatorial rings, as previously reported in studies in which the sacculus of *Escherichia coli* and *Bacillus subtilis* (Yao et al., 1999; Bramkamp and van Baarle, 2009), of *Streptococci* and *Enterococci* (Wheeler et al., 2011; Dover et al., 2015), and *Staphylococcus aureus* (Turner et al., 2010) was analysed. With very few exceptions, the chemical and biological bacterial components studied that contribute to cell mechanics are related with the peptidoglycan layer of the cell envelope and changes in its structure (Persat et al., 2015; Auer and Weibel, 2017; Trivedi et al., 2018). Apparently, the peptidoglycan is made

up of circumferential oriented bands of material interspersed with a more porous network (Turner et al., 2013). Peptidoglycan is the largest component of the bacterial cell wall determining the shape and preserving its integrity (Turner et al., 2014). Its elastic nature helps withstand stretching forces caused by bacterial turgor pressure. The reported bands were proposed to define regions with different availability for insertion of new peptidoglycan (Turner et al., 2013; Turner et al., 2016). Interestingly, indentation tests performed in liquid show that the surface elasticity modulus, decreases significantly from the early isolate to the last clonal variant. However, the highly antibiotic resistant intermediary isolate IST4113 (Coutinho et al., 2011a) expressed maximal values. Higher rigidity and increased elasticity was recently reported to be associated with a lower outer membrane permeability which may lead to the reduction of antibiotic diffusion into the cells (Uzoechi and Abu-Lail, 2019a).

The suggested remodelling of cell surface of the three clonal variants examined in this study during long term infection was already anticipated based on the results of the comparison of genomic expression of these same clonal variants using transcriptomic and quantitative proteomic analyses (Madeira et al., 2011; Mira et al., 2011; Madeira et al., 2013). These studies have shown differences in the level of expression of genes/proteins involved in the biogenesis of cell envelope and outer membrane in the three variants, among the several hundred of genes found to be differentially transcribed in the late isolates compared to the early isolate. These genome-wide expression results reflect a marked reprogramming of genomic expression at different levels (Madeira et al., 2011; Mira et al., 2011; Madeira et al., 2013), including the alteration of bacterial cell surface that contributes to the intrinsic and acquired resistance of Bcc bacteria to antibiotics. Remarkably, it was found that the late isolates are significantly more resistant to a wide range of antibiotics, with isolate IST4113 displaying the higher resistance levels (Coutinho et al., 2011a; Coutinho et al., 2011b; Mira et al., 2011; Madeira et al., 2013). Recent genomic studies strengthened the concept that cell wall remodelling relates with the alteration of bacterial mechanical properties (Auer et al., 2016; Trivedi et al., 2018). For example, *E. coli* mutants deleted for genes encoding proteins associated with cell-wall synthesis exhibit different stiffness defects<sup>35</sup> and the accumulation of the peptidoglycan *D*-Alanine residues is tightly regulated in *P. aeruginosa* since their accumulation reduces peptidoglycan cross-linking and cell stiffness (Trivedi et al., 2018).

In this study, we clearly observed a consistent and progressive pattern of decrease of the height and the increase of the width/length ratio of *B. cenocepacia* cells during long term infection, both in air and in liquid environments. Although the *B. cenocepacia* clonal isolates examined may not be representative of the expected population heterogeneity present at each isolation time in the CF lung, the consistency of the pattern strongly suggests that *B. cenocepacia* underwent convergent evolution towards the minimization of bacterial size during infection. Moreover, besides the decrease of the size of *B. cenocepacia* cells during infection, the bacterium underwent a cell shape evolution from the typical rod

form of the species to a cell morphology closer to the spherical form of cocci. The referred pattern was observed for both the planktonic and biofilm growth mode. This same pattern was described before for two nasopharyngeal bacterial pathogens during adaptation to human mucosa and the authors hypothesized that this transition was selected to reduce the cell surface sensitivity to immune attacks given that the ratio surface/volume is smaller than that of bacilli (Veyrier et al., 2015). Consistent with this hypothesis, several studies have shown the relevance of cell surface size when bacterial cells are facing immune attacks, small microbial size allowing a more efficient evasion of host defences (Dalia and Weiser, 2011; Weiser, 2013; Yang et al., 2016). Microbial cell size appears to be an important pathogenesis factor and minimization of bacterial size was demonstrated to be a mechanism used for example by *Streptococcus pneumoniae* to circumvent complement-mediated killing by the host (Dalia and Weiser, 2011). The cell shape modification from rods to cocci-like form has been very-recently reported to occur in response to antibiotic stress in multi-drug resistant *E. coli* (Uzoehi and Abu-Lail, 2019b) and has been genetically and biochemically demonstrated to occur during the prolonged antibiotic selective pressure that is extensively and aggressively administered to CF patients chronically infected with *P. aeruginosa* bacteria (Yang et al., 2016). Remarkably, in the particular case of the CF patient from whom the examined isolates were obtained, the clinical situation was significantly deteriorated being hospitalized and submitted to intensive intravenous antibiotic therapy immediately before IST4113 isolation (Cunha et al., 2003; Coutinho et al., 2011a).

Many Gram-negative pathogens alter their characteristic rod-shaped forms to smaller coccoid-like forms after incubation for days to weeks in fresh or salt water and in nutrient poor environments bacteria tend to be much smaller in size than those grown in laboratory cultures (Yang et al., 2016). Free-living cells tend to be smaller in nutrient poor environments because the acquisition of nutrients relies on diffusion and capture of molecules at the surface of the cell. In *E. coli*, cell size was found to be reduced by a factor of 3 in response to nutrient starvation (Vadia et al., 2017), *E. coli* adjusting size and growing larger and faster in nutrient-rich media compared with nutrient-poor media (Yao et al., 2012; Vadia et al., 2017; Dai et al., 2018; Du Toit, 2019). During the continuous and rapid deterioration of lung function, as the disease progresses, the oxygen concentration levels in the CF airways suffer a marked decrease (Lorè et al., 2012; Winstanley et al., 2016). Responses of *P. aeruginosa* to oxygen limitation indicate that this species growth in the CF lung is by aerobic respiration (Alvarez-Ortega and Harwood, 2007) and the same metabolism was proposed for *B. cenocepacia* (Coutinho et al., 2011a; Mira et al., 2011; Pessi et al., 2013). According to the hospital records, when the early isolate IST439 was obtained, the FEV1 value (the forced expiratory value in the first second) was 22% but no further values of FEV1 are available due to the subsequent severe deterioration of pulmonary function (Correia et al., 2008; Coutinho et al., 2011a). The fact that the late variants appear to grow more efficiently under microaerophilic conditions, while the early isolate exhibits the most efficient growth when in aerobiosis, supports the hypothesis of an adaptation of the late variants to severe oxygen depletion. The more

adapted growth of the late isolates to oxygen-limitation is also consistent with the biomass increase of the biofilms resulting from growth during the first hours (4 and 6 hours) following initial bacteria adhesion. After 12-48 hours of growth, with the maturation of the biofilms formed, other mechanisms take over as it is the case of exopolysaccharide (EPS) production capacity (Coutinho et al., 2011a; Fazli et al., 2014; Valentini et al., 2018), the biomass concentration of the biofilms formed correlating well with the levels of EPS produced by each clonal variant (Coutinho et al., 2011a). Differences observed between IST4113 and IST4134 growth curves under both microaerophilic and aerobic conditions are likely the result of IST113 resistance to multiple antibiotics, resistance to which a fitness cost is associated (Olivares et al., 2014; Olivares Pacheco et al., 2017).

In summary, independently of the selective pressures that drive *B. cenocepacia* cell size and shape alterations during chronic infection of the lungs, it is likely that the adaptive evolution registered in this study may lead to a better performance under limiting oxygen concentration, to more efficient nutrient acquisition and to evasion of the host complement deposition, favouring persistent infection and pathogenesis. The positive correlation observed between cell shape change and elasticity modulus indicates that elasticity of the cell wall may play a key role in this adaptation process. Results from former genome wide expression analyses and extensive phenotyping of the isolates here examined have provided clues that strongly suggest a genetic adaptation to the challenges exerted by the immune system, antimicrobial therapy and nutrient and oxygen availability (Madeira et al., 2011; Mira et al., 2011; Madeira et al., 2013). The shape and size evolution observed in this study is considered part of such metabolic reprogramming that leads to *B. cenocepacia* persistence in the CF lung (Madeira et al., 2011; Mira et al., 2011; Madeira et al., 2013). Understanding the underlying adaptation mechanisms is essential also for an improved therapeutic outcome of long term infections in CF patients.



## **6 Final discussion**

---



## 6. Final discussion

Chronic respiratory infections are still the main cause of premature death among cystic fibrosis (CF) patients. The CF lung is an hostile environment to the microorganisms involved in these long-term infections, characterised by the action of the human host immune responses, the continuous presence of high antibiotic concentrations, high levels of oxidative stress and low oxygen concentration due to the degradation of the pulmonary function (Moriarty et al., 2007; Palmer et al., 2007; Reid et al., 2007; Williams et al., 2007). The outcome of the adaptive evolution of each strain of different *Burkholderia cepacia* complex (Bcc) species resulting from exposure to those environmental stresses during chronic infection is a topic of high interest in the context of CF research and bacterial pathogenomics. In all the comparative genome-wide studies reported so far, the information obtained for the evolutionary patterns of pulmonary bacterial pathogens in CF was for individual strains of different species when exposed to a specific CF lung environment and host immune responses. In this thesis work, the results obtained from a retrospective study concern the alternations observed in the genome sequences of the most prevalent and feared Bcc species *B. cenocepacia* and *B. multivorans* that co-inhabited the same host-selective environment for a period of time of at least 3 years. Both species evolved as the result of mutation rates of 2.08 (*B. cenocepacia*) and 2.27 (*B. multivorans*) SNPs/year during the course of the chronic co-infection in the CF-lung. These values are considered to lead to a fast genotypic and phenotypic diversification within-the CF lung-host (Markussen et al., 2014; Didelot et al., 2016). The obtained data suggest that *B. cenocepacia* adaptive evolution involved periods of diversification dominated by positive selection, as previously suggested by others (Lee et al., 2017; Nunvar et al., 2017) while *B. multivorans* evolution involved periods of diversification dominated by positive selection followed by periods of relatively neutral evolution, as reported before for this particular species (Silva et al., 2016). Based on this study design, the suggested specificity of the adaptive evolution mechanisms previously proposed to occur in these two Bcc species was here confirmed considering that the registered differences were also observed when exposed to the same CF environment during a chronic co-infection.

As a function of the CF lung environment stressing conditions that were present during the chronic co-infection, different genes were found to be mutated. Among those *B. cenocepacia* and *B. multivorans* genes are those encoding oxidative stress regulatory proteins and heavy metal-sensing proteins, suggesting that the host immune system has a fundamental role in driving the evolution of both species during chronic CF infection. This type of selection driven by the host immune system was however not reported for *B. dolosa* or *P. aeruginosa* evolution within CF-patient (Lieberman et al., 2011; Lieberman et al., 2014; Winstanley et al., 2016). Other mutated genes found in *B. cenocepacia* or *B. multivorans* during co-infection are involved in defense mechanisms and their regulation and to the response to prolonged-antibiotic pressures. Increased resistance to antibiotics is also a well-described evolutionary

trait associated with many CF pathogens and genes related to antibiotic resistance are among the most frequently reported genes to be under selective pressure (Jeannot et al., 2008; Lieberman et al., 2011; Lieberman et al., 2014; Silva et al., 2016; Winstanley et al., 2016; Nunvar et al., 2017).

The results obtained from the comparative genomics analysis carried out in this thesis work, together with other equivalent published studies, also indicated that the LPS OAg moiety undergoes alterations during chronic infection, presumably associated with bacterial adaptation to biofilm lifestyle (Lieberman et al., 2011; Traverse et al., 2013), and to the selective pressure from antibiotics and immune evasion (Maldonado et al., 2016). Due to the limited knowledge on this subject, in the present thesis work we have characterized a novel hybrid genetic LPS OAg locus in the eleven *B. cenocepacia* clonal variants that were examined in the above-mentioned comparative genomic analysis. This novel hybrid OAg biosynthetic locus, comprising genes with homology to *B. multivorans* ATCC 17616 genes, encode enzymes and proteins required for biosynthesis of novel OAg repeating units whose chemical structure for the first isolate retrieved from the CF patient was also demonstrated in this work, as the result of collaboration with the group of Prof. A Molinaro, at University of Napoli Federico II, Complesso Universitario Monte Santangelo, Napoli, Italy. The origin of the *B. multivorans*-like region within the *B. cenocepacia* OAg biosynthetic locus is unknown, but its GC content differs significantly from the GC content of the entire cluster and core genome. Such difference in the GC contents is considered an indicator for the presence of a putative genetic island and mobile element (Hacker and Carniel, 2001; Zhang et al., 2014). Since all of the 10-subsequent clonal *B. cenocepacia* variants collected from the same patient until death with cepacia syndrome do not produce the OAg, this result reinforces the idea that the OAg is lost during *B. cenocepacia* adaptation over chronic lung infection (Maldonado et al., 2016).

Although the Bcc is a group of 24 closely related species, it also remained to be elucidated whether the LPS OAg loss may confer an advantage to the different Bcc species over the course of infection and can be considered common general response occurring in all the Bcc species or if it is a trait specific to *B. cenocepacia*. Since it is essential to obtain comprehensive knowledge about the pathogenesis of the various Bcc species involved in acute or chronic infections in different patients in order to fight the associated infections, a systematic retrospective and longitudinal screening was performed in this work to understand whether the loss of OAg expression during CF infection can be considered a general phenomenon in different Bcc species favouring its chronicity. For this, 357 isolates retrieved from 19 chronically infected patients receiving care at a central hospital in Lisbon were screened. These isolates represent 21 Bcc strains of six/seven Bcc species/lineages, frequently or rarely isolated from CF patients worldwide, *B. cenocepacia* *recA* lineages IIIA and IIIB, *B. multivorans*, *B. dolosa*, *B. stabilis*, *B. cepacia* and *B. contaminans*. Among these six Bcc species, the two most prevalent and feared species, *B. cenocepacia* and *B. multivorans*, showed a tendency to lose the OAg along chronic infection. Also,

*B. cenocepacia* *recA* lineage IIIA strains known to lead to particularly destructive infections (Drevinek and Mahenthiralingam, 2010; Scoffone et al., 2017) exhibit the most frequent OAg loss, compared with lineage IIIB. For the first time, it was shown that *B. cepacia* and *B. contaminans* do not lose the OAg even during long term infections that last for periods of time as extended as 9.6 and 15.2 years, respectively. Moreover, *B. stabilis* strains exhibit a stable phenotype concerning the presence or absence of the OAg over the infection period, consistent with the description of *B. stabilis* sp. Nov, as having a relatively stable genome (Vandamme et al., 2000). For the single patient chronically infected for 5.5 years with the more distantly related Bcc species *B. dolosa*, the OAg moiety was absent from the beginning of the infection until the patient death, consistent with the LOS structure established before for the same strain IST4208 in a collaborative work with A. Molinaro laboratory, Italy (Lorenzo et al., 2013). It is still not clear why different Bcc species/strains differ in their ability to persist in the CF lung and pathogenic potential and this topic requires more in-depth studies.

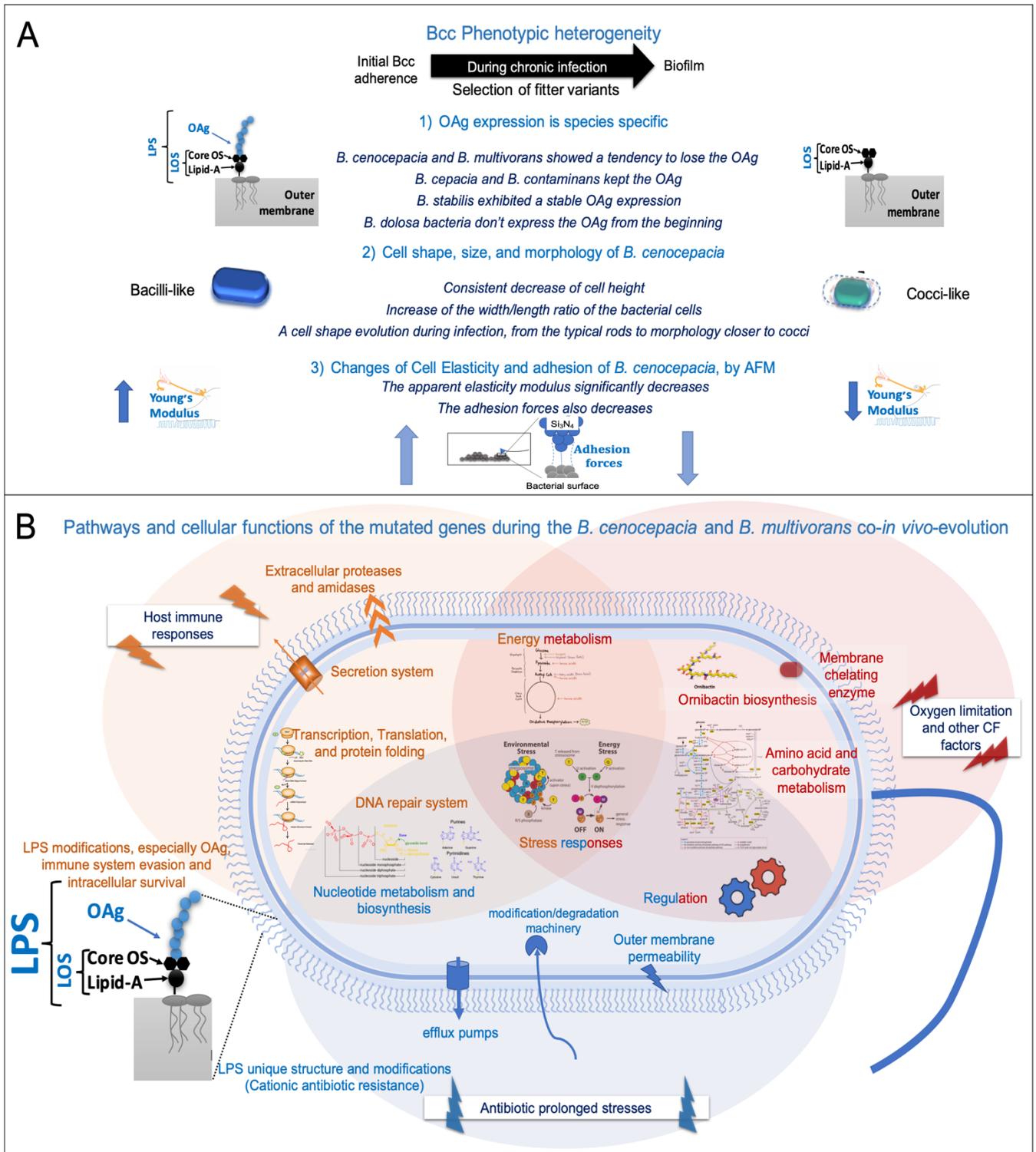
Results also suggest that the observed frequency of the OAg expression switch is not only species dependent but also dependent on the duration of the infection and of the level of deterioration of the lung function. This observation is consistent with the fact that the CF lung is a very hostile environment resulting in marked oxidative stress that causes increased mutation rates (Galli et al., 2012). During chronic pulmonary infection, the OAg units at the Bcc cell surfaces are important for cell evasion from the host immune responses and an OAg deficient *B. cenocepacia* strain was found to be more susceptible to phagocytic internalization (Saldias et al., 2009; Kotrange et al., 2011) while, the loss of OAg expression in *P. aeruginosa* did not confer phagocytic resistance *in vitro* (Demirdjian et al., 2017). Recent published studies are consistent with the results obtained in the present comparative genomic analysis, indicating that the OAg biosynthetic locus is under selective pressure, and the consequent modifications contribute to the evasion of the host immune defences (Murray et al., 2003; Yang et al., 2011a; Maldonado et al., 2016; McCarthy et al., 2017; Faure et al., 2018; Neiger et al., 2019). Moreover, the absence of the OAg was shown to increase Bcc survival in eukaryotic cells as amoebae, epithelial cells and human macrophages (Saldias et al., 2009; Saldias and Valvano, 2009; Kotrange et al., 2011; Maldonado et al., 2016). In agreement with these roles of the OAg to subvert the host's immune function, three of the *B. cenocepacia* isolates examined in this work and retrieved from patient J showed that the late variants, IST4113 and IST4134, were significantly more internalized exhibiting increased survival within human dendritic cells than the early isolate IST439, the sole isolate that expresses the OAg (Cabral et al., 2017). In addition, the absence of the OAg in these same isolates reduces *B. cenocepacia* virulence potential assessed by using the *Galleria mellonella* infection model (Moreira et al., 2017). Such decrease in the virulence potential of early *B. cenocepacia* clonal variants exhibiting the OAg compared with late clonal variants lacking the OAg was also observed for two other chronically infected patients (AB and AN, also examined in the thesis work) with *B. cenocepacia* (Moreira et al., 2017).

The loss or modification of the OAg appears to play an important role during the infection process, in particular in the colonization step (adhesion) and ability to overcome host defence mechanisms (King et al., 2009; Cunneen et al., 2011; Maldonado et al., 2016). Over the last years Atomic force microscopy (AFM) emerged as an essential tool for the characterization of the bacterial adhesion and the nanomechanics of live systems (Touhami et al., 2003; Muller et al., 2009; Costa et al., 2014). Hence, AFM was used in this thesis work to characterize the adhesion and mechanical properties of the Bcc clonal variants retrieved during chronic infection by studying the surface morphology and mapping the mechanical properties of the widely-studied *B. cenocepacia* clonal variants obtained from patient J. In this part of the thesis, it was found that the ability of the early isolate to adhere to the AFM Si<sub>3</sub>N<sub>4</sub> tip was significantly higher than the estimated ability of the late variants that lack the OAg and exhibit similar adhesion values. These results are consistent with the concept that the variability in the LPS OAg affects bacterial adhesion being an important factor in *B. cenocepacia* adaptation to chronic infection (Saldias et al., 2009; Maldonado et al., 2016; Ranf, 2016; Hassan et al., 2017; Hassan et al., 2019). Interestingly, the surface elasticity modulus was found to significantly decrease from the early isolate to the last clonal variant. However, the highly antibiotic resistant intermediary isolate IST4113 (Coutinho et al., 2011a) expressed maximum values. Higher rigidity and increased elasticity was recently reported to be associated with a lower outer membrane permeability, resulting in the reduction of antibiotic diffusion into the cell (Uzoehi and Abu-Lail, 2019a). Consistent with this idea, the genome-wide expression results previously obtained for these three clonal variants reflect a marked reprogramming of genomic expression at different levels (Madeira et al., 2011; Mira et al., 2011; Madeira et al., 2013), in particular the cell envelope that contributes to the antibiotic resistance of Bcc bacteria.

For the first time, the present thesis works clearly describe a consistent and a progressive pattern of decrease of the height and the increase of the width/length ratio of *B. cenocepacia* cells during long term infection and the bacterium underwent a cell shape evolution from the typical rod form of the species to a cell morphology closer to the spherical form of cocci. Such changes in the cell shape were described before for two nasopharyngeal bacterial pathogens during adaptation to human mucosa and it was then hypothesized that this transition was selected to reduce the cell surface sensitivity to immune attacks given that the ratio surface/volume is smaller than that of bacilli (Veyrier et al., 2015). Consistent with this hypothesis, several studies have shown the relevance of cell surface size when bacterial cells are facing immune attacks since a smaller microbial size allows a more efficient evasion of host defences (Dalia and Weiser, 2011; Weiser, 2013; Yang et al., 2016). Also, the present work showed that the late variants appear to grow more efficiently under microaerophilic conditions, while the early isolate exhibits the most efficient growth when in aerobiosis, supporting the hypothesis of an adaptation of the late variants to severe oxygen depletion. Altogether, and independently of the selective pressures that drive *B. cenocepacia* cell size and shape alterations during chronic infection of the lungs,

it is likely that the registered adaptive evolution may lead to a better performance under limiting oxygen concentration, to a more efficient nutrient acquisition and to evasion of the host complement deposition, favouring persistent infection and pathogenesis.

In conclusion, this thesis work supports the notion that Bcc bacteria are human opportunistic pathogens highly adaptable to the CF lung environment where they are exposed to prolonged antibiotic pressures, to the host immune system and to the continuous decrease of oxygen availability. This, mutation-inducing environment drives genetic and phenotypic diversification during chronic infection, resulting in the emergence of a heterogenous population better suited to survive in such challenging environment. The results obtained in this thesis suggest species-specific adaptive strategies and a number of common responses and phenotypic and genotypic alterations, increasing the current understanding of the adaptive mechanisms underlying Bcc adaptation to the CF lungs. The knowledge that has been gathered during this thesis work led us to propose the following working model (Fig. 6.1).



**Figure 6.1** | Proposed model for the mechanisms underlying clonal diversification occurring in Bcc bacteria during long-term chronic infection inside the CF lung, based on the experimental evidences gathered during this PhD. A) Phenotypic changes with a focus on *B. cenocepacia*. B) Pathways and cellular functions related with the mutated genes during the *B. cenocepacia* and *B. multivorans* co-*in vivo*-evolution as a function of CF-environmental factors (Orange – host immune responses, Blue – prolonged antibiotic stresses and Red – oxygen limitation and others).

## 7 References

---



## 7. References

- Aaron, S.D., Ferris, W., Henry, D.A., Speert, D.P., and Macdonald, N.E. (2000). Multiple combination bactericidal antibiotic testing for patients with cystic fibrosis infected with *Burkholderia cepacia*. *American Journal of Respiratory and Critical Care Medicine* 161(4 Pt 1), 1206-1212. doi: 10.1164/ajrcm.161.4.9907147.
- Abu-Lail, L.I., Liu, Y., Atabek, A., and Camesano, T.A. (2007). Quantifying the adhesion and interaction forces between *Pseudomonas aeruginosa* and natural organic matter. *Environmental Science and Technology* 41(23), 8031-8037.
- Agnoli, K., Frauenknecht, C., Freitag, R., Schwager, S., Jenul, C., Vergunst, A., et al. (2014). The third replicon of members of the *Burkholderia cepacia* Complex, plasmid pC3, plays a role in stress tolerance. *Applied and Environmental Microbiology* 80(4), 1340-1348. doi: 10.1128/AEM.03330-13.
- Agnoli, K., Freitag, R., Gomes, M.C., Jenul, C., Suppiger, A., Mannweiler, O., et al. (2017). The Use of Synthetic Hybrid Strains to Determine the Role of Replicon 3 in Virulence of the *Burkholderia cepacia* Complex. *Applied and Environmental Microbiology*. doi: 10.1128/AEM.00461-17.
- Agnoli, K., Lowe, C.A., Farmer, K.L., Husnain, S.I., and Thomas, M.S. (2006). The ornibactin biosynthesis and transport genes of *Burkholderia cenocepacia* are regulated by an extracytoplasmic function sigma factor which is a part of the Fur regulon. *Journal of Bacteriology* 188(10), 3631-3644. doi: 10.1128/jb.188.10.3631-3644.2006.
- Agnoli, K., Schwager, S., Uehlinger, S., Vergunst, A., Viteri, D.F., Nguyen, D.T., et al. (2012). Exposing the third chromosome of *Burkholderia cepacia* complex strains as a virulence plasmid. *Molecular Microbiology* 83(2), 362-378. doi: 10.1111/j.1365-2958.2011.07937.x.
- Agodi, A., Barchitta, M., Giannino, V., Collura, A., Pensabene, T., Garlaschi, M.L., et al. (2002). *Burkholderia cepacia* complex in cystic fibrosis and non-cystic fibrosis patients: identification of a cluster of epidemic lineages. *Journal of Hospital Infection* 50(3), 188-195. doi: 10.1053/jhin.2001.1160.
- Ahimou, F., Touhami, A., and Dufrene, Y.F. (2003). Real-time imaging of the surface topography of living yeast cells by atomic force microscopy. *Yeast* 20(1), 25-30. doi: 10.1002/yea.923.
- Ahn, Y., Kim, J.M., Ahn, H., Lee, Y.J., LiPuma, J.J., Hussong, D., et al. (2014). Evaluation of liquid and solid culture media for the recovery and enrichment of *Burkholderia cenocepacia* from distilled water. *Journal of Industrial Microbiology and Biotechnology* 41(7), 1109-1118. doi: 10.1007/s10295-014-1442-3.
- Ahn, Y., Kim, J.M., Kweon, O., Kim, S.J., Jones, R.C., Woodling, K., et al. (2016). Intrinsic Resistance of *Burkholderia cepacia* Complex to Benzalkonium Chloride. *MBio* 7(6). doi: 10.1128/mBio.01716-16.
- Ahrazem, O., Prieto, A., Leal, J., Jimenez-Barbero, J., and Bernabe, M. (2002). Fungal cell wall galactomannan isolated from *Apodus deciduus*. *Carbohydrate Research* 337(16), 1503-1506. doi: 10.1016/S0008-6215(02)00184-2.
- Ahrazem, O., Prieto, A., Leal, J.A., Inmaculada Gimenez-Abian, M., Jimenez-Barbero, J., and Bernabe, M. (2007). Fungal cell wall polysaccharides isolated from *Discula destructiva* spp. *Carbohydrate Research* 342(8), 1138-1143. doi: 10.1016/j.carres.2007.02.011.
- Ali, N.O., Bignon, J., Rapoport, G., and Debarbouille, M. (2001). Regulation of the acetoin catabolic pathway is controlled by sigma L in *Bacillus subtilis*. *Journal of Bacteriology* 183(8), 2497-2504. doi: 10.1128/JB.183.8.2497-2504.2001.
- Alikhan, N.F., Petty, N.K., Ben Zakour, N.L., and Beatson, S.A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12, 402. doi: 10.1186/1471-2164-12-402.
- Allison, D.P., Mortensen, N.P., Sullivan, C.J., and Doktycz, M.J. (2010). Atomic force microscopy of biological samples. *Wiley Interdisciplinary Reviews : Nanomedicine and Nanobiotechnology* 2(6), 618-634. doi: 10.1002/wnan.104.
- Allison, D.P., Sullivan, C.J., Mortensen, N.P., Retterer, S.T., and Doktycz, M. (2011). Bacterial immobilization for imaging by atomic force microscopy. *Journal of Visualized Experiments* (54). doi: 10.3791/2880.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* 215(3), 403-410. doi: 10.1016/S0022-2836(05)80360-2.
- Alvarez-Ortega, C., and Harwood, C.S. (2007). Responses of *Pseudomonas aeruginosa* to low oxygen indicate that growth in the cystic fibrosis lung is by aerobic respiration. *Molecular Microbiology* 65(1), 153-165. doi: 10.1111/j.1365-2958.2007.05772.x.
- Amir, A., Babaeipour, F., McIntosh, D.B., Nelson, D.R., and Jun, S. (2014). Bending forces plastically deform growing bacterial cell walls. *PNAS, Proceedings of the National Academy of Sciences* 111(16), 5778-5783. doi: 10.1073/pnas.1317497111.
- Anderson, M.S., Bulawa, C.E., and Raetz, C.R. (1985). The biosynthesis of gram-negative endotoxin. Formation of lipid A precursors from UDP-GlcNAc in extracts of *Escherichia coli*. *The Journal of Biological Chemistry* 260(29), 15536-15541.
- Anderson, M.S., and Raetz, C.R. (1987). Biosynthesis of lipid A precursors in *Escherichia coli*. A cytoplasmic acyltransferase that converts UDP-N-acetylglucosamine to UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine. *The Journal of Biological Chemistry* 262(11), 5159-5169.

- Angeloni, L. (2016). Microbial Cells Force Spectroscopy by Atomic Force Microscopy: A Review. *Nanoscience and Nanometrology* 2(1), 30. doi: 10.11648/j.nsnm.20160201.13.
- Aravind, L., and Koonin, E.V. (1999). Novel predicted RNA-binding domains associated with the translation machinery. *Journal of Molecular Evolution* 48(3), 291-302.
- Atabek, A., and Camesano, T.A. (2007). Atomic force microscopy study of the effect of lipopolysaccharides and extracellular polymers on adhesion of *Pseudomonas aeruginosa*. *Journal of Bacteriology* 189(23), 8503-8509. doi: 10.1128/JB.00769-07.
- Atabek, A., Liu, Y., Pinzon-Arango, P.A., and Camesano, T.A. (2008). Importance of LPS structure on protein interactions with *Pseudomonas aeruginosa*. *Colloids and Surfaces B: Biointerfaces* 67(1), 115-121. doi: 10.1016/j.colsurfb.2008.08.013.
- Auer, G.K., Lee, T.K., Rajendram, M., Cesar, S., Miguel, A., Huang, K.C., et al. (2016). Mechanical Genomics Identifies Diverse Modulators of Bacterial Cell Stiffness. *Cell Systems* 2(6), 402-411. doi: 10.1016/j.cels.2016.05.006.
- Auer, G.K., and Weibel, D.B. (2017). Bacterial Cell Mechanics. *Biochemistry* 56(29), 3710-3724. doi: 10.1021/acs.biochem.7b00346.
- Babu, M., Bundalovic-Torma, C., Calmettes, C., Phanse, S., Zhang, Q., Jiang, Y., et al. (2018). Global landscape of cell envelope protein complexes in *Escherichia coli*. *Nature Biotechnology* 36(1), 103-112. doi: 10.1038/nbt.4024.
- Bach, E., Sant'Anna, F.H., Magrich Dos Passos, J.F., Balsanelli, E., de Baura, V.A., Pedrosa, F.O., et al. (2017). Detection of misidentifications of species from the *Burkholderia cepacia* complex and description of a new member, the soil bacterium *Burkholderia catarinensis* sp. nov. *Pathogens and Disease* 75(6). doi: 10.1093/femspd/ftx076.
- Baldwin, A., Mahenthiralingam, E., Drevinek, P., Pope, C., Waive, D.J., Henry, D.A., et al. (2008). Elucidating global epidemiology of *Burkholderia multivorans* in cases of cystic fibrosis by multilocus sequence typing. *Journal of Clinical Microbiology* 46(1), 290-295. doi: 10.1128/JCM.01818-07.
- Baldwin, A., Mahenthiralingam, E., Thickett, K.M., Honeybourne, D., Maiden, M.C., Govan, J.R., et al. (2005). Multilocus sequence typing scheme that provides both species and strain differentiation for the *Burkholderia cepacia* complex. *Journal of Clinical Microbiology* 43(9), 4665-4673. doi: 10.1128/JCM.43.9.4665-4673.2005.
- Baldwin, A., Sokol, P.A., Parkhill, J., and Mahenthiralingam, E. (2004). The *Burkholderia cepacia* epidemic strain marker is part of a novel genomic island encoding both virulence and metabolism-associated genes in *Burkholderia cenocepacia*. *Infection and Immunity* 72(3), 1537-1547. doi: 10.1128/iai.72.3.1537-1547.2004.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19(5), 455-477. doi: 10.1089/cmb.2012.0021.
- Barrick, J.E., Colburn, G., Deatherage, D.E., Traverse, C.C., Strand, M.D., Borges, J.J., et al. (2014). Identifying structural variation in haploid microbial genomes from short-read resequencing data using breseq. *BMC Genomics* 15, 1039. doi: 10.1186/1471-2164-15-1039.
- Barth, A.L., and Pitt, T.L. (1995). Auxotrophy of *Burkholderia (Pseudomonas) cepacia* from cystic fibrosis patients. *Journal of Clinical Microbiology* 33(8), 2192-2194. PMC228364.
- Baym, M., Kryazhinskiy, S., Lieberman, T.D., Chung, H., Desai, M.M., and Kishony, R. (2015). Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One* 10(5), e0128036. doi: 10.1371/journal.pone.0128036.
- Beaulaurier, J., Zhang, X.S., Zhu, S., Sebra, R., Rosenbluh, C., Deikus, G., et al. (2015). Single molecule-level detection and long read-based phasing of epigenetic variations in bacterial methylomes. *Nature Communications* 6, 7438. doi: 10.1038/ncomms8438.
- Beaussart, A., Baker, A.E., Kuchma, S.L., El-Kirat-Chatel, S., O'Toole, G.A., and Dufrene, Y.F. (2014). Nanoscale adhesion forces of *Pseudomonas aeruginosa* type IV Pili. *ACS Nano* 8(10), 10723-10733. doi: 10.1021/nn5044383.
- Bender, J.K., Wille, T., Blank, K., Lange, A., and Gerlach, R.G. (2013). LPS structure and PhoQ activity are important for *Salmonella* Typhimurium virulence in the *Galleria mellonella* infection model [corrected]. *PLoS One* 8(8), e73287. doi: 10.1371/journal.pone.0073287.
- Bentley, S.D., and Parkhill, J. (2015). Genomic perspectives on the evolution and spread of bacterial pathogens. *Proceedings of the Royal Society B: Biological Sciences* 282(1821), 20150488. doi: 10.1098/rspb.2015.0488.
- Bernhardt, S.A., Spilker, T., Coffey, T., and LiPuma, J.J. (2003). *Burkholderia cepacia* complex in cystic fibrosis: frequency of strain replacement during chronic infection. *Clinical Infectious Diseases* 37(6), 780-785. doi: 10.1086/377541.
- Berry, M.C., McGhee, G.C., Zhao, Y., and Sundin, G.W. (2009). Effect of a *waaL* mutation on lipopolysaccharide composition, oxidative stress survival, and virulence in *Erwinia amylovora*. *FEMS Microbiology Letters* 291(1), 80-87. doi: 10.1111/j.1574-6968.2008.01438.x.
- Bertelli, C., Laird, M.R., Williams, K.P., Simon Fraser University Research Computing, G., Lau, B.Y., Hoad, G., et al. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Research* 45(W1), W30-W35. doi: 10.1093/nar/gkx343.
- Biddick, R., Spilker, T., Martin, A., and LiPuma, J.J. (2003). Evidence of transmission of *Burkholderia cepacia*, *Burkholderia multivorans* and *Burkholderia dolosa* among persons with cystic fibrosis. *FEMS Microbiology Letters* 228(1), 57-62.

- Bilwes, A.M., Alex, L.A., Crane, B.R., and Simon, M.I. (1999). Structure of CheA, a signal-transducing histidine kinase. *Cell* 96(1), 131-141. doi: 10.1016/s0092-8674(00)80966-6.
- Binnig, G., Quate, C.F., and Gerber, C. (1986). Atomic force microscope. *Physical Review Letters* 56(9), 930-933. doi: 10.1103/PhysRevLett.56.930.
- Bobik, C., Meilhoc, E., and Batut, J. (2006). FixJ: a major regulator of the oxygen limitation response and late symbiotic functions of *Sinorhizobium meliloti*. *Journal of Bacteriology* 188(13), 4890-4902. doi: 10.1128/JB.00251-06.
- Boetzer, M., Henkel, C.V., Jansen, H.J., Butler, D., and Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27(4), 578-579. doi: 10.1093/bioinformatics/btq683.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15), 2114-2120. doi: 10.1093/bioinformatics/btu170.
- Bowden, S.D., Hale, N., Chung, J.C., Hodgkinson, J.T., Spring, D.R., and Welch, M. (2013). Surface swarming motility by *Pectobacterium atrosepticum* is a latent phenotype that requires O antigen and is regulated by quorum sensing. *Microbiology* 159(Pt 11), 2375-2385. doi: 10.1099/mic.0.070748-0.
- Bramkamp, M., and van Baarle, S. (2009). Division site selection in rod-shaped bacteria. *Current Opinion in Microbiology* 12(6), 683-688. doi: 10.1016/j.mib.2009.10.002.
- Brisse, S., Verduin, C.M., Milatovic, D., Fluit, A., Verhoef, J., Laevens, S., et al. (2000). Distinguishing species of the *Burkholderia cepacia* complex and *Burkholderia gladioli* by automated ribotyping. *Journal of Clinical Microbiology* 38(5), 1876-1884. PMC86613.
- Burkholder, W.H. (1950). Sour skin, a bacterial rot of Onion bulbs. *Phytopathology* 40(1), 115-117.
- Busscher, H.J., Norde, W., and van der Mei, H.C. (2008). Specific molecular recognition and nonspecific contributions to bacterial interaction forces. *Applied and Environmental Microbiology* 74(9), 2559-2564. doi: 10.1128/AEM.02839-07.
- Butt, A.T., and Thomas, M.S. (2018). Iron Acquisition Mechanisms and Their Role in the Virulence of Burkholderia Species (vol 7, 460, 2017). *Frontiers in Cellular and Infection Microbiology* 8. doi: 10.3389/fcimb.2018.00305.
- Butt, H.J., Cappella, B., and Kappl, M. (2005). Force measurements with the atomic force microscope: Technique, interpretation and applications. *Surface Science Reports* 59(1-6), 1-152. doi: 10.1016/j.surfrep.2005.08.003.
- Bylund, J., Burgess, L.A., Cescutti, P., Ernst, R.K., and Speert, D.P. (2006). Exopolysaccharides from *Burkholderia cenocepacia* inhibit neutrophil chemotaxis and scavenge reactive oxygen species. *The Journal of Biological Chemistry* 281(5), 2526-2532. doi: 10.1074/jbc.M510692200.
- Cabral, M., Pereira, M., Silva, Z., Iria, I., Coutinho, C., Lopes, A., et al. (2017). Using dendritic cells to evaluate how *Burkholderia cenocepacia* clonal isolates from a chronically infected cystic fibrosis patient subvert immune functions. *Medical Microbiology and Immunology* 206(2), 111-123. doi: 10.1007/s00430-016-0488-4.
- Caraher, E., Duff, C., Mullen, T., Mc Keon, S., Murphy, P., Callaghan, M., et al. (2007). Invasion and biofilm formation of *Burkholderia dolosa* is comparable with *Burkholderia cenocepacia* and *Burkholderia multivorans*. *Journal of Cystic Fibrosis* 6(1), 49-56. doi: 10.1016/j.jcf.2006.05.007.
- Cardona, S.T., Mueller, C.L., and Valvano, M.A. (2006). Identification of essential operons with a rhamnase-inducible promoter in *Burkholderia cenocepacia*. *Applied Environmental Microbiology* 72(4), 2547-2555. doi: 10.1128/AEM.72.4.2547-2555.2006.
- Cardona, S.T., and Valvano, M.A. (2005). An expression vector containing a rhamnase-inducible promoter provides tightly regulated gene expression in *Burkholderia cenocepacia*. *Plasmid* 54(3), 219-228. doi: 10.1016/j.plasmid.2005.03.004.
- Carrier, A., Agnoli, K., Pessi, G., Suppiger, A., Jenul, C., Schmid, N., et al. (2014). Genome Sequence of *Burkholderia cenocepacia* H111, a Cystic Fibrosis Airway Isolate. *Genome Announcements* 2(2). doi: 10.1128/genomeA.00298-14.
- Carver, T., Berriman, M., Tivey, A., Patel, C., Bohme, U., Barrell, B.G., et al. (2008). Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics* 24(23), 2672-2676. doi: 10.1093/bioinformatics/btn529.
- Carver, T.J., Rutherford, K.M., Berriman, M., Rajandream, M.A., Barrell, B.G., and Parkhill, J. (2005). ACT: the Artemis Comparison Tool. *Bioinformatics* 21(16), 3422-3423. doi: 10.1093/bioinformatics/bti553.
- Cerf, A., Cau, J.C., Vieu, C., and Dague, E. (2009). Nanomechanical properties of dead or alive single-patterned bacteria. *Langmuir* 25(10), 5731-5736. doi: 10.1021/la9004642.
- Chen, J.S., Witzmann, K.A., Spilker, T., Fink, R.J., and LiPuma, J.J. (2001). Endemicity and inter-city spread of *Burkholderia cepacia* genomovar III in cystic fibrosis. *The Journal of Pediatrics* 139(5), 643-649. doi: 10.1067/mpd.2001.118430.
- Chiarini, L., Cescutti, P., Drigo, L., Impallomeni, G., Herasimenka, Y., Bevivino, A., et al. (2004). Exopolysaccharides produced by *Burkholderia cenocepacia recA* lineages IIIA and IIIB. *Journal of Cystic Fibrosis* 3(3), 165-172. doi: 10.1016/j.jcf.2004.04.004.
- Chin, C.S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., et al. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature Methods* 10(6), 563-569. doi: 10.1038/nmeth.2474.
- Chiou, Y.W., Lin, H.K., Tang, M.J., Lin, H.H., and Yeh, M.L. (2013). The influence of physical and physiological cues on atomic force microscopy-based cell stiffness assessment. *PLoS One* 8(10), e77384. doi: 10.1371/journal.pone.0077384.

- Chiu, C.H., Ostry, A., and Speert, D.P. (2001). Invasion of murine respiratory epithelial cells in vivo by *Burkholderia cepacia*. *Journal of Medical Microbiology* 50(7), 594-601. doi: 10.1099/0022-1317-50-7-594.
- Chua, S.L., Ding, Y., Liu, Y., Cai, Z., Zhou, J., Swarup, S., et al. (2016). Reactive oxygen species drive evolution of pro-biofilm variants in pathogens by modulating cyclic-di-GMP levels. *Open Biology* 6(11). doi: 10.1098/rsob.160162.
- Chua, S.L., Ding, Y., Liu, Y., Cai, Z., Zhou, J., Swarup, S., et al. (2017). Correction to 'Reactive oxygen species drive evolution of pro-biofilm variants in pathogens by modulating cyclic-di-GMP levels'. *Open Biology* 7(9). doi: 10.1098/rsob.170197.
- Chung, H.S., and Raetz, C.R. (2011). Dioxygenases in *Burkholderia ambifaria* and *Yersinia pestis* that hydroxylate the outer Kdo unit of lipopolysaccharide. *PNAS, Proceedings of the National Academy of Sciences* 108(2), 510-515. doi: 10.1073/pnas.1016462108.
- Chung, H.S., Yang, E.G., Hwang, D., Lee, J.E., Guan, Z., and Raetz, C.R. (2014). Kdo hydroxylase is an inner core assembly enzyme in the Ko-containing lipopolysaccharide biosynthesis. *Biochemical and Biophysical Research Communications* 452(3), 789-794. doi: 10.1016/j.bbrc.2014.08.153.
- Cieri, M.V., Mayer-Hamblett, N., Griffith, A., and Burns, J.L. (2002). Correlation between an in vitro invasion assay and a murine model of *Burkholderia cepacia* lung infection. *Infection and Immunity* 70(3), 1081-1086. doi: 10.1128/iai.70.3.1081-1086.2002.
- Cigana, C., Curcuru, L., Leone, M.R., Ierano, T., Lore, N.I., Bianconi, I., et al. (2009). *Pseudomonas aeruginosa* exploits lipid A and mucopeptides modification as a strategy to lower innate immunity during cystic fibrosis lung infection. *PLoS One* 4(12), e8439. doi: 10.1371/journal.pone.0008439.
- Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* 6(2), 80-92. doi: 10.4161/fly.19695.
- Ciucanu, I., and Kerek, F. (1984). A Simple and Rapid Method for the Permethylation of Carbohydrates. *Carbohydrate Research* 131(2), 209-217. doi: 10.1016/0008-6215(84)85242-8.
- Coburn, B., Wang, P.W., Diaz Caballero, J., Clark, S.T., Brahma, V., Donaldson, S., et al. (2015). Lung microbiota across age and disease stage in cystic fibrosis. *Scientific Reports* 5, 10241. doi: 10.1038/srep10241.
- Coenye, T., and LiPuma, J.J. (2002). Multilocus restriction typing: a novel tool for studying global epidemiology of *Burkholderia cepacia* complex infection in cystic fibrosis. *The Journal of Infectious Diseases* 185(10), 1454-1462. doi: 10.1086/340279.
- Coenye, T., Spilker, T., Van Schoor, A., LiPuma, J.J., and Vandamme, P. (2004). Recovery of *Burkholderia cenocepacia* strain PHDC from cystic fibrosis patients in Europe. *Thorax* 59(11), 952-954. doi: 10.1136/thx.2003.019810.
- Coenye, T., and Vandamme, P. (2003). Diversity and significance of *Burkholderia* species occupying diverse ecological niches. *Environmental Microbiology* 5(9), 719-729.
- Coenye, T., Vandamme, P., Govan, J.R., and LiPuma, J.J. (2001). Taxonomy and identification of the *Burkholderia cepacia* complex. *Journal of Clinical Microbiology* 39(10), 3427-3436. doi: 10.1128/JCM.39.10.3427-3436.2001.
- Compant, S., Nowak, J., Coenye, T., Clement, C., and Ait Barka, E. (2008). Diversity and occurrence of *Burkholderia* spp. in the natural environment. *FEMS Microbiology Reviews* 32(4), 607-626. doi: 10.1111/j.1574-6976.2008.00113.x.
- Conway, B.A., Chu, K.K., Bylund, J., Altman, E., and Speert, D.P. (2004). Production of exopolysaccharide by *Burkholderia cenocepacia* results in altered cell-surface interactions and altered bacterial clearance in mice. *The Journal of Infectious Diseases* 190(5), 957-966. doi: 10.1086/423141.
- Cooper, V.S. (2014). The origins of specialization: insights from bacteria held 25 years in captivity. *PLoS Biology* 12(2), e1001790. doi: 10.1371/journal.pbio.1001790.
- Cooper, V.S., Staples, R.K., Traverse, C.C., and Ellis, C.N. (2014). Parallel evolution of small colony variants in *Burkholderia cenocepacia* biofilms. *Genomics* 104(6 Pt A), 447-452. doi: 10.1016/j.ygeno.2014.09.007.
- Corbett, C.R., Burntack, M.N., Kooi, C., Woods, D.E., and Sokol, P.A. (2003). An extracellular zinc metalloprotease gene of *Burkholderia cepacia*. *Microbiology* 149(Pt 8), 2263-2271. doi: 10.1099/mic.0.26243-0.
- Correia, S., Nascimento, C., Pereira, L., Cunha, M.V., Sá-Correia, I., and Barreto, C. (2008). Infecção respiratória por bactérias do complexo *cepacia*: Evolução clínica em doentes com fibrose quística. *Revista Portuguesa de Pneumologia* 14, 5-26.
- Costa, L., Rodrigues, M.S., Benseny-Cases, N., Mayeux, V., Chevrier, J., and Comin, F. (2014). Spectroscopic investigation of local mechanical impedance of living cells. *PLoS One* 9(7), e101687. doi: 10.1371/journal.pone.0101687.
- Coutinho, C.P., Barreto, C., Pereira, L., Lito, L., Melo Cristino, J., and Sá-Correia, I. (2015). Incidence of *Burkholderia contaminans* at a cystic fibrosis centre with an unusually high representation of *Burkholderia cepacia* during 15 years of epidemiological surveillance. *Journal of Medical Microbiology* 64(8), 927-935. doi: 10.1099/jmm.0.000094.
- Coutinho, C.P., de Carvalho, C.C., Madeira, A., Pinto-de-Oliveira, A., and Sa-Correia, I. (2011a). *Burkholderia cenocepacia* phenotypic clonal variation during a 3.5-year colonization in the lungs of a cystic fibrosis patient. *Infection and Immunity* 79(7), 2950-2960. doi: 10.1128/IAI.01366-10.
- Coutinho, C.P., Dos Santos, S.C., Madeira, A., Mira, N.P., Moreira, A.S., and Sá-Correia, I. (2011b). Long-term colonization of the cystic fibrosis lung by *Burkholderia cepacia* complex bacteria: epidemiology, clonal variation, and genome-wide expression alterations. *Frontiers in Cellular and Infection Microbiology* 1, 12. doi: 10.3389/fcimb.2011.00012.

- Cox, A.D., and Wilkinson, S.G. (1989). Structures of the O-specific polymers from the lipopolysaccharides of the reference strains for *Pseudomonas cepacia* serogroups O3 and O5. *Carbohydrate Research* 195(1), 123-129. doi: 10.1016/0008-6215(89)85095-5.
- Craig, F.F., Coote, J.G., Parton, R., Freer, J.H., and Gilmour, N.J. (1989). A plasmid which can be transferred between *Escherichia coli* and *Pasteurella haemolytica* by electroporation and conjugation. *Journal of General Microbiology* 135(11), 2885-2890. doi: 10.1099/00221287-135-11-2885.
- Cullen, L., and McClean, S. (2015). Bacterial Adaptation during Chronic Respiratory Infections. *Pathogens* 4(1), 66-89. doi: 10.3390/pathogens4010066.
- Cunha, M.V., Leitão, J.H., Mahenthiralingam, E., Vandamme, P., Lito, L., Barreto, C., et al. (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. doi: 10.1128/JCM.41.9.4113-4120.2003.
- Cunha, M.V., Pinto-de-Oliveira, A., Meirinhos-Soares, L., Salgado, M.J., Melo-Cristino, J., Correia, S., et al. (2007). Exceptionally high representation of *Burkholderia cepacia* among *B. cepacia* complex isolates recovered from the major Portuguese cystic fibrosis center. *Journal of Clinical Microbiology* 45(5), 1628-1633. doi: 10.1128/JCM.00234-07.
- Cunha, M.V., Sousa, S.A., Leitao, J.H., Moreira, L.M., Videira, P.A., and Sá-Correia, I. (2004). Studies on the involvement of the exopolysaccharide produced by cystic fibrosis-associated isolates of the *Burkholderia cepacia* complex in biofilm formation and in persistence of respiratory infections. *Journal of Clinical Microbiology* 42(7), 3052-3058. doi: 10.1128/JCM.42.7.3052-3058.2004.
- Cunneen, M.M., Reeves, P.R., Valvano, M.A., Furlong, S.E., Patel, K.B., Knirel, Y.A., et al. (2011). "Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis and Interaction with Host Cells", (eds.) Y.A. Knirel & M.A. Valvano. (Vienna: Springer Vienna).
- Cuzzi, B., Cescutti, P., Furlan, L., Lagatolla, C., Sturiale, L., Garozzo, D., et al. (2012). Investigation of bacterial resistance to the immune system response: cepacian depolymerisation by reactive oxygen species. *Innate Immunity* 18(4), 661-671. doi: 10.1177/1753425911435954.
- D'Souza, J.M., Samuel, G.N., and Reeves, P.R. (2005). Evolutionary origins and sequence of the *Escherichia coli* O4 O-antigen gene cluster. *FEMS Microbiology Letters* 244(1), 27-32. doi: 10.1016/j.femsle.2005.01.012.
- Dai, X., Shen, Z., Wang, Y., and Zhu, M. (2018). *Sinorhizobium meliloti*, a Slow-Growing Bacterium, Exhibits Growth Rate Dependence of Cell Size under Nutrient Limitation. *mSphere* 3(6). doi: 10.1128/mSphere.00567-18.
- Dalia, A.B., and Weiser, J.N. (2011). Minimization of bacterial size allows for complement evasion and is overcome by the agglutinating effect of antibody. *Cell Host and Microbe* 10(5), 486-496. doi: 10.1016/j.chom.2011.09.009.
- Darling, A.E., Mau, B., and Perna, N.T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5(6), e11147. doi: 10.1371/journal.pone.0011147.
- De Smet, B., Mayo, M., Peeters, C., Zlosnik, J.E., Spilker, T., Hird, T.J., et al. (2015). *Burkholderia stagnalis* sp. nov. and *Burkholderia territorii* sp. nov., two novel *Burkholderia cepacia* complex species from environmental and human sources. *International Journal of Systematic and Evolutionary Microbiology* 65(7), 2265-2271. doi: 10.1099/ijs.0.000251.
- De Soyza, A., Ellis, C.D., Khan, C.M., Corris, P.A., and Demarco de Hormaeche, R. (2004). *Burkholderia cenocepacia* lipopolysaccharide, lipid A, and proinflammatory activity. *American Journal of Respiratory and Critical Care Medicine* 170(1), 70-77. doi: 10.1164/rccm.200304-592OC.
- De Soyza, A., Silipo, A., Lanzetta, R., Govan, J.R., and Molinaro, A. (2008). Chemical and biological features of *Burkholderia cepacia* complex lipopolysaccharides. *Innate Immunity* 14(3), 127-144. doi: 10.1177/1753425908093984.
- Dedekova, K., Kalferstova, L., Strnad, H., Vavrova, J., and Drevinek, P. (2013). Novel diagnostic PCR assay for *Burkholderia cenocepacia* epidemic strain ST32 and its utility in monitoring infection in cystic fibrosis patients. *Journal of Cystic Fibrosis* 12(5), 475-481. doi: 10.1016/j.jcf.2012.12.007.
- Del Campo, M., Kaya, Y., and Ofengand, J. (2001). Identification and site of action of the remaining four putative pseudouridine synthases in *Escherichia coli*. *Rna* 7(11), 1603-1615. PMC1370202.
- Delcour, A.H. (2009). Outer membrane permeability and antibiotic resistance. *Biochimica et Biophysica Acta (BBA)* 1794(5), 808-816. doi: 10.1016/j.bbapap.2008.11.005.
- Demirdjian, S., Schutz, K., Wargo, M.J., Lam, J.S., and Berwin, B. (2017). The effect of loss of O-antigen ligase on phagocytic susceptibility of motile and non-motile *Pseudomonas aeruginosa*. *Molecular Immunology* 92, 106-115. doi: 10.1016/j.molimm.2017.10.015.
- Deng, Y., Sun, M., and Shaevitz, J.W. (2011). Direct measurement of cell wall stress stiffening and turgor pressure in live bacterial cells. *Physical Review Letters* 107(15), 158101. doi: 10.1103/PhysRevLett.107.158101.
- Depoorter, E., Bull, M.J., Peeters, C., Coenye, T., Vandamme, P., and Mahenthiralingam, E. (2016). *Burkholderia*: an update on taxonomy and biotechnological potential as antibiotic producers. *Applied Microbiology and Biotechnology* 100(12), 5215-5229. doi: 10.1007/s00253-016-7520-x.
- Deretic, V., Schurr, M.J., Boucher, J.C., and Martin, D.W. (1994). Conversion of *Pseudomonas aeruginosa* to mucoidy in cystic fibrosis: environmental stress and regulation of bacterial virulence by alternative sigma factors. *Journal of Bacteriology* 176(10), 2773-2780. PMC205429.

- Desmarais, S.M., De Pedro, M.A., Cava, F., and Huang, K.C. (2013). Peptidoglycan at its peaks: how chromatographic analyses can reveal bacterial cell wall structure and assembly. *Molecular Immunology* 89(1), 1-13. doi: 10.1111/mmi.12266.
- Dettman, J.R., Rodrigue, N., Aaron, S.D., and Kassen, R. (2013). Evolutionary genomics of epidemic and nonepidemic strains of *Pseudomonas aeruginosa*. *PNAS, Proceedings of the National Academy of Sciences* 110(52), 21065-21070. doi: 10.1073/pnas.1307862110.
- Di Lorenzo, F., Kubik, L., Oblak, A., Lore, N.I., Cigana, C., Lanzetta, R., et al. (2015). Activation of Human Toll-like Receptor 4 (TLR4). Myeloid Differentiation Factor 2 (MD-2) by Hypoacylated Lipopolysaccharide from a Clinical Isolate of *Burkholderia cenocepacia*. *The Journal of Biological Chemistry* 290(35), 21305-21319. doi: 10.1074/jbc.M115.649087.
- Didelot, X., Walker, A.S., Peto, T.E., Crook, D.W., and Wilson, D.J. (2016). Within-host evolution of bacterial pathogens. *Nature Reviews Microbiology* 14(3), 150-162. doi: 10.1038/nrmicro.2015.13.
- Dillon, M.M., Sung, W., Lynch, M., and Cooper, V.S. (2015). The Rate and Molecular Spectrum of Spontaneous Mutations in the GC-Rich Multichromosome Genome of *Burkholderia cenocepacia*. *Genetics* 200(3), 935-946. doi: 10.1534/genetics.115.176834.
- Dillon, M.M., Sung, W., Sebra, R., Lynch, M., and Cooper, V.S. (2017). Genome-Wide Biases in the Rate and Molecular Spectrum of Spontaneous Mutations in *Vibrio cholerae* and *Vibrio fischeri*. *Molecular Biology and Evolution* 34(1), 93-109. doi: 10.1093/molbev/msw224.
- Dimitriadis, E.K., Horkay, F., Maresca, J., Kachar, B., and Chadwick, R.S. (2002). Determination of elastic moduli of thin layers of soft material using the atomic force microscope. *Biophysical Journal* 82(5), 2798-2810. doi: 10.1016/s0006-3495(02)75620-8.
- Doktycz, M.J., Sullivan, C.J., Hoyt, P.R., Pelletier, D.A., Wu, S., and Allison, D.P. (2003). AFM imaging of bacteria in liquid media immobilized on gelatin coated mica surfaces. *Ultramicroscopy* 97(1-4), 209-216. doi: 10.1016/S0304-3991(03)00045-7.
- Döring, G., Parameswaran, I.G., and Murphy, T.F. (2011). Differential adaptation of microbial pathogens to airways of patients with cystic fibrosis and chronic obstructive pulmonary disease. *FEMS Microbiology Reviews* 35(1), 124-146. doi: 10.1111/j.1574-6976.2010.00237.x.
- Dover, R.S., Bitler, A., Shimoni, E., Trieu-Cuot, P., and Shai, Y. (2015). Multiparametric AFM reveals turgor-responsive net-like peptidoglycan architecture in live *streptococci*. *Nature Communications* 6, 7193. doi: 10.1038/ncomms8193.
- Drenkard, E., and Ausubel, F.M. (2002). *Pseudomonas* biofilm formation and antibiotic resistance are linked to phenotypic variation. *Nature* 416(6882), 740-743. doi: 10.1038/416740a.
- Drevinek, P., Holden, M.T., Ge, Z., Jones, A.M., Ketchell, I., Gill, R.T., et al. (2008). Gene expression changes linked to antimicrobial resistance, oxidative stress, iron depletion and retained motility are observed when *Burkholderia cenocepacia* grows in cystic fibrosis sputum. *BMC Infect Dis* 8, 121. doi: 10.1186/1471-2334-8-121.
- Drevinek, P., and Mahenthiralingam, E. (2010). *Burkholderia cenocepacia* in cystic fibrosis: epidemiology and molecular mechanisms of virulence. *Clinical Microbiology and Infection* 16(7), 821-830. doi: 10.1111/j.1469-0691.2010.03237.x.
- Du Toit, A. (2019). Growth capacity and cell size. *Nature Reviews Microbiology* 17(1), 2. doi: 10.1038/s41579-018-0124-y.
- Duff, C., Murphy, P.G., Callaghan, M., and McClean, S. (2006). Differences in invasion and translocation of *Burkholderia cepacia* complex species in polarised lung epithelial cells in vitro. *Microbial Pathogenesis* 41(4-5), 183-192. doi: 10.1016/j.micpath.2006.07.005.
- Dufrène, Y. (2011). *Life at the nanoscale : atomic force microscopy of live cells*. Singapore: Pan Stanford Publishing.
- Dufrene, Y.F. (2004). Using nanotechniques to explore microbial surfaces. *Nature Reviews Microbiology* 2(6), 451-460. doi: 10.1038/nrmicro905.
- Dufrene, Y.F. (2014). Atomic force microscopy in microbiology: new structural and functional insights into the microbial cell surface. *MBio* 5(4), e01363-01314. doi: 10.1128/mBio.01363-14.
- Dufrene, Y.F. (2015). Sticky microbes: forces in microbial cell adhesion. *Trends in Microbiology* 23(6), 376-382. doi: 10.1016/j.tim.2015.01.011.
- Eaton, P., Fernandes, J.C., Pereira, E., Pintado, M.E., and Xavier Malcata, F. (2008). Atomic force microscopy study of the antibacterial effects of chitosans on *Escherichia coli* and *Staphylococcus aureus*. *Ultramicroscopy* 108(10), 1128-1134. doi: 10.1016/j.ultramic.2008.04.015.
- Ederer, G.M., and Matsen, J.M. (1972). Colonization and infection with *Pseudomonas cepacia*. *The Journal of Infectious Diseases* 125(6), 613-618. doi: 10.1093/infdis/125.6.613.
- El-Kirat-Chatel, S., Mil-Homens, D., Beaussart, A., Fialho, A.M., and Dufrene, Y.F. (2013). Single-molecule atomic force microscopy unravels the binding mechanism of a *Burkholderia cenocepacia* trimeric autotransporter adhesin. *Molecular Microbiology* 89(4), 649-659. doi: 10.1111/mmi.12301.
- Elbourne, A., Chapman, J., Gelmi, A., Cozzolino, D., Crawford, R.J., and Truong, V.K. (2019). Bacterial-nanostructure interactions: The role of cell elasticity and adhesion forces. *Journal of Colloid and Interface Science* 546, 192-210. doi: 10.1016/j.jcis.2019.03.050.

- Farrell, P.M. (2008). The prevalence of cystic fibrosis in the European Union. *Journal of Cystic Fibrosis* 7(5), 450-453. doi: 10.1016/j.jcf.2008.03.007.
- Fathy Mohamed, Y., Hamad, M., Ortega, X.P., and Valvano, M.A. (2017). The LpxL acyltransferase is required for normal growth and penta-acylation of lipid A in *Burkholderia cenocepacia*. *Molecular Microbiology* 104(1), 144-162. doi: 10.1111/mmi.13618.
- Fathy Mohamed, Y., Scott, N.E., Molinaro, A., Creuzenet, C., Ortega, X., Lertmemongkolchai, G., et al. (2019). A general protein O-glycosylation machinery conserved in *Burkholderia* species improves bacterial fitness and elicits glycan immunogenicity in humans. *Journal of Biological Chemistry* 294(36), 13248-13268. doi: 10.1074/jbc.RA119.009671.
- Faure, E., Kwong, K., and Nguyen, D. (2018). *Pseudomonas aeruginosa* in Chronic Lung Infections: How to Adapt Within the Host?. *Frontiers in Immunology* 9, 2416. doi: 10.3389/fimmu.2018.02416.
- Fazli, M., Almlblad, H., Rybtke, M.L., Givskov, M., Eberl, L., and Tolker-Nielsen, T. (2014). Regulation of biofilm formation in *Pseudomonas* and *Burkholderia* species. *Environmental Microbiology* 16(7), 1961-1981. doi: 10.1111/1462-2920.12448.
- Fehlner-Gardiner, C.C., Hopkins, T.M., and Valvano, M.A. (2002). Identification of a general secretory pathway in a human isolate of *Burkholderia vietnamiensis* (formerly *B. cepacia* complex genomovar V) that is required for the secretion of hemolysin and phospholipase C activities. *Microbial Pathogenesis* 32(5), 249-254. doi: 10.1006/mpat.2002.0503.
- Fernandes, J.C., Eaton, P., Gomes, A.M., Pintado, M.E., and Xavier Malcata, F. (2009). Study of the antibacterial effects of chitosans on *Bacillus cereus* (and its spores) by atomic force microscopy imaging and nanoindentation. *Ultramicroscopy* 109(8), 854-860. doi: 10.1016/j.ultramic.2009.03.015.
- Ferreira, A.S., Leitao, J.H., Silva, I.N., Pinheiro, P.F., Sousa, S.A., Ramos, C.G., et al. (2010). Distribution of cepacian biosynthesis genes among environmental and clinical *Burkholderia* strains and role of cepacian exopolysaccharide in resistance to stress conditions. *Applied Environmental Microbiology* 76(2), 441-450. doi: 10.1128/AEM.01828-09.
- Ferreira, M.R., Gomes, S.C., and Moreira, L.M. (2019). Mucoid switch in *Burkholderia cepacia* complex bacteria: Triggers, molecular mechanisms and implications in pathogenesis. *Adv Appl Microbiol* 107, 113-140. doi: 10.1016/bs.aambs.2019.03.001.
- Figurski, D.H., and Helinski, D.R. (1979). Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *PNAS, Proceedings of the National Academy of Sciences* 76(4), 1648-1652. PMC383447.
- Flannagan, R.S., and Valvano, M.A. (2008). *Burkholderia cenocepacia* requires RpoE for growth under stress conditions and delay of phagolysosomal fusion in macrophages. *Microbiology* 154(Pt 2), 643-653. doi: 10.1099/mic.0.2007/013714-0.
- Foti, J.J., Devadoss, B., Winkler, J.A., Collins, J.J., and Walker, G.C. (2012). Oxidation of the guanine nucleotide pool underlies cell death by bactericidal antibiotics. *Science* 336(6079), 315-319. doi: 10.1126/science.1219192.
- Francius, G., Polyakov, P., Merlin, J., Abe, Y., Ghigo, J.M., Merlin, C., et al. (2011). Bacterial surface appendages strongly impact nanomechanical and electrokinetic properties of *Escherichia coli* cells subjected to osmotic stress. *PLoS One* 6(5), e20066. doi: 10.1371/journal.pone.0020066.
- Gaboriaud, F., Dague, E., Baille, S., Jorand, F., Duval, J., and Thomas, F. (2006). Multiscale dynamics of the cell envelope of *Shewanella putrefaciens* as a response to pH change. *Colloids and Surfaces B: Biointerfaces* 52(2), 108-116.
- Galli, F., Battistoni, A., Gambari, R., Pompella, A., Bragonzi, A., Pilolli, F., et al. (2012). Oxidative stress and antioxidant therapy in cystic fibrosis. *Biochimica et Biophysica Acta (BBA)* 1822(5), 690-713. doi: 10.1016/j.bbadis.2011.12.012.
- Gan, L., Chen, S., and Jensen, G.J. (2008). Molecular organization of Gram-negative peptidoglycan. *PNAS, Proceedings of the National Academy of Sciences* 105(48), 18953-18957. doi: 10.1073/pnas.0808035105.
- Ganesan, S., and Sajjan, U.S. (2011). Host evasion by *Burkholderia cenocepacia*. *Frontiers in Cellular and Infection Microbiology* 1, 25. doi: 10.3389/fcimb.2011.00025.
- Gavara, N. (2017). A beginner's guide to atomic force microscopy probing for cell mechanics. *Microscopy Research and Technique* 80(1), 75-84. doi: 10.1002/jemt.22776.
- Geftic, S.G., Heymann, H., and Adair, F.W. (1979). Fourteen-year survival of *Pseudomonas cepacia* in a salts solution preserved with benzalkonium chloride. *Applied and Environmental Microbiology* 37(3), 505-510. PMC243245.
- George, A.M., Jones, P.M., and Middleton, P.G. (2009). Cystic fibrosis infections: treatment strategies and prospects. *FEMS Microbiology Letters* 300(2), 153-164. doi: 10.1111/j.1574-6968.2009.01704.x.
- Gislason, A.S., Turner, K., Domaratzki, M., and Cardona, S.T. (2017). Comparative analysis of the *Burkholderia cenocepacia* K56-2 essential genome reveals cell envelope functions that are uniquely required for survival in species of the genus *Burkholderia*. *Microbial Genomics* 3(11). doi: 10.1099/mgen.0.000140.
- Govan, J.R., Brown, A.R., and Jones, A.M. (2007). Evolving epidemiology of *Pseudomonas aeruginosa* and the *Burkholderia cepacia* complex in cystic fibrosis lung infection. *Future Microbiology* 2(2), 153-164. doi: 10.2217/17460913.2.2.153.
- Govan, J.R., Brown, P.H., Maddison, J., Doherty, C.J., Nelson, J.W., Dodd, M., et al. (1993). Evidence for transmission of *Pseudomonas cepacia* by social contact in cystic fibrosis. *Lancet* 342(8862), 15-19. doi: 10.1016/0140-6736(93)91881-1.

- Govan, J.R., and Deretic, V. (1996). Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. *Microbiological Reviews* 60(3), 539-574. PMC239456.
- Graindorge, A., Menard, A., Monnez, C., and Cournoyer, B. (2012). Insertion sequence evolutionary patterns highlight convergent genetic inactivations and recent genomic island acquisitions among epidemic *Burkholderia cenocepacia*. *Journal of Medical Microbiology* 61(Pt 3), 394-409. doi: 10.1099/jmm.0.036822-0.
- Greenberg, D.E., Goldberg, J.B., Stock, F., Murray, P.R., Holland, S.M., and Lipuma, J.J. (2009). Recurrent *Burkholderia* infection in patients with chronic granulomatous disease: 11-year experience at a large referral center. *Clinical Infectious Diseases* 48(11), 1577-1579. doi: 10.1086/598937.
- Gronow, S., Noah, C., Blumenthal, A., Lindner, B., and Brade, H. (2003). Construction of a deep-rough mutant of *Burkholderia cepacia* ATCC 25416 and characterization of its chemical and biological properties. *Journal of Biological Chemistry* 278(3), 1647-1655. doi: 10.1074/jbc.M206942200.
- Guo, F.B., Xiong, L., Zhang, K.Y., Dong, C., Zhang, F.Z., and Woo, P.C. (2017). Identification and analysis of genomic islands in *Burkholderia cenocepacia* AU 1054 with emphasis on pathogenicity islands. *BMC Microbiology* 17(1), 73. doi: 10.1186/s12866-017-0986-6.
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29(8), 1072-1075. doi: 10.1093/bioinformatics/btt086.
- Hacker, J., and Carniel, E. (2001). Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. *EMBO Reports* 2(5), 376-381. doi: 10.1093/embo-reports/kve097.
- Hall, C.L., and Lee, V.T. (2018). Cyclic-di-GMP regulation of virulence in bacterial pathogens. *Wiley Interdiscip Rev RNA* 9(1). doi: 10.1002/wrna.1454.
- Hamad, M.A., Di Lorenzo, F., Molinaro, A., and Valvano, M.A. (2012). Aminoarabinose is essential for lipopolysaccharide export and intrinsic antimicrobial peptide resistance in *Burkholderia cenocepacia*. *Molecular Microbiology* 85(5), 962-974. doi: 10.1111/j.1365-2958.2012.08154.x.
- Hansma, P.K., Cleveland, J.P., Radmacher, M., Walters, D.A., Hillner, P.E., Bezantilla, M., et al. (1994). Tapping mode atomic force microscopy in liquids. *Applied Physics Letters* 64(13), 1738-1740. doi: 10.1063/1.111795.
- Harrison, F. (2007). Microbial ecology of the cystic fibrosis lung. *Microbiology* 153(Pt 4), 917-923. doi: 10.1099/mic.0.2006/004077-0.
- Hassan, A.A., Coutinho, C.P., and Sá-Correia, I. (2019). *Burkholderia cepacia* complex species differ in the frequency of variation of the lipopolysaccharide O-antigen expression during cystic fibrosis chronic respiratory infection. *Frontiers in Cellular and Infection Microbiology* 9(273). doi: 10.3389/fcimb.2019.00273.
- Hassan, A.A., Maldonado, R.F., Dos Santos, S.C., Di Lorenzo, F., Silipo, A., Coutinho, C.P., et al. (2017). Structure of O-Antigen and Hybrid Biosynthetic Locus in *Burkholderia cenocepacia* Clonal Variants Recovered from a Cystic Fibrosis Patient. *Frontiers in Microbiology* 8, 1027. doi: 10.3389/fmicb.2017.01027.
- Hatziagorou, E., Orenti, A., Drevinek, P., Kashirskaya, N., Mei-Zahav, M., De Boeck, K., et al. (2019). Changing epidemiology of the respiratory bacteriology of patients with cystic fibrosis-data from the European cystic fibrosis society patient registry. *Journal of Cystic Fibrosis*. doi: 10.1016/j.jcf.2019.08.006.
- Hedge, J., and Wilson, D.J. (2016). Practical Approaches for Detecting Selection in Microbial Genomes. *PLoS Computational Biology* 12(2), e1004739. doi: 10.1371/journal.pcbi.1004739.
- Heras, J., Dominguez, C., Mata, E., Pascual, V., Lozano, C., Torres, C., et al. (2015). GelJ--a tool for analyzing DNA fingerprint gel images. *BMC Bioinformatics* 16, 270. doi: 10.1186/s12859-015-0703-0.
- Herrmann, M., Schneck, E., Gutschmann, T., Brandenburg, K., and Tanaka, M. (2015). Bacterial lipopolysaccharides form physically cross-linked, two-dimensional gels in the presence of divalent cations. *Soft Matter* 11(30), 6037-6044. doi: 10.1039/c5sm01002k.
- Hill, P.J., Scordo, J.M., Arcos, J., Kirkby, S.E., Wewers, M.D., Wozniak, D.J., et al. (2017). Modifications of *Pseudomonas aeruginosa* cell envelope in the cystic fibrosis airway alters interactions with immune cells. *Scientific Reports* 7(1), 4761. doi: 10.1038/s41598-017-05253-9.
- Hoboth, C., Hoffmann, R., Eichner, A., Henke, C., Schmoltdt, S., Imhof, A., et al. (2009). Dynamics of adaptive microevolution of hypermutable *Pseudomonas aeruginosa* during chronic pulmonary infection in patients with cystic fibrosis. *The Journal of Infectious Diseases* 200(1), 118-130. doi: 10.1086/599360.
- Hogardt, M., and Heesemann, J. (2010). Adaptation of *Pseudomonas aeruginosa* during persistence in the cystic fibrosis lung. *International Journal of Medical Microbiology* 300(8), 557-562. doi: 10.1016/j.ijmm.2010.08.008.
- Holden, M.T., Seth-Smith, H.M., Crossman, L.C., Sebahia, M., Bentley, S.D., Cerdeno-Tarraga, A.M., et al. (2009). The genome of *Burkholderia cenocepacia* J2315, an epidemic pathogen of cystic fibrosis patients. *Journal of Bacteriology* 191(1), 261-277. doi: 10.1128/JB.01230-08.
- Hu, B., Perepelov, A.V., Liu, B., Shevelev, S.D., Guo, D., Senchenkova, S.N., et al. (2010). Structural and genetic evidence for the close relationship between *Escherichia coli* O71 and *Salmonella enterica* O28 O-antigens. *FEMS Immunology and Medical Microbiology* 59(2), 161-169. doi: 10.1111/j.1574-695X.2010.00676.x.

- Huang, M., Oppermann-Sanio, F.B., and Steinbuchel, A. (1999). Biochemical and molecular characterization of the *Bacillus subtilis* acetoin catabolic pathway. *Journal of Bacteriology* 181(12), 3837-3841. PMC93865.
- Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C., et al. (2017). Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Molecular Biology and Evolution* 34(8), 2115-2122. doi: 10.1093/molbev/msx148.
- Hughes, J.E., Stewart, J., Barclay, G.R., and Govan, J.R. (1997). Priming of neutrophil respiratory burst activity by lipopolysaccharide from *Burkholderia cepacia*. *Infection and Immunity* 65(10), 4281-4287. PMC175614.
- Hull, J., Vervaart, P., Grimwood, K., and Phelan, P. (1997). Pulmonary oxidative stress response in young children with cystic fibrosis. *Thorax* 52(6), 557-560. doi: 10.1136/thx.52.6.557.
- Hunt, T.A., Kooi, C., Sokol, P.A., and Valvano, M.A. (2004). Identification of *Burkholderia cenocepacia* genes required for bacterial survival in vivo. *Infection and Immunity* 72(7), 4010-4022. doi: 10.1128/IAI.72.7.4010-4022.2004.
- Hutchison, M.L., Bonell, E.C., Poxton, I.R., and Govan, J.R. (2000). Endotoxic activity of lipopolysaccharides isolated from emergent potential cystic fibrosis pathogens. *FEMS Immunology and Medical Microbiology* 27(1), 73-77. doi: 10.1111/j.1574-695X.2000.tb01414.x.
- Isles, A., Maclusky, I., Corey, M., Gold, R., Prober, C., Fleming, P., et al. (1984). *Pseudomonas cepacia* infection in cystic fibrosis: an emerging problem. *The Journal of Pediatrics* 104(2), 206-210. doi: 10.1016/s0022-3476(84)80993-2.
- Isshiki, Y., Kawahara, K., and Zahringer, U. (1998). Isolation and characterisation of disodium (4-amino-4-deoxy-beta-L-arabinopyranosyl)-(1-->8)-(D-glycero-alpha-D-talo-oct-2-ulopyranosylona te)- (2-->4)-(methyl 3-deoxy-D-manno-oct-2-ulopyranosid)onate from the lipopolysaccharide of *Burkholderia cepacia*. *Carbohydrate Research* 313(1), 21-27. doi: 10.1016/S0008-6215(98)00179-7.
- Ivanov, I.E., Kintz, E.N., Porter, L.A., Goldberg, J.B., Burnham, N.A., and Camesano, T.A. (2011). Relating the physical properties of *Pseudomonas aeruginosa* lipopolysaccharides to virulence by atomic force microscopy. *Journal of Bacteriology* 193(5), 1259-1266. doi: 10.1128/JB.01308-10.
- Jeannot, K., Elsen, S., Kohler, T., Attree, I., van Delden, C., and Plesiat, P. (2008). Resistance and virulence of *Pseudomonas aeruginosa* clinical strains overproducing the MexCD-OprJ efflux pump. *Antimicrobial Agents and Chemotherapy* 52(7), 2455-2462. doi: 10.1128/AAC.01107-07.
- Jia, B., Raphenya, A.R., Alcock, B., Waglechner, N., Guo, P., Tsang, K.K., et al. (2017). CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research* 45(D1), D566-D573. doi: 10.1093/nar/gkw1004.
- Johnson, S.L., Bishop-Lilly, K.A., Ladner, J.T., Daligault, H.E., Davenport, K.W., Jaissle, J., et al. (2016). Correction for Johnson et al., Complete Genome Sequences for 59 *Burkholderia* Isolates, Both Pathogenic and Near Neighbor. *Genome Announcements* 4(2). doi: 10.1128/genomeA.00313-16.
- Johnson, S.L., Bishop-Lilly, K.A., Ladner, J.T., Daligault, H.E., Davenport, K.W., Jaissle, J., et al. (2015). Complete genome sequences for 59 *burkholderia* isolates, both pathogenic and near neighbor. *Genome Announcements* 3(2). doi: 10.1128/genomeA.00159-15.
- Johnson, W.M., Tyler, S.D., and Rozee, K.R. (1994). Linkage analysis of geographic and clinical clusters in *Pseudomonas cepacia* infections by multilocus enzyme electrophoresis and ribotyping. *Journal of Clinical Microbiology* 32(4), 924-930. PMC263164.
- Johnston, R.B., Jr. (2001). Clinical aspects of chronic granulomatous disease. *Current Opinion in Hematology* 8(1), 17-22. doi: 10.1097/00062752-200101000-00004.
- Jolley, K.A., and Maiden, M.C.J. (2010). BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 11(1), 595. doi: 10.1186/1471-2105-11-595.
- Jones, A.M., Dodd, M.E., Govan, J.R., Barcus, V., Doherty, C.J., Morris, J., et al. (2004). *Burkholderia cenocepacia* and *Burkholderia multivorans*: influence on survival in cystic fibrosis. *Thorax* 59(11), 948-951. doi: 10.1136/thx.2003.017210.
- Jones, A.M., Dodd, M.E., and Webb, A.K. (2001). *Burkholderia cepacia*: current clinical issues, environmental controversies and ethical dilemmas. *European Respiratory Journal* 17(2), 295-301. doi: 10.1183/09031936.01.17202950.
- Kaas, R.S., Leekitcharoenphon, P., Aarestrup, F.M., and Lund, O. (2014). Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS One* 9(8), e104984. doi: 10.1371/journal.pone.0104984.
- Kalish, L.A., Waltz, D.A., Dovey, M., Potter-Bynoe, G., McAdam, A.J., Lipuma, J.J., et al. (2006). Impact of *Burkholderia dolosa* on lung function and survival in cystic fibrosis. *American Journal of Respiratory and Critical Care Medicine* 173(4), 421-425. doi: 10.1164/rccm.200503-344OC.
- Kamio, Y., and Nikaido, H. (1976). Outer membrane of *Salmonella typhimurium*: accessibility of phospholipid head groups to phospholipase c and cyanogen bromide activated dextran in the external medium. *Biochemistry* 15(12), 2561-2570.
- Kenna, D.T.D., Lilley, D., Coward, A., Martin, K., Perry, C., Pike, R., et al. (2017). Prevalence of *Burkholderia* species, including members of *Burkholderia cepacia* complex, among UK cystic and non-cystic fibrosis patients. *Journal of Medical Microbiology* 66(4), 490-501. doi: 10.1099/jmm.0.000458.

- Kim, J.M., Ahn, Y., LiPuma, J.J., Hussong, D., and Cerniglia, C.E. (2015). Survival and susceptibility of *Burkholderia cepacia* complex in chlorhexidine gluconate and benzalkonium chloride. *Journal of Industrial Microbiology and Biotechnology* 42(6), 905-913. doi: 10.1007/s10295-015-1605-x.
- Kim, S.H., Jia, W., Parreira, V.R., Bishop, R.E., and Gyles, C.L. (2006). Phosphoethanolamine substitution in the lipid A of *Escherichia coli* O157 : H7 and its association with PmrC. *Microbiology* 152(Pt 3), 657-666. doi: 10.1099/mic.0.28692-0.
- King, J.D., Kocincova, D., Westman, E.L., and Lam, J.S. (2009). Review: Lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa*. *Innate Immunity* 15(5), 261-312. doi: 10.1177/1753425909106436.
- Kintz, E., and Goldberg, J.B. (2008). Regulation of lipopolysaccharide O antigen expression in *Pseudomonas aeruginosa*. *Future Microbiology* 3(2), 191-203. doi: 10.2217/17460913.3.2.191.
- Kintz, E., Heiss, C., Black, I., Donohue, N., Brown, N., Davies, M.R., et al. (2017). *Salmonella enterica* Serovar Typhi Lipopolysaccharide O-antigen Modifications Impact on Serum Resistance and Antibody Recognition. *Infection and Immunity* 85(4). doi: 10.1128/IAI.01021-16.
- Kittelberger, R., and Hilbink, F. (1993). Sensitive silver-staining detection of bacterial lipopolysaccharides in polyacrylamide gels. *Journal of Biochemical and Biophysical Methods* 26(1), 81-86. doi: 10.1016/0165-022X(93)90024-I.
- Komatsu, H., Imura, Y., Ohori, A., Nagata, Y., and Tsuda, M. (2003). Distribution and organization of auxotrophic genes on the multichromosomal genome of *Burkholderia multivorans* ATCC 17616. *Journal of Bacteriology* 185(11), 3333-3343. doi: 10.1128/JB.185.11.3333-3343.2003.
- Kooi, C., Subsin, B., Chen, R., Pohorelic, B., and Sokol, P.A. (2006). *Burkholderia cenocepacia* ZmpB is a broad-specificity zinc metalloprotease involved in virulence. *Infection and Immunity* 74(7), 4083-4093. doi: 10.1128/iai.00297-06.
- Kotrange, S., Kopp, B., Akhter, A., Abdelaziz, D., Abu Khweek, A., Caution, K., et al. (2011). *Burkholderia cenocepacia* O polysaccharide chain contributes to caspase-1-dependent IL-1beta production in macrophages. *Journal of Leukocyte Biology* 89(3), 481-488. doi: 10.1189/jlb.0910513.
- Kryazhimskiy, S., and Plotkin, J.B. (2008). The population genetics of dN/dS. *PLoS Genetics* 4(12), e1000304. doi: 10.1371/journal.pgen.1000304.
- Lamothe, J., Huynh, K.K., Grinstein, S., and Valvano, M.A. (2007). Intracellular survival of *Burkholderia cenocepacia* in macrophages is associated with a delay in the maturation of bacteria-containing vacuoles. *Cellular Microbiology* 9(1), 40-53. doi: 10.1111/j.1462-5822.2006.00766.x.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9(4), 357-359. doi: 10.1038/nmeth.1923.
- Lee, A.H., Flibotte, S., Sinha, S., Paiero, A., Ehrlich, R.L., Balashov, S., et al. (2017). Phenotypic diversity and genotypic flexibility of *Burkholderia cenocepacia* during long-term chronic infection of cystic fibrosis lungs. *Genome Research* 27(4), 650-662. doi: 10.1101/gr.213363.116.
- Lee, B., Haagensen, J.A., Ciofu, O., Andersen, J.B., Hoiby, N., and Molin, S. (2005). Heterogeneity of biofilms formed by nonmucoid *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis. *Journal of Clinical Microbiology* 43(10), 5247-5255. doi: 10.1128/JCM.43.10.5247-5255.2005.
- Lee, H., Hsu, F.F., Turk, J., and Groisman, E.A. (2004). The PmrA-regulated pmrC gene mediates phosphoethanolamine modification of lipid A and polymyxin resistance in *Salmonella enterica*. *Journal of Bacteriology* 186(13), 4124-4133. doi: 10.1128/JB.186.13.4124-4133.2004.
- Lee, T.W., Verhey, T.B., Antiperovitch, P.A., Atamanyuk, D., Desroy, N., Oliveira, C., et al. (2013). Structural-functional studies of *Burkholderia cenocepacia* D-glycero-beta-D-manno-heptose 7-phosphate kinase (HldA) and characterization of inhibitors with antibiotic adjuvant and antivirulence properties. *J Med Chem* 56(4), 1405-1417. doi: 10.1021/jm301483h.
- Lefebvre, M., and Valvano, M. (2001). In vitro resistance of *Burkholderia cepacia* complex isolates to reactive oxygen species in relation to catalase and superoxide dismutase production. *Microbiology* 147(Pt 1), 97-109. doi: 10.1099/00221287-147-1-97.
- Leitao, J.H., Sousa, S.A., Cunha, M.V., Salgado, M.J., Melo-Cristino, J., Barreto, M.C., et al. (2008). Variation of the antimicrobial susceptibility profiles of *Burkholderia cepacia* complex clonal isolates obtained from chronically infected cystic fibrosis patients: a five-year survey in the major Portuguese treatment center. *European Journal of Clinical Microbiology and Infectious Diseases* 27(11), 1101-1111. doi: 10.1007/s10096-008-0552-0.
- Lerouge, I., and Vanderleyden, J. (2002). O-antigen structural variation: mechanisms and possible roles in animal/plant-microbe interactions. *FEMS Microbiology Reviews* 26(1), 17-47. doi: 10.1111/j.1574-6976.2002.tb00597.x.
- Lessie, T.G., Hendrickson, W., Manning, B.D., and Devereux, R. (1996). Genomic complexity and plasticity of *Burkholderia cepacia*. *FEMS Microbiology Letters* 144(2-3), 117-128. doi: 10.1111/j.1574-6968.1996.tb08517.x.
- Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26(5), 589-595. doi: 10.1093/bioinformatics/btp698.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16), 2078-2079. doi: 10.1093/bioinformatics/btp352.

- Lieberman, T.D., Flett, K.B., Yelin, I., Martin, T.R., McAdam, A.J., Priebe, G.P., et al. (2014). Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures. *Nature Genetics* 46(1), 82-87. doi: 10.1038/ng.2848.
- Lieberman, T.D., Michel, J.B., Aingaran, M., Potter-Bynoe, G., Roux, D., Davis, M.R., Jr., et al. (2011). Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nature Genetics* 43(12), 1275-1280. doi: 10.1038/ng.997.
- Lipuma, J.J. (2010). The changing microbial epidemiology in cystic fibrosis. *Clinical Microbiology Reviews* 23(2), 299-323. doi: 10.1128/CMR.00068-09.
- LiPuma, J.J., Dasen, S.E., Nielson, D.W., Stern, R.C., and Stull, T.L. (1990). Person-to-person transmission of *Pseudomonas cepacia* between patients with cystic fibrosis. *Lancet* 336(8723), 1094-1096. doi: 10.1016/0140-6736(90)92571-x.
- LiPuma, J.J., Spilker, T., Coenye, T., and Gonzalez, C.F. (2002). An epidemic *Burkholderia cepacia* complex strain identified in soil. *Lancet* 359(9322), 2002-2003. doi: 10.1016/S0140-6736(02)08836-0.
- LiPuma, J.J., Spilker, T., Gill, L.H., Campbell, P.W., 3rd, Liu, L., and Mahenthalingam, E. (2001). Disproportionate distribution of *Burkholderia cepacia* complex species and transmissibility markers in cystic fibrosis. *American Journal of Respiratory and Critical Care Medicine* 164(1), 92-96. doi: 10.1164/ajrccm.164.1.2011153.
- Liu, B., Zheng, D., Jin, Q., Chen, L., and Yang, J. (2019). VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Research* 47(D1), D687-D692. doi: 10.1093/nar/gky1080.
- Liu, L., Spilker, T., Coenye, T., and LiPuma, J.J. (2003). Identification by subtractive hybridization of a novel insertion element specific for two widespread *Burkholderia cepacia* genomovar III strains. *Journal of Clinical Microbiology* 41(6), 2471-2476. doi: 10.1128/jcm.41.6.2471-2476.2003.
- Liu, Y., and Camesano, T.A. (2008). "Immobilizing Bacteria for Atomic Force Microscopy Imaging or Force Measurements in Liquids," in *Microbial Surfaces*. American Chemical Society, 163-188.
- Lombardi, A., Andreozzi, C., Pavone, V., Triglione, V., Angiolini, L., and Caccia, P. (2013). Evaluation of the oligosaccharide composition of commercial follicle stimulating hormone preparations. *Electrophoresis* 34(16), 2394-2406. doi: 10.1002/elps.201300045.
- Lopez-Causape, C., Rojo-Molinero, E., Macia, M.D., and Oliver, A. (2015). The problems of antibiotic resistance in cystic fibrosis and solutions. *Expert Review of Respiratory Medicine* 9(1), 73-88. doi: 10.1586/17476348.2015.995640.
- Lorè, N.I., Cigana, C., De Fino, I., Riva, C., Juhas, M., Schwager, S., et al. (2012). Cystic fibrosis-niche adaptation of *Pseudomonas aeruginosa* reduces virulence in multiple infection hosts. *PLoS One* 7(4), e35648. doi: 10.1371/journal.pone.0035648.
- Lorenzo, F.D., Sturiale, L., Palmigiano, A., Lembo-Fazio, L., Paciello, I., Coutinho, C.P., et al. (2013). Chemistry and biology of the potent endotoxin from a *Burkholderia dolosa* clinical isolate from a cystic fibrosis patient. *ChemBioChem* 14(9), 1105-1115. doi: 10.1002/cbic.201300062.
- Loutet, S.A., Flanagan, R.S., Kooi, C., Sokol, P.A., and Valvano, M.A. (2006). A complete lipopolysaccharide inner core oligosaccharide is required for resistance of *Burkholderia cenocepacia* to antimicrobial peptides and bacterial survival in vivo. *Journal of Bacteriology* 188(6), 2073-2080. doi: 10.1128/JB.188.6.2073-2080.2006.
- Loutet, S.A., and Valvano, M.A. (2010). A decade of *Burkholderia cenocepacia* virulence determinant research. *Infection and Immunity* 78(10), 4088-4100. doi: 10.1128/IAI.00212-10.
- Lyczak, J.B., Cannon, C.L., and Pier, G.B. (2002). Lung infections associated with cystic fibrosis. *Clinical Microbiology Reviews* 15(2), 194-222. doi: 10.1128/cmr.15.2.194-222.2002.
- Madeira, A., dos Santos, S.C., Santos, P.M., Coutinho, C.P., Tyrrell, J., McClean, S., et al. (2013). Proteomic profiling of *Burkholderia cenocepacia* clonal isolates with different virulence potential retrieved from a cystic fibrosis patient during chronic lung infection. *PLoS One* 8(12), e83065. doi: 10.1371/journal.pone.0083065.
- Madeira, A., Santos, P.M., Coutinho, C.P., Pinto-de-Oliveira, A., and Sá-Correia, I. (2011). Quantitative proteomics (2-D DIGE) reveals molecular strategies employed by *Burkholderia cenocepacia* to adapt to the airways of cystic fibrosis patients under antimicrobial therapy. *Proteomics* 11(7), 1313-1328. doi: 10.1002/pmic.201000457.
- Mahenthalingam, E., Baldwin, A., and Dowson, C.G. (2008). *Burkholderia cepacia* complex bacteria: opportunistic pathogens with important natural biology. *Journal of Applied Microbiology* 104(6), 1539-1551. doi: 10.1111/j.1365-2672.2007.03706.x.
- Mahenthalingam, E., Baldwin, A., and Vandamme, P. (2002). *Burkholderia cepacia* complex infection in patients with cystic fibrosis. *Journal of Medical Microbiology* 51(7), 533-538. doi: 10.1099/0022-1317-51-7-533.
- Mahenthalingam, E., Bischof, J., Byrne, S.K., Radomski, C., Davies, J.E., Av-Gay, Y., et al. (2000a). DNA-Based diagnostic approaches for identification of *Burkholderia cepacia* complex, *Burkholderia vietnamiensis*, *Burkholderia multivorans*, *Burkholderia stabilis*, and *Burkholderia cepacia* genomovars I and III. *Journal of Clinical Microbiology* 38(9), 3165-3173. PMC87345.
- Mahenthalingam, E., Campbell, M.E., Henry, D.A., and Speert, D.P. (1996). Epidemiology of *Burkholderia cepacia* infection in patients with cystic fibrosis: analysis by randomly amplified polymorphic DNA fingerprinting. *Journal of Clinical Microbiology* 34(12), 2914-2920. PMC229433.

- Mahenthalingam, E., Coenye, T., Chung, J.W., Speert, D.P., Govan, J.R., Taylor, P., et al. (2000b). Diagnostically and experimentally useful panel of strains from the *Burkholderia cepacia* complex. *Journal of Clinical Microbiology* 38(2), 910-913. PMC86244.
- Mahenthalingam, E., Urban, T.A., and Goldberg, J.B. (2005). The multifarious, multireplicon *Burkholderia cepacia* complex. *Nature Reviews Microbiology* 3(2), 144-156. doi: 10.1038/nrmicro1085.
- Maldonado, R.F., Sá-Correia, I., and Valvano, M.A. (2016). Lipopolysaccharide modification in Gram-negative bacteria during chronic infection. *FEMS Microbiology Reviews* 40(4), 480-493. doi: 10.1093/femsre/fuw007.
- Manno, G., Dalmastrì, C., Tabacchioni, S., Vandamme, P., Lorini, R., Minicucci, L., et al. (2004). Epidemiology and clinical course of *Burkholderia cepacia* complex infections, particularly those caused by different *Burkholderia cenocepacia* strains, among patients attending an Italian Cystic Fibrosis Center. *Journal of Clinical Microbiology* 42(4), 1491-1497. PMC387599.
- Markussen, T., Marvig, R.L., Gomez-Lozano, M., Aanaes, K., Burleigh, A.E., Hoiby, N., et al. (2014). Environmental heterogeneity drives within-host diversification and evolution of *Pseudomonas aeruginosa*. *MBio* 5(5), e01592-01514. doi: 10.1128/mBio.01592-14.
- Marolda, C.L., Lahiry, P., Vines, E., Saldias, S., and Valvano, M.A. (2006). Micromethods for the characterization of lipid A-core and O-antigen lipopolysaccharide. *Methods in Molecular Biology* 347, 237-252. doi: 10.1385/1-59745-167-3:237.
- Marolda, C.L., Welsh, J., Dafoe, L., and Valvano, M.A. (1990). Genetic analysis of the O7-polysaccharide biosynthesis region from the *Escherichia coli* O7:K1 strain VW187. *Journal of Bacteriology* 172(7), 3590-3599. doi: 10.1128/jb.172.7.3590-3599.1990.
- Marquez, L., Jones, K.N., Whaley, E.M., Koy, T.H., Revell, P.A., Taylor, R.S., et al. (2017). An Outbreak of *Burkholderia cepacia* Complex Infections Associated with Contaminated Liquid Docusate. *Infection Control and Hospital Epidemiology* 38(5), 567-573. doi: 10.1017/ice.2017.11.
- Marr, N., Tirsoaga, A., Blanot, D., Fernandez, R., and Caroff, M. (2008). Glucosamine found as a substituent of both phosphate groups in *Bordetella* lipid A backbones: role of a BvgAS-activated ArnT ortholog. *Journal of Bacteriology* 190(12), 4281-4290. doi: 10.1128/JB.01875-07.
- Marteyn, B., West, N.P., Browning, D.F., Cole, J.A., Shaw, J.G., Palm, F., et al. (2010). Modulation of *Shigella* virulence in response to available oxygen in vivo. *Nature* 465(7296), 355-358. doi: 10.1038/nature08970.
- Martina, P., Bettiol, M., Vescina, C., Montanaro, P., Mannino, M.C., Prieto, C.I., et al. (2013). Genetic diversity of *Burkholderia contaminans* isolates from cystic fibrosis patients in Argentina. *Journal of Clinical Microbiology* 51(1), 339-344. doi: 10.1128/JCM.02500-12.
- Martina, P., Leguizamón, M., Prieto, C.I., Sousa, S.A., Montanaro, P., Draghi, W.O., et al. (2018). *Burkholderia puraquae* sp. nov., a novel species of the *Burkholderia cepacia* complex isolated from hospital settings and agricultural soils. *International Journal of Systematic and Evolutionary Microbiology* 68(1), 14-20. doi: 10.1099/ijsem.0.002293.
- Marvig, R.L., Dolce, D., Sommer, L.M., Petersen, B., Ciofu, O., Campana, S., et al. (2015). Within-host microevolution of *Pseudomonas aeruginosa* in Italian cystic fibrosis patients. *BMC Microbiol* 15, 218. doi: 10.1186/s12866-015-0563-9.
- Marvig, R.L., Johansen, H.K., Molin, S., and Jelsbak, L. (2013). Genome analysis of a transmissible lineage of *Pseudomonas aeruginosa* reveals pathoadaptive mutations and distinct evolutionary paths of hypermutators. *PLoS Genetics* 9(9), e1003741. doi: 10.1371/journal.pgen.1003741.
- McCarthy, R.R., Mazon-Moya, M.J., Moscoso, J.A., Hao, Y., Lam, J.S., Bordi, C., et al. (2017). Cyclic-di-GMP regulates lipopolysaccharide modification and contributes to *Pseudomonas aeruginosa* immune evasion. *Nature Microbiology* 2, 17027. doi: 10.1038/nmicrobiol.2017.27.
- McClellan, S., and Callaghan, M. (2009). *Burkholderia cepacia* complex: epithelial cell-pathogen confrontations and potential for therapeutic intervention. *Journal of Medical Microbiology* 58(Pt 1), 1-12. doi: 10.1099/jmm.0.47788-0.
- Mechaly, A.E., Sassoon, N., Betton, J.M., and Alzari, P.M. (2014). Segmental helical motions and dynamical asymmetry modulate histidine kinase autophosphorylation. *PLoS Biology* 12(1), e1001776. doi: 10.1371/journal.pbio.1001776.
- Medina-Pascual, M.J., Valdezate, S., Carrasco, G., Villalon, P., Garrido, N., and Saez-Nieto, J.A. (2015). Increase in isolation of *Burkholderia contaminans* from Spanish patients with cystic fibrosis. *Clinical Microbiology and Infection* 21(2), 150-156. doi: 10.1016/j.cmi.2014.07.014.
- Medina-Pascual, M.J., Valdezate, S., Villalon, P., Garrido, N., Rubio, V., and Saez-Nieto, J.A. (2012). Identification, molecular characterisation and antimicrobial susceptibility of genomovars of the *Burkholderia cepacia* complex in Spain. *European Journal of Clinical Microbiology and Infectious Diseases* 31(12), 3385-3396. doi: 10.1007/s10096-012-1707-6.
- Merhej, V., Georgiades, K., and Raoult, D. (2013). Postgenomic analysis of bacterial pathogens repertoire reveals genome reduction rather than virulence factors. *Briefings in Functional Genomics* 12(4), 291-304. doi: 10.1093/bfgp/elt015.
- Miethke, M., and Marahiel, M.A. (2007). Siderophore-based iron acquisition and pathogen control. *Microbiology and Molecular Biology Reviews* 71(3), 413-451. doi: 10.1128/MMBR.00012-07.
- Miller, E., Garcia, T., Hultgren, S., and Oberhauser, A.F. (2006). The mechanical properties of *E. coli* type 1 pili measured by atomic force microscopy techniques. *Biophysical Journal* 91(10), 3848-3856. doi: 10.1529/biophysj.106.088989.

- Miller, R.R., Hird, T.J., Tang, P., and Zlosnik, J.E. (2015). Whole-Genome Sequencing of Three Clonal Clinical Isolates of *B. cenocepacia* from a Patient with Cystic Fibrosis. *PLoS One* 10(11), e0143472. doi: 10.1371/journal.pone.0143472.
- Mira, N.P., Madeira, A., Moreira, A.S., Coutinho, C.P., and Sá-Correia, I. (2011). Genomic expression analysis reveals strategies of *Burkholderia cenocepacia* to adapt to cystic fibrosis patients' airways and antimicrobial therapy. *PLoS One* 6(12), e28831. doi: 10.1371/journal.pone.0028831.
- Moendarbary, E., and Harris, A.R. (2014). Cell mechanics: principles, practices, and prospects. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 6(5), 371-388. PMC4309479.
- Moehring, R.W., Lewis, S.S., Isaacs, P.J., Schell, W.A., Thomann, W.R., Althaus, M.M., et al. (2014). Outbreak of bacteremia due to *Burkholderia contaminans* linked to intravenous fentanyl from an institutional compounding pharmacy. *JAMA Internal Medicine* 174(4), 606-612. doi: 10.1001/jamainternmed.2013.13768.
- Molinaro, A., De Castro, C., Lanzetta, R., Evidente, A., Parrilli, M., and Holst, O. (2002). Lipopolysaccharides possessing two L-glycero-D-manno-heptopyranosyl- $\alpha$ -(1 $\rightarrow$ 5)-3-deoxy-D-manno-oct-2-ulopyranosonic acid moieties in the core region. The structure of the core region of the lipopolysaccharides from *Burkholderia caryophylli*. *The Journal of Biological Chemistry* 277(12), 10058-10063. doi: 10.1074/jbc.M110283200.
- Moreira, A.S., Coutinho, C.P., Azevedo, P., Lito, L., Melo-Cristino, J., and Sá-Correia, I. (2014). *Burkholderia dolosa* phenotypic variation during the decline in lung function of a cystic fibrosis patient during 5.5 years of chronic colonization. *Journal of Medical Microbiology* 63(Pt 4), 594-601. doi: 10.1099/jmm.0.069849-0.
- Moreira, A.S., Lourenco, A.B., and Sá-Correia, I. (2016). (1)H-NMR-Based Endometabolome Profiles of *Burkholderia cenocepacia* Clonal Variants Retrieved from a Cystic Fibrosis Patient during Chronic Infection. *Frontiers in Microbiology* 7(2024), 2024. doi: 10.3389/fmicb.2016.02024.
- Moreira, A.S., Mil-Homens, D., Sousa, S.A., Coutinho, C.P., Pinto-de-Oliveira, A., Ramos, C.G., et al. (2017). Variation of *Burkholderia cenocepacia* virulence potential during cystic fibrosis chronic lung infection. *Virulence* 8(6), 782-796. doi: 10.1080/21505594.2016.1237334.
- Moreno-Cencerrado, A., Iturri, J., Pecorari, I., M, D.M.V., Sbaizero, O., and Toca-Herrera, J.L. (2017). Investigating cell-substrate and cell-cell interactions by means of single-cell-probe force spectroscopy. *Microscopy Research and Technique* 80(1), 124-130. doi: 10.1002/jemt.22706.
- Moriarty, T.F., McElnay, J.C., Elborn, J.S., and Tunney, M.M. (2007). Sputum antibiotic concentrations: implications for treatment of cystic fibrosis lung infection. *Pediatric Pulmonology* 42(11), 1008-1017. doi: 10.1002/ppul.20671.
- Mowat, E., Paterson, S., Fothergill, J.L., Wright, E.A., Ledson, M.J., Walshaw, M.J., et al. (2011). *Pseudomonas aeruginosa* population diversity and turnover in cystic fibrosis chronic infections. *Ajrcm - American Journal of Respiratory and Critical Care Medicine* 183(12), 1674-1679. doi: 10.1164/rccm.201009-1430OC.
- Mullen, T., Markey, K., Murphy, P., McClean, S., and Callaghan, M. (2007). Role of lipase in *Burkholderia cepacia* complex (Bcc) invasion of lung epithelial cells. *European Journal of Clinical Microbiology and Infectious Diseases* 26(12), 869-877. doi: 10.1007/s10096-007-0385-2.
- Muller, D.J., Amrein, M., and Engel, A. (1997). Adsorption of biological molecules to a solid support for scanning probe microscopy. *Journal of Structural Biology* 119(2), 172-188. doi: 10.1006/jsbi.1997.3875.
- Muller, D.J., Helenius, J., Alsteens, D., and Dufrene, Y.F. (2009). Force probing surfaces of living cells to molecular resolution. *Nature Chemical Biology* 5(6), 383-390. doi: 10.1038/nchembio.181.
- Murray, G.L., Attridge, S.R., and Morona, R. (2003). Regulation of *Salmonella typhimurium* lipopolysaccharide O antigen chain length is required for virulence; identification of FepE as a second Wzz. *Molecular Microbiology* 47(5), 1395-1406. doi: DOI 10.1046/j.1365-2958.2003.03383.x.
- Needham, B.D., and Trent, M.S. (2013). Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis. *Nature Reviews Microbiology* 11(7), 467-481. doi: 10.1038/nrmicro3047.
- Nei, M., and Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* 3(5), 418-426. doi: 10.1093/oxfordjournals.molbev.a040410.
- Neiger, M.R., Gonzalez, J.F., Gonzalez-Escobedo, G., Kuck, H., White, P., and Gunn, J.S. (2019). Pathoadaptive alteration of *Salmonella* biofilm formation in response to the gallbladder environment. *Journal of Bacteriology* 201(14). doi: 10.1128/JB.00774-18.
- Nishiyama, E., Ohtsubo, Y., Nagata, Y., and Tsuda, M. (2010). Identification of *Burkholderia multivorans* ATCC 17616 genes induced in soil environment by in vivo expression technology. *Environmental Microbiology* 12(9), 2539-2558. doi: 10.1111/j.1462-2920.2010.02227.x.
- Nunvar, J., Capek, V., Fiser, K., Fila, L., and Drevinek, P. (2017). What matters in chronic *Burkholderia cenocepacia* infection in cystic fibrosis: Insights from comparative genomics. *PLoS Pathogens* 13(12), e1006762. doi: 10.1371/journal.ppat.1006762.
- Nunvar, J., Kalferstova, L., Bloodworth, R.A., Kolar, M., Degrossi, J., Lubovich, S., et al. (2016). Understanding the Pathogenicity of *Burkholderia contaminans*, an Emerging Pathogen in Cystic Fibrosis. *PLoS One* 11(8), e0160975. doi: 10.1371/journal.pone.0160975.

- O'Grady, E.P., and Sokol, P.A. (2011). *Burkholderia cenocepacia* differential gene expression during host-pathogen interactions and adaptation to the host environment. *Frontiers in Cellular and Infection Microbiology* 1, 15. doi: 10.3389/fcimb.2011.00015.
- O'Sullivan, L.A., and Mahenthiralingam, E. (2005). Biotechnological potential within the genus *Burkholderia*. *Letters in Applied Microbiology* 41(1), 8-11. doi: 10.1111/j.1472-765X.2005.01758.x.
- Oberhardt, M.A., Goldberg, J.B., Hogardt, M., and Papin, J.A. (2010). Metabolic network analysis of *Pseudomonas aeruginosa* during chronic cystic fibrosis lung infection. *Journal of Bacteriology* 192(20), 5534-5548. doi: 10.1128/JB.00900-10.
- Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405(6784), 299-304. doi: 10.1038/35012500.
- Okuda, S., Sherman, D.J., Silhavy, T.J., Ruiz, N., and Kahne, D. (2016). Lipopolysaccharide transport and assembly at the outer membrane: the PEZ model. *Nature Reviews Microbiology* 14(6), 337-345. doi: 10.1038/nrmicro.2016.25.
- Olagnon, C., Monjaras Feria, J., Grünwald-Gruber, C., Blaukopf, M., Valvano, M.A., and Kosma, P. (2019). Synthetic Phosphodiester-Linked 4-Amino-4-deoxy-l-arabinose Derivatives Demonstrate that ArnT is an Inverting Aminoarabosyl Transferase. *ChemBioChem* n/a(n/a). doi: 10.1002/cbic.201900349.
- Olivares, J., Alvarez-Ortega, C., and Martinez, J.L. (2014). Metabolic compensation of fitness costs associated with overexpression of the multidrug efflux pump MexEF-OprN in *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy* 58(7), 3904-3913. doi: 10.1128/AAC.00121-14.
- Olivares Pacheco, J., Alvarez-Ortega, C., Alcalde Rico, M., and Martinez, J.L. (2017). Metabolic Compensation of Fitness Costs Is a General Outcome for Antibiotic-Resistant *Pseudomonas aeruginosa* Mutants Overexpressing Efflux Pumps. *MBio* 8(4). doi: 10.1128/mBio.00500-17.
- Ong, K.S., Aw, Y.K., Lee, L.H., Yule, C.M., Cheow, Y.L., and Lee, S.M. (2016). *Burkholderia paludis* sp nov., an Antibiotic-Siderophore Producing Novel *Burkholderia cepacia* Complex Species, Isolated from Malaysian Tropical Peat Swamp Soil. *Frontiers in Microbiology* 7, 2046. doi: 10.3389/fmicb.2016.02046.
- Oppy, C.C., Jebeli, L., Kuba, M., Oates, C.V., Strugnell, R., Edgington-Mitchell, L.E., et al. (2019). Loss of O-Linked Protein Glycosylation in *Burkholderia cenocepacia* Impairs Biofilm Formation and Siderophore Activity and Alters Transcriptional Regulators. *mSphere* 4(6). doi: 10.1128/mSphere.00660-19.
- Ortega, X., Hunt, T.A., Loutet, S., Vinion-Dubiel, A.D., Datta, A., Choudhury, B., et al. (2005). Reconstitution of O-specific lipopolysaccharide expression in *Burkholderia cenocepacia* strain J2315, which is associated with transmissible infections in patients with cystic fibrosis. *Journal of Bacteriology* 187(4), 1324-1333. doi: 10.1128/JB.187.4.1324-1333.2005.
- Ortega, X., Silipo, A., Saldias, M.S., Bates, C.C., Molinaro, A., and Valvano, M.A. (2009). Biosynthesis and structure of the *Burkholderia cenocepacia* K56-2 lipopolysaccharide core oligosaccharide: truncation of the core oligosaccharide leads to increased binding and sensitivity to polymyxin B. *The Journal of Biological Chemistry* 284(32), 21738-21751. doi: 10.1074/jbc.M109.008532.
- Ortega, X.P., Cardona, S.T., Brown, A.R., Loutet, S.A., Flannagan, R.S., Campopiano, D.J., et al. (2007). A putative gene cluster for aminoarabinose biosynthesis is essential for *Burkholderia cenocepacia* viability. *Journal of Bacteriology* 189(9), 3639-3644. doi: 10.1128/JB.00153-07.
- Osiro, D., Filho, R.B., Assis, O.B., Jorge, L.A., and Colnago, L.A. (2012). Measuring bacterial cells size with AFM. *Brazilian Journal of Microbiology* 43(1), 341-347. doi: 10.1590/S1517-838220120001000040.
- Ota, T., and Nei, M. (1994). Variance and covariances of the numbers of synonymous and nonsynonymous substitutions per site. *Molecular Biology and Evolution* 11(4), 613-619. doi: 10.1093/oxfordjournals.molbev.a040140.
- Ovchinnikova, E.S., Krom, B.P., Harapanahalli, A.K., Busscher, H.J., and van der Mei, H.C. (2013). Surface thermodynamic and adhesion force evaluation of the role of chitin-binding protein in the physical interaction between *Pseudomonas aeruginosa* and *Candida albicans*. *Langmuir* 29(15), 4823-4829. doi: 10.1021/la400554g.
- Palmer, K.L., Aye, L.M., and Whiteley, M. (2007). Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum. *Journal of Bacteriology* 189(22), 8079-8087. doi: 10.1128/JB.01138-07.
- Panmanee, W., Gomez, F., Witte, D., Pancholi, V., Britigan, B.E., and Hasset, D.J. (2008). The peptidoglycan-associated lipoprotein OprL helps protect a *Pseudomonas aeruginosa* mutant devoid of the transactivator OxyR from hydrogen peroxide-mediated killing during planktonic and biofilm culture. *Journal of Bacteriology* 190(10), 3658-3669. doi: 10.1128/JB.00022-08.
- Parke, J.L., and Gurian-Sherman, D. (2001). Diversity of the *Burkholderia cepacia* complex and implications for risk assessment of biological control strains. *Annual Review of Phytopathology* 39, 225-258. doi: 10.1146/annurev.phyto.39.1.225.
- Peeters, C., Cooper, V.S., Hatcher, P.J., Verheyde, B., Carlier, A., and Vandamme, P. (2017). Comparative genomics of *Burkholderia multivorans*, a ubiquitous pathogen with a highly conserved genomic structure. *PLoS One* 12(4), e0176191. doi: 10.1371/journal.pone.0176191.

- Peeters, C., Zlosnik, J.E., Spilker, T., Hird, T.J., LiPuma, J.J., and Vandamme, P. (2013). *Burkholderia pseudomultivorans* sp. nov., a novel *Burkholderia cepacia* complex species from human respiratory samples and the rhizosphere. *Systematic and Applied Microbiology* 36(7), 483-489. doi: 10.1016/j.syam.2013.06.003.
- Peeters, E., Sass, A., Mahenthiralingam, E., Nelis, H., and Coenye, T. (2010). Transcriptional response of *Burkholderia cenocepacia* J2315 sessile cells to treatments with high doses of hydrogen peroxide and sodium hypochlorite. *BMC Genomics* 11, 90. doi: 10.1186/1471-2164-11-90.
- Persat, A., Nadell, C.D., Kim, M.K., Ingremeau, F., Siryaporn, A., Drescher, K., et al. (2015). The mechanical world of bacteria. *Cell* 161(5), 988-997. doi: 10.1016/j.cell.2015.05.005.
- Pessi, G., Braunwalder, R., Grunau, A., Omasits, U., Ahrens, C.H., and Eberl, L. (2013). Response of *Burkholderia cenocepacia* H111 to Micro-Oxia. *PLOS ONE* 8(9), e72939. doi: 10.1371/journal.pone.0072939.
- Pfeundschuh, M., Martinez-Martin, D., Mulvihill, E., Wegmann, S., and Muller, D.J. (2014). Multiparametric high-resolution imaging of native proteins by force-distance curve-based AFM. *Nature Protocols* 9(5), 1113-1130. doi: 10.1038/nprot.2014.070.
- Piantini, U., Sorensen, O.W., and Ernst, R.R. (1982). Multiple Quantum Filters for Elucidating NMR Coupling Networks. *Journal of the American Chemical Society* 104(24), 6800-6801. doi: 10.1021/ja00388a062.
- Pier, G.B. (2007). *Pseudomonas aeruginosa* lipopolysaccharide: a major virulence factor, initiator of inflammation and target for effective immunity. *International Journal of Medical Microbiology* 297(5), 277-295. doi: 10.1016/j.ijmm.2007.03.012.
- Power, R.F., Linnane, B., Martin, R., Power, N., Harnett, P., Casserly, B., et al. (2016). The first reported case of *Burkholderia contaminans* in patients with cystic fibrosis in Ireland: from the Sargasso Sea to Irish Children. *BMC Pulmonary Medicine* 16(1), 57. doi: 10.1186/s12890-016-0219-z.
- Price, E.P., Sarovich, D.S., Mayo, M., Tuanyok, A., Drees, K.P., Kaestli, M., et al. (2013). Within-host evolution of *Burkholderia pseudomallei* over a twelve-year chronic carriage infection. *MBio* 4(4), 1-10. doi: 10.1128/mBio.00388-13.
- Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 5(3), e9490. doi: 10.1371/journal.pone.0009490.
- Qin, Z., Zhang, J., Hu, Y., Chi, Q., Mortensen, N.P., Qu, D., et al. (2009). Organic compounds inhibiting *S. epidermidis* adhesion and biofilm formation. *Ultramicroscopy* 109(8), 881-888. doi: 10.1016/j.ultramic.2009.03.040.
- Raetz, C.R., and Whitfield, C. (2002). Lipopolysaccharide endotoxins. *Annual Review of Biochemistry* 71, 635-700. doi: 10.1146/annurev.biochem.71.110601.135414.
- Ramos, J.L., Martínez-Bueno, M., Molina-Henares, A.J., Terán, W., Watanabe, K., Zhang, X., et al. (2005). The TetR family of transcriptional repressors. *Microbiology and Molecular Biology Reviews : MMBR* 69(2), 326-356. doi: 10.1128/MMBR.69.2.326-356.2005.
- Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R., and Wüthrich, K. (1983). Improved spectral resolution in cosy 1H NMR spectra of proteins via double quantum filtering. *Biochemical and Biophysical Research Communications* 117(2), 479-485. doi: 10.1016/0006-291X(83)91225-1.
- Ranf, S. (2016). Immune Sensing of Lipopolysaccharide in Plants and Animals: Same but Different. *PLoS Pathogens* 12(6), e1005596. doi: 10.1371/journal.ppat.1005596.
- Ratjen, F., and Doring, G. (2003). Cystic fibrosis. *Lancet* 361(9358), 681-689. doi: 10.1016/S0140-6736(03)12567-6.
- Rau, M.H., Marvig, R.L., Ehrlich, G.D., Molin, S., and Jelsbak, L. (2012). Deletion and acquisition of genomic content during early stage adaptation of *Pseudomonas aeruginosa* to a human host environment. *Environmental Microbiology* 14(8), 2200-2211. doi: 10.1111/j.1462-2920.2012.02795.x.
- Reeves, P.R., Cunneen, M.M., Liu, B., and Wang, L. (2013). Genetics and evolution of the *Salmonella* galactose-initiated set of o antigens. *PLoS One* 8(7), e69306. doi: 10.1371/journal.pone.0069306.
- Reid, D.W., Misso, N., Aggarwal, S., Thompson, P.J., and Walters, E.H. (2007). Oxidative stress and lipid-derived inflammatory mediators during acute exacerbations of cystic fibrosis. *Respirology* 12(1), 63-69. doi: 10.1111/j.1440-1843.2006.00962.x.
- Reik, R., Spilker, T., and 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. doi: 10.1128/JCM.43.6.2926-2928.2005.
- Reyes, R.E., González, C.R., Jiménez, R.C., Herrera, M.O., and Andrade, A.A. (2012). "Mechanisms of O-Antigen Structural Variation of Bacterial Lipopolysaccharide (LPS)," in *The Complex World of Polysaccharides*, ed. D.N. Karunaratne. (Rijeka: InTech), Ch. 03.
- Reyrat, J.M., David, M., Blonski, C., Boistard, P., and Batut, J. (1993). Oxygen-regulated in vitro transcription of *Rhizobium meliloti* nifA and fixK genes. *Journal of Bacteriology* 175(21), 6867-6872. doi: 10.1128/jb.175.21.6867-6872.1993.
- Rhodes, K.A., and Schweizer, H.P. (2016). Antibiotic resistance in *Burkholderia* species. *Drug Resistance Updates* 28, 82-90. doi: 10.1016/j.drug.2016.07.003.

- Richau, J.A., Leitao, J.H., Correia, M., Lito, L., Salgado, M.J., Barreto, C., et al. (2000). Molecular typing and exopolysaccharide biosynthesis of *Burkholderia cepacia* isolates from a Portuguese cystic fibrosis center. *Journal of Clinical Microbiology* 38(4), 1651-1655. PMC86514.
- Rietschel, E.T. (1976). Absolute-Configuration of 3-Hydroxy Fatty-Acids Present in Lipopolysaccharides from Various Bacterial Groups. *European Journal of Biochemistry* 64(2), 423-428. doi: 10.1111/j.1432-1033.1976.tb10318.x.
- Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., et al. (2011). Integrative genomics viewer. *Nature Biotechnology* 29(1), 24-26. doi: 10.1038/nbt.1754.
- Rocha, E.P., Smith, J.M., Hurst, L.D., Holden, M.T., Cooper, J.E., Smith, N.H., et al. (2006). Comparisons of dN/dS are time dependent for closely related bacterial genomes. *Journal of Theoretical Biology* 239(2), 226-235. doi: 10.1016/j.jtbi.2005.08.037.
- Roszniewski, B., McClean, S., and Drulis-Kawa, Z. (2018). *Burkholderia cenocepacia* Prophages-Prevalence, Chromosome Location and Major Genes Involved. *Viruses* 10(6). doi: 10.3390/v10060297.
- Roux, D., Weatherholt, M., Clark, B., Gadjeva, M., Renaud, D., Scott, D., et al. (2017). Immune Recognition of the Epidemic Cystic Fibrosis Pathogen *Burkholderia dolosa*. *Infect Immun* 85(6). doi: 10.1128/IAI.00765-16.
- Rowbotham, N.J., Palser, S.C., Smith, S.J., and Smyth, A.R. (2019). Infection prevention and control in cystic fibrosis: a systematic review of interventions. *Expert Review of Respiratory Medicine* 13(5), 425-434. doi: 10.1080/17476348.2019.1595594.
- Roze, K.R., Haase, D., Macdonald, N.E., and Johnson, W.M. (1994). Comparison by extended ribotyping of *Pseudomonas cepacia* isolated from cystic fibrosis patients with acute and chronic infections. *Diagnostic Microbiology and Infectious Disease* 20(4), 181-186. doi: 10.1016/0732-8893(94)90001-9.
- Rubio, T., Oyanedel, D., Labreuche, Y., Toulza, E., Luo, X., Bruto, M., et al. (2019). Species-specific mechanisms of cytotoxicity toward immune cells determine the successful outcome of *Vibrio* infections. *PNAS, Proceedings of the National Academy of Sciences* 116(28), 14238-14247. doi: 10.1073/pnas.1905747116.
- Saar-Dover, R., Bitler, A., Nezer, R., Shmuel-Galia, L., Firon, A., Shimoni, E., et al. (2012). D-alanylation of lipoteichoic acids confers resistance to cationic peptides in group B. *streptococcus* by increasing the cell wall density. *PLoS Pathogens* 8(9), e1002891. doi: 10.1371/journal.ppat.1002891.
- Sader, J.E., Borgani, R., Gibson, C.T., Haviland, D.B., Higgins, M.J., Kilpatrick, J.I., et al. (2016). A virtual instrument to standardise the calibration of atomic force microscope cantilevers. *Review of Scientific Instruments* 87(9), 093711. doi: 10.1063/1.4962866.
- Sahu, K., Bansal, H., Mukherjee, C., Sharma, M., and Gupta, P.K. (2009). Atomic force microscopic study on morphological alterations induced by photodynamic action of Toluidine Blue O in *Staphylococcus aureus* and *Escherichia coli*. *J Photochem Photobiol B* 96(1), 9-16. doi: 10.1016/j.jphotobiol.2009.03.008.
- Saito, K., Ito, E., Hosono, K., Nakamura, K., Imai, K., Iizuka, T., et al. (2003). The uncoupling of oxygen sensing, phosphorylation signalling and transcriptional activation in oxygen sensor FixL and FixJ mutants. *Molecular Microbiology* 48(2), 373-383.
- Sajjan, S.U., Carmody, L.A., Gonzalez, C.F., and LiPuma, J.J. (2008). A type IV secretion system contributes to intracellular survival and replication of *Burkholderia cenocepacia*. *Infection and Immunity* 76(12), 5447-5455. doi: 10.1128/IAI.00451-08.
- Sajjan, U.S., and Forstner, J.F. (1993). Role of a 22-kilodalton pilin protein in binding of *Pseudomonas cepacia* to buccal epithelial cells. *Infection and Immunity* 61(8), 3157-3163. PMC280983.
- Saldias, M.S., Lamothe, J., Wu, R., and Valvano, M.A. (2008). *Burkholderia cenocepacia* requires the RpoN sigma factor for biofilm formation and intracellular trafficking within macrophages. *Infection and Immunity* 76(3), 1059-1067. doi: 10.1128/IAI.01167-07.
- Saldias, M.S., Ortega, X., and Valvano, M.A. (2009). *Burkholderia cenocepacia* O antigen lipopolysaccharide prevents phagocytosis by macrophages and adhesion to epithelial cells. *Journal of Medical Microbiology* 58(Pt 12), 1542-1548. doi: 10.1099/jmm.0.013235-0.
- Saldias, M.S., and Valvano, M.A. (2009). Interactions of *Burkholderia cenocepacia* and other *Burkholderia cepacia* complex bacteria with epithelial and phagocytic cells. *Microbiology* 155(Pt 9), 2809-2817. doi: 10.1099/mic.0.031344-0.
- Samant, S., Lee, H., Ghassemi, M., Chen, J., Cook, J.L., Mankin, A.S., et al. (2008). Nucleotide biosynthesis is critical for growth of bacteria in human blood. *PLoS Pathogens* 4(2), e37. doi: 10.1371/journal.ppat.0040037.
- Sass, A., Marchbank, A., Tullis, E., Lipuma, J.J., and Mahenthalingam, E. (2011). Spontaneous and evolutionary changes in the antibiotic resistance of *Burkholderia cenocepacia* observed by global gene expression analysis. *BMC Genomics* 12, 373. doi: 10.1186/1471-2164-12-373.
- Sass, A.M., Schmerk, C., Agnoli, K., Norville, P.J., Eberl, L., Valvano, M.A., et al. (2013). The unexpected discovery of a novel low-oxygen-activated locus for the anoxic persistence of *Burkholderia cenocepacia*. *The ISME Journal* 7(8), 1568-1581. doi: 10.1038/ismej.2013.36.
- Sawana, A., Adeolu, M., and Gupta, R.S. (2014). Molecular signatures and phylogenomic analysis of the genus *Burkholderia*: proposal for division of this genus into the emended genus *Burkholderia* containing pathogenic organisms and a new

- genus *Paraburkholderia* gen. nov. harboring environmental species. *Frontiers in Genetics* 5, 429. doi: 10.3389/fgene.2014.00429.
- Schaefer, M.M., Liao, T.L., Boisvert, N.M., Roux, D., Yoder-Himes, D., and Priebe, G.P. (2017). An Oxygen-Sensing Two-Component System in the *Burkholderia cepacia* Complex Regulates Biofilm, Intracellular Invasion, and Pathogenicity. *PLoS Pathogen* 13(1), e1006116. doi: 10.1371/journal.ppat.1006116.
- Schillers, H., Rianna, C., Schape, J., Luque, T., Doschke, H., Walte, M., et al. (2017). Standardized Nanomechanical Atomic Force Microscopy Procedure (SNAP) for Measuring Soft and Biological Samples. *Scientific Reports* 7(1), 5117. doi: 10.1038/s41598-017-05383-0.
- Schmerk, C.L., and Valvano, M.A. (2013). *Burkholderia multivorans* survival and trafficking within macrophages. *Journal of Medical Microbiology* 62(Pt 2), 173-184. doi: 10.1099/jmm.0.051243-0.
- Schwab, U., Abdullah, L.H., Perlmutter, O.S., Albert, D., Davis, C.W., Arnold, R.R., et al. (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. doi: 10.1128/IAI.01876-14.
- Schwager, S., Agnoli, K., Kothe, M., Feldmann, F., Givskov, M., Carlier, A., et al. (2013). Identification of *Burkholderia cenocepacia* strain H111 virulence factors using nonmammalian infection hosts. *Infection and Immunity* 81(1), 143-153. doi: 10.1128/IAI.00768-12.
- Scoffone, V.C., Chiarelli, L.R., Trespidi, G., Mentasti, M., Riccardi, G., and Buroni, S. (2017). *Burkholderia cenocepacia* Infections in Cystic Fibrosis Patients: Drug Resistance and Therapeutic Approaches. *Frontiers in Microbiology* 8, 1592. doi: 10.3389/fmicb.2017.01592.
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30(14), 2068-2069. doi: 10.1093/bioinformatics/btu153.
- Sfeir, M.M. (2018). *Burkholderia cepacia* complex infections: More complex than the bacterium name suggest. *Journal of Infection* 77(3), 166-170. doi: 10.1016/j.jinf.2018.07.006.
- Shimomura, H., Matsuura, M., Saito, S., Hirai, Y., Isshiki, Y., and Kawahara, K. (2001). Lipopolysaccharide of *Burkholderia cepacia* and its unique character to stimulate murine macrophages with relative lack of interleukin-1 $\beta$ -inducing ability. *Infection and Immunity* 69(6), 3663-3669. doi: 10.1128/IAI.69.6.3663-3669.2001.
- Shinoy, M., Dennehy, R., Coleman, L., Carberry, S., Schaffer, K., Callaghan, M., et al. (2013). Immunoproteomic analysis of proteins expressed by two related pathogens, *Burkholderia multivorans* and *Burkholderia cenocepacia*, during human infection. *PLoS One* 8(11), e80796. doi: 10.1371/journal.pone.0080796.
- Shommu, N.S., Vogel, H.J., and Storey, D.G. (2015). Potential of metabolomics to reveal *Burkholderia cepacia* complex pathogenesis and antibiotic resistance. *Frontiers in Microbiology* 6, 668. doi: 10.3389/fmicb.2015.00668.
- Silhavy, T.J., Kahne, D., and Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harbor Perspectives in Biology* 2(5), a000414. doi: 10.1101/cshperspect.a000414.
- Silipo, A., Molinaro, A., Cescutti, P., Bedini, E., Rizzo, R., Parrilli, M., et al. (2005). Complete structural characterization of the lipid A fraction of a clinical strain of *B. cepacia* genomovar I lipopolysaccharide. *Glycobiology* 15(5), 561-570. doi: 10.1093/glycob/cwi029.
- Silipo, A., Molinaro, A., Ierano, T., De Soyza, A., Sturiale, L., Garozzo, D., et al. (2007). The complete structure and pro-inflammatory activity of the lipooligosaccharide of the highly epidemic and virulent gram-negative bacterium *Burkholderia cenocepacia* ET-12 (strain J2315). *Chemistry* 13(12), 3501-3511. doi: 10.1002/chem.200601406.
- Silva, I.N., Ferreira, A.S., Becker, J.D., Zlosnik, J.E., Speert, D.P., He, J., et al. (2011). Mucoid morphotype variation of *Burkholderia multivorans* during chronic cystic fibrosis lung infection is correlated with changes in metabolism, motility, biofilm formation and virulence. *Microbiology* 157(Pt 11), 3124-3137. doi: 10.1099/mic.0.050989-0.
- Silva, I.N., Santos, P.M., Santos, M.R., Zlosnik, J.E., Speert, D.P., Buskirk, S.W., et al. (2016). Long-Term Evolution of *Burkholderia multivorans* during a Chronic Cystic Fibrosis Infection Reveals Shifting Forces of Selection. *mSystems* 1(3), e00029-00016. doi: 10.1128/mSystems.00029-16.
- Skerker, J.M., Prasol, M.S., Perchuk, B.S., Biondi, E.G., and Laub, M.T. (2005). Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. *PLoS biology* 3(10), e334. doi: 10.1371/journal.pbio.0030334.
- Smit, J., Kamio, Y., and Nikaido, H. (1975). Outer membrane of *Salmonella typhimurium*: chemical analysis and freeze-fracture studies with lipopolysaccharide mutants. *Journal of Bacteriology* 124(2), 942-958. PMC235985.
- Smith, D.L., Gumery, L.B., Smith, E.G., Stableforth, D.E., Kaufmann, M.E., and Pitt, T.L. (1993). Epidemic of *Pseudomonas cepacia* in an adult cystic fibrosis unit: evidence of person-to-person transmission. *Journal of Clinical Microbiology* 31(11), 3017-3022. PMC266193.
- Smolyakov, G., Formosa-Dague, C., Severac, C., Duval, R.E., and Dague, E. (2016). High speed indentation measures by FV, QI and QNM introduce a new understanding of bionanomechanical experiments. *Micron* 85, 8-14. doi: 10.1016/j.micron.2016.03.002.
- Sokol, P.A., Darling, P., Lewenza, S., Corbett, C.R., and Kooi, C.D. (2000). Identification of a siderophore receptor required for ferric ornibactin uptake in *Burkholderia cepacia*. *Infection and Immunity* 68(12), 6554-6560. doi: 10.1128/iai.68.12.6554-6560.2000.

- Sokol, P.A., Darling, P., Woods, D.E., Mahenthiralingam, E., and Kooi, C. (1999). Role of ornibactin biosynthesis in the virulence of *Burkholderia cepacia*: characterization of pvdA, the gene encoding L-ornithine N(5)-oxygenase. *Infection and Immunity* 67(9), 4443-4455. PMC96763.
- Sousa, S.A., Feliciano, J.R., Pita, T., Guerreiro, S.I., and Leitao, J.H. (2017). *Burkholderia cepacia* Complex Regulation of Virulence Gene Expression: A Review. *Genes (Basel)* 8(1). doi: 10.3390/genes8010043.
- Sousa, S.A., Ramos, C.G., and Leitao, J.H. (2011). *Burkholderia cepacia* Complex: Emerging Multihost Pathogens Equipped with a Wide Range of Virulence Factors and Determinants. *International Journal of Microbiology* 2011, 1-9. doi: 10.1155/2011/607575.
- Speert, D.P., Henry, D., Vandamme, P., Corey, M., and Mahenthiralingam, E. (2002). Epidemiology of *Burkholderia cepacia* complex in patients with cystic fibrosis, Canada. *Emerging Infectious Diseases journal (EID Journal)* 8(2), 181-187. doi: 10.3201/eid0802.010163.
- Spencer, D.H., Kas, A., Smith, E.E., Raymond, C.K., Sims, E.H., Hastings, M., et al. (2003). Whole-genome sequence variation among multiple isolates of *Pseudomonas aeruginosa*. *Journal of Bacteriology* 185(4), 1316-1325. doi: 10.1128/JB.185.4.1316-1325.2003.
- Spiewak, H.L., Shastri, S., Zhang, L., Schwager, S., Eberl, L., Vergunst, A.C., et al. (2019). *Burkholderia cenocepacia* utilizes a type VI secretion system for bacterial competition. *Microbiologyopen*, e774. doi: 10.1002/mbo3.774.
- Spilker, T., Baldwin, A., Bumford, A., Dowson, C.G., Mahenthiralingam, E., and LiPuma, J.J. (2009). Expanded multilocus sequence typing for *burkholderia* species. *Journal of Clinical Microbiology* 47(8), 2607-2610. doi: 10.1128/JCM.00770-09.
- Sriramulu, D.D., Lunsdorf, H., Lam, J.S., and Romling, U. (2005). Microcolony formation: a novel biofilm model of *Pseudomonas aeruginosa* for the cystic fibrosis lung. *Journal of Medical Microbiology* 54(Pt 7), 667-676. doi: 10.1099/jmm.0.45969-0.
- States, D.J., Haberkorn, R.A., and Ruben, D.J. (1982). A two-dimensional nuclear overhauser experiment with pure absorption phase in four quadrants. *Journal of Magnetic Resonance (1969)* 48(2), 286-292. doi: 10.1016/0022-2364(82)90279-7.
- Stern, A.S., Li, K.B., and Hoch, J.C. (2002). Modern spectrum analysis in multidimensional NMR spectroscopy: comparison of linear-prediction extrapolation and maximum-entropy reconstruction. *Journal of the American Chemical Society* 124(9), 1982-1993. doi: 10.1021/ja011669o.
- Stewart, P.S. (2014). Biophysics of biofilm infection. *Pathogens and Disease* 70(3), 212-218. doi: 10.1111/2049-632X.12118.
- Strauss, J., Burnham, N.A., and Camesano, T.A. (2009). Atomic force microscopy study of the role of LPS O-antigen on adhesion of *E. coli*. *Journal of Molecular Recognition* 22(5), 347-355. doi: 10.1002/jmr.955.
- Sturiale, L., Garozzo, D., Silipo, A., Lanzetta, R., Parrilli, M., and Molinaro, A. (2005). New conditions for matrix-assisted laser desorption/ionization mass spectrometry of native bacterial R-type lipopolysaccharides. *Rapid Communications in Mass Spectrometry* 19(13), 1829-1834. doi: 10.1002/rcm.1994.
- Summer, E.J., Gonzalez, C.F., Carlisle, T., Mebane, L.M., Cass, A.M., Savva, C.G., et al. (2004). *Burkholderia cenocepacia* phage BcepMu and a family of Mu-like phages encoding potential pathogenesis factors. *Journal of Molecular Biology* 340(1), 49-65. doi: 10.1016/j.jmb.2004.04.053.
- Suppiger, A., Schmid, N., Aguilar, C., Pessi, G., and Eberl, L. (2013). Two quorum sensing systems control biofilm formation and virulence in members of the *Burkholderia cepacia* complex. *Virulence* 4(5), 400-409. doi: 10.4161/viru.25338.
- Tatusova, T., Ciufu, S., Fedorov, B., O'Neill, K., and Tolstoy, I. (2014). RefSeq microbial genomes database: new representation and annotation strategy. *Nucleic Acids Research* 42(Database issue), D553-559. doi: 10.1093/nar/gkt1274.
- Tatusova, T., Ciufu, S., Fedorov, B., O'Neill, K., and Tolstoy, I. (2015). RefSeq microbial genomes database: new representation and annotation strategy. *Nucleic Acids Research* 43(7), 3872. doi: 10.1093/nar/gkv278.
- Teri, A., Sottotetti, S., Biffi, A., Girelli, D., D'Accico, M., Arghittu, M., et al. (2018). Molecular typing of *Burkholderia cepacia* complex isolated from patients attending an Italian Cystic Fibrosis Centre. *New Microbiologica* 41(2), 141-144.
- Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* 14(2), 178-192. doi: 10.1093/bib/bbs017.
- Thrane, S.W., Taylor, V.L., Freschi, L., Kukavica-Ibrulj, I., Boyle, B., Laroche, J., et al. (2015). The Widespread Multidrug-Resistant Serotype O12 *Pseudomonas aeruginosa* Clone Emerged through Concomitant Horizontal Transfer of Serotype Antigen and Antibiotic Resistance Gene Clusters. *MBio* 6(5), e01396-01315. doi: 10.1128/mBio.01396-15.
- Toguchi, A., Siano, M., Burkart, M., and Harshey, R.M. (2000). Genetics of swarming motility in *Salmonella enterica* serovar typhimurium: critical role for lipopolysaccharide. *Journal of Bacteriology* 182(22), 6308-6321. doi: 10.1128/JB.182.22.6308-6321.2000.
- Tomich, M., Herfst, C.A., Golden, J.W., and Mohr, C.D. (2002). Role of flagella in host cell invasion by *Burkholderia cepacia*. *Infection and Immunity* 70(4), 1799-1806. doi: 10.1128/iai.70.4.1799-1806.2002.
- Tomomori, C., Tanaka, T., Dutta, R., Park, H., Saha, S.K., Zhu, Y., et al. (1999). Solution structure of the homodimeric core domain of *Escherichia coli* histidine kinase EnvZ. *Nature Structural and Molecular Biology* 6(8), 729-734. doi: 10.1038/11495.

- Torre, B., Ricci, D., and Braga, P.C. (2011). How the atomic force microscope works? *Methods in Molecular Biology* 736, 3-18. doi: 10.1007/978-1-61779-105-5\_1.
- Touhami, A., Nysten, B., and Dufrene, Y.F. (2003). Nanoscale mapping of the elasticity of microbial cells by atomic force microscopy. *Langmuir* 19(11), 4539-4543. doi: 10.1021/la034136x.
- Traverse, C.C., Mayo-Smith, L.M., Poltak, S.R., and Cooper, V.S. (2013). Tangled bank of experimentally evolved *Burkholderia* biofilms reflects selection during chronic infections. *PNAS, Proceedings of the National Academy of Sciences* 110(3), E250-259. doi: 10.1073/pnas.1207025110.
- Trent, M.S., Ribeiro, A.A., Lin, S., Cotter, R.J., and Raetz, C.R. (2001). An inner membrane enzyme in *Salmonella* and *Escherichia coli* that transfers 4-amino-4-deoxy-L-arabinose to lipid A: induction on polymyxin-resistant mutants and role of a novel lipid-linked donor. *The Journal of Biological Chemistry* 276(46), 43122-43131. doi: 10.1074/jbc.M106961200.
- Trivedi, R.R., Crooks, J.A., Auer, G.K., Pendry, J., Foik, I.P., Siryaporn, A., et al. (2018). Mechanical Genomic Studies Reveal the Role of d-Alanine Metabolism in *Pseudomonas aeruginosa* Cell Stiffness. *MBio* 9(5). doi: 10.1128/mBio.01340-18.
- Turner, R.D., Hobbs, J.K., and Foster, S.J. (2016). Atomic Force Microscopy Analysis of Bacterial Cell Wall Peptidoglycan Architecture. *Methods in Molecular Biology* 1440, 3-9. doi: 10.1007/978-1-4939-3676-2\_1.
- Turner, R.D., Hurd, A.F., Cadby, A., Hobbs, J.K., and Foster, S.J. (2013). Cell wall elongation mode in Gram-negative bacteria is determined by peptidoglycan architecture. *Nature Communication* 4, 1496. doi: 10.1038/ncomms2503.
- Turner, R.D., Ratcliffe, E.C., Wheeler, R., Golestanian, R., Hobbs, J.K., and Foster, S.J. (2010). Peptidoglycan architecture can specify division planes in *Staphylococcus aureus*. *Nature Communication* 1, 26. doi: 10.1038/ncomms1025.
- Turner, R.D., Vollmer, W., and Foster, S.J. (2014). Different walls for rods and balls: the diversity of peptidoglycan. *Molecular Microbiology* 91(5), 862-874. doi: 10.1111/mmi.12513.
- Tuson, H.H., Auer, G.K., Renner, L.D., Hasebe, M., Tropini, C., Salick, M., et al. (2012). Measuring the stiffness of bacterial cells from growth rates in hydrogels of tunable elasticity. *Molecular Immunology* 84(5), 874-891. doi: 10.1111/j.1365-2958.2012.08063.x.
- Uzoechi, S.C., and Abu-Lail, N.I. (2019a). Changes in cellular elasticities and conformational properties of bacterial surface biopolymers of multidrug-resistant *Escherichia coli* (MDR-*E. coli*) strains in response to ampicillin. *The Cell Surface* 5, 100019. doi: 10.1016/j.tcs.2019.100019.
- Uzoechi, S.C., and Abu-Lail, N.I. (2019b). The Effects of beta-Lactam Antibiotics on Surface Modifications of Multidrug-Resistant *Escherichia coli*: A Multiscale Approach. *Microscopy and Microanalysis* 25(1), 135-150. doi: 10.1017/S1431927618015696.
- Vadia, S., Tse, J.L., Lucena, R., Yang, Z., Kellogg, D.R., Wang, J.D., et al. (2017). Fatty Acid Availability Sets Cell Envelope Capacity and Dictates Microbial Cell Size. *Current Biology* 27(12), 1757-1767 e1755. doi: 10.1016/j.cub.2017.05.076.
- Valentini, M., Gonzalez, D., Mavridou, D.A., and Filloux, A. (2018). Lifestyle transitions and adaptive pathogenesis of *Pseudomonas aeruginosa*. *Current Opinion in Microbiology* 41, 15-20. doi: 10.1016/j.mib.2017.11.006.
- Valvano, M.A. (2003). Export of O-specific lipopolysaccharide. *Frontiers in Bioscience* 8, s452-471. doi: 10.2741/1079.
- Valvano, M.A. (2015a). "Genetics and biosynthesis of lipopolysaccharide," in *Molecular Medical Microbiology*, eds. Y.-W. Tang, M. Sussman, L. Dongyou, I. Poxton & J. Schwartzman. Second ed: Elsevier), 55-89.
- Valvano, M.A. (2015b). Intracellular survival of *Burkholderia cepacia* complex in phagocytic cells. *Canadian Journal of Microbiology* 61(9), 607-615. doi: 10.1139/cjm-2015-0316.
- Valvano, M.A., Furlong, S.E., and Patel, K.B. (2011). "Genetics, Biosynthesis and Assembly of O-Antigen," in *Bacterial Lipopolysaccharides*, eds. Y.A. Knirel & M.A. Valvano. (Springer-Verlag/Wien 2011: Springer), 275-310.
- van den Bogaart, G., Hermans, N., Krasnikov, V., and Poolman, B. (2007). Protein mobility and diffusive barriers in *Escherichia coli*: consequences of osmotic stress. *Molecular Immunology* 64(3), 858-871. doi: 10.1111/j.1365-2958.2007.05705.x.
- Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics* 43, 11 10 11-33. doi: 10.1002/0471250953.bi1110s43.
- van Hartingsveldt, J., and Stouthamer, A.H. (1973). Mapping and characterization of mutants of *Pseudomonas aeruginosa* affected in nitrate respiration in aerobic or anaerobic growth. *Journal of General Microbiology* 74(1), 97-106. doi: 10.1099/00221287-74-1-97.
- Vandamme, P., and Dawyndt, P. (2011). Classification and identification of the *Burkholderia cepacia* complex: Past, present and future. *Systematic and Applied Microbiology* 34(2), 87-95. doi: 10.1016/j.syapm.2010.10.002.
- Vandamme, P., Holmes, B., Coenye, T., Goris, J., Mahenthalingam, E., LiPuma, J.J., et al. (2003). *Burkholderia cenocepacia* sp. nov. -- a new twist to an old story. *Research in Microbiology* 154(2), 91-96. doi: 10.1016/S0923-2508(03)00026-3.
- Vandamme, P., Holmes, B., Vancanneyt, M., Coenye, T., Hoste, B., Coopman, R., et al. (1997). Occurrence of multiple genomovars of *Burkholderia cepacia* in cystic fibrosis patients and proposal of *Burkholderia multivorans* sp. nov. *International journal of systematic bacteriology* 47(4), 1188-1200. doi: 10.1099/00207713-47-4-1188.

- Vandamme, P., Mahenthalingam, E., Holmes, B., Coenye, T., Hoste, B., De Vos, P., et al. (2000). Identification and population structure of *Burkholderia stabilis* sp. nov. (formerly *Burkholderia cepacia* genomovar IV). *Journal of Clinical Microbiology* 38(3), 1042-1047. PMC86333.
- Vandeplasseche, E., Sass, A., Lemarcq, A., Dandekar, A.A., Coenye, T., and Crabbe, A. (2019). In vitro evolution of *Pseudomonas aeruginosa* AA2 biofilms in the presence of cystic fibrosis lung microbiome members. *Scientific Reports* 9(1), 12859. doi: 10.1038/s41598-019-49371-y.
- Vander Wauven, C., Pierard, A., Kley-Raymann, M., and Haas, D. (1984). *Pseudomonas aeruginosa* mutants affected in anaerobic growth on arginine: evidence for a four-gene cluster encoding the arginine deiminase pathway. *Journal of Bacteriology* 160(3), 928-934. PMC215798.
- Vanlaere, E., Baldwin, A., Gevers, D., Henry, D., De Brandt, E., LiPuma, J.J., et al. (2009). Taxon K, a complex within the *Burkholderia cepacia* complex, comprises at least two novel species, *Burkholderia contaminans* sp. nov. and *Burkholderia lata* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* 59(Pt 1), 102-111. doi: 10.1099/ijs.0.001123-0.
- Varga, J.J., Losada, L., Zelazny, A.M., Kim, M., McCarrison, J., Brinkac, L., et al. (2013). Draft Genome Sequences of *Burkholderia cenocepacia* ET12 Lineage Strains K56-2 and BC7. *Genome Announcements* 1(5). doi: 10.1128/genomeA.00841-13.
- Vargas-Pinto, R., Gong, H., Vahabikashi, A., and Johnson, M. (2013). The effect of the endothelial cell cortex on atomic force microscopy measurements. *Biophysical Journal* 105(2), 300-309. doi: 10.1016/j.bpj.2013.05.034.
- Vermis, K., Coenye, T., Mahenthalingam, E., Nelis, H.J., and Vandamme, P. (2002). Evaluation of species-specific *recA*-based PCR tests for genomovar level identification within the *Burkholderia cepacia* complex. *Journal of Medical Microbiology* 51(11), 937-940.
- Veyrier, F.J., Biais, N., Morales, P., Belkacem, N., Guilhen, C., Ranjeva, S., et al. (2015). Common Cell Shape Evolution of Two Nasopharyngeal Pathogens. *PLoS Genetics* 11(7), e1005338. doi: 10.1371/journal.pgen.1005338.
- Viberg, L.T., Sarovich, D.S., Kidd, T.J., Geake, J.B., Bell, S.C., Currie, B.J., et al. (2017). Within-Host Evolution of *Burkholderia pseudomallei* during Chronic Infection of Seven Australasian Cystic Fibrosis Patients. *MBio* 8(2). doi: 10.1128/mBio.00356-17.
- Vinion-Dubiel, A.D., and Goldberg, J.B. (2003). Lipopolysaccharide of *Burkholderia cepacia* complex. *Journal Endotoxin Research* 9(4), 201-213. doi: 10.1179/096805103225001404.
- Vinogradov, E.V., Bock, K., Petersen, B.O., Holst, O., and Brade, H. (1997). The structure of the carbohydrate backbone of the lipopolysaccharide from *Acinetobacter* strain ATCC 17905. *European Journal of Biochemistry* 243(1-2), 122-127.
- Vinogradov, E.V., Duus, J.O., Brade, H., and Holst, O. (2002a). The structure of the carbohydrate backbone of the lipopolysaccharide from *Acinetobacter baumannii* strain ATCC 19606. *European Journal of Biochemistry* 269(2), 422-430.
- Vinogradov, E.V., Lindner, B., Kocharova, N.A., Senchenkova, S.N., Shashkov, A.S., Knirel, Y.A., et al. (2002b). The core structure of the lipopolysaccharide from the causative agent of plague, *Yersinia pestis*. *Carbohydrate Research* 337(9), 775-777.
- Vinogradov, E.V., Petersen, B.O., Thomas-Oates, J.E., Duus, J., Brade, H., and Holst, O. (1998). Characterization of a novel branched tetrasaccharide of 3-deoxy-D-manno-oct-2-ulopyranosonic acid. The structure of the carbohydrate backbone of the lipopolysaccharide from *Acinetobacter baumannii* strain nctc 10303 (atcc 17904). *The Journal of Biological Chemistry* 273(43), 28122-28131.
- Vollmer, W., and Bertsche, U. (2008). Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*. *Biochimica et Biophysica Acta (BBA)* 1778(9), 1714-1734. doi: 10.1016/j.bbamem.2007.06.007.
- Wagner, R., Moon, R., Pratt, J., Shaw, G., and Raman, A. (2011). Uncertainty quantification in nanomechanical measurements using the atomic force microscope. *Nanotechnology* 22(45), 455703. doi: 10.1088/0957-4484/22/45/455703.
- Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9(11), e112963. doi: 10.1371/journal.pone.0112963.
- Wang, S., Arellano-Santoyo, H., Combs, P.A., and Shaevitz, J.W. (2010). Actin-like cytoskeleton filaments contribute to cell mechanics in bacteria. *PNAS, Proceedings of the National Academy of Sciences* 107(20), 9182-9185. doi: 10.1073/pnas.0911517107.
- Wang, X., Ribeiro, A.A., Guan, Z., and Raetz, C.R. (2009). Identification of undecaprenyl phosphate-beta-D-galactosamine in *Francisella novicida* and its function in lipid A modification. *Biochemistry* 48(6), 1162-1172. doi: 10.1021/bi802211k.
- Warren, A.E., Boulianne-Larsen, C.M., Chandler, C.B., Chiotti, K., Kroll, E., Miller, S.R., et al. (2011). Genotypic and phenotypic variation in *Pseudomonas aeruginosa* reveals signatures of secondary infection and mutator activity in certain cystic fibrosis patients with chronic lung infections. *Infection and Immunity* 79(12), 4802-4818. doi: 10.1128/IAI.05282-11.
- Weber, C.F., and King, G.M. (2017). Volcanic Soils as Sources of Novel CO-Oxidizing *Paraburkholderia* and *Burkholderia*: *Paraburkholderia hiiakae* sp. nov., *Paraburkholderia metrosideri* sp. nov., *Paraburkholderia paradisi* sp. nov.,

- Paraburkholderia peleeae* sp. nov., and *Burkholderia alpina* sp. nov. a Member of the *Burkholderia cepacia* Complex. *Frontiers in Microbiology* 8, 207. doi: 10.3389/fmicb.2017.00207.
- Weiser, J.N. (2013). The battle with the host over microbial size. *Current Opinion in Microbiology* 16(1), 59-62. doi: 10.1016/j.mib.2013.01.001.
- Westphal, O., and Jann, K. (1965). "Bacterial lipopolysaccharides: Extraction with phenol-water and further applications of the procedure," in *Methods in Carbohydrate Chemistry*, eds. R.L. Whistler & M.L. Wolfrom. (New York: Academic Press), 83-91.
- Wheeler, R., Mesnage, S., Boneca, I.G., Hobbs, J.K., and Foster, S.J. (2011). Super-resolution microscopy reveals cell wall dynamics and peptidoglycan architecture in ovococcal bacteria. *Molecular Microbiology* 82(5), 1096-1109. doi: 10.1111/j.1365-2958.2011.07871.x.
- Whitby, P.W., VanWagoner, T.M., Taylor, A.A., Seale, T.W., Morton, D.J., LiPuma, J.J., et al. (2006). Identification of an RTX determinant of *Burkholderia cenocepacia* J2315 by subtractive hybridization. *Journal of Medical Microbiology* 55(Pt 1), 11-21. doi: 10.1099/jmm.0.46138-0.
- Whiteson, K.L., Meinardi, S., Lim, Y.W., Schmieder, R., Maughan, H., Quinn, R., et al. (2014). Breath gas metabolites and bacterial metagenomes from cystic fibrosis airways indicate active pH neutral 2,3-butanedione fermentation. *The ISME Journal* 8(6), 1247-1258. doi: 10.1038/ismej.2013.229.
- Whitfield, C., and Trent, M.S. (2014). Biosynthesis and export of bacterial lipopolysaccharides. *Annual Review of Biochemistry* 83, 99-128. doi: 10.1146/annurev-biochem-060713-035600.
- Williams, H.D., Zlosnik, J.E., and Ryall, B. (2007). Oxygen, cyanide and energy generation in the cystic fibrosis pathogen *Pseudomonas aeruginosa*. *Advances in Microbial Physiology* 52, 1-71. doi: 10.1016/S0065-2911(06)52001-6.
- Winsor, G.L., Khaira, B., Van Rossum, T., Lo, R., Whiteside, M.D., and Brinkman, F.S. (2008). The *Burkholderia* Genome Database: facilitating flexible queries and comparative analyses. *Bioinformatics* 24(23), 2803-2804. doi: 10.1093/bioinformatics/btn524.
- Winstanley, C., O'Brien, S., and Brockhurst, M.A. (2016). *Pseudomonas aeruginosa* Evolutionary Adaptation and Diversification in Cystic Fibrosis Chronic Lung Infections. *Trends in Microbiology* 24(5), 327-337. doi: 10.1016/j.tim.2016.01.008.
- Wolanin, P.M., Thomason, P.A., and Stock, J.B. (2002). Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome biology* 3(10), REVIEWS3013. doi: 10.1186/gb-2002-3-10-reviews3013.
- Woods, C.W., Bressler, A.M., LiPuma, J.J., Alexander, B.D., Clements, D.A., Weber, D.J., et al. (2004). Virulence associated with outbreak-related strains of *Burkholderia cepacia* complex among a cohort of patients with bacteremia. *Clinical Infectious Diseases* 38(9), 1243-1250. doi: 10.1086/383313.
- Worlitzsch, D., Tarran, R., Ulrich, M., Schwab, U., Cekici, A., Meyer, K.C., et al. (2002). Effects of reduced mucus oxygen concentration in airway *Pseudomonas* infections of cystic fibrosis patients. *The Journal of Clinical Investigation* 109(3), 317-325. doi: 10.1172/JCI13870.
- Yabuuchi, E., Kosako, Y., Oyaizu, H., Yano, I., Hotta, H., Hashimoto, Y., et al. (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. *Microbiology and Immunology* 36(12), 1251-1275. doi: 10.1111/j.1348-0421.1992.tb02129.x.
- Yamamoto, K., Hirao, K., Oshima, T., Aiba, H., Utsumi, R., and Ishihama, A. (2005). Functional characterization in vitro of all two-component signal transduction systems from *Escherichia coli*. *The Journal of Biological Chemistry* 280(2), 1448-1456. doi: 10.1074/jbc.M410104200.
- Yang, C.H., Chen, Y.C., Peng, S.Y., Tsai, A.P., Lee, T.J., Yen, J.H., et al. (2018). An engineered arginine-rich alpha-helical antimicrobial peptide exhibits broad-spectrum bactericidal activity against pathogenic bacteria and reduces bacterial infections in mice. *Scientific Reports* 8(1), 14602. doi: 10.1038/s41598-018-32981-3.
- Yang, D.C., Blair, K.M., and Salama, N.R. (2016). Staying in Shape: the Impact of Cell Shape on Bacterial Survival in Diverse Environments. *Microbiology and Molecular Biology Reviews* 80(1), 187-203. doi: 10.1128/MMBR.00031-15.
- Yang, J.H., Spilker, T., and 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. doi: 10.1016/j.diagmicrobio.2005.08.020.
- Yang, L., Jelsbak, L., Marvig, R.L., Damkiaer, S., Workman, C.T., Rau, M.H., et al. (2011a). Evolutionary dynamics of bacteria in a human host environment. *PNAS, Proceedings of the National Academy of Sciences* 108(18), 7481-7486. doi: 10.1073/pnas.1018249108.
- Yang, L., Jelsbak, L., and Molin, S. (2011b). Microbial ecology and adaptation in cystic fibrosis airways. *Environmental Microbiology* 13(7), 1682-1689. doi: 10.1111/j.1462-2920.2011.02459.x.
- Yao, X., Jericho, M., Pink, D., and Beveridge, T. (1999). Thickness and elasticity of gram-negative murein sacculi measured by atomic force microscopy. *Journal of Bacteriology* 181(22), 6865-6875. PMC94159.
- Yao, X., Walter, J., Burke, S., Stewart, S., Jericho, M.H., Pink, D., et al. (2002). Atomic force microscopy and theoretical considerations of surface properties and turgor pressures of bacteria. *Colloids and Surfaces B: Biointerfaces* 23(2), 213-230. doi: 10.1016/S0927-7765(01)00249-1.

- Yao, Z., Davis, R.M., Kishony, R., Kahne, D., and Ruiz, N. (2012). Regulation of cell size in response to nutrient availability by fatty acid biosynthesis in *Escherichia coli*. *PNAS, Proceedings of the National Academy of Sciences* 109(38), E2561-2568. doi: 10.1073/pnas.1209742109.
- Younes, J.A., van der Mei, H.C., van den Heuvel, E., Busscher, H.J., and Reid, G. (2012). Adhesion forces and coaggregation between vaginal *staphylococci* and *lactobacilli*. *PLoS One* 7(5), e36917. doi: 10.1371/journal.pone.0036917.
- Zankari, E. (2014). Comparison of the web tools ARG-ANNOT and ResFinder for detection of resistance genes in bacteria. *Antimicrobial Agents and Chemotherapy* 58(8), 4986. doi: 10.1128/AAC.02620-14.
- Zemanick, E.T., and Hoffman, L.R. (2016). Cystic Fibrosis: Microbiology and Host Response. *Pediatr Clin North Am* 63(4), 617-636. doi: 10.1016/j.pcl.2016.04.003.
- Zerbino, D.R. (2010). Using the Velvet de novo assembler for short-read sequencing technologies. *Current Protocols in Bioinformatics* Chapter 11, Unit 11 15. doi: 10.1002/0471250953.bi1105s31.
- Zerbino, D.R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* 18(5), 821-829. doi: 10.1101/gr.074492.107.
- Zhang, R., Ou, H.Y., Gao, F., and Luo, H. (2014). Identification of Horizontally-transferred Genomic Islands and Genome Segmentation Points by Using the GC Profile Method. *Current Genomics* 15(2), 113-121. doi: 10.2174/1389202915999140328163125.
- Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology* 7(1-2), 203-214. doi: 10.1089/10665270050081478.
- Zhong, Q., Inniss, D., Kjoller, K., and Elings, V.B. (1993). Fractured polymer/silica fiber surface studied by tapping mode atomic force microscopy. *Surface Science* 290(1), L688-L692. doi: 10.1016/0039-6028(93)90582-5.
- Zlosnik, J.E., Costa, P.S., Brant, R., Mori, P.Y., Hird, T.J., Fraenkel, M.C., et al. (2011). Mucoïd and nonmucoïd *Burkholderia cepacia* complex bacteria in cystic fibrosis infections. *American Journal of Respiratory and Critical Care Medicine* 183(1), 67-72. doi: 10.1164/rccm.201002-0203OC.
- Zlosnik, J.E., Hird, T.J., Fraenkel, M.C., Moreira, L.M., Henry, D.A., and Speert, D.P. (2008). Differential mucoïd exopolysaccharide production by members of the *Burkholderia cepacia* complex. *Journal of Clinical Microbiology* 46(4), 1470-1473. doi: 10.1128/JCM.02273-07.
- Zlosnik, J.E., Mori, P.Y., To, D., Leung, J., Hird, T.J., and Speert, D.P. (2014). Swimming motility in a longitudinal collection of clinical isolates of *Burkholderia cepacia* complex bacteria from people with cystic fibrosis. *PLoS One* 9(9), e106428. doi: 10.1371/journal.pone.0106428.
- Zlosnik, J.E., and Speert, D.P. (2010). The role of mucoïdity in virulence of bacteria from the *Burkholderia cepacia* complex: a systematic proteomic and transcriptomic analysis. *The Journal of Infectious Diseases* 202(5), 770-781. doi: 10.1086/655663.
- Zlosnik, J.E., Zhou, G., Brant, R., Henry, D.A., Hird, T.J., Mahenthalingam, E., et al. (2015). *Burkholderia species* infections in patients with cystic fibrosis in British Columbia, Canada. 30 years' experience. *Annals of the American Thoracic Society* 12(1), 70-78. doi: 10.1513/AnnalsATS.201408-395OC.

## **8**      **Supplementary Data**

---



## 8. Supplementary Data

### 8.1. Supplementary notes on materials and methods

#### *8.1.1. Data related to chapter II*

##### *8.1.1.1. Genomic DNA sequencing, De novo assembly, and annotation*

###### *8.1.1.1.1. Illumina sequencing by CD Genomics*

We used a whole-genome shotgun sequencing strategy and Illumina Genome Analyser sequencing technology. A 100 bp paired-end run was performed with the strains *Burkholderia cenocepacia* IST439, IST4113, IST4129 and IST4134. Genomic DNA was sheared by a nebulizer to generate DNA fragments for the Illumina Paired-End (PE) Sequencing method. DNA libraries (20 ng/μl) were constructed by ligating the specific oligonucleotides (Illumina adapters) designed for PE sequencing to both ends of DNA fragments with the TA cloning method. The ligated DNA was then size-selected on a 2% agarose gel. DNA fragments of ~ 500 bp were excised from the preparative portion of the gel. DNA was then recovered using a Qiagen gel extraction kit and was PCR amplified to produce the final DNA library. Five picomoles of DNA from each strain were loaded onto two lanes of the sequencing chip, and the clusters were generated on the cluster generation station of the GAIIX using the Illumina cluster generation kit. Bacteriophage X174 DNA was used as a control. In the case of paired-end reads, distinct adaptors from Illumina were ligated to each end with PCR primers that allowed reading of each end as separate runs. The sequencing reaction was run for 100 cycles (tagging, imaging, and cleavage of one terminal base at a time), and four images of each tile on the chip were taken in different wavelengths for exciting each base-specific fluorophore. For paired-end reads, data were collected as two sets of matched 100-bp reads. Reads for each of the indexed samples were then separated using a custom Perl script. Image analysis and base calling were done using the Illumina GA Pipeline software.

#### **Adapters sequence:**

5' P-GATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG  
5' ACACTCTTCCCTACACGACGCTCTTCCGATCT

#### 8.1.1.1.2. PacBio sequencing by a real-time (SMRT) Pacific Biosciences

*B. cenocepacia* IST439 was later-sequenced to generate a complete assembly by using a combination of single molecule, real-time (SMRT) Pacific Biosciences – PacBio reads and Illumina 100-bp paired-end reads. In brief, genomic DNA (gDNA) was prepared using the Qiagen Genomic-Tip Kit (20/G) from overnight cultures of *B. cenocepacia* IST439 grown in LB at 37°C using manufacturer’s instructions. Importantly, this kit uses gravity filtration to purify gDNA, which limits shearing and increases the average fragment size of the resulting gDNA sample. Long insert library preparation and SMRT sequencing was performed on IST439 gDNA at the Icahn School of Medicine at Mount Sinai according to the manufacturer’s instructions, as described previously (Beaulaurier et al., 2015). Briefly, libraries were size selected using Sage Science Blue Pippin 0.75% agarose cassettes to enrich for long-reads, and were assessed for quantity and insert size using an Agilent DNA 12,000 gel chip. Primers, polymerases, and magnetic beads were loaded to generate a completed SMRTbell library, which was run in a single SMRT cell of a Pacific Biosciences RSII sequencer at a concentration of 75 pM for 180 minutes. We used the hierarchical genome-assembly process workflow (HGAP3) to generate a completed assembly of *B. cenocepacia* IST439 and polished our assembly using the Quiver algorithm (Chin et al., 2013).

#### 8.1.1.2. Variant calling and SNP/INDEL detections

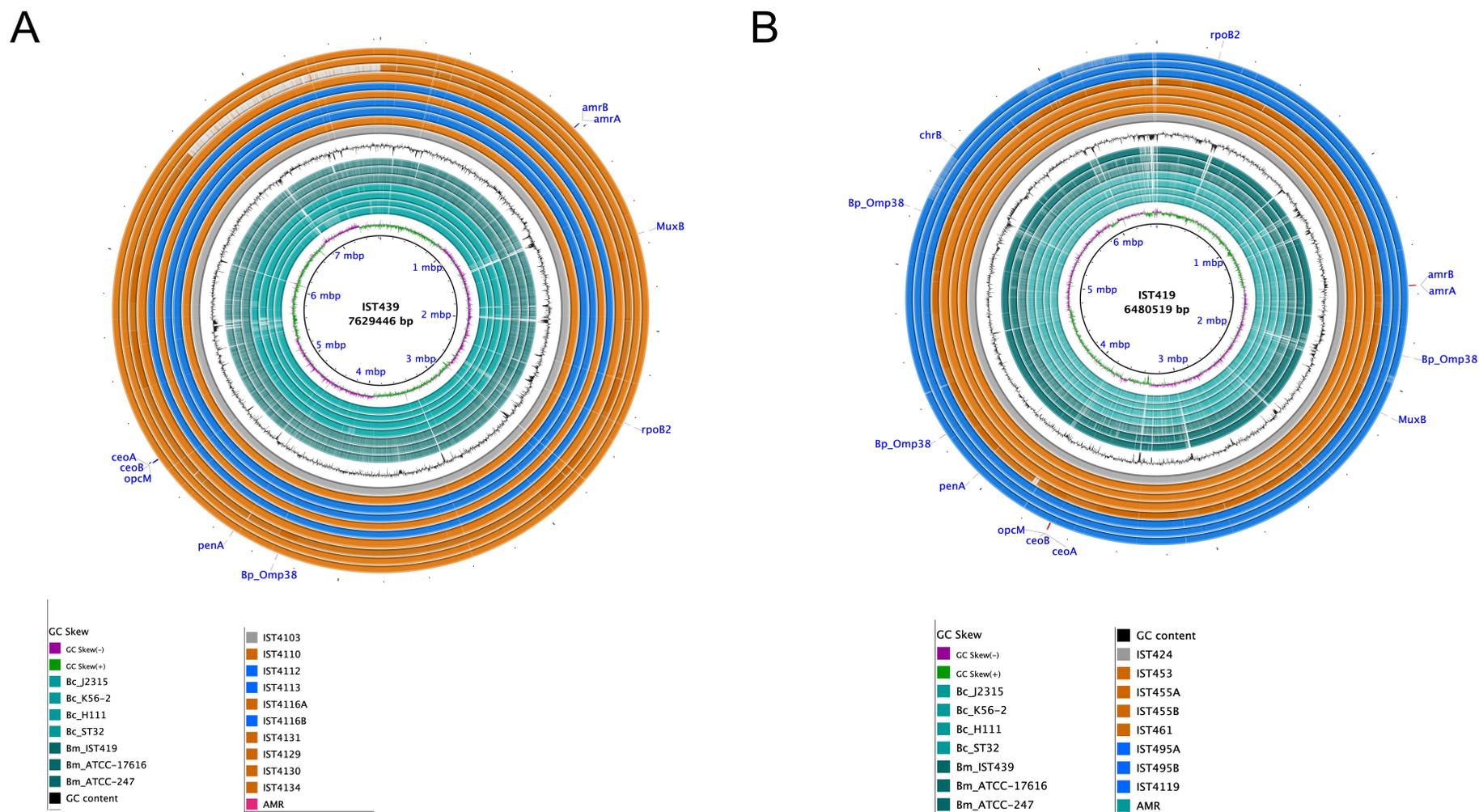
Trimmed paired-end reads were mapped against the reference complete genome of *B. cenocepacia* IST439 (for all *B. cenocepacia* clonal variants) and against the reference draft genome sequence of *B. multivorans* IST419 for its clonal variants using BWA-MEM packages of Burrows-Wheeler Aligner (BWA v.0.7.10) (Li and Durbin, 2010) and NovoAlign (www.novocraft.com). Variants [Single nucleotide polymorphisms (SNPs) and insertion-deletion mutations (INDELs)] were called as described previously by using two independent, commonly used, standard variant calling pipelines; GATK and SAMtools/BCFtools toolbox (Li et al., 2009; Van der Auwera et al., 2013; Dillon et al., 2015; Dillon et al., 2017). Briefly, we used SAMtools to convert the SAM alignment files produced for each clonal isolate to mpileup format (Li et al., 2009), then in-house perl scripts were used to produce the forward and reverse read alignments for each position in each line. Next, a three-step process was used to detect putative polymorphisms. First, a base for each individual isolate was called if a site was covered by at least two forward and two reverse reads, and at least 80% of those reads identified the same base. Otherwise, the site was not analysed. Second, an ancestral consensus was called as the base with the highest support among reads across all clonal isolates, as long as there were at least three isolates with

sufficient coverage to identify a base. Lastly, at sites where both an individual line base and ancestral consensus were identified, individual line bases were compared to the ancestral base, and if they were different, a putative base-substitution mutation was identified. This analysis was carried out independently with the alignments generated by BWA and Novoalign, and putative SNPs were considered genuine only if both pipelines independently identified the mutation. Finally, the called variants of both pipelines (the above in-house perl scripts and GATK) were combined and filtered with the SAMtools/BCFtools toolbox, v1.9 (Li et al., 2009). In order to avoid false positive calls, maximum read depth (<2000) was considered to avoid any further false duplications and the call confidence was concerned based on RMS mapping quality  $\geq 30$  and the minimum read depth  $>5$ .

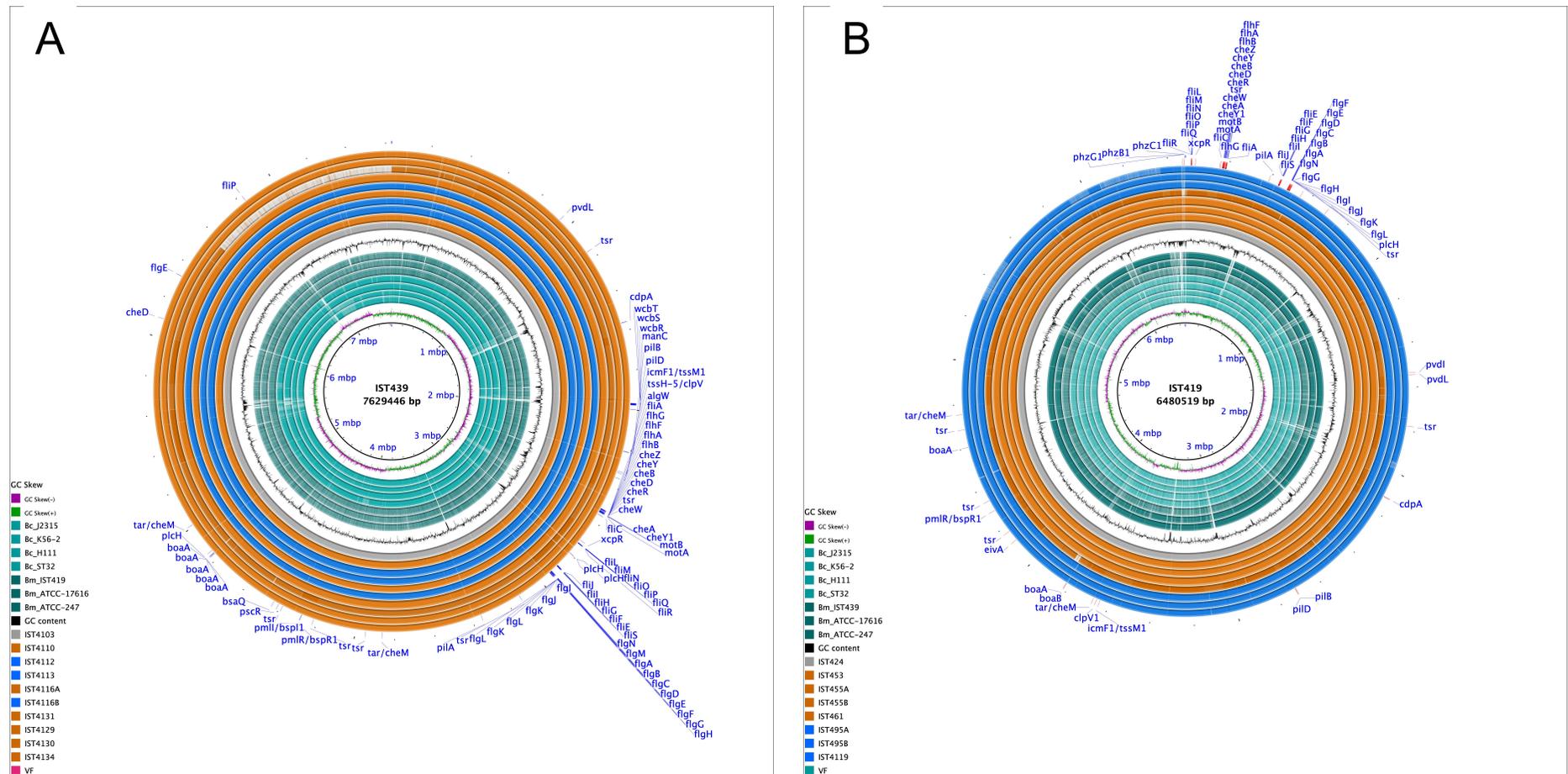
All indels identified in this study were also detected using similar requirements to those previously described (Dillon et al., 2015), taking into account that inherent difficulties with gaps and repeat elements can reduce agreement in the alignment of single reads using short-read alignment algorithms, even in the case of true indels. All putative short-indels that were independently identified with the alignments were considered genuine only if both pipelines (the above in-house perl scripts and GATK) independently identified the mutation. To perform a quality control of the above method, a parallel comparison was performed using Breseq (Barrick et al., 2014), and similar predictions were predominantly observed in the case of base-substitutions, with some discrepancies in indels. All putative SNPs/INDELS were then manually inspected and evaluated using the Integrative Genomics Viewer – IGV (Robinson et al., 2011; Thorvaldsdottir et al., 2013), and discarded if the BWA and Novoalign-produced alignments did not provide enough confidence (poor coverage as aforementioned). Thus, we are confident that nearly all base-substitution and indels identified in this study were genuine events that arose during the in-host evolutionary process. Finally, Functional annotation of the called variants was performed by using SnpEff v3.1 (Cingolani et al., 2012) with manual BLAST verification against the NCBI Microbes genome database.

## 8.2. Supplementary Figures

### 8.2.1. Data related to chapter II

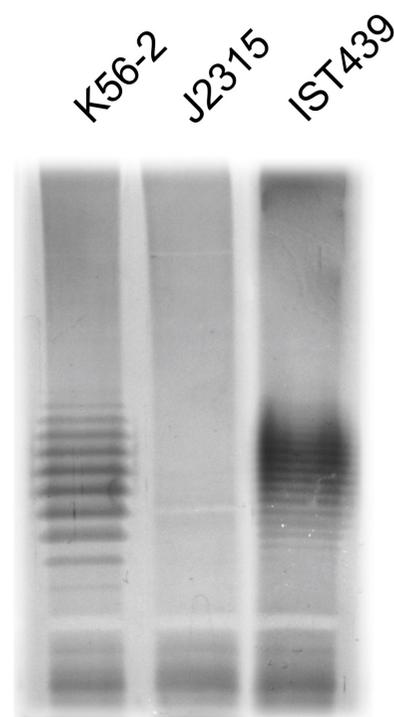


**Figure S2.1** | Antibiotic resistance associated genes – AMR in (A) *B. cenocepacia* IST439 and (B) *B. multivorans* IST419. The inner circles denote the genomes of reference ancestor isolates and the corresponding clonal variants, ordered as indicated in the legend; Bc – *B. cenocepacia* and Bm – *B. multivorans* isolates. Blue and orange colors denote different clades as indicated in Fig. 2.1. All clonal variants were chronologically ordered based on isolation date. Visualizations were carried out by BRIG (Alikhan et al., 2011). For more details, see tables S2.6 and S2.7.

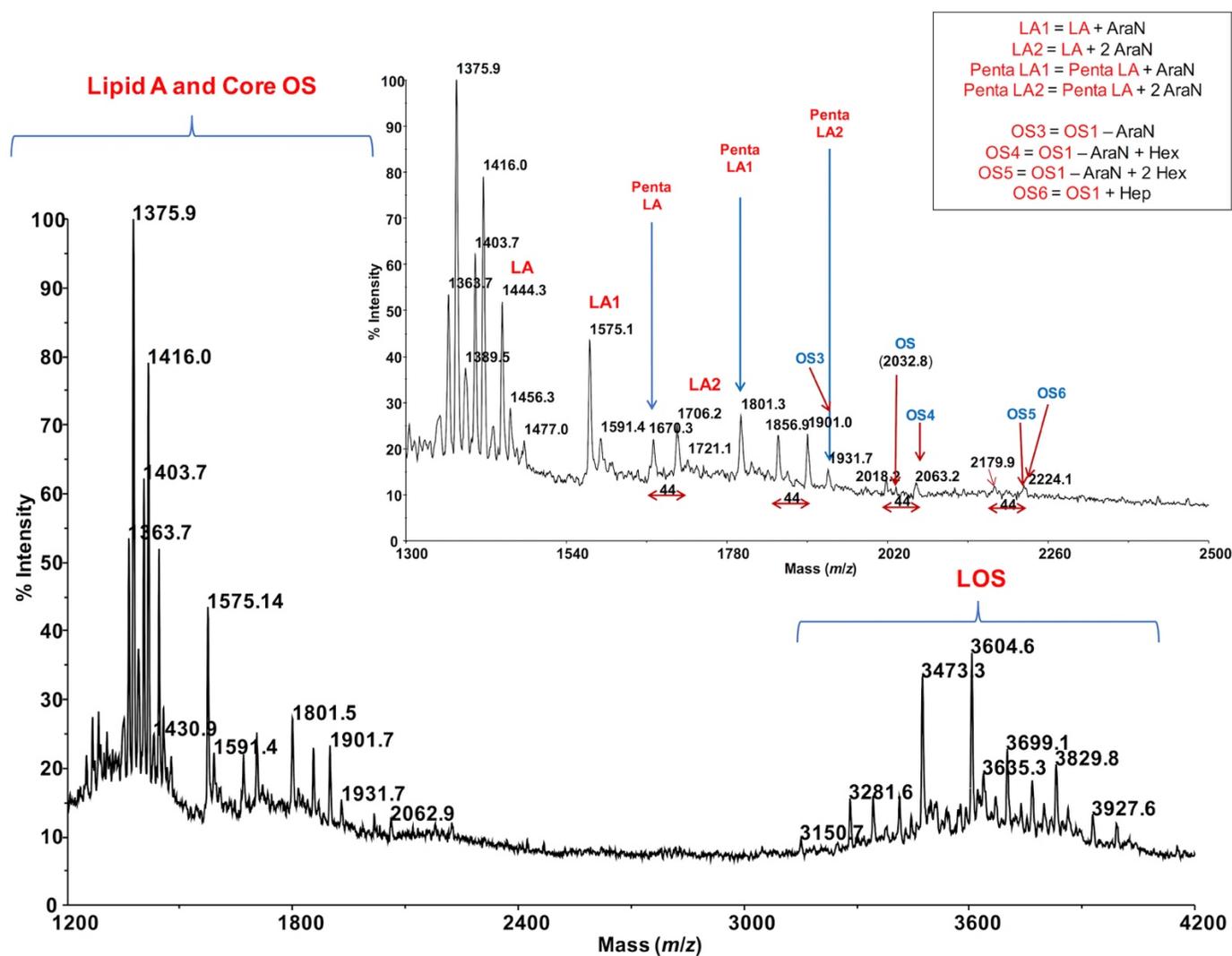


**Figure S2.2** | Virulence factors associated genes (VF) in (A) *B. cenocepacia* IST439 and (B) *B. multivorans* IST419. The inner circles denote the genomes of reference ancestor isolates and the corresponding clonal variants, ordered as indicated in the legend; Bc – *B. cenocepacia* and Bm – *B. multivorans* isolates. Blue and orange colors denote different clades as indicated in Fig. 2.1. All clonal variants were chronologically ordered based on isolation date. Visualizations were carried out by BRIG (Alikhan et al., 2011). For more details, see tables S2.6 and S2.7.

## 8.2.2. Data related to chapter III



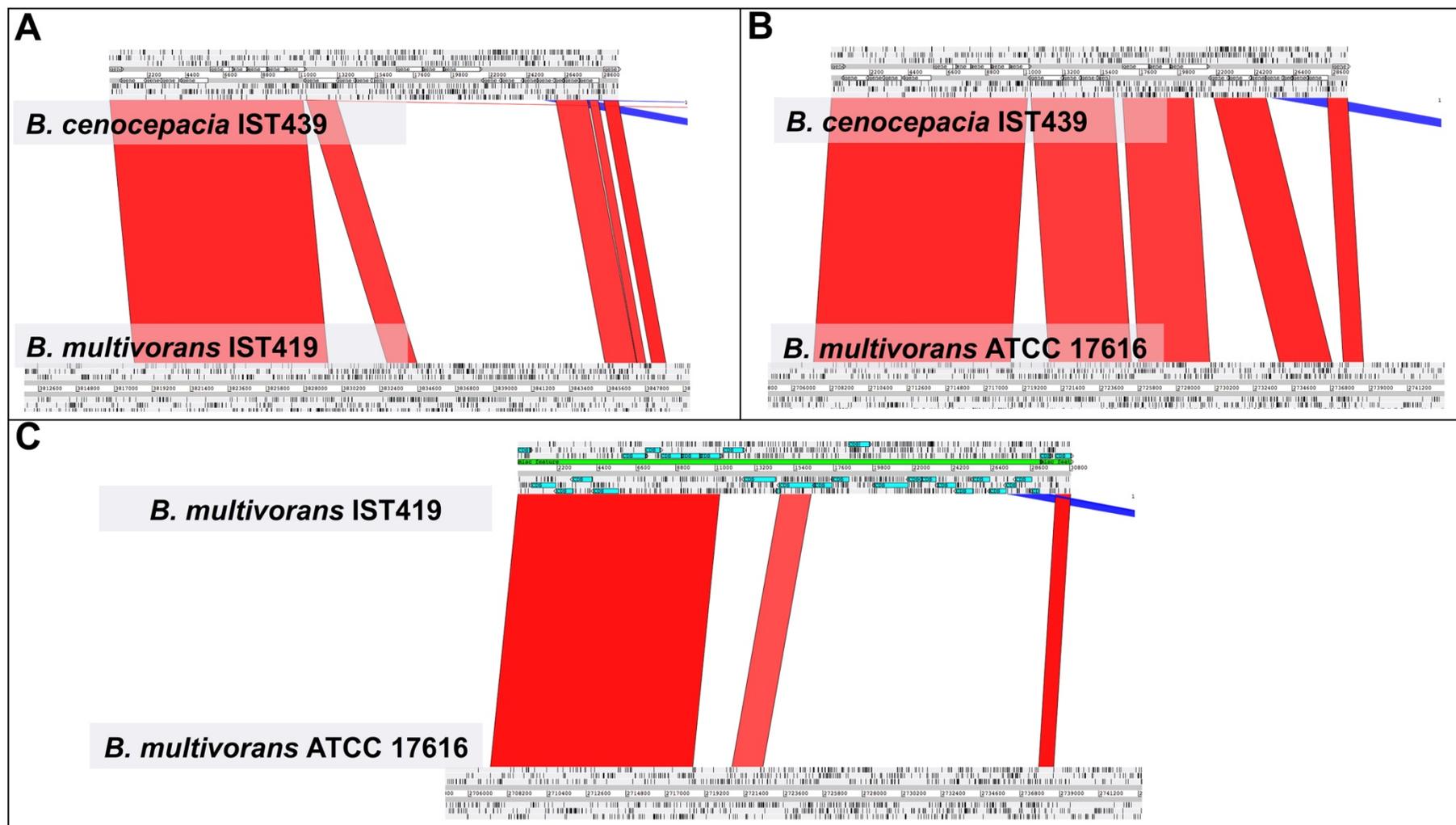
**Figure S3.1** | Electrophoretic profiles of the LPS extracted from *Burkholderia cenocepacia* IST439 isolate and the (*recA* lineage IIIA) ET-12 reference strains K56-2 and J2315. LPS loading was standardized based on culture optical density. Samples were run in 14% polyacrylamide gels in a Tricine-SDS system and developed by silver staining.



**Figure S3.2** | The MALDI mass spectrum of the intact IST4113 LPS (mass range 1200-420000 Da). In the inset, the magnification of the mass-region 1300-2500 Da is also shown.

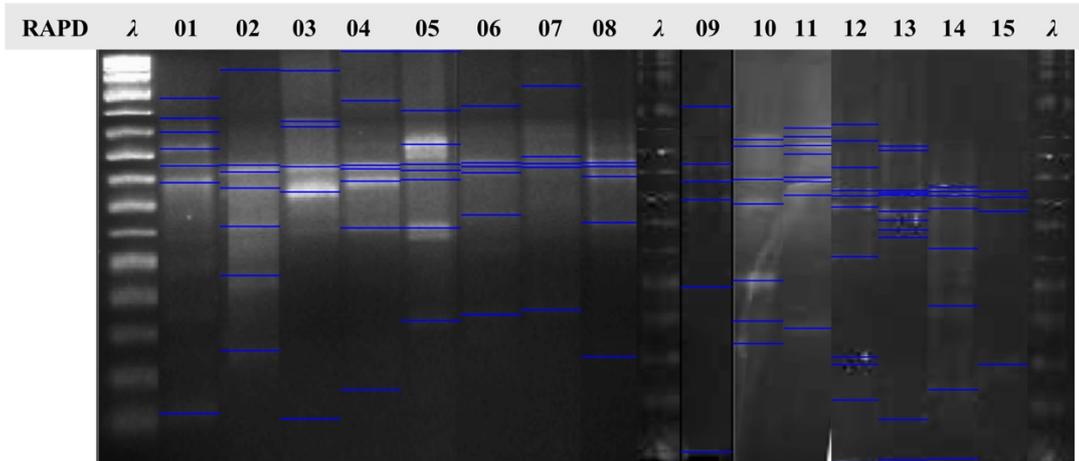
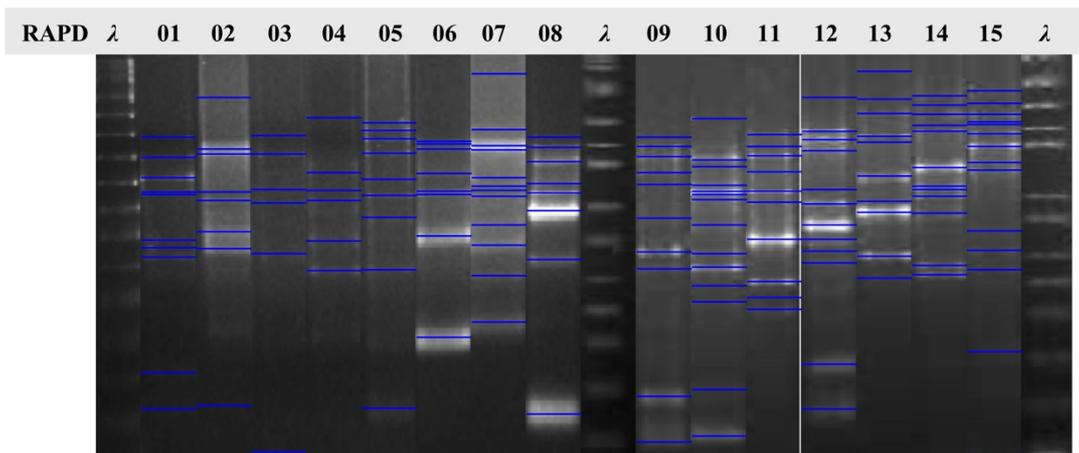
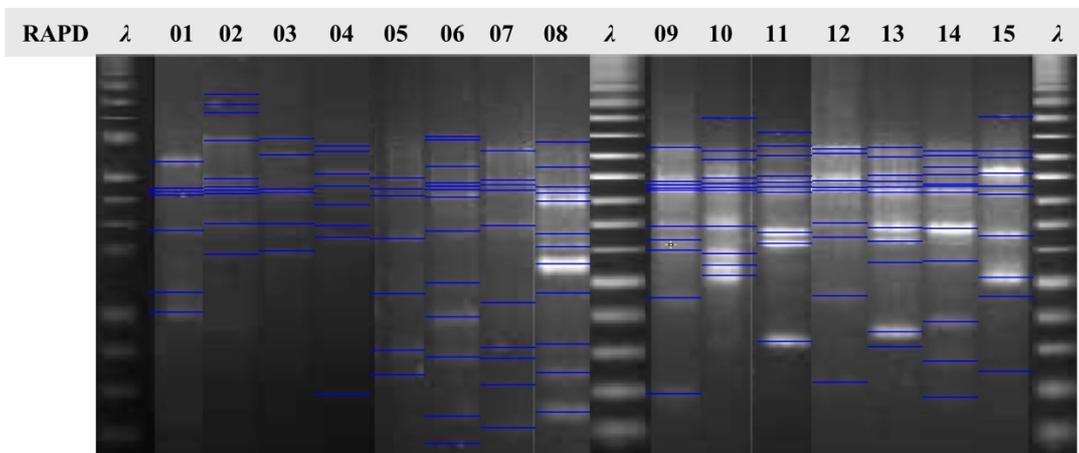


**Figure S3.3** | Comparative genomic of the OAg cluster is visualized by Artemis/ACT among *B. cenocepacia* IST439 and the reference strains J2315 and k56-2.



**Figure S3.4** | Comparative visualization of the OAg biosynthetic clusters in binary mode between *B. cenocepacia* IST439 and *B. multivorans* IST419 (A), between *B. cenocepacia* IST439 and *B. multivorans* ATCC 17616 (B), and between *B. multivorans* isolates IST419 and ATCC 17616 (C) are shown.

## 8.2.3. Data related to chapter IV

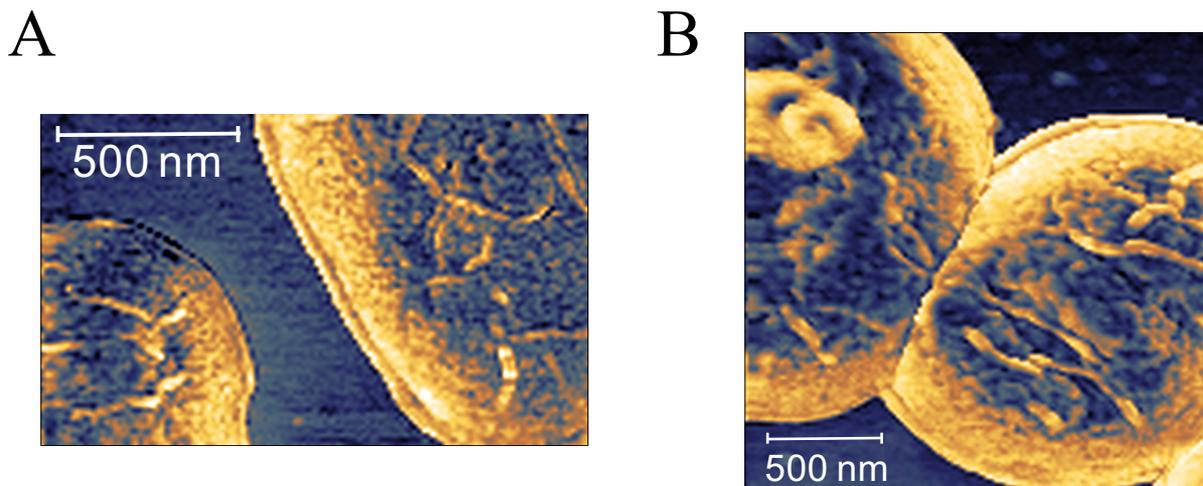
**A****B****C**

**Figure S4.1** | Representative profiles based on Random Amplified Polymorphic DNA (RAPD) analysis of the Bcc isolates examined in this study retrieved from different CF patients during several

years of chronic infection. The polymorphisms shown are based on the amplification by RAPD primers 270 (A), 208 (B), and 272 (C) in the same sequential order. The polymorphisms originated were analyzed using the software GelJ to detect the banding patterns. These profiles correspond to the 15-different representative selected isolates of *B. cenocepacia* IIIA (RAPD 01-08), *B. cenocepacia* IIIB (RAPD 09-11), *B. cepacia* (RAPD 12 and 13), and *B. stabilis* (RAPD 14 and 15).

**RAPD profiles for isolates:** 01 – IST416, 02 – IST439, 03 – IST462, 04 – IST432, 05 – IST4121, 06 – IST4240a, 07 – IST4272, 08 – IST4197, 09 – IST435, 10 – IST438, 11 – IST466, 12 – IST4152, 13 – IST4546, 14 – IST413, 15 – IST412 and  $\lambda$  – 1 Kb plus DNA ladder.

#### 8.2.4. Data related to chapter V



**Figure S5.1** | The adhesion force maps (A –unattached and B –attached cells) of bacterial surfaces for IST4113 and IST4134, respectively, showed that the Nano-scale surface architectures have less adhesion among the rest of the surfaces (z-range = 4.88 nN).

## 8.3. Supplementary Tables

## 8.3.1. Data related to chapter II

**Table S2.1** | List of the whole genome sequences of Bcc isolates used in the present study for comparative genomic analyses. The isolation year and country are indicated as well as the multi-locus sequence typing (MLST). CS – cepacia syndrome. <sup>a</sup> – un-known data.

Strain	Isolation Year	Isolation Country	Isolation Source	MLST sequence type	MLST clonal complex	Status	Accession	Reference
<i>Burkholderia cenocepacia</i> IST439	1999	Portugal	CF sputum-CS	ST218	CC31	complete	XXXX	This study
<i>Burkholderia cenocepacia</i> IST4103	2001	Portugal	CF sputum-CS	ST218	CC31	draft	XXXX	This study
<i>Burkholderia cenocepacia</i> IST4110	2001	Portugal	CF sputum-CS	ST218	CC31	draft	XXXX	This study
<i>Burkholderia cenocepacia</i> IST4112	2001	Portugal	CF sputum-CS	ST218	CC31	draft	XXXX	This study
<i>Burkholderia cenocepacia</i> IST4113	2001	Portugal	CF sputum-CS	ST218	CC31	draft	XXXX	This study
<i>Burkholderia cenocepacia</i> IST4116A	2002	Portugal	CF sputum-CS	ST218	CC31	draft	XXXX	This study
<i>Burkholderia cenocepacia</i> IST4116B	2002	Portugal	CF sputum-CS	ST218	CC31	draft	XXXX	This study
<i>Burkholderia cenocepacia</i> IST4131	2002	Portugal	CF sputum-CS	ST218	CC31	draft	XXXX	This study
<i>Burkholderia cenocepacia</i> IST4129	2002	Portugal	CF sputum-CS	ST218	CC31	draft	XXXX	This study
<i>Burkholderia cenocepacia</i> IST4130	2002	Portugal	CF sputum-CS	ST218	CC31	draft	XXXX	This study
<i>Burkholderia cenocepacia</i> IST4134	2002	Portugal	CF sputum-CS	ST218	CC31	draft	XXXX	This study
<i>Burkholderia cenocepacia</i> J2315	1989	UK	CF sputum-CS	ST28	CC31	complete	GCA_000009485.1	(Holden et al., 2009)
<i>Burkholderia cenocepacia</i> K56-2	1999	Canada	CF sputum-CS	ST227	CC31	draft	GCA_000981305.1	(Varga et al., 2013)
<i>Burkholderia cenocepacia</i> BC7	prior to 1992	Canada	CF sputum-CS	ST28	CC31	draft	GCA_000333135.2	(Varga et al., 2013)

<i>Burkholderia cenocepacia</i> H111	1993	Germany	CF sputum	ST1506	- <sup>a</sup>	complete	GCA_000236215.4	(Carlier et al., 2014)
<i>Burkholderia cenocepacia</i> ST32	1997	Czech Republic	CF sputum-CS	ST32	CC31	complete	GCA_001484665.1	(Dedeckova et al., 2013)
<i>Burkholderia cenocepacia</i> VC1254	1985	Canada	CF sputum	ST32	CC31	complete	GCA_001999925.1	(Lee et al., 2017)
<i>Burkholderia cenocepacia</i> VC2307	1987	Canada	CF sputum	ST210	CC31	complete	GCA_001999805.1	(Lee et al., 2017)
<i>Burkholderia dolosa</i> AU0158	- <sup>a</sup>	USA	CF patient	ST472	- <sup>a</sup>	complete	GCA_000959505.1	(Johnson et al., 2015; Johnson et al., 2016)
<i>Burkholderia multivorans</i> ATCC 17616	- <sup>a</sup>	USA	Soil isolate	ST21	- <sup>a</sup>	complete	GCA_000018505.1	(Komatsu et al., 2003)
<i>Burkholderia multivorans</i> ATCC BAA-247	- <sup>a</sup>	Belgium	CF sputum	ST650	- <sup>a</sup>	complete	GCA_000959525.1	(Johnson et al., 2015; Johnson et al., 2016)
<i>Burkholderia multivorans</i> IST419	1998	Portugal	CF sputum-CS	ST836	- <sup>a</sup>	draft	XXXX	This study
<i>Burkholderia multivorans</i> IST424	1998	Portugal	CF sputum-CS	ST836	- <sup>a</sup>	draft	XXXX	This study
<i>Burkholderia multivorans</i> IST453	1999	Portugal	CF sputum-CS	ST836	- <sup>a</sup>	draft	XXXX	This study
<i>Burkholderia multivorans</i> IST455A	2000	Portugal	CF sputum-CS	ST836	- <sup>a</sup>	draft	XXXX	This study
<i>Burkholderia multivorans</i> IST455B	2000	Portugal	CF sputum-CS	ST836	- <sup>a</sup>	draft	XXXX	This study
<i>Burkholderia multivorans</i> IST461	2000	Portugal	CF sputum-CS	ST836	- <sup>a</sup>	draft	XXXX	This study
<i>Burkholderia multivorans</i> IST495A	2001	Portugal	CF sputum-CS	ST836	- <sup>a</sup>	draft	XXXX	This study
<i>Burkholderia multivorans</i> IST495B	2001	Portugal	CF sputum-CS	ST836	- <sup>a</sup>	draft	XXXX	This study
<i>Burkholderia multivorans</i> IST4119	2002	Portugal	CF sputum-CS	ST836	- <sup>a</sup>	draft	XXXX	This study

**Table S2.2** | Metadata information of the studied clonal variants sequentially ordered based on isolation date; details of genomic assembly and functional annotation are shown as well as the *in silico* MLST profiles. For clarity, the obtained data were divided into the two following tables.

Patient	Isolate signature	Bcc species	Isolation date	Months after the beginning (exact number) considering IST419 as the first	<i>In silico</i> MLST profiles							
					Bcc complex profile/sequence type	atpD	altB	gyrB	recA	lepA	phaC	trpB
Patient J who died with cepacia syndrome	IST419	<i>B. multivorans</i>	2/26/98	0	836	9	50	169	81	409	96	133
	IST424	<i>B. multivorans</i>	6/4/98	3.223684211	836	9	50	169	81	409	96	133
	IST439	<i>B. cenocepacia</i> IIIA	1/30/99	11.11842105	218	132	157	186	14	11	6	79
	IST453	<i>B. multivorans</i>	7/19/99	16.71052632	836	9	50	169	81	409	96	133
	IST455A	<i>B. multivorans</i>	2/1/00	23.19078947	836	9	50	169	81	409	96	133
	IST455B	<i>B. multivorans</i>	2/1/00	23.19078947	836	9	50	169	81	409	96	133
	IST461	<i>B. multivorans</i>	4/4/00	25.26315789	836	9	50	169	81	409	96	133
	IST495A	<i>B. multivorans</i>	5/29/01	39.07894737	836	9	50	169	81	409	96	133
	IST495B	<i>B. multivorans</i>	5/29/01	39.07894737	836	9	50	169	81	409	96	133
	IST4103	<i>B. cenocepacia</i> IIIA	7/24/01	40.92105263	218	132	157	186	14	11	6	79
	IST4110	<i>B. cenocepacia</i> IIIA	9/25/01	42.99342105	218	132	157	186	14	11	6	79
	IST4112	<i>B. cenocepacia</i> IIIA	10/11/01	43.51973684	218	132	157	186	14	11	6	79
	IST4113	<i>B. cenocepacia</i> IIIA	11/6/01	44.375	218	132	157	186	14	11	6	79
	IST4119	<i>B. multivorans</i>	1/22/02	46.90789474	836	9	50	169	81	409	96	133
	IST4116A	<i>B. cenocepacia</i> IIIA	2/11/02	47.56578947	218	132	157	186	14	11	6	79
	IST4116B	<i>B. cenocepacia</i> IIIA	2/11/02	47.56578947	218	132	157	186	14	11	6	79
	IST4131	<i>B. cenocepacia</i> IIIA	2/26/02	48.05921053	218	132	157	186	14	11	6	79
IST4129	<i>B. cenocepacia</i> IIIA	3/26/02	48.98026316	218	132	157	186	14	11	6	79	
IST4130	<i>B. cenocepacia</i> IIIA	5/14/02	50.59210526	218	132	157	186	14	11	6	79	
IST4134	<i>B. cenocepacia</i> IIIA	7/2/02	52.20394737	218	132	157	186	14	11	6	79	

Isolate signature	Genomic and assembly features					Annotation features						
	No. of contigs/Scaffolds	size of Largest contig	Genome size (bp)	GC content (%)	N50	No. of contigs/Scaffolds	no. of bases	no. of CDS	no. of genes	no. of misc_RNA	no. of tmRNA	no. of tRNA
IST419	76	1301193	6480509	67.13	522894	76	6480509	5793	5879	23	1	62
IST424	52	569474	6456751	67.15	332277	52	6456751	5764	5844	23	1	56
IST439	3	3562400	7629436	67.45	3006387	3	7629436	6848	6960	26	1	85
IST453	52	672568	6454872	67.15	287972	52	6454872	5743	5824	23	1	57
IST455A	54	580992	6454417	67.15	287947	54	6454417	5768	5846	22	1	55
IST455B	56	586375	6452626	67.15	369692	56	6452626	5768	5847	22	1	56
IST461	25	6456554	6563785	67.02	6456554	25	6563785	5910	6004	19	1	74
IST495A	29	3303134	6588994	66.99	3303134	29	6588994	5947	6047	21	1	78
IST495B	68	516717	6446002	67.15	264147	68	6446002	5763	5838	23	1	51
IST4103	70	833489	7530272	67.45	228859	70	7530272	6783	6891	31	1	76
IST4110	69	833746	7515247	67.5	238796	69	7515247	6761	6868	32	1	74
IST4112	56	1001610	7518224	67.5	238702	56	7518224	6761	6866	33	1	71
IST4113	58	717001	7522815	67.48	261934	58	7522815	6764	6871	33	1	73
IST4119	85	326140	5815930	67.16	131227	85	5815930	5182	5260	25	1	52
IST4116A	53	1280982	7520933	67.5	271233	53	7520933	6761	6861	31	1	68
IST4116B	51	834294	7520477	67.51	279425	51	7520477	6756	6856	31	1	68
IST4131	46	1002283	7521433	67.5	294867	46	7521433	6758	6861	31	1	71
IST4129	34	916088	6495866	67.41	447237	34	6495866	5870	5972	29	1	72
IST4130	54	1001482	7519966	67.5	278379	54	7519966	6760	6863	32	1	70
IST4134	54	834246	7525013	67.48	261934	54	7525013	6770	6875	31	1	73

**Table S2.3** | Published genomic islands (Holden et al., 2009; Graindorge et al., 2012; Nunvar et al., 2017) present in *B. cenocepacia* strains J2315 and ST32 and their eventual presence in *B. cenocepacia* IST439 and *B. multivorans* IST419. For clarity, the obtained data were divided into the two following pages.

Chromosome	Genomic island	size (Kb)	Original Coordinates in J2315		integration site	GC %	Presence (+)/absence (-) in Bc_IST439	Presence (+)/absence (-) in Bm_IST419	Functional note
			START	STOP					
Chromosome 1									
	<b>BcenGI1</b>	25	100361	125265	tRNA Arg	68.5	-	-	Prophage, contains a group II intron; similar to $\phi$ K96243 from <i>B. pseudomallei</i> K96243
	<b>BcenGI2</b>	16.4	188011	204429	tRNA Ala	55.7	-	-	ICE element similar to GI11 in <i>B. pseudomallei</i> K96243
	<b>BcenGI3</b>	13.1	449978	463082	-	56.5	-	-	Miscellaneous island, possible remnant, contains type I restriction modification system (BCAL0414, BCAL0418, and BCAL0420)
	<b>BcenGI4</b>	4.2	1032360	1036537	tRNA Met	42.8	-	-	Miscellaneous island, contains three CDSs of possible plasmid origin; similar island in <i>B. vietnamiensis</i> G4
	<b>BcenGI5</b>	92.8	1222566	1315385	-	60.3	-	-	Miscellaneous island, contains miscellaneous metabolism, including glycosyltransferase (BCAL1147), polysaccharide deacetylase (BCAL1148), putative O-antigen acetylase (BCAL1191), possible hydroxybenzoate degradation components (BCAL1151 to BCAL1162), fusaric acid resistance-family transporter (BCAL1176 to BCAL1178), and glycerate kinase (BCAL1181)
	<b>BcenGI6</b>	43.2	1402882	1437088	tRNA Arg	58.8	-	-	Miscellaneous island
	<b>BcenGI7</b>	37.7	1728919	1766584	BCAL1558	62.9	-	-	Prophage, Mu-like
	<b>BcenGI8</b>	121.9	2738065	2859923	BCAL2601	61.4	-	-	Miscellaneous island, contains miscellaneous metabolism including D-lactate dehydrogenase (BCAL2487) and putative O-antigen acetylase (BCAL2519), glutathione S-transferase (BCAL2539), 3- isopropylmalate dehydratase (BCAL2542 and BCAL2543), MFS transporter (BCAL2525 and BCAL2545), ABC transporter (BCAL2573 to BCAL2576), and nitrilase (BCAL2585)
	<b>BcenGI9</b>	16.3	3241588	3257909	tmRNA	62.8	-	-	Miscellaneous island, phage origins, possible phage remnant
	<b>BcenGI10</b>	6.3	3368693	3375006	tRNA Gly	59.4	-	-	Miscellaneous island, inserted into divergent region
Chromosome 2									
	<b>BcenGI11 - entire cci_GI</b>	44.1	290291	334378	-	62	+	+	Cenocepacia island, contains arsenic resistance (BCAM0233 to BCAM0235), antibiotic resistance (BCAM0237A), ion and sulfate family transporter (BCAM0238 and BCAM0281), and stress response CDSs (BCAM0276 and BCAM0278)
	<b>BcenGI12</b>	46.8	1140183	1186975	tRNA Sec	63.2	+	+	Prophage
	<b>BcenGI13</b>	46.2	2091701	2137922	BCAM1874	54.3	+	+	Prophage
Chromosome 3									
	<b>BcenGI14</b>	36.7	573207	609954	-	62.8	+	+	Prophage BcepMu

Chromosome 1		length	start	end					
	<b>GiST32-01</b>	36,238	3,811,949	25,440	-	-	-	-	Genes as transposases, hypothetical proteins
	<b>GiST32-02</b>	91,371	261,509	352,880	-	-	-	-	Genes as metabolic enzymes (various), conjugal transfer proteins, transcriptional regulators, MGE nucleases, hypothetical proteins
	<b>GiST32-03</b>	24,474	1,933,309	1,957,783	-	-	-	-	Genes as transposases, hypothetical proteins
	<b>GiST32-04</b>	28,516	3,049,973	3,078,489	-	-	-	-	Genes as type I R-M system, transposases, hypothetical proteins
	<b>GiST32-05</b>	113,931	3,288,957	3,402,888	-	-	-	-	Genes as metabolic enzymes (various), transporters, conjugal transfer proteins, transcriptional regulators, MGE nucleases, hypothetical proteins
	<b>GiST32-06</b>	33,426	3,531,047	3,564,473	-	-	-	-	Genes as carbohydrate metabolism proteins, transposases, hypothetical proteins
Chromosome 2									
	<b>GiST32-07</b>	15,256	49,520	64,776	-	-	-	-	Genes as transcriptional regulators, transposases, hypothetical proteins
	<b>GiST32-08</b>	12,279	574,223	586,502	-	-	-	-	Genes as hypothetical proteins
	<b>GiST32-09</b>	21,087	1,170,849	1,191,936	-	-	-	-	Genes as transposases, hypothetical proteins
	<b>GiST32-10</b>	44,101	1,728,875	1,772,976	-	-	-	-	Genes as cation transporters, conjugal transfer proteins, transcriptional regulators, hypothetical proteins
	<b>GiST32-11</b>	92,710	2,544,158	2,636,868	-	-	-	-	Genes as metabolic enzymes (various), transporters, transcriptional regulators, hypothetical proteins
Chromosome 3									
	<b>GiST32-12</b>	183,688	96,353	280,041	-	-	+	±	Genes as metabolic enzymes (various), transporters, transcriptional regulators, hypothetical proteins
	<b>GiST32-13</b>	21,693	371,242	392,935	-	-	+	±	Genes as metabolic enzymes (various), transporters
	<b>GiST32-14</b>	14,254	975,331	989,585	-	-	-	-	Genes as transposases
Plasmid									
	<b>GiST32-15</b>	77,659	142,679	28,394	-	-	-	-	Genes as transporters, histidine kinases, transposases, hypothetical proteins
	<b>GiST32-16</b>	27,948	44,927	72,875	-	-	-	-	Genes as copper metabolism proteins, transporters, transcriptional regulators, MGE nucleases, hypothetical proteins

All the Genomic islands, previously published (Holden et al., 2009; Graindorge et al., 2012; Nunvar et al., 2017), from *B. cenocepacia* strains J2315 and ST32 and their presence in our reference clones of *B. cenocepacia* IST439 and *B. multivorans* IST419. Hits above 90% query length and/or 95% sequence identity were considered as present (+); Hits below 40% query length and/or 70% sequence identity were considered non-reliable and discarded (absence: -); Hits within 40-90% query length and/or within 95-70% sequence identity were highlighted as (±). All of these results were confirmed by comparative visualization done by MAUVE and ACT/Artemis.

**Table S2.4** | BcenST218\_GIs as Genomic islands of *B. cenocepacia* IST439 and their presence in *B. multivorans* IST419, *B. cenocepacia* J2315 and ST32, *B. multivorans* ATCC\_17616 and ATCC\_BAA-247.

GI prediction/verification method	Chromosome	Genomic island	size (Kb)	Original Coordinates in IST439		integration site	GC %	Presence (+)/absence (-) in Bm_IST419	Presence (+)/absence (-) in Bc_J2315	Presence (+)/absence (-) in Bc_ST32	Presence (+)/absence (-) in Bm_ATCC_17616	Presence (+)/absence (-) in Bm_ATCC_BAA-247	Functional note
				START	STOP								
	Chromosome 1												
IslandViewer, manually curated/confirmed		<b>BcenST218_GI1</b>	29.69	86458	116148	tRNA Arg and tRNA Met	63.5	-	-	+	-	-	Miscellaneous island, contains flanked regions homologous to <i>B. cenocepacia</i> J2315 and 13.6 Kb unique Prophage CP4-57 that contains IntA Prophage integrase and its AlpA regulatory proteins as well as other genes encode 9 hypothetical proteins,
IslandViewer, manually curated/confirmed		<b>BcenST218_GI2</b>	52.11	1328969	1381083	tRNA Arg	63.5	-	-	-	-	-	Miscellaneous island, contains flanked regions homologous to <i>B. cenocepacia</i> J2315 and 49.4 Kb unique Miscellaneous DNA region that contains putative formate transporter 1, FocA, Ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE, and Nitrate regulatory protein, NasR, putative integrase as well as other genes encode 20 hypothetical proteins,
IslandViewer, manually curated/confirmed		<b>BcenST218_GI3</b>	31.29	1388801	1420087	tRNA Pro	57.5	-	-	-	-	-	Miscellaneous region that contains 27.6 unique regions which has 12 genes of putative proteins, two assembly and one integrase proteins of Prophage CPS-53 and Putative transposon Tn552 DNA-invertase bin3
IslandViewer, manually curated/confirmed		<b>BcenST218_GI4</b>	59.19	1942750	2001939	tRNA Lys	61.5	-	±	±	-	-	Miscellaneous region that contains some homologous genes to J2315 but involves 21.8 Kb unique region. It involves several genes encoding proteins associated with surface polysaccharides and its biosynthesis and other regulatory proteins. It has several transposases and IS2 transposase TnpB.
IslandViewer, manually curated/confirmed		<b>BcenST218_GI5</b>	42.7	2563426	2606123	tRNA Arg	64.8	-	-	-	-	±	Unique bacteriophage P2, GpU, GPD, GpE, and Tail_P2_I
IslandViewer, manually curated/confirmed		<b>BcenST218_GI6</b>	21.69	2753693	2775382	tRNA Val	57.8	-	±	±	-	-	Mobile genomic endonuclease - MGE nucleases and proteins of putative transmembrane anti-sigma factor
IslandViewer, manually curated/confirmed		<b>BcenST218_GI7</b>	20.64	3329766	3350403	tRNA Met	60	±	±	±	-	-	Miscellaneous island, contains flanked regions homologous to <i>B. cenocepacia</i> J2315 and 5.7 Kb unique region Phage proteins, Succinyl-CoA ligases, transcriptional regulators, and other hypothetical membrane proteins
	Chromosome 2												
Manually prediction due to its specificity		<b>BcenST218_GI8</b>	18.55	5453264	5471815	-	67.2	-	-	-	-	-	Unique region that involves genes associated with Transcription regulation and metabolism of organic compounds and inorganic compounds
IslandViewer, manually curated/confirmed		<b>BcenST218_GI9</b>	53.7	5593173	5646850	-	61	-	±	±	-	-	IS2 transposase TnpB and other Transposases inserted within several Type IV secretion system protein and other porin proteins
	Chromosome 3												
Manually prediction due to its specificity		<b>BcenST218_GI10</b>	19.8	6788287	6808083	-	67.4	-	-	-	-	-	Unique region that involves genes associated with metabolism, ABC transporters, MDR proteins, transcriptions regulators and hypothetical proteins
Manually prediction due to its specificity		<b>BcenST218_GI11</b>	41.33	6810787	6852114	-	62.5	-	-	-	-	-	Unique region that involves genes associated with metabolism, Flagellar biosynthesis/assembly proteins, and hypothetical membrane proteins
Manually prediction due to its specificity		<b>BcenST218_GI12</b>	191.9	6941321	7133249	-	69.4	-	-	+	±	±	Unique region that involves genes associated with metabolism, Transcription, transporters, and hypothetical proteins
Manually prediction due to its specificity		<b>BcenST218_GI13</b>	12.22	7227847	7240067	-	66.8	+	-	+	+	+	Unique region that involves genes associated with metabolism, Transcription, transporters, and hypothetical proteins
Manually prediction due to its specificity		<b>BcenST218_GI14</b>	21.55	7508984	7530535	-	67.8	-	-	+	-	-	Unique region that involves genes associated with metabolism, Transcription regulators, efflux transporters, and hypothetical proteins

**Table S2.5** | BmST836\_GIs as Genomic islands of *B. multivorans* IST419 and their presence in *B. cenocepacia* IST439, *B. cenocepacia* J2315 and ST32, *B. multivorans* ATCC\_17616 and ATCC\_BAA-247. For clarity, the obtained data were divided into the two following pages.

GI prediction/verification method	Chromosome	Genomic island	size (Kb)	Original Coordinates in IST419		integration site	GC %	Presence (+)/absence (-) in Bc_IST439	Presence (+)/absence (-) in Bc_J2315	Presence (+)/absence (-) in Bc_ST32	Presence (+)/absence (-) in Bm_ATCC_17616	Presence (+)/absence (-) in Bm_ATCC_BAA-247	Functional note
				START	STOP								
	Chromosome 1												
IslandViewer, manually curated/confirmed		<b>BmST836_GI1</b>	38.53	349334	387865	tRNA Thr, tRNA Gly, and tRNA Met	59.6	-	±	±	-	-	Miscellaneous island, contains 27.6 Kb unique region with several transposases and Prophage, genes encode/act as recombination/transcriptional regulator, transporters, histidine kinase, Transporters, DNA-repair enzymes, type VI secretion-associated protein, cytochrome C related proteins, copper resistance protein CopC, others involved in metabolisms and several hypothetical proteins
IslandViewer, manually curated/confirmed		<b>BmST836_GI2</b>	11.45	866447	877900	tRNA Ser	56.9	-	-	-	-	-	Bacteriophage GP29 together with LysR family transcriptional regulator and other hypothetical proteins
IslandViewer, manually curated/confirmed		<b>BmST836_GI3</b>	19.32	1018078	1037398	-	61.3	-	-	-	-	±	Prophage CP4-57 with lambda family phage portal proteins, virulence-associated E family proteins, XRE family transcriptional regulator, other hypothetical proteins
IslandViewer, manually curated/confirmed		<b>BmST836_GI4</b>	17.6	1050200	1067800	tRNA Ser	59.5	-	-	-	-	-	Bacteriophage P2, GPU, Tail_P2_I, and hypothetical proteins
IslandViewer, manually curated/confirmed		<b>BmST836_GI5</b>	28.98	1170793	1199777	tRNA Leu	61.4	-	-	-	-	-	Prophage that has 10.2 Kb unique DNA involving transposases, XRE family transcriptional regulator, NAD-dependent epimerase/dehydratase, 2'-5' RNA ligase, and other hypothetical proteins
IslandViewer, manually curated/confirmed		<b>BmST836_GI6</b>	17.28	1728503	1745784	tRNA Val	61.3	-	-	-	-	-	Miscellaneous island, contains 10.3 Kb unique Prophage and region with MerR family transcriptional regulator, antibiotic biosynthesis monooxygenase, lipoproteins involved in metabolisms and other hypothetical proteins
IslandViewer, manually curated/confirmed		<b>BmST836_GI7</b>	31549	1913167	1944716	tRNA Pro and tRNA Arg	61.1	-	-	-	-	±	Miscellaneous island, with LysR family transcriptional regulator, NAD-dependent epimerase/dehydratase, other transposases and genes involved in metabolisms and other hypothetical proteins
IslandViewer, manually curated/confirmed		<b>BmST836_GI8</b>	14.04	1951739	1965780	tRNA Pro	62.5	-	-	-	-	-	Putative prophage that has putative bacteriophage proteins, putative exported phage proteins, Phage antirepressor protein KilAC domain protein, Arc-like DNA binding domain protein, mannosyl-glycoprotein endo-beta-N-acetylglucosamidase, and several hypothetical proteins
IslandViewer, manually curated/confirmed		<b>BmST836_GI9</b>	12.5	1975107	1987609	tRNA Pro	60.6	-	-	-	-	-	Putative prophage that has putative exported phage proteins, Phage antirepressor proteins, exonuclease VIII, integrase family protein, CsbD family protein, entericidin EcnAB, catalase and several hypothetical proteins

IslandViewer, manually curated/confirmed		<b>BmST836_GI10</b>	17.2	2408556	2425760	-	58.1	-	-	-	-	-	Miscellaneous island, contains miscellaneous metabolism, including glycosyltransferases, polysaccharide deacetylases, putative O-antigen acetylase, ABC-2 type transporter, polysaccharide biosynthesis protein CapD, General stress protein A, and two hypothetical proteins
IslandViewer, manually curated/confirmed		<b>BmST836_GI11</b>	31.44	3011171	3042608	tRNA Leu	60.6	-	-	-	-	-	Miscellaneous island, has transposase IS4 and Prophage proteins, and contains enoyl-CoA hydratase, LysR family transcriptional regulators, N,N-dimethylformamide beta subunit, aminotransferases, FMN-dependent alpha-hydroxy acid dehydrogenases, major facilitator transporters, galactonate dehydratase, Stress responsive A/B Barrel Domain protein, and several hypothetical proteins
IslandViewer, manually curated/confirmed		<b>BmST836_GI12</b>	20.05	3339286	3359339	-	57.3	-	-	-	±	±	bacteriophage GP29 associated proteins with 4 ImpA family type VI secretion-associated proteins, DotU family type IV/VI secretion system protein, and two hypothetical proteins
	Chromosome 2												
IslandViewer, manually curated/confirmed		<b>BmST836_GI13</b>	44.53	3808183	3852711	-	64.6	-	-	-	-	-	Miscellaneous island with two transposases, and contains miscellaneous DNA, including OmpA/MotB domain-containing protein, NADPH-dependent FMN reductase, ATP-dependent proteases ATP-binding subunit ClpX, S-(hydroxymethyl)glutathione dehydrogenase/class III alcohol dehydrogenases, AraC family transcriptional regulators, acetyl-CoA acetyltransferases, LysR family transcriptional regulators, glucose-methanol-choline oxidoreductase, and other hypothetical proteins
IslandViewer, manually curated/confirmed		<b>BmST836_GI14</b>	13.12	5211528	5224649	-	64.2	-	-	-	-	-	Prophage
IslandViewer, manually curated/confirmed		<b>BmST836_GI15</b>	30.28	5321097	5351374	-	61.8	-	-	-	-	-	Miscellaneous island with Bacteriophage P2; GpU, GpE, and other related genes, Transposases; IS116/IS110/IS902 family protein, transposase IS3/IS911 family protein, transposase mutator type, transposase IS66, and IS66 Orf2 family protein, tRNA_anti-like protein, XRE family transcriptional regulator, transcriptional activator Ogr/delta, zinc finger CHC2-family protein, LysR family transcriptional regulator, others associated with metabolisms and hypothetical proteins
IslandViewer, manually curated/confirmed		<b>BmST836_GI16</b>	22.42	5737734	5760149	-	64.7	-	-	-	-	-	Miscellaneous island with transposases mutator type, adhesin, PpiC-type peptidyl-prolyl cis-trans isomerases, and hypothetical proteins
	Chromosome 3												
Manually prediction due to its specificity ... it has several contigs; starting from Sc 24 to the end		<b>BmST836_GI17</b>	117.7	6362788	6480500			-	-	-	-	-	Miscellaneous island with several contigs that might be false prediction

**Table S2.6** | Antibiotic resistance (AMR) and virulence factors (VF) associated genes in *B. cenocepacia* IST439 detected by the presented database.

Gene	Start	End	Coverage	Gaps	% Coverage	% Identity	Database	Accession	Product
									NC_001611.1
<i>pvdL</i>	946540	946767	4382- 4612/13029	43562	1.73	75.54	vfdb	NP_251114	peptide synthase PvdL [pyoverdine (IA001)] [Pseudomonas aeruginosa PAO1]
<i>tsr</i>	1160269	1160945	842- 1517/2022	43653	33.28	76.76	vfdb	YP_109900	methyl-accepting chemotaxis protein I [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>cdpA</i>	1558032	1559774	1-1743/1743	43624	99.77	86.2	vfdb	YP_107884	cyclic di-GMP phosphodiesterase [CdpA (VF0432)] [Burkholderia pseudomallei K96243]
<i>wcbT</i>	1966195	1967508	1-1314/1320	0/0	99.55	87.29	vfdb	YP_109382	acyl-CoA transferase [Capsule I (VF0436)] [Burkholderia pseudomallei K96243]
<i>wcbS</i>	1967573	1968489	2-918/918	0/0	99.89	90.08	vfdb	YP_109383	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase [Capsule I (VF0436)] [Burkholderia pseudomallei K96243]
<i>wcbR</i>	1968487	1976121	7-7641/7641	11994	99.71	88.5	vfdb	YP_109384	capsular polysaccharide biosynthesis fatty acid synthase [Capsule I (VF0436)] [Burkholderia pseudomallei K96243]
<i>manC</i>	1996528	1997942	1-1415/1428	43498	99.02	83.4	vfdb	YP_109405	GDP-mannose pyrophosphorylase [Capsule I (VF0436)] [Burkholderia pseudomallei K96243]
<i>pilB</i>	2210238	2211422	73- 1260/1260	43788	93.41	78.09	vfdb	YP_109603	type IV pilus assembly protein [Type IV pili (VF0431)] [Burkholderia pseudomallei K96243]
<i>pilD</i>	2212761	2213593	104-930/930	43532	88.82	80.1	vfdb	YP_109605	type IV prepilin leader peptide type M1 [Type IV pili (VF0431)] [Burkholderia pseudomallei K96243]
<i>icmF1/tssM1</i>	2314770	2314961	142- 336/3306	43533	5.72	75.25	vfdb	NP_248767	type VI secretion system protein IcmF1 [HSI-I (VF0334)] [Pseudomonas aeruginosa PAO1]
<i>tssH-5/clpV</i>	2322923	2323606	580- 1263/3039	43791	22.15	78.71	vfdb	YP_111509	Clp-type ATPase chaperone protein [T6SS-1 (VF0429)] [Burkholderia pseudomallei K96243]
<i>algW</i>	2349683	2350358	253- 919/1170	15/27	56.24	75.47	vfdb	NP_253136	AlgW protein [Alginate regulation (CVF523)] [Pseudomonas aeruginosa PAO1]
<i>fliA</i>	2521145	2521876	1-732/732	0/0	100	89.75	vfdb	YP_109887	flagellar biosynthesis sigma factor [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliG</i>	2521958	2522720	1-763/813	0/0	93.85	81	vfdb	YP_109888	flagellar biosynthesis protein FliG [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliF</i>	2522713	2524512	1-1752/1752	45/192	95.89	76.66	vfdb	YP_109889	flagellar biosynthesis regulator FliF [Flagella (VF0430)] [Burkholderia pseudomallei K96243]

<i>flhA</i>	2524509	2526595	14- 2103/2103	43468	99.24	91.1	vfdb	YP_109890	flagellar biosynthesis protein FlhA [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flhB</i>	2526630	2527817	2-1190/1218	43657	97.13	86.93	vfdb	YP_109891	flagellar biosynthesis protein FlhB [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>cheZ</i>	2530940	2531671	1-729/729	43505	99.59	82.31	vfdb	YP_109895	chemotaxis regulator CheZ [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>cheY</i>	2531674	2532098	3-429/429	43498	99.07	92.51	vfdb	YP_109896	chemotaxis protein CheY [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>cheB</i>	2532137	2533227	2-1104/1104	43569	98.73	88.32	vfdb	YP_109897	chemotaxis-specific methyltransferase [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>cheD</i>	2533380	2533983	1-613/705	43694	85.11	87.68	vfdb	YP_109898	chemoreceptor glutamine deamidase CheD [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>cheR</i>	2534093	2534917	56-880/948	0/0	87.03	86.55	vfdb	YP_109899	chemotaxis protein methyltransferase [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>tsr</i>	2535350	2536934	1-1584/2022	43527	78.29	83.73	vfdb	YP_109900	methyl-accepting chemotaxis protein I [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>cheW</i>	2536977	2537496	7-523/528	43468	97.92	91.15	vfdb	YP_109901	chemotaxis protein CheW [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>cheA</i>	2537543	2539796	1-2236/2244	30/80	98.26	85.08	vfdb	YP_109902	chemotaxis two-component sensor kinase CheA [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>cheY1</i>	2539849	2540207	1-359/381	43498	93.96	78.89	vfdb	YP_109903	chemotaxis two-component response regulator CheY1 [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>motB</i>	2540298	2541251	1-954/1023	0/0	93.26	88.05	vfdb	YP_109904	flagellar motor protein MotB [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>motA</i>	2541265	2542125	1-861/861	0/0	100	92.45	vfdb	YP_109905	flagellar motor protein MotA [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliC</i>	2553769	2554923	1-1167/1167	17/44	97.6	78.19	vfdb	YP_109915	flagellin [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>xcpR</i>	2712256	2713136	322- 1202/1509	43754	57.85	75.14	vfdb	NP_251793	general secretion pathway protein E [xcp secretion system (VF0084)] [Pseudomonas aeruginosa PAO1]
<i>fliL</i>	2728654	2729197	1-543/543	43655	99.26	80.29	vfdb	YP_106651	flagellar basal body protein FliL [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliM</i>	2729220	2730218	1-999/999	0/0	100	91.79	vfdb	YP_106652	flagellar motor switch protein FliM [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliN</i>	2730211	2730699	1-498/498	43570	97.59	85.03	vfdb	YP_106653	flagellar motor switch protein FliN [Flagella (VF0430)] [Burkholderia pseudomallei K96243]

<i>fliO</i>	2730798	2731250	1-468/495	13/31	89.9	79.62	vfdb	YP_106654	flagellar protein FliO [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliP</i>	2731308	2732068	1-761/762	43498	99.74	88.85	vfdb	YP_106655	flagellar biosynthesis protein FliP [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliQ</i>	2732089	2732361	1-273/273	0/0	100	89.01	vfdb	YP_106656	flagellar biosynthesis protein FliQ [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliR</i>	2732393	2733142	1-750/783	43559	95.53	84.84	vfdb	YP_106657	flagellar biosynthetic protein FliR [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>plcH</i>	2795906	2796124	898-1116/2193	43559	9.9	76.47	vfdb	NP_249535	hemolytic phospholipase C precursor [PLC (VF0092)] [Pseudomonas aeruginosa PAO1]
<i>plcH</i>	2796262	2796454	1182-1374/2193	43502	8.66	76.53	vfdb	NP_249535	hemolytic phospholipase C precursor [PLC (VF0092)] [Pseudomonas aeruginosa PAO1]
<i>fliJ</i>	2883619	2884065	1-447/453	43559	98.23	82.63	vfdb	YP_106854	flagellar protein FliJ [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliI</i>	2884071	2885598	21-1572/1572	43705	97.07	88.35	vfdb	YP_106855	flagellum-specific ATP synthase FliI [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliH</i>	2885609	2886292	1-681/681	43629	99.27	83.16	vfdb	YP_106856	flagellar assembly protein H [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliG</i>	2886285	2887280	1-996/996	0/0	100	92.57	vfdb	YP_106857	flagellar motor switch protein G [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliF</i>	2887270	2889033	1-1797/1797	13/39	98	87	vfdb	YP_106858	flagellar M-ring protein FliF [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliE</i>	2889293	2889626	12-342/342	43468	96.78	88.02	vfdb	YP_106859	flagellar hook-basal body complex protein FliE [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliS</i>	2889866	2890288	1-423/435	0/0	97.24	83.69	vfdb	YP_106860	flagellar protein FliS [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgN</i>	2925655	2926076	1-422/441	43498	95.46	81.32	vfdb	YP_106894	flagella synthesis protein FlgN [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgM</i>	2926172	2926369	125-322/345	43498	57.1	76.38	vfdb	YP_106895	negative regulator of flagellin synthesis [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgA</i>	2926600	2927334	23-764/765	43661	95.56	82.57	vfdb	YP_106896	flagellar basal body P-ring biosynthesis protein FlgA [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgB</i>	2928015	2928500	1-483/492	43527	98.17	86.21	vfdb	YP_106897	flagellar basal-body rod protein FlgB [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgC</i>	2928542	2928961	1-420/426	0/0	98.59	90	vfdb	YP_106898	flagellar basal body rod protein FlgC [Flagella (VF0430)] [Burkholderia pseudomallei K96243]

<i>flgD</i>	2929122	2929734	196-808/852	0/0	71.95	86.79	vfdb	YP_106899	flagellar basal-body rod modification protein FlgD [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgE</i>	2929797	2931038	1-1242/1242	43660	99.44	85.67	vfdb	YP_106900	flagellar hook protein FlgE [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgF</i>	2931058	2931814	1-760/762	43468	99.34	85.92	vfdb	YP_106901	flagellar basal body rod protein FlgF [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgG</i>	2931851	2932639	1-789/789	0/0	100	92.02	vfdb	YP_106902	flagellar basal body rod protein FlgG [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgH</i>	2932698	2933349	72-723/723	0/0	90.18	88.96	vfdb	YP_106903	flagellar L-ring protein precursor FlgH [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgI</i>	2933455	2934539	107-1194/1194	43468	90.87	93.47	vfdb	YP_106904	flagellar P-ring protein precursor FlgI [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgJ</i>	2934556	2935517	14-936/936	27699	96.69	80.1	vfdb	YP_106905	flagellar rod assembly protein/muramidase FlgJ [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgK</i>	2936683	2938047	1-1374/2004	43665	67.86	81.65	vfdb	YP_106907	flagellar hook-associated protein 1 FlgK [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgL</i>	2938022	2938280	596-851/1233	43499	20.76	75.68	vfdb	YP_106908	flagellar hook-associated protein 3 FlgL [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgK</i>	2938231	2938614	1624-2004/2004	43468	19.01	86.72	vfdb	YP_106907	flagellar hook-associated protein 1 FlgK [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgL</i>	2938632	2939863	1-1229/1233	43468	99.68	79.79	vfdb	YP_106908	flagellar hook-associated protein 3 FlgL [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>tsr</i>	3133122	3133814	817-1509/2022	43626	34.03	79.66	vfdb	YP_109900	methyl-accepting chemotaxis protein I [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>pilA</i>	3350590	3351093	12-515/519	43500	96.72	77.87	vfdb	YP_107407	type IV fimbrial pilin protein [Type IV pili (VF0431)] [Burkholderia pseudomallei K96243]
NC_001612.1									
<i>tsr</i>	384152	384814	842-1504/2022	43622	32.64	77.63	vfdb	YP_109900	methyl-accepting chemotaxis protein I [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>tsr</i>	656026	656720	818-1512/2022	43813	34.03	77.07	vfdb	YP_109900	methyl-accepting chemotaxis protein I [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>pmlR/bspR1</i>	720570	721285	1-716/720	43498	99.31	82.15	vfdb	YP_110896	N-acylhomoserine lactone dependent regulatory protein [Quorum-sensing (VF0433)] [Burkholderia pseudomallei K96243]
<i>pmlI/bspI1</i>	722013	722588	1-579/612	43562	93.79	80.21	vfdb	YP_110894	N-acylhomoserine lactone synthase [Quorum-sensing (VF0433)] [Burkholderia pseudomallei K96243]

<i>tsr</i>	826684	827348	827-1494/2022	43723	32.59	77	vfdb	YP_109900	methyl-accepting chemotaxis protein I [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>pscR</i>	849712	849918	414-620/654	0/0	31.65	76.33	vfdb	NP_250384	type III secretion system protein PscR [TTSS (VF0083)] [Pseudomonas aeruginosa PAO1]
<i>bsaQ</i>	876808	877382	190-764/2073	43732	27.16	75.13	vfdb	YP_111549	Type III secretion system protein BsaQ [Bsa T3SS (VF0428)] [Burkholderia pseudomallei K96243]
<i>boaA</i>	1265013	1265719	2587-3293/4962	0/0	14.25	78.78	vfdb	YP_110805	surface-exposed protein autotransporter [BoaA (VF0434)] [Burkholderia pseudomallei K96243]
<i>boaA</i>	1265394	1266235	2452-3293/4962	0/0	16.97	79.69	vfdb	YP_110805	surface-exposed protein autotransporter [BoaA (VF0434)] [Burkholderia pseudomallei K96243]
<i>boaA</i>	1269717	1270258	2452-2993/4962	43656	10.82	78.43	vfdb	YP_110805	surface-exposed protein autotransporter [BoaA (VF0434)] [Burkholderia pseudomallei K96243]
<i>boaA</i>	1270147	1270258	2495-2606/4962	0/0	2.26	84.82	vfdb	YP_110805	surface-exposed protein autotransporter [BoaA (VF0434)] [Burkholderia pseudomallei K96243]
<i>boaA</i>	1271049	1271312	2482-2736/4962	13089	4.88	76.53	vfdb	YP_110805	surface-exposed protein autotransporter [BoaA (VF0434)] [Burkholderia pseudomallei K96243]
<i>plcH</i>	1285985	1286185	82-282/2193	0/0	9.17	75.62	vfdb	NP_249535	hemolytic phospholipase C precursor [PLC (VF0092)] [Pseudomonas aeruginosa PAO1]
<i>cheD</i>	2530816	2530979	1102-1265/1674	0/0	9.8	77.44	vfdb	YP_001006778	methyl-accepting chemotaxis protein CheD [peritrichous flagella (AI145)] [Yersinia enterocolitica subsp. enterocolitica 8081]
<i>flgE</i>	2763359	2764597	1-1239/1242	43595	99.36	83.92	vfdb	YP_106900	flagellar hook protein FlgE [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
NC_001613.1									
<i>fliP</i>	261648	261832	572-756/768	43498	23.96	76.34	vfdb	NP_250137	flagellar biosynthetic protein FliP [Flagella (VF0273)] [Pseudomonas aeruginosa PAO1]

**Table S2.7** | Antibiotic resistance (AMR) and virulence factors (VF) associated genes in *B. multivorans* IST419 detected by the presented database.

Gene	Sequence	Start	End	Coverage	Gaps	% Coverage	% Identity	Database	Accession	Product
<i>pilA</i>	sc10_si175006	143456	143967	3-517/519	43594	98.07	80.12	vfdb	YP_107407	type IV fimbrial pilin protein [Type IV pili (VF0431)] [Burkholderia pseudomallei K96243]
<i>tsr</i>	sc1_si1301000	38336	39052	817-1533/2022	43748	35.21	77.98	vfdb	YP_109900	methyl-accepting chemotaxis protein I [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>cdpA</i>	sc1_si1301000	377883	379615	1-1733/1743	43559	99.31	86.63	vfdb	YP_107884	cyclic di-GMP phosphodiesterase [CdpA (VF0432)] [Burkholderia pseudomallei K96243]
<i>pilB</i>	sc1_si1301000	969497	970658	87-1248/1260	43685	91.9	80.79	vfdb	YP_109603	type IV pilus assembly protein [Type IV pili (VF0431)] [Burkholderia pseudomallei K96243]
<i>pilD</i>	sc1_si1301000	972004	972835	105-930/930	43626	88.6	80.58	vfdb	YP_109605	type IV prepilin leader peptide type M1 [Type IV pili (VF0431)] [Burkholderia pseudomallei K96243]
<i>pmlR/bspR1</i>	sc2_si731746	2340	3057	1-718/720	43498	99.58	82.34	vfdb	YP_110896	N-acylhomoserine lactone dependent regulatory protein [Quorum-sensing (VF0433)] [Burkholderia pseudomallei K96243]
<i>tsr</i>	sc2_si731746	53141	53821	827-1509/2022	13/16	33.33	76.23	vfdb	YP_109900	methyl-accepting chemotaxis protein I [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>boaA</i>	sc2_si731746	315757	315918	4732-4893/4962	0/0	3.26	78.4	vfdb	YP_110805	surface-exposed protein autotransporter [BoaA (VF0434)] [Burkholderia pseudomallei K96243]
<i>tsr</i>	sc2_si731746	397812	398472	842-1502/2022	43622	32.54	76.51	vfdb	YP_109900	methyl-accepting chemotaxis protein I [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>boaB</i>	sc4_si537795	66229	66560	1516-1832/4821	43539	6.58	75.6	vfdb	YP_108306	autotransporter protein [BoaB (VF0435)] [Burkholderia pseudomallei K96243]
<i>boaB</i>	sc4_si537795	67285	67430	1516-1667/4821	43471	3.03	75.66	vfdb	YP_108306	autotransporter protein [BoaB (VF0435)] [Burkholderia pseudomallei K96243]

<b><i>boaA</i></b>	sc4_si537795	68833	69019	3517- 3703/4962	43500	3.73	78.84	vfdb	YP_110805	surface-exposed protein autotransporter [BoaA (VF0434)] [Burkholderia pseudomallei K96243]
<b><i>tsr</i></b>	sc4_si537795	423623	424324	789- 1490/2022	14/20	34.22	76.12	vfdb	YP_109900	methyl-accepting chemotaxis protein I [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>phzG1</i></b>	sc53_si597	249	597	297- 645/645	0/0	54.11	99.71	vfdb	NP_252906	phenazine biosynthesis protein PhzG pyridoxamine 5'-phosphate oxidase [Phenazines biosynthesis (CVF536)] [Pseudomonas aeruginosa PAO1]
<b><i>icmF1/tssM1</i></b>	sc5_si522894	493091	493295	129- 335/3306	43500	6.17	75	vfdb	NP_248767	type VI secretion system protein IcmF1 [HSI-I (VF0334)] [Pseudomonas aeruginosa PAO1]
<b><i>clpV1</i></b>	sc5_si522894	509125	509822	593- 1290/2709	43559	25.69	75.43	vfdb	NP_248780	type VI secretion system AAA+ family ATPase [HSI-I (VF0334)] [Pseudomonas aeruginosa PAO1]
<b><i>phzB1</i></b>	sc66_si534	1	228	262- 489/489	0/0	46.63	99.12	vfdb	NP_252900	phenazine biosynthesis protein PhzB [Phenazines biosynthesis (CVF536)] [Pseudomonas aeruginosa PAO1]
<b><i>phzC1</i></b>	sc66_si534	252	534	1- 283/1218	0/0	23.23	100	vfdb	NP_252901	phenazine biosynthesis protein PhzC [Phenazines biosynthesis (CVF536)] [Pseudomonas aeruginosa PAO1]
<b><i>pvdI</i></b>	sc6_si465875	339121	339336	536- 751/15450	43595	1.37	75.57	vfdb	NP_251092	peptide synthase [pyoverdine (IA001)] [Pseudomonas aeruginosa PAO1]
<b><i>pvdL</i></b>	sc6_si465875	348227	348488	4392- 4653/1302 9	43498	2	76.05	vfdb	NP_251114	peptide synthase PvdL [pyoverdine (IA001)] [Pseudomonas aeruginosa PAO1]
<b><i>fliJ</i></b>	sc7_si450623	11876	12310	13- 447/453	43498	95.81	83.72	vfdb	YP_106854	flagellar protein FliJ [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>fliI</i></b>	sc7_si450623	12367	13734	193- 1560/1572	0/0	87.02	90.86	vfdb	YP_106855	flagellum-specific ATP synthase FliI [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>fliH</i></b>	sc7_si450623	13941	14618	1-681/681	43564	99.12	84.8	vfdb	YP_106856	flagellar assembly protein H [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>fliG</i></b>	sc7_si450623	14611	15606	1-996/996	0/0	100	92.17	vfdb	YP_106857	flagellar motor switch protein G [Flagella (VF0430)] [Burkholderia pseudomallei K96243]

<i>fliF</i>	sc7_si450623	15596	17368	1- 1797/1797	14/50	97.94	85.47	vfdb	YP_106858	flagellar M-ring protein FliF [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliE</i>	sc7_si450623	17628	17961	12- 342/342	43468	96.78	86.53	vfdb	YP_106859	flagellar hook-basal body complex protein FliE [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliS</i>	sc7_si450623	18147	18557	1-411/435	0/0	94.48	83.21	vfdb	YP_106860	flagellar protein FliS [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgN</i>	sc7_si450623	56767	57188	1-422/441	43498	95.46	83.69	vfdb	YP_106894	flagella synthesis protein FlgN [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgA</i>	sc7_si450623	57673	58310	122- 765/765	43502	83.4	84.16	vfdb	YP_106896	flagellar basal body P-ring biosynthesis protein FlgA [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgB</i>	sc7_si450623	59041	59526	1-483/492	43529	97.97	87.27	vfdb	YP_106897	flagellar basal-body rod protein FlgB [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgC</i>	sc7_si450623	59570	59989	1-420/426	0/0	98.59	88.81	vfdb	YP_106898	flagellar basal body rod protein FlgC [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgD</i>	sc7_si450623	60147	60757	196- 806/852	0/0	71.71	86.74	vfdb	YP_106899	flagellar basal-body rod modification protein FlgD [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgE</i>	sc7_si450623	60822	62063	1- 1242/1242	43660	99.44	85.75	vfdb	YP_106900	flagellar hook protein FlgE [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgF</i>	sc7_si450623	62084	62840	1-760/762	43468	99.34	87.11	vfdb	YP_106901	flagellar basal body rod protein FlgF [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgG</i>	sc7_si450623	62880	63668	1-789/789	0/0	100	92.9	vfdb	YP_106902	flagellar basal body rod protein FlgG [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgH</i>	sc7_si450623	63709	64376	50- 723/723	43532	92.25	89.63	vfdb	YP_106903	flagellar L-ring protein precursor FlgH [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgI</i>	sc7_si450623	64398	65554	26- 1194/1194	43756	96.65	90.02	vfdb	YP_106904	flagellar P-ring protein precursor FlgI [Flagella (VF0430)] [Burkholderia pseudomallei K96243]

<i>flgJ</i>	sc7_si450623	65578	66520	24- 936/936	14154	97.12	80.68	vfdb	YP_106905	flagellar rod assembly protein/muramidase FlgJ [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgK</i>	sc7_si450623	67672	69651	1- 2004/2004	14946	98.4	81.11	vfdb	YP_106907	flagellar hook-associated protein 1 FlgK [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>flgL</i>	sc7_si450623	69669	70902	1- 1231/1233	43631	99.35	78.63	vfdb	YP_106908	flagellar hook-associated protein 3 FlgL [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>plcH</i>	sc7_si450623	125934	126083	133- 282/2193	0/0	6.84	76.67	vfdb	NP_249535	hemolytic phospholipase C precursor [PLC (VF0092)] [Pseudomonas aeruginosa PAO1]
<i>tsr</i>	sc7_si450623	263770	264457	826- 1509/2022	43691	33.58	79.08	vfdb	YP_109900	methyl-accepting chemotaxis protein I [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliR</i>	sc8_si246735	25736	26501	1-766/783	43498	97.7	83.7	vfdb	YP_106657	flagellar biosynthetic protein FliR [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliQ</i>	sc8_si246735	26525	26797	1-273/273	0/0	100	88.64	vfdb	YP_106656	flagellar biosynthesis protein FliQ [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliP</i>	sc8_si246735	26816	27562	15- 761/762	43498	97.9	90.78	vfdb	YP_106655	flagellar biosynthesis protein FliP [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliO</i>	sc8_si246735	27618	28099	2-495/495	13424	94.95	81.42	vfdb	YP_106654	flagellar protein FliO [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliN</i>	sc8_si246735	28190	28568	120- 498/498	0/0	76.1	92.88	vfdb	YP_106653	flagellar motor switch protein FliN [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliM</i>	sc8_si246735	28662	29660	1-999/999	0/0	100	92.69	vfdb	YP_106652	flagellar motor switch protein FliM [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>fliL</i>	sc8_si246735	29683	30223	1-543/543	43760	97.79	79.02	vfdb	YP_106651	flagellar basal body protein FliL [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<i>xcpR</i>	sc8_si246735	45358	46240	320- 1202/1509	14/20	57.85	75.03	vfdb	NP_251793	general secretion pathway protein E [xcp secretion system (VF0084)] [Pseudomonas aeruginosa PAO1]
<i>fliC</i>	sc8_si246735	158634	159791	1- 1167/1167	21/53	97.34	78.13	vfdb	YP_109915	flagellin [Flagella (VF0430)] [Burkholderia pseudomallei K96243]

<b><i>motA</i></b>	sc8_si246735	169582	170442	1-861/861	0/0	100	92.45	vfdb	YP_109905	flagellar motor protein MotA [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>motB</i></b>	sc8_si246735	170456	171409	1-954/1023	0/0	93.26	88.78	vfdb	YP_109904	flagellar motor protein MotB [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>cheYI</i></b>	sc8_si246735	171579	171937	1-359/381	43502	93.44	79.28	vfdb	YP_109903	chemotaxis two-component response regulator CheY1 [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>cheA</i></b>	sc8_si246735	171990	174235	1-2234/2244	28/74	98.17	85.77	vfdb	YP_109902	chemotaxis two-component sensor kinase CheA [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>cheW</i></b>	sc8_si246735	174284	174817	7-528/528	43477	98.86	89.89	vfdb	YP_109901	chemotaxis protein CheW [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>tsr</i></b>	sc8_si246735	174858	176412	1-1555/2022	43528	76.81	84.14	vfdb	YP_109900	vfdb~tsr~YP_109900 (tsr) methyl-accepting chemotaxis protein I [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>cheR</i></b>	sc8_si246735	176896	177750	56-908/948	43622	89.77	86.93	vfdb	YP_109899	chemotaxis protein methyltransferase [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>cheD</i></b>	sc8_si246735	177824	178476	1-662/705	43576	91.77	83.08	vfdb	YP_109898	chemoreceptor glutamine deamidase CheD [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>cheB</i></b>	sc8_si246735	178622	179712	2-1104/1104	43536	98.82	88.67	vfdb	YP_109897	chemotaxis-specific methylesterase [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>cheY</i></b>	sc8_si246735	179752	180175	4-429/429	43498	98.83	90.14	vfdb	YP_109896	chemotaxis protein CheY [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>cheZ</i></b>	sc8_si246735	180177	180908	1-729/729	43499	100	83.33	vfdb	YP_109895	chemotaxis regulator CheZ [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>flhB</i></b>	sc8_si246735	183167	184376	2-1218/1218	43757	98.85	87.24	vfdb	YP_109891	flagellar biosynthesis protein FlhB [Flagella (VF0430)] [Burkholderia pseudomallei K96243]

<b><i>flhA</i></b>	sc8_si246735	184373	186475	1- 2103/2103	43502	99.86	91.26	vfdb	YP_109890	flagellar biosynthesis protein FlhA [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>flhF</i></b>	sc8_si246735	186472	188250	1- 1752/1752	43/139	96.8	77.77	vfdb	YP_109889	flagellar biosynthesis regulator FlhF [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>flhG</i></b>	sc8_si246735	188243	188996	1-754/813	43498	92.62	79.87	vfdb	YP_109888	flagellar biosynthesis protein FlhG [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>fliA</i></b>	sc8_si246735	189093	189824	1-732/732	0/0	100	90.71	vfdb	YP_109887	flagellar biosynthesis sigma factor [Flagella (VF0430)] [Burkholderia pseudomallei K96243]
<b><i>chrB</i></b>	sc3_si559217	182380	182588	529- 737/939	0/0	22.26	76.56	card	A7J11_05670	23S rRNA (guanine(748)-N(1))-methyltransferase ChrB
<b><i>Nocardia_ri fampin_resi stant_beta- subunit_of_ RNA_poly merase_(rp oB2)</i></b>	sc10_si175006	7869	8226	2980- 3337/3489	43498	10.23	77.44	card	AP006618.1:4 835200- 4838689	Due to gene duplication the genomes of Nocardia species include both rifampin-sensitive beta-subunit of RNA polymerase (rpoB) and rifampin-resistant beta-subunit of RNA polymerase (rpoB2) genes with ~88% similarity between the two gene products. Expression of the rpoB2 variant results in replacement of rifampin sensitivity with rifampin resistance.
<b><i>Burkholder ia_pseudom allei_Omp3 8</i></b>	sc13_si111755	50013	50197	178- 359/1122	43468	16.22	75.14	card	AY312416:1- 1123	Heterologous expression of Burkholderia pseudomallei Omp38 (BpsOmp38) in Omp-deficient E. coli host cells lowers their permeability and in consequence their antimicrobial susceptibility to penicillin G cefoxitin ceftazidime and imipenem.
<b><i>ceoA</i></b>	sc15_si76403	14047	15264	1- 1218/1218	43498	99.92	92.86	card	U97042:1- 1219	ceoA is a periplasmic linker subunit of the CeoAB-OpcM efflux pump
<b><i>ceoB</i></b>	sc15_si76403	15310	18394	1- 3084/3084	43466	100	95.36	card	U97042:1264- 4348	ceoB is a cytoplasmic membrane component of the CeoAB-OpcM efflux pump

<b><i>opcM</i></b>	sc15_si76403	18506	20044	1- 1536/1536	43590	99.93	90.52	card	U38944.1:1- 1537	OpcM is an outer membrane factor protein found in Burkholderia cepacia. It is part of the CeoAB-OpcM complex.
<b><i>Burkholderia pseudomallei_Omp38</i></b>	sc1_si1301000	119310	119477	196- 360/1122	43468	14.71	75	card	AY312416:1- 1123	Heterologous expression of Burkholderia pseudomallei Omp38 (BpsOmp38) in Omp-deficient E. coli host cells lowers their permeability and in consequence their antimicrobial susceptibility to penicillin G cefoxitin ceftazidime and imipenem.
<b><i>MuxB</i></b>	sc1_si1301000	373131	373658	2521- 3048/3132	43559	16.79	77.36	card	NC_002516.2: 2850883- 2854015	MuxB is one of the two necessary RND components in the Pseudomonas aeruginosa efflux pump system MuxABC-OpmB.
<b><i>penA</i></b>	sc4_si537795	302102	302990	1-919/939	43646	94.68	79.87	card	A7J11_03272	PenA family class A beta-lactamase
<b><i>Burkholderia pseudomallei_Omp38</i></b>	sc4_si537795	512248	513336	1- 1122/1122	17/59	95.9	79.38	card	AY312416:1- 1123	Heterologous expression of Burkholderia pseudomallei Omp38 (BpsOmp38) in Omp-deficient E. coli host cells lowers their permeability and in consequence their antimicrobial susceptibility to penicillin G cefoxitin ceftazidime and imipenem.
<b><i>amrB</i></b>	sc6_si465875	366440	369554	1- 3115/3132	43563	99.33	87.62	card	NC_006350.1: 2147818- 2150950	amrB is the membrane fusion protein of the AmrAB-OprM multidrug efflux complex.
<b><i>amrA</i></b>	sc6_si465875	369570	370741	23- 1200/1200	43789	97.08	77.89	card	NC_006350.1: 2150966- 2152166	amrA is the efflux pump subunit of the AmrAB-OprM multidrug efflux complex. amrA corresponds to 1 locus in Pseudomonas aeruginosa PAO1 and 1 locus in Pseudomonas aeruginosa LESB58.

**Table S2.8** | List of mutations found among *B. cenocepacia* clonal variants. All mutated genes were grouped to predicted gene functions (COG). Isolates are chronologically ordered as indicated in Fig. 2.1. For clarity, the obtained data were divided into the following pages.

A – GATK pipeline and B – SAMtools/BCFtools toolbox.

Gene locus	Gene prediction, Annotation	COG category	COG Symbol	<i>B. cenocepacia</i> J2315 homolog	<i>B. multivorans</i> ATCC_17616 homolog	Mutation effect	Predicted functional effect and impact	Isolates														
								IST4103	IST4110	IST4112	IST4113	IST4116A	IST4116B	IST4131	IST4128	IST4130	IST4134					
Cellular Processes and Signaling																						
IST439_01755	<i>wbil/capD/wbqV</i> , UDP-N-acetyl-alpha-D-glucosamine C6 dehydratase / nucleotide sugar epimerase/dehydratase	Cell wall/membrane/envelope biogenesis	M	BCAL3119	Bmul_2504 / BMULJ_RS03625	Leu493Pro	CDS: missense variant, moderate			SNP AB	SNP AB		SNP AB									
IST439_01755	<i>wbil/capD</i> , UDP-N-acetyl-alpha-D-glucosamine C6 dehydratase / nucleotide sugar epimerase/dehydratase		M	BCAL3119	Bmul_2504 / BMULJ_RS03625	Glu487Lys	CDS: missense variant, moderate		SNP AB			SNP AB		SNP AB	SNP AB	SNP AB	SNP AB					
IST439_01871	<i>wbbL/mshA_1</i> , D-inositol 3-phosphate glycosyltransferase		M	-	-	Tyr150His	CDS: missense variant, moderate			SNP AB	SNP AB											
IST439_05843	<i>bceR/cotSA</i> , Spore coat protein SA		M	BCAM1008	Bmul_4608 / BMULJ_RS19505	Phe362Val	CDS: missense variant, moderate	SNP AB														
IST439_02709	<i>flhK</i> , Flagellar hook-length control protein	Cell motility	N	BCAL0520	-	Ala96_Thr97del (-5bp)	CDS: frameshift variant, moderate		INDEL A													
IST439_02709	<i>flhK</i> , Flagellar hook-length control protein		N	BCAL0520	-	Thr99_Ala100del (-5bp)	CDS: frameshift variant, moderate		INDEL B													
IST439_01743	<i>ureE</i> , Urease accessory protein UreE	Post-translational modification, protein turnover, and chaperones	O	BCAL3107	Bmul_2491 / BMULJ_RS03690	Asp201_His202 del	CDS: frameshift variant, high							INDEL A								
IST439_01499	<i>zraR_2</i> , Transcriptional regulatory protein ZraR / sigma-54 interacting response regulator protein	Signal transduction mechanisms	T	BCAL1011	Bmul_2164 / BMULJ_RS05385	Leu363Pro	CDS: missense variant, moderate											SNP AB				
IST439_06597	<i>fgrL/yycG_3</i> , Sensor histidine kinase YycG/Two-component regulatory system		T	BCAS0619	Bmul_5359 / BMULJ_RS30790	Val176Met	CDS: missense variant, moderate		SNP AB			SNP AB		SNP AB		SNP AB	SNP AB					
IST439_06759	<i>todS</i> , Sensor histidine kinase TodS		T	BCAS0708	-	Thr161Ile	CDS: missense variant, moderate											SNP AB				
IST439_01028	<i>emrA</i> , Multidrug export protein EmrA	Defense mechanisms	V	BCAL1511	Bmul_1751 / BMULJ_RS07395	Ala254Gly	CDS: missense variant, moderate	SNP AB														
IST439_01475	Putative (ABC transporter) multidrug export ATP-binding/permease membrane protein		V	BCAL1039	Bmul_2146 / BMULJ_RS05495	Gln347*	CDS: stop gained, High	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB				
IST439_04198	<i>penA</i> , Beta-lactamase precursor		V	BCAM2165	Bmul_3689 / BMULJ_RS24095	Asp262Gly	CDS: missense variant, moderate			SNP AB	SNP AB		SNP AB									

Information Storage and Processing																		
IST439_00705	<i>hisS</i> , Histidine-tRNA ligase / histidyl-tRNA synthetase	Translation, ribosomal structure and biogenesis	J	BCAL1883	Bmul_1464 / BMULJ_RS08840	Arg204Cys	CDS: missense variant, moderate					SNP AB		SNP AB	SNP AB			
IST439_00910	<i>bpeR/envR/tetR</i> , TetR family regulatory protein / HTH-type transcriptional regulator TtgR	Transcription	K	BCAL2823	Bmul_0687	Cys63Arg	CDS: missense variant, moderate			SNP AB	SNP AB		SNP AB					
IST439_00910	<i>bpeR/envR/tetR</i> , TetR family regulatory protein / HTH-type transcriptional regulator TtgR		K	BCAL2823	Bmul_0687	Gly196Asp	CDS: missense variant, moderate		SNP AB			SNP AB		SNP AB	SNP AB	SNP AB	SNP AB	
IST439_00910	<i>bpeR/envR/tetR</i> , TetR family regulatory protein / HTH-type transcriptional regulator TtgR		K	BCAL2823	Bmul_0687	Lys113fs (-57bp)	CDS: frameshift variant, high	INDEL B										
IST439_00921	<i>rbsR</i> , Ribose operon repressor		K	BCAL1660	Bmul_1628 / BMULJ_RS08025	Val287Asp	CDS: missense variant, moderate	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB
IST439_01925	<i>oxyR</i> , Hydrogen peroxide-inducible genes activator / oxidative stress regulatory protein		K	BCAL3301	Bmul_2661 / BMULJ_RS02880	Ser53Arg	CDS: missense variant, moderate		SNP AB			SNP AB		SNP AB				
IST439_02295	<i>rpoC</i> , DNA-directed RNA polymerase subunit beta		K	BCAL0227	Bmul_0242 / BMULJ_RS15055	Tyr669Cys	CDS: missense variant, moderate	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB
IST439_02296	<i>rpoB</i> , DNA-directed RNA polymerase subunit beta		K	BCAL0226	Bmul_0241 / BMULJ_RS15060	Ser613Pro	CDS: missense variant, moderate								SNP AB			
IST439_02593	<i>prtR</i> , putative transmembrane anti-sigma factor		K	BCAL3479	Bmul_0008 / BMULJ_RS00310	Asp15Tyr	CDS: missense variant, moderate	SNP AB										
IST439_04244	<i>acoR_2</i> , Acetoin catabolism regulatory protein		K	BCAM2211	Bmul_3641 / BMULJ_RS24350	Pro599_Ala601del	CDS: frameshift variant, moderate			INDEL AB	INDEL AB		INDEL AB					
IST439_05266	<i>cynR-13</i> , LysR family regulatory protein / HTH-type transcriptional regulator CynR		K	BCAM0421	Bmul_5195 / BMULJ_RS16620	Lys259Glu	CDS: missense variant, moderate		SNP AB			SNP AB		SNP AB				
IST439_05755	<i>rpoD/sigE</i> , RNA polymerase sigma factor RpoD		K	BCAM0918	Bmul_4813 / BMULJ_RS18520	Arg167Thr	CDS: missense variant, moderate	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB
IST439_06446	<i>YdcR_5</i> , putative HTH-type transcriptional regulator YdcR / GntR family		K	BCAS0024 / BCAM0920	Bmul_6146 / BMULJ_RS26905	Leu300Gln	CDS: missense variant, moderate		SNP AB									
IST439_06699	<i>rutR_3</i> , HTH-type transcriptional regulator RutR / TetR family		K	BCAS0007	Bmul_6158 / BMULJ_RS26845	Thr49Ala	CDS: missense variant, moderate					SNP AB						
IST439_06667	<i>yhaO/NOPBF</i> , putative metallophosphoesterase YhaO / putative calcineurin-like phosphoesterase family protein	Replication, recombination and repair	L	BCAS0748	Bmul_5337 / BMULJ_RS30925	Gly153Trp	CDS: missense variant, moderate		SNP AB		SNP AB		SNP AB		SNP AB	SNP AB	SNP AB	

Metabolism													
IST439_01911	<i>hmpA</i> , Flavohemoprotein	Energy production and conversion	C	BCAL3285	Bmul_2649 / BMULJ_RS02940	Asp337Glu	CDS: missense variant, moderate	SNP AB		SNP AB	SNP AB	SNP AB	
IST439_02838	2-oxoacid ferredoxin oxidoreductase		C	BCAL0650	Bmul_0357 / BMULJ_RS14475	Val1062Leu	CDS: missense variant, moderate	SNP AB		SNP AB	SNP AB	SNP AB	
IST439_00975	<i>fliY_1</i> , Cystine-binding periplasmic protein	Amino acid transport and metabolism	E	BCAL1610	Bmul_1679 / BMULJ_RS07760	Asp13Gly	CDS: missense variant, moderate	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB
IST439_01981	<i>artP/glnQ</i> , Arginine transporter ATP-binding subunit / Glutamine transport ATP-binding protein GlnQ/ABC transporter-like protein		E	-	Bmul_2711	Gln136*	CDS: stop gained, High	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB
IST439_05824	<i>asd</i> , Aspartate-semialdehyde dehydrogenase		E	BCAM0986	Bmul_4626 / BMULJ_RS19415	Leu105His/Pro	CDS: missense variant, moderate						SNP AB
IST439_01622	<i>add/ADA</i> , Adenine deaminase	Nucleotide transport and metabolism	F	BCAL2989	Bmul_2376 / BMULJ_RS04320	Val116Ile	CDS: missense variant, moderate					SNP AB	
IST439_01686	<i>pgl</i> , 6-phosphogluconolactonase	Carbohydrate transport and metabolism	G	BCAL3043	Bmul_2432 / BMULJ_RS03990	Ala113Gly	CDS: missense variant, moderate	SNP AB					
IST439_05703	<i>BceL/ymfM_4</i> , major facilitator superfamily protein / Inner membrane transport protein YnfM		G	BCAM0865	Bmul_4909 / BMULJ_RS18035	Thr117Ala	CDS: missense variant, moderate	SNP AB		SNP AB	SNP AB	SNP AB	SNP AB
IST439_06741	<i>dhaK</i> , Dihydroxyacetone kinase		G	BCAS0051	Bmul_6137 / BMULJ_RS26950	Ala235Val	CDS: missense variant, moderate		SNP AB				
IST439_00216	<i>sulP/dauA/cynT</i> , putative sulfate transporter / C4-dicarboxylic acid transporter DauA	Inorganic ion transport and metabolism	P	BCAL2353	Bmul_1019 / BMULJ_RS11160	Leu61Pro	CDS: missense variant, moderate	SNP AB		SNP AB	SNP AB	SNP AB	SNP AB
IST439_00883	<i>orba</i> , Ornibactin receptor / TonB-dependent siderophore receptor		P	BCAL1700	Bmul_1594 / BMULJ_RS08185	Ser382Leu	CDS: missense variant, moderate		SNP AB	SNP AB	SNP AB		
IST439_05233	<i>phoA_2</i> , Alkaline phosphatase 4 precursor		P	BCAM0390	-	Arg458Gly	CDS: missense variant, moderate		SNP B		SNP B		
IST439_00884	<i>pvdA</i> , L-ornithine 5-monooxygenase	Secondary metabolites biosynthesis, transport, and catabolism	Q	BCAL1699	Bmul_1595 / BMULJ_RS08180	Ser450fs (-1bp, G)	CDS: frameshift variant, high			INDEL B			
IST439_00884	<i>pvdA</i> , L-ornithine 5-monooxygenase		Q	BCAL1699	Bmul_1595 / BMULJ_RS08180	Val449fs (-1a)	CDS: frameshift variant, high		INDEL AB	INDEL AB	INDEL AB		
IST439_00884	<i>pvdA</i> , L-ornithine 5-monooxygenase		Q	BCAL1699	Bmul_1595 / BMULJ_RS08180	Val449Leu	CDS: missense variant, moderate		SNP AB	SNP AB	SNP AB		
IST439_01356	<i>mdcE</i> , Malonate decarboxylase gamma subunit / Malonyl-S-ACP:biotin-protein carboxyltransferase MADD		Q	BCAL1243	Bmul_2047 / BMULJ_RS05985	Asp67Glu	CDS: missense variant, moderate	SNP AB		SNP AB	SNP AB	SNP AB	SNP AB
IST439_02887	<i>mIaE</i> , putative phospholipid ABC transporter permease protein MlaE		Q	BCAL0697	Bmul_0403 / BMULJ_RS14240	Leu31_Ala32del	CDS: frameshift variant, moderate				INDEL AB		
IST439_04949	<i>pcaC</i> , 4-Carboxymuconolactone decarboxylase family protein		Q	BCAM0062	Bmul_5284 / BMULJ_RS16180	Ser52Arg	CDS: missense variant, moderate	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB
IST439_06912	<i>tauD/tfdA</i> , Taurine catabolism dioxygenase TauD, TfdA family		Q	BCAS0205	-	Arg137Pro	CDS: missense variant, moderate	SNP AB					

Poorly characterized																
		General function prediction only / undefined	R													
IST439_00207	Hypothetical protein / putative transmembrane protein	Function unknown	S	BCAL2362	-	Ala213fs (-40bp)	CDS: frameshift variant, high					INDEL AB	INDEL B			
IST439_00383	<i>rpfE</i> , Hypothetical protein / putative Regulatory (signal peptide protein), RpfE		S	BCAL2187	Bmul_1155 / BMULJ_RS10430	Glu305_Leu316 del (-35bp)	CDS: frameshift variant, moderate	INDEL B	INDEL B	INDEL B	INDEL B	INDEL L B	INDEL B	INDEL B	INDEL B	INDEL B
IST439_00383	<i>rpfE</i> , Hypothetical protein / putative Regulatory (signal peptide protein), RpfE		S	BCAL2187	Bmul_1155 / BMULJ_RS10430	Leu301_Ala312 del (-35)	CDS: frameshift variant, moderate	INDEL A	INDEL A	INDEL A	INDEL A	INDEL L A	INDEL A	INDEL A	INDEL A	INDEL A
IST439_00764	Hypothetical protein		S	BCAL1825	Bmul_1508 / BMULJ_RS08615	Val35Ala	CDS: missense variant, moderate						SNP AB			
IST439_00951	Hypothetical Cytosolic protein		S	BCAL1630	Bmul_1658 / BMULJ_RS07875	Asp140Asn	CDS: missense variant, moderate	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB				
IST439_01003	Hypothetical membrane protein / Putative Tad-like Flp pilus-assembly		S	BCAL1535	-	Leu232Gln	CDS: missense variant, moderate	SNP AB								
IST439_01261	Hypothetical protein		S	-	-	Thr137Ile	CDS: missense variant, moderate	SNP AB								
IST439_01472	Hypothetical membrane protein		S	BCAL1042	Bmul_2143 / BMULJ_RS05510	His130Pro	CDS: missense variant, moderate						SNP AB			
IST439_01761	Hypothetical protein / HAD superfamily hydrolase-like protein involved in OAg biosynthetic process		S	-	Bmul_2510	Thr116Pro	CDS: missense variant, moderate	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB				
IST439_02890	Putative lipoprotein / ABC transporter auxiliary component-like protein		S	BCAL0700	Bmul_0406 / BMULJ_RS14225	Leu197fs (-2bp_GC)	CDS: frameshift variant, high					INDEL L AB		INDEL AB	INDEL AB	
IST439_03324	<i>lptF/yjgP</i> , Putative permease / Lipopolysaccharide export system permease protein LptF		S	BCAL2677	Bmul_0827 / BMULJ_RS12115	Thr67Met	CDS: missense variant, moderate						SNP AB	SNP AB	SNP AB	
IST439_03496	<i>ycbX</i> , Putative sulfure-carrier protein / MOSC domain-containing protein		S	BCAM1384	Bmul_4341 / BMULJ_RS20830	Thr124Ala	CDS: missense variant, moderate	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB				
IST439_05619	Hypothetical protein		S	-	-	Gln1070Arg	CDS: missense variant, moderate	SNP AB								
IST439_06941	<i>ytfT</i> , Putative ABC transporter / Inner membrane ABC transporter permease protein YtfT		S	BCAS0231	Bmul_5961 / BMULJ_RS27845	Ala275Val	CDS: missense variant, moderate						SNP B		SNP B	

Intergenic																				
IST439_00155--- IST439_00156	<i>rhlE2</i> , putative ATP-dependent RNA helicase 2 --- <i>gluQ</i> , glutamyl-Q tRNA(Asp) synthetase	Replication, recombination and repair --- Translation, ribosomal structure and biogenesis	L/J	BCAL2412 --- BCAL2411	Bmul_0961 --- Bmul_0962		IG: (16/160 bp upstream CDS), modifier	INDEL B												
IST439_02253--- IST439_02254	<i>yedY</i> , putative sulfite oxidase subunit YedY --- <i>ccsA</i> , putative cytochrome c biogenesis protein	Function unknown --- Post-translational modification, protein turnover, and chaperones	S/O	BCAL0269 --- BCAL0268	Bmul_0284 --- Bmul_0283		IG: (96/63 bp upstream CDS), modifier	SNP AB												
IST439_02458--- IST439_02459	<i>gdsL</i> , hypothetical lipolytic enzyme, GdsL --- <i>ydgJ</i> , putative oxidoreductase YdgJ, Gfo idh MocA family	Amino acid transport and metabolism --- Cell wall/membrane/envelope biogenesis	E/M	BCAL0078 --- BCAL0079			IG: (16/160 bp upstream CDS), modifier			INDEL A	INDEL A		INDEL A							
IST439_02458--- IST439_02459	<i>gdsL</i> , hypothetical lipolytic enzyme, GdsL --- <i>ydgJ</i> , putative oxidoreductase YdgJ, Gfo idh MocA family	Amino acid transport and metabolism --- Cell wall/membrane/envelope biogenesis	E/M	BCAL0078 --- BCAL0079			IG: (16/160 bp upstream CDS), modifier			INDEL B	INDEL B		INDEL B							
IST439_02458--- IST439_02459	<i>gdsL</i> , hypothetical lipolytic enzyme, GdsL --- <i>ydgJ</i> , putative oxidoreductase YdgJ, Gfo idh MocA family	Amino acid transport and metabolism --- Cell wall/membrane/envelope biogenesis	E/M	BCAL0078 --- BCAL0079			IG: (16/160 bp upstream CDS), modifier						INDEL A							
IST439_03477--- IST439_03478	<i>hpaR</i> , homoprotocatechuate degradative operon repressor/Organic hydroperoxide resistance transcriptional regulator --- <i>hpaG</i> , 5-carboxymethyl-2-hydroxyruconate delta-isomerase/ Homoprotocatechuate catabolism bifunctional isomerase/decarboxylase	Transcription --- Secondary metabolites biosynthesis, transport, and catabolism	K/Q	BCAM1365 --- BCAM1366			IG: (146/8 bp upstream CDS), modifier	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB	SNP AB
IST439_04393--- IST439_04394	Putative aminoacylate hydrolase RutD --- Endo-1,4-beta-xylanase Z precursor	Function unknown --- Function unknown	S/S	BCAM2357 --- BCAM2358			IG: (56/2 bp upstream CDS), modifier		SNP A			SNP A		SNP A						
IST439_04393--- IST439_04394	Putative aminoacylate hydrolase RutD --- Endo-1,4-beta-xylanase Z precursor	Function unknown --- Function unknown	S/S	BCAM2357 --- BCAM2358			IG: (56/2 bp upstream CDS), modifier		SNP B			SNP B		SNP B						
IST439_04885--- IST439_04886	NADPH-dependent FMN reductase --- <i>ecfJ</i> , RNA polymerase sigma factor	Function unknown --- Transcription	S/K	BCAM2840 --- BCAM0001			IG: (272/97 bp upstream CDS), modifier		SNP AB											
IST439_04885--- IST439_04886	NADPH-dependent FMN reductase --- <i>ecfJ</i> , RNA polymerase sigma factor	Function unknown --- Transcription	S/K	BCAM2840 --- BCAM0001			IG: (300/69 bp upstream CDS), modifier		SNP AB			SNP AB		SNP AB						



**Table S2.9** | List of mutations found among *B. multivorans* clonal variants. All mutated genes were grouped to predicted gene functions (COG). Isolates are chronologically ordered as indicated in Fig. 2.1. For clarity, the obtained data were divided into the following pages.

A – GATK pipeline and B – SAMtools/BCFtools toolbox.

Gene locus	Gene prediction, Annotation	COG category	COG Symbol	<i>B. cenocepacia</i> J2315 homolog	<i>B. multivorans</i> ATCC_17616 homolog	Mutation effect	Predicted functional effect and impact	IST424 IST453 IST455A IST455B IST461 IST495A IST495B IST4119										
Cellular Processes and Signaling																		
IST419_02260	<i>wbil/capD/wbqV</i> , UDP-N-acetyl-alpha-D-glucosamine C6 dehydratase / nucleotide sugar epimerase/dehydratase/polysaccharide biosynthesis protein CapD	Cell wall/membrane/envelope biogenesis	M	BCAL3119	Bmul_2504 / BMULJ_RS03625	Met69fs (+64bp)	CDS: frameshift variant, high	INDEL B	INDEL B	INDEL AB	INDEL B							
IST419_02264	<i>gspA/cps1D</i> , General stress protein A/Glycosyl transferase family protein		M	-	-	His96Asp	CDS: missense variant, moderate											SNP AB
IST419_02356	<i>rmlB/rflB</i> , dTDP-glucose 4,6-dehydratase		M	BCAL3135	Bmul_2596 / BMULJ_RS03175	Tyr192His	CDS: missense variant, moderate											SNP AB
IST419_02515	<i>mpl</i> , UDP-N-acetylmuramate--L-alanyl-gamma-D-glutamyl-meso-diaminopimelate ligase		M	BCAL3416	Bmul_2775 / BMULJ_RS02315	Val314Gly	CDS: missense variant, moderate											SNP AB
IST419_03735	patatin/Putative patatin-like phospholipase		M	BCAM2172	Bmul_3684 / BMULJ_RS24120	Gln196*	CDS: stop gained, High											SNP AB
IST419_00471	<i>flgD</i> , flagellar basal body rod modification protein		Cell motility	N	BCAL0566	Bmul_3018 / BMULJ_RS01060	Ser219Ala	CDS: missense variant, moderate										SNP AB
IST419_04098	Serine protease/peptidase S1 and S6 chymotrypsin/Hap	Post-translational modification, protein turnover, and chaperones	O	-	Bmul_4138 / BMULJ_RS21865	Ala442_Gly443ins GlyAlaAlaGlyAla	CDS: frameshift variant, moderate										INDEL AB	
IST419_05079	<i>osmC</i> , Organic hydroperoxide resistance protein/OsmC family protein		O	BCAM0896	Bmul_4881 / BMULJ_RS18175	Gln120Pro	CDS: missense variant, moderate											SNP AB
IST419_00702	<i>hprK</i> , HPr kinase/phosphorylase	Signal transduction mechanisms	T	BCAL0809	Bmul_0521 / BMULJ_RS13640	Ser168Asn	CDS: missense variant, moderate	SNP AB	SNP AB	SNP AB	SNP AB							
IST419_01938	diguanylate phosphodiesterase		T	BCAL1100	Bmul_2085 / BMULJ_RS05790	Lys2Asn	CDS: missense variant, moderate											SNP AB
IST419_05375	<i>fgrL/yycG_3</i> , Sensor histidine kinase YycG/Two-component regulatory system/integral membrane sensor signal transduction histidine kinase		T	BCAS0619	Bmul_5359 / BMULJ_RS30790	Leu181Phe	CDS: missense variant, moderate											SNP AB
IST419_03607	ABC-2 type transporter	Defense mechanisms	V	BCAM2351	Bmul_3534 / BMULJ_RS24900	Met1*	CDS: stop gained, High	SNP AB	SNP AB	SNP AB	SNP AB							
IST419_05223	<i>fhaB/palA/hyla</i> , filamentous hemagglutinin outer membrane protein/hemolysin	muraceanar tracking, secretion, and vesicular transport	U	-	-	Pro115Leu	CDS: missense variant, moderate										SNP AB	

Information Storage and Processing																			
IST419_01084	<i>rsuA/rluF</i> , pseudouridine synthase, Rsu / RNA-binding S4 domain-containing protein	Translation, ribosomal structure and biogenesis	J	BCAL2094	BMULJ_01993 / Bmul_1254	Val285Ala	CDS: missense variant, moderate	SNP A	SNP A	SNP A	SNP A	SNP A	SNP A						
IST419_01084	<i>rsuA/rluF</i> , pseudouridine synthase, Rsu / RNA-binding S4 domain-containing protein		J	BCAL2094	BMULJ_01993 / Bmul_1254	Leu280Pro	CDS: missense variant, moderate	SNP A	SNP A										
IST419_01084	<i>rsuA/rluF</i> , pseudouridine synthase, Rsu / RNA-binding S4 domain-containing protein		J	BCAL2094	BMULJ_01993 / Bmul_1254	GlnAlaProLeuLysAlaAlaSerAlaHisArgIleAlaThrProArgProProAlaLysAlaArgSerValAlaValSerAlaSerValAlaArgArgProThrAlaProHisAlaAlaThrThrThrMetHisValArgValAla	CDS: frameshift variant, moderate	INDEL B					INDEL B						
IST419_01084	<i>rsuA/rluF</i> , pseudouridine synthase, Rsu / RNA-binding S4 domain-containing protein		J	BCAL2094	BMULJ_01993 / Bmul_1254	Ala266Val	CDS: missense variant, moderate		SNP A										SNP A
IST419_01084	<i>rsuA/rluF</i> , pseudouridine synthase, Rsu / RNA-binding S4 domain-containing protein		J	BCAL2094	BMULJ_01993 / Bmul_1254	Pro264Arg	CDS: missense variant, moderate		SNP A										SNP A
IST419_01084	<i>rsuA/rluF</i> , pseudouridine synthase, Rsu / RNA-binding S4 domain-containing protein		J	BCAL2094	BMULJ_01993 / Bmul_1254	ArgHisValAlaThrThrThrMetHisValArgValAlaArgAlaProProLysAlaAlaSerAlaHisArgIleAlaThrProArgProProAlaLysAlaArgSerValAlaValSerAlaSerValAlaAlaArgArgProThrAla	CDS: frameshift variant, moderate	INDEL B		INDEL B	INDEL B		INDEL B						
IST419_02422	<i>yafY</i> , helix-turn-helix type 11 domain-containing protein		Transcription	K	BCAL3322	Bmul_2679 / BMULJ_RS02790	Gly92Ala	CDS: missense variant, moderate											SNP A
IST419_03081	<i>fixK</i> , Crp/FNR family transcriptional regulator	K		BCAM0049	Bmul_5297 / BMULJ_RS16115	Ala11fs (-8bp)	CDS: frameshift variant, high			INDEL AB	INDEL AB								
IST419_03081	<i>fixK</i> , Crp/FNR family transcriptional regulator	K		BCAM0049	Bmul_5297 / BMULJ_RS16115	Pro12fs (-8bp)	CDS: frameshift variant, high			INDEL AB									
IST419_00350	putative phage-related protein / YqaJ-like viral recombinase domain	Replication, recombination and repair	L	BCAL1199	-	Thr54Ala	CDS: missense variant, moderate											SNP B	
IST419_02115	DNA gyrase subunit A		L	BCAL2957	Bmul_2264 / BMULJ_RS04880	Ser83Arg	CDS: missense variant, moderate											SNP AB	
IST419_04560	<i>orfA</i> , putative transposase		L	BCAM1930	-	Glu103Gln	CDS: missense variant, moderate											SNP B	
IST419_04562	putative integrase		L	BCAL2755	AI46_RS28850	Met1fs (+21bp)	CDS: frameshift variant, moderate						INDEL B					SNP B	
IST419_04865	transposase IS3/IS911 family protein		L	-	-	Met1*	CDS: stop gained, High											SNP B	
IST419_04865	transposase IS3/IS911 family protein		L	-	-	Glu99Asp	CDS: missense variant, moderate						SNP AB					SNP AB	
IST419_04866	integrase catalytic subunit		L	BCAS0656	Bmul_5774	His49Gln	CDS: missense variant, moderate						SNP AB					SNP AB	
IST419_04866	integrase catalytic subunit		L	BCAS0656	Bmul_5774	Thr69Ser	CDS: missense variant, moderate						SNP B					SNP B	

Metabolism																		
IST419_02576	molybdopterin oxidoreductase	Energy production and conversion	C	BCAL3475	Bmul_2849 / BMULJ_RS01945	Lys360Gln	CDS: missense variant, moderate	SNP AB	SNP AB	SNP AB	SNP AB							
IST419_03612	<i>fdhA</i> , formaldehyde dehydrogenase		C	BCAM2333	Bmul_3539 / BMULJ_RS24870	Ala160Val	CDS: missense variant, moderate				SNP AB							
IST419_03962	<i>betB/gbsA</i> , betaine-aldehyde dehydrogenase		C	BCAM1844	Bmul_3992 / BMULJ_RS22590	Gly105Asp	CDS: missense variant, moderate										SNP AB	
IST419_05773	<i>tfdB</i> , monooxygenase FAD-binding protein		C,H	-	Bmul_2460 / BMULJ_RS03850	Leu190Val	CDS: missense variant, moderate	SNP AB	SNP AB	SNP AB	SNP AB							
IST419_00066	<i>livF</i> , ABC transporter-like protein		E	BCAL0016	Bmul_0084 / BMULJ_RS15895	Val185Leu	CDS: missense variant, moderate					SNP AB	SNP AB					
IST419_00071	<i>livK</i> , extracellular ligand-binding receptor		E	BCAL0020	Bmul_0089 / BMULJ_RS15875	Gln139Arg	CDS: missense variant, moderate					SNP AB	SNP AB					
IST419_00093	<i>putA</i> , trifunctional transcriptional regulator/proline dehydrogenase/pyrroline-5-carboxylate dehydrogenase		E	BCAL0042	Bmul_0114 / BMULJ_RS15755	Pro1142Ala	CDS: missense variant, moderate											SNP AB
IST419_01537	<i>icyB/yecS</i> , polar amino acid ABC transporter inner membrane subunit	Amino acid transport and metabolism	E	BCAL1609	Bmul_1680 / BMULJ_RS07755	Ser165_Ile167dup	CDS: frameshift variant, moderate					INDEL AB	INDEL AB					
IST419_01971	<i>astA</i> , argi+S137nine/ornithine succinyltransferase subunit alpha		E	BCAL1060	Bmul_2123 / BMULJ_RS05605	Leu213Arg	CDS: missense variant, moderate				SNP AB							
IST419_02455	<i>gltL/aatP</i> , ABC transporter-like protein		E	-	Bmul_2711 / BMULJ_RS02630	Asp168Asn	CDS: missense variant, moderate	SNP AB	SNP AB	SNP AB	SNP AB							
IST419_02455	<i>gltL/aatP</i> , ABC transporter-like protein		E	-	Bmul_2711 / BMULJ_RS02630	Gln55*	CDS: stop gained, High											SNP AB
IST419_02769	<i>argH/ASL</i> , argininosuccinate lyase		E	BCAL2638	Bmul_0865 / BMULJ_RS11925	His163Leu	CDS: missense variant, moderate											SNP AB
IST419_03581	<i>soxA</i> , sarcosine oxidase subunit alpha		E	BCAM2389	Bmul_3506 / BMULJ_RS25045	Trp385Arg	CDS: missense variant, moderate					SNP AB	SNP AB					
IST419_01882	<i>carB</i> , carbamoyl phosphate synthase large subunit		Nucleotide transport and metabolism	F	BCAL1262	Bmul_2028 / BMULJ_RS06075	Ala945Val	CDS: missense variant, moderate										
IST419_01882	<i>carB</i> , carbamoyl phosphate synthase large subunit	F		BCAL1262	Bmul_2028 / BMULJ_RS06075	Cys769Trp	CDS: missense variant, moderate											SNP AB
IST419_01882	<i>carB</i> , carbamoyl phosphate synthase large subunit	F		BCAL1262	Bmul_2028 / BMULJ_RS06075	Cys769Arg	CDS: missense variant, moderate											SNP AB
IST419_02163	<i>relA/spoT</i> , (p)ppGpp synthetase I SpoT/RelA	F		BCAL3010	Bmul_2399 / BMULJ_RS04175	Ile493Met	CDS: missense variant, moderate					SNP AB	SNP AB					
IST419_04369	<i>codA</i> , N-isopropylammelide isopropylaminohydrolase	F		BCAM1457	Bmul_4281 / BMULJ_RS21125	Ala85Ser	CDS: missense variant, moderate											SNP AB
IST419_02029	<i>fabF2</i> , 3-oxoacyl-(acyl carrier protein) synthase II	Lipid transport and metabolism	I	BCAL0996	Bmul_2179 / BMULJ_RS05310	Gly145Ala	CDS: missense variant, moderate					SNP AB	SNP AB	SNP AB				
IST419_02029	<i>fabF2</i> , 3-oxoacyl-(acyl carrier protein) synthase II		I	BCAL0996	Bmul_2179 / BMULJ_RS05310	Phe74Tyr	CDS: missense variant, moderate		SNP AB	SNP AB	SNP AB	SNP AB						
IST419_05562	acetyl-CoA synthetase		I	-	Bmul_5976	Ala257Thr	CDS: missense variant, moderate	SNP B		SNP B	SNP B							
IST419_05562	acetyl-CoA synthetase		I	-	Bmul_5976	Glu484Gln	CDS: missense variant, moderate		SNP B			SNP B		SNP B	SNP AB			

IST419_00635	<i>ptsH</i> , phosphotransferase system, phosphocarrier protein HPr	Carbohydrate transport and metabolism	G	BCAL0735	Bmul_0442 / BMULJ_RS14050	Gln3fs (-1a)	CDS: frameshift variant, high				INDEL AB								
IST419_03450	<i>fucU</i> , transport protein RbsD/FucU		G	BCAM2509	Bmul_3411 / BMULJ_RS25550	Phe115Leu	CDS: missense variant, moderate												SNP AB
IST419_03737	amylase-1,6-galactosidase		G	BCAM2168	Bmul_3686 / BMULJ_RS24110	Glu133Asp	CDS: missense variant, moderate			SNP AB									
IST419_04557	<i>pepM</i> , phosphoenolpyruvate phosphomutase		G	BCAM1245	Bmul_4471 / BMULJ_RS20195	Asp448Glu	CDS: missense variant, moderate												SNP AB
IST419_04936	malto-oligosyltrehalose trehalohydrolase		G	-	Bmul_5487 / BMULJ_RS30125	Asp274Asn	CDS: missense variant, moderate												SNP AB
IST419_04114	hydrophobe/amphiphile efflux-1 (HAE1) family protein / multidrug efflux RND transporter permease subunit	Inorganic ion transport and metabolism	P	BCAL2821	Bmul_0689	Arg619Gly	CDS: missense variant, moderate			SNP AB									
IST419_04995	binding-protein-dependent transport system inner membrane protein		P	BCAM0954	Bmul_4663 / BMULJ_RS19235	Ile174Leu	CDS: missense variant, moderate												SNP AB
IST419_03373	hypothetical protein	Secondary metabolites biosynthesis, transport, and catabolism	Q	BCAL0362	Bmul_2900 / BMULJ_RS01670	His225Tyr	CDS: missense variant, moderate			SNP AB	SNP AB	SNP AB	SNP AB						
IST419_05081	PHB depolymerase family esterase		Q	BCAL0362	Bmul_2900 / BMULJ_RS01670	Ala272Ser	CDS: missense variant, moderate				SNP AB	SNP AB	SNP AB						
Poorly characterized																			
IST419_00681	<i>pagL</i> , Lipid A deacylase PagL precursor	Function unknown	S	BCAL0788	Bmul_0487 / BMULJ_RS13820	His181fs (+69bp)	CDS: frameshift variant, high											INDEL B	
IST419_00836	<i>rlmJ/comJ</i> , Ribosomal RNA large subunit methyltransferase J		S	BCAL2306	Bmul_1063 / BMULJ_RS10915	Leu161Val	CDS: missense variant, moderate												SNP AB
IST419_01128	hypothetical protein		S	-	-	Glu91fs (+39)	CDS: frameshift variant, high												INDEL B
IST419_01285	<a href="#">acoX</a> , ATP-NAD/AcoX kinase		S	BCAL1913	Bmul_1432	Gly82_Val131del (-49bp)	CDS: frameshift variant, moderate												INDEL B
IST419_01797	putative hemolysin-like protein		S	BCAL1281	Bmul_1997 / BMULJ_RS06225	His83fs (-9bp)	CDS: frameshift variant, high												INDEL L AB
IST419_01831	hypothetical protein		S	-	Bmul_1835 / BMULJ_RS31735	Arg120fs (+1bp)	CDS: frameshift variant, high												INDEL B
IST419_01831	hypothetical protein		S	-	Bmul_1835 / BMULJ_RS31735	Glu118fs (+28bp)	CDS: frameshift variant, high												INDEL B
IST419_02028	hypothetical protein		S	BCAL2873	Bmul_2178 / BMULJ_RS05315	Ala13Gly	CDS: missense variant, moderate												SNP AB
IST419_02386	<i>tldD</i> , peptidase U62 modulator of DNA gyrase		S	BCAL3281	Bmul_2645 / BMULJ_RS02960	Asn10Lys	CDS: missense variant, moderate				SNP AB	SNP AB	SNP AB						
IST419_02810	<i>lptG</i> , YjgP/YjgQ family permease		S	BCAL2678	Bmul_0826 / BMULJ_RS12120	Val168Leu	CDS: missense variant, moderate												SNP AB
IST419_02834	hypothetical protein / Hypothetical phage protein		S	BCAM1900	-	Gln80Leu	CDS: missense variant, moderate												SNP B
IST419_02834	hypothetical protein / Hypothetical phage protein		S	BCAM1900	-	Gln81Glu	CDS: missense variant, moderate												SNP B

IST419_02834	hypothetical protein / Hypothetical phage protein	Function unknown	S	BCAM1900	-	Ser84fs (+1bp)	CDS: frameshift variant, high									INDEL AB	INDEL AB			
IST419_02834	hypothetical protein / Hypothetical phage protein		S	BCAM1900	-	Pro83_Ser84insThr (+3bp)	CDS: frameshift variant, moderate										INDEL AB	INDEL AB		
IST419_02834	hypothetical protein / Hypothetical phage protein		S	BCAM1900	-	Ser84fs (+2bp)	CDS: frameshift variant, high										INDEL AB	INDEL AB		
IST419_02834	hypothetical protein / Hypothetical phage protein		S	BCAM1900	-	Ser86fs (+8bp)	CDS: frameshift variant, high										INDEL AB			
IST419_02834	hypothetical protein / Hypothetical phage protein		S	BCAM1900	-	Ser86fs (+34bp)	CDS: frameshift variant, high										INDEL AB			
IST419_03517	hypothetical protein		S	-	-	Met26_Pro50del (-74bp)	CDS: frameshift variant, moderate										INDEL B	INDEL B		
IST419_03746	hypothetical protein		S	BCAM2154	Bmul_3695 / BMULJ_RS24065	Val64Met	CDS: missense variant, moderate												SNP AB	
IST419_03850	Sel1 domain-containing protein		S	BCAM2008	Bmul_3794 / BMULJ_RS23580	Ala72Pro	CDS: missense variant, moderate												SNP AB	
IST419_04769	<i>mcp</i> , hypothetical protein		S	BCAM2689	Bmul_3293 / BMULJ_RS26125	Gly238_Lys239ins LysProGluAlaAla	CDS: frameshift variant, moderate	INDEL A		INDEL A	INDEL A									
IST419_05096	hypothetical protein		S	BCAM0878	Bmul_4898 / BMULJ_RS18090	Ala115Glu	CDS: missense variant, moderate		SNP AB											
IST419_05247	IgA FC receptor precursor		S	-	-	Pro302_Val307del (-17bp)	CDS: frameshift variant, moderate	INDEL A												
IST419_05247	IgA FC receptor precursor		S	-	-	Asp153_Pro170del (-51bp)	CDS: frameshift variant, moderate			INDEL B	INDEL B	INDEL B								
IST419_05247	IgA FC receptor precursor		S	-	-	Glu150Asp	CDS: missense variant, moderate		SNP A											
IST419_05508	hypothetical protein		S	-	-	Ala10Gly	CDS: missense variant, moderate										SNP AB	SNP AB		
Intergenic																				
IST419_00140--- IST419_00141	BadF/BadG/BcrA/BcrD type ATPase --- cof family hydrolase	Carbohydrate transport and metabolism --- Function unknown	G/S	-	BMULJ_03109 --- Bmul_0156	IG: (193/161 bp upstream CDS), modifier											SNP B			
IST419_00140--- IST419_00141	BadF/BadG/BcrA/BcrD type ATPase --- cof family hydrolase	Carbohydrate transport and metabolism --- Function unknown	G/S	-	BMULJ_03109 --- Bmul_0156	IG: (241/113 bp upstream CDS), modifier				SNP B								SNP B		
IST419_00140--- IST419_00141	BadF/BadG/BcrA/BcrD type ATPase --- cof family hydrolase	Carbohydrate transport and metabolism --- Function unknown	G/S	-	BMULJ_03109 --- Bmul_0156	IG: (242/112 bp upstream CDS), modifier				SNP B								SNP B		
IST419_00866--- IST419_00867	phosphoesterase DHHA1 --- hypothetical protein	Carbohydrate transport and metabolism --- Function unknown	S/G	-	Bmul_1092 --- BMULJ_02164	IG: (87/183 bp upstream CDS), modifier													SNP AB	
IST419_01136- CHR_END	-	Function unknown	S	-	-	IG, modifier													SNP A	
IST419_01136- CHR_END	-	Function unknown	S	-	-	IG, modifier													SNP B	



IST419_03270-- IST419_03271	aldehyde dehydrogenase --- LysR family transcriptional regulator	Energy production and conversion --- Transcription	C/K	-	-		IG: (102/282 bp upstream CDS), modifier					SNP AB					
IST419_03270-- IST419_03271	aldehyde dehydrogenase --- LysR family transcriptional regulator	Energy production and conversion --- Transcription	C/K	-	-		IG: (104/280 bp upstream CDS), modifier					SNP A					
IST419_03270-- IST419_03271	aldehyde dehydrogenase --- LysR family transcriptional regulator	Energy production and conversion --- Transcription	C/K	-	-		IG: (108/276 bp upstream CDS), modifier					SNP A					
IST419_03270-- IST419_03271	aldehyde dehydrogenase --- LysR family transcriptional regulator	Energy production and conversion --- Transcription	C/K	-	-		IG: (120/264 bp upstream CDS), modifier					SNP A					
IST419_03270-- IST419_03271	aldehyde dehydrogenase --- LysR family transcriptional regulator	Energy production and conversion --- Transcription	C/K	-	-		IG: (121/263 bp upstream CDS), modifier					SNP A					
IST419_03270-- IST419_03271	aldehyde dehydrogenase --- LysR family transcriptional regulator	Energy production and conversion --- Transcription	C/K	-	-		IG: (127/257 bp upstream CDS), modifier					SNP A					
IST419_03286-- IST419_03287	activator of Hsp90 ATPase 1 family protein --- N-acetyltransferase GCN5	Function unknown --- Transcription	S/K	-	BMULJ_05205 --- Bmul_3322		IG: (313/187 bp upstream CDS), modifier			SNP A							
IST419_03294-- IST419_03295	17 kDa surface antigen --- hypothetical protein / allantoin catabolism protein	Function unknown --- Function unknown	S/S	-	Bmul_3329 --- Bmul_3330		IG: (934/315 bp upstream CDS), modifier					SNP AB	SNP AB				
IST419_03470-- IST419_03471	hypothetical protein --- phosphate transporter	Inorganic ion transport and function unknown	S/P	-	Bmul_3431 --- Bmul_3432		IG: (600/950 bp upstream CDS), modifier										SNP A
IST419_03470-- IST419_03471	hypothetical protein --- phosphate transporter	Inorganic ion transport and function unknown	S/P	-	Bmul_3431 --- Bmul_3432		IG: (687/863 bp upstream CDS), modifier					SNP A					
IST419_03861-- IST419_03862	3-ketoacyl-ACP reductase --- short chain dehydrogenase	Lipid transport and metabolism --- Function	I/S	BCAM2001 --- BCAM2002	BMULJ_04714 --- BMULJ_04713		IG: (374/338 bp upstream CDS), modifier	SNP B	SNP B		SNP B						SNP B
IST419_03861-- IST419_03862	3-ketoacyl-ACP reductase --- short chain dehydrogenase	Lipid transport and metabolism --- Function	I/S	BCAM2001 --- BCAM2002	BMULJ_04714 --- BMULJ_04713		IG: (374/338 bp upstream CDS), modifier	SNP B			SNP B						
IST419_04463-- IST419_04464	Crp/FNR family transcriptional regulator --- acyl carrier protein	signal transduction mechanisms --- Lipid transport and metabolism	T/I	-	Bmul_4390 --- BMULJ_04116		IG: (102/68 bp upstream CDS), modifier					SNP AB	SNP B				
IST419_04561-- IST419_04562	putative integrase --- putative integrase	Replication, recombination and repair	L/L	BCAM1929 --- BCAL2755	-		IG: (24/3 bp upstream CDS), modifier					SNP B					
IST419_04561-- IST419_04562	putative integrase --- putative integrase	Replication, recombination and repair	L/L	BCAM1929 --- BCAL2755	-		IG: (25/2 bp upstream CDS), modifier			SNP B		SNP B					
IST419_04769-- IST419_04770	hypothetical protein --- hypothetical protein	Function unknown --- Function unknown	S/S	-	-		IG: (65/110 bp upstream CDS), modifier					SNP B					
IST419_04864-- IST419_04865	LysR family transcriptional regulator --- transposase IS3/IS911 family protein	Replication, recombination and repair	K/L	-	-		IG: (126/15 bp upstream CDS), modifier										SNP AB
IST419_04909-- IST419_04910	Low affinity iron permease --- FRG domain-containing protein	Function unknown --- Function unknown	S/S	-	Bmul_5514 --- BMULJ_05983		IG: (219/141 bp upstream CDS), modifier										SNP A
IST419_04909-- IST419_04910	Low affinity iron permease --- FRG domain-containing protein	Function unknown --- Function unknown	S/S	-	Bmul_5514 --- BMULJ_05983		IG: (226/134 bp upstream CDS), modifier			SNP A							
IST419_04909-- IST419_04910	Low affinity iron permease --- FRG domain-containing protein	Function unknown --- Function unknown	S/S	-	Bmul_5514 --- BMULJ_05983		IG: (240/120 bp upstream CDS), modifier					SNP AB					



IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain	Replication, recombination and repair	L	BCAL1199	-	Arg305Arg	CDS: synonymous variant, low	SNP B			SNP B	SNP B							
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	Ala304Ala	CDS: synonymous variant, low	SNP B				SNP B	SNP B						
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	Gln275Gln	CDS: synonymous variant, low	SNP B				SNP B							
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	Thr138Thr	CDS: synonymous variant, low				SNP B							SNP B	
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	Pro79Pro	CDS: synonymous variant, low											SNP B	SNP AB
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	Ala65Ala	CDS: synonymous variant, low											SNP B	SNP AB
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	Gly63Gly	CDS: synonymous variant, low											SNP B	
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	Thr62Thr	CDS: synonymous variant, low											SNP B	
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	Leu55Leu	CDS: synonymous variant, low												SNP B
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	Leu48Leu	CDS: synonymous variant, low												SNP B
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	Asp28Asp	CDS: synonymous variant, low												SNP B
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	Glu23Glu	CDS: synonymous variant, low												SNP B
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	Pro14Pro	CDS: synonymous variant, low												SNP B
IST419_00350	putative phage-related protein / Y qaj-like viral recombinase domain		L	BCAL1199	-	His5His	CDS: synonymous variant, low	SNP B											SNP B
IST419_00352	hypothetical protein		Function unknown	S	BCAL2507	-	Asn174Asn	CDS: synonymous variant, low				SNP AB	SNP AB	SNP AB					
IST419_00352	hypothetical protein	S		BCAL2507	-	Gly173Gly	CDS: synonymous variant, low				SNP AB	SNP AB	SNP AB						
IST419_00352	hypothetical protein	S		BCAL2507	-	Leu168Leu	CDS: synonymous variant, low				SNP AB	SNP AB	SNP AB						
IST419_00352	hypothetical protein	S		BCAL2507	-	Thr125Thr	CDS: synonymous variant, low				SNP B		SNP B						
IST419_00352	hypothetical protein	S		BCAL2507	-	Gln106Gln	CDS: synonymous variant, low				SNP AB	SNP AB	SNP AB						
IST419_00352	hypothetical protein	S		BCAL2507	-	Phe105Phe	CDS: synonymous variant, low				SNP AB	SNP AB	SNP AB						
IST419_01281	helix-hairpin-helix DNA-binding motif-containing protein	Replication, recombination and repair	L	-	Bmul_1428	Gly70Gly	CDS: synonymous variant, low											SNP AB	SNP AB
IST419_01652	electron-transferring-flavoprotein dehydrogenase	Energy production and conversion	C	-	Bmul_1788	Glu186Glu	CDS: synonymous variant, low												SNP B
IST419_01652	electron-transferring-flavoprotein dehydrogenase		C	-	Bmul_1788	Leu202Leu	CDS: synonymous variant, low	SNP B				SNP B							
IST419_01652	electron-transferring-flavoprotein dehydrogenase		C	-	Bmul_1788	Ile297Ile	CDS: synonymous variant, low	SNP B				SNP B							
IST419_01652	electron-transferring-flavoprotein dehydrogenase		C	-	Bmul_1788	Ala299Ala	CDS: synonymous variant, low	SNP B											



## 8.3.2. Data related to chapter III

**Table S3.1** | Fasta file of the OAg cluster**>IST439\_01731 3-deoxy-D-manno-octulosonic-acid transferase**

ATGCTGAGGGCGATCTATCGCGCTGTGGTGCTCGTCGCGCCGGCCGCGGTATCCGGCTCTACGTGCGCTCGCGCAAGGAGCGGGTTATCGCGAGCATATCGG  
 CGAGCGCTTCGGTCATGTCGCGGGCCGCTCGCGGACGACCCGCGCCGCTGATCTGGGTGCATGCGGTGTCGGTCGCGGAGACGCGCCGCGCAGCCGCTGATCG  
 ATGCGCTGATGCGCGCGCGCCCCGATGCGCGCATCTGCTACGCACATGACGCCGAGCGGCCGCGGACCCGGCAACAGATCTTCGGCGATCGCGTGTGCGCTGC  
 TACCTGCCGTACGACATGCCGGGCGGGTGCGGCGCTTCTGCGTGCATGGCGGCCGACGCTCGGCCTCGTGATGGAACCGAGGTGTGGCCGACGCTGATCGACGA  
 GTGTCGCCGCGCGGACGTGCCGCTGGTGTGACCAATGCGCGGATGTCCGCGGCTTGGTTCGGGCGCGCGGGCAAGTTTCGGCGCGGCGACGCGCGACGTGTTCCGGC  
 GCTTCTCGCGCGTGTCTCGCGCAGAGCCCCGGCCGATGCGGAGCGGCTGACGTCGCTCGGCGCCCCGAACGTGACCGTGTCTCGGCAACCTGAAGTTCGACATGACGACG  
 CCGCCGAACTCGCGGCACGCGGCCATGCGTGGCGCGACGCGATCGGCGCGCGGCCGGTGTGGGTCGCGCCAGCACGCGCGAGAACGAAGAGGGCGCTGGTGTTC  
 AGCGTTCGCGGAAATGCGTACGCCCCGGCGCGTGTGGTGTCTCGTCCGCGTCATCCGCGAGCGCTTCGCCGAGGTTCGAGGCGCTCGTCGCGCGCGGGCGGGTCAAG  
 TGCGTGCGGCGCTCGGTGTGGGCGGCCGATGCGGCCGCGCTCGCAGCGGGCCGGCCGGCGGCCGACCCGCTGCCGACGACGTCACGGTGTGCTCGGCGATTTCGAT  
 GGGCGAGCTCGGCGCGTACTACGCGCGGCCGACATCGCGTTCATCGGCGGCAGCCTGTGCGCTCGGCGGGCAGAACCTGATCGAAGCGTGCGCGGTCCGGCGTG  
 CGGTGCTGATCGGGCCGACGTGTTCAACTTCACGCAGGCGACCCGCGATGCGGTTCGCGGCCGGTTCGCGCGATGCAGGTTCGAGGATCCGCTCGATCTCGCACACGTG  
 CTCGACGCGCTGTTCCGCCACAACGCGCGGCCGATCGCGATGGGCGCGGCCGGCGCGGCTTCGATCGCGTACCCGCGCGCGACCCGCGCGCACGGTTCGACGTGT  
 CGCCGCGCTGTTGCCGCTGCCGCAACCGGCGCGCGTGCCTGCGCTGCCGGACGTGCAGGACGCTCCGGGCGACGACGCATAG

**>IST439\_01732 hypothetical protein**

ATGAGCGAATCCCAGATCATCGAAGTCCCGTCCGCCACTGGAGCGGACACAACCTGTGACGCGCCGCGGAGCAGTTGCTGGCCGCGGTTCGAGGAAGGCAAGGTGC  
 TGTATTTCCCGCACCTGCGCTTCGCGATCGAAGGCGGCGAGGAAGCGCTGCTCGATCCGCGCTCGCCGATCCGAAACGCAAGAATCAGCCTCGCGCCGAACGGC  
 GCGCGCTCGCCGGCGTGTGCGGACAGCGTCACGCAGTCGGCCGTGCGCGCGCTCGTCGCGCGCTTCCAGCAGCAGGCCGGCACGCTCGTCGACGGCTCTTTCC  
 GAATACCGCGGAAGTGCAGTGCAGCGCGCCGACGAGCCTGCGGCTGATGCAGGTGAAACGCGCCAGACGTCGTGGCGCAAGGACGACAGCCGGCTGCACGTGACG  
 CGTTCCTCGTCCGCGCCGAACACTACGGCGAGCGCATCTGCGCGTGTTCACGAACGTGAACCCGGCCGGCGCGCCGCGCGTGTGGCGTGTGCGGAGCCGTTTCGAGGAC  
 GTCGCGAAGCGCTTCTGCCGAAGATCCGGCCGAGTTCGCCGGCTCGGCGTGGTGTGTAACCTGTGACGTCGACGAAATCGCCGCGCAGCGCATAACGACCATCT  
 GATGCTGAACCTGCACGACGGGATGAAGGCCGACCTCGACTACCAGAAGACGTGTCCGCGAGCAACGATGCCGTTTCCGCCGGGCGAGCGTGTGGATATGTTTCTCCG  
 ATCAGACTTCGACGCTGTGATGTCGGGCCAGTTCATGCTCGAGCAGACCTTCTTCTGCGGTCGACGCGATGGTTCGCCCGCAATGCGCGCCGCTCGGCATTCTCG  
 AACGCCTGACGGGCAGGGCGCTGGTTGA

**>IST439\_01733 lipopolysaccharide heptosyltransferase I**

GTGCAAAAGATCCTGATCGTGTGCGTGTGTCGTCGCTGGGCGACGTCGTGCACAACATGCCGGTATCGCCGATATCCGGCGCCGCCACCCCGATGCGCAGATCGACTG  
 GCTCGTCGAGGAAAGCTTCGTCGACCTCGTACGGCTCGTCGACGGCGTGCGCAACGTGCTGCCGTTCTCGCTGCGCCGCTGGCGCAAGAAGCCGTTCTCGGGGGCCAC  
 GTGGCGGAGATCCGCGCGTTCGCCCGGCGCTCGCGGCCGAGCAGTACGACCTCGTGATCGACTGCCAGGGCCTCATCAAGACCCGCTGGGTTCGCGAGCTGGGCCC  
 GCGGCCGCTCGTCGGGCTCGGCAACCCGACCCGACGGCGCCGGCTACGAGTGGCCGGTGCCTTCTTCTACCGCAAGCGCGTGCAGTTCGCGCCGCGCACGCACGTC  
 GTCGAGCGCTCGCGCCAGTTGGTTCGCGGCCGCGCTGGACGACCCGGCGCCAACCCCGGCCGACCCGTCGAGTTCGGCCTCGACACGCGGGGCGGACGCGCTCGCGGT  
 GGCCGCGCTCGGCTGAACTTGCCTGCGGTGCGTACGTCGTTGTCACCGCGACGTCGCGCGCCGACAAGCAGTGGCCGGACGCCGCGTGGATCGAGCTCGGCCAGG  
 CGTGTGCGCGCGCGCGCTCGCTCGCTGCGGTGGGCAACGACGCGGAACCGCGGACAGCGAGCGGCTCGCGAAGGAATTCGGTTCGCGCGCGATCGTGCC  
 GCCGAAGCTGTGCTGCCGGCCGTTGGTTCGGGCTGATCGACGGCGCGGCCGCGACGGTTCGGGTTGATACAGGTCTGGTTCACATCGCGGCCGCGCTGAAGCGCCCGA  
 CGGTCGAACTGTACAATTTCCGCGACGGCCTGGCGGACCCGGCGGCTACTGGTTCGCCGAACGTGTCATCTCGGCACGGCGGGGAGCCCCGTCGATCGCGCAG  
 GTGAAGTCGGCGCTCGCGGGCTTCGGCCTCCTGTA

**>IST439\_01734 phosphomannomutase**

ATGATCTCCCAATCCATCTTCAAGGCATATGACATTCGTGGCGTGGTTCGGCAAGACCCTCGACGTCGACACGGCGCGGGGATCGGCCGGGCATTTCGGCAGCGAAGT  
GCGTGCAGGGCGGGGATGCGGTGCTGTCGCGCGGACGGCCGCTGTCCGGACCGGAGCTGGTTCGGCGCGCTGGCCGACGGCCTGCGTGCGGCCGGGGTTCGAC  
GTCGTGACGTCGGCATGGTTCGGACGCCGGTTCGGTTATTTTCGCGGCGAACGTGCCGCTCGCGCTGAAGGGCGGGAGCGCCGCGTTCGATTTCGTGCATCGTTCGAC  
GGGACGCCACAACCCGCCCCGACTACAACGGCTTCAAGATGGTGTGTCGCGGGCGCCGCAATCTACGGCGAGCAGATCCAGGGCGTGTACCGCCGCATCGTCGACGAGC  
GGTTCGAGACGGGACGGCACGTATGAAGCAATCGACGTCGCGGATCAGTACATCGCGCGCATCGTCGGCGACGTGAAGCTCGCGCGCCGCTGAAGCTCGTTCGTCGTC  
GACGCCGGCAACGGCGTCGCGGGCCCGCTCGCGACGCGCCTGTTCAAGGGCGCTCGGCTGCGAACTCGTCGAGCGTTACCGACATCGACGGCACGGCATTCCCGAATCA  
CCACCCGATCCCGCCACCCGAAAAACCTGCAGGACGTGATCCAGGGCGTGAAGGACCCGATGCGGAGCTCGGTTTCGCGTTCGACGGCGACGGCGACCGTCTCG  
GCGTCGTCACGAAGGACGGCCAGATCATCTATCCGGACCGCCAGCTGATGCTGTTTCGCGAAAGAGTGTGTCGCGCAATCCGGGCGCCAGATCATCTACGACGTG  
AAGTGCACGCGCCACCTCGCGCAGTGGGTGAAGTGAAGGGCGGGAGCCGCTGATGTGAAGACCGGCCATTTCGTCGTGAAGGCGAAGCTGCGCGAGACGGGG  
CACCGCTGGCCGGCGAAATGAGCGGCCACGTGTTCTCAAGGATCGCTGGTACGGCTTCGACGACGGCCTGTACACGGGCGCGCGCCTGCTGGAAATTCTCGCGAAG  
ACGGCCGATCCGAGCGCGTGTCAACAGCCTGCCGGACGCGATGAGCACGCCGAACTGCAGCTCTGGCTCGACGAAGGGCGAGAATTTCCGCTGATCGAGCAGTT  
GCAGAAGGAAGCGAAGTTCGACGGCGCCGACGAAGTCGTGACGATCGACGGCCTGCGTGTGAGTACCCGGACGGCTTCGGCCTCGCGGTTTCGTCGAACACGACG  
CCGTCGTGATGCGCTTCGAAGCCGAGACGCAGGAAGGGCTCAAGCGCATCCAGGAGGACTTCCGCCGCGTGTGACGGCCGCGAAGCCGGACGTCAAGCTGC  
CGTTCTGA

**>IST439\_01735 membrane protein**

ATGCTGAAGCGCTTCGGCAACCCGGACGTCGCGAAGGCCGTCGCGAACCTCGTCTGGCTGGGGCTCGAACGGCTCACCCAGATCGGGCTCGCGATCGCGATCAGCGG  
CCTGCTGGCCCGTTATTTTCGGGCCAGATGTGTTTCGGCAAATGGCAGTATGCGAATACGCTTCTCTGGTACTGGCGCCGCTCACCTGGGTGTGCGGCGCGGAAATCCT  
CGTCCCACCATCGTCCAGCGCCCGCCGCGCAGCTCGGCGCGGTGCTCGGCAGCGCGTTTCGTGCTGCGCATCGCCGTGTCGGCCCGCGCTCGCCGCGACCTGGAT  
CGCGATCGCCGCGGGCGCCTTCGACCCGCTGGTTCGGCGCGATGCTCGCGGGCCTCGCGGTACGATGGTGTTCGCGAGCCGTTTCGTGCGCGTGTATCAACGCGTGGT  
GCAGAGCATGACCTACAGCAAGCCGACGTCGTCACCAGCATGGTCAACCGCTCGCAAGGGCTGCTCGTCTGGTGTGGTTCGCGCGGGCCGCGCCCGCCCGCC  
GTTTCGCTGGCTGGGGCGCTGGAAGCCCGCCGATCGGCTGCTGCTACTACCGCCATCGTAACGGCGGTGCGCTCGGCTGGACAGTTTCGACAAGCCGC  
TGTTCCGGCACTTTGCAACGGCCGGCACCGTGTTCGGCTCGGCCTCATCTGCATGTACCTGTTCTGAAGCTCGACCCGCTGATGCTCGAGCGCCACGTGTGCTTCGC  
CGATCTCGGCCGCTATTCGGCCGCGCAGCAGCTCAACGAGAAGTGGATCACGCTCGCGCTGATGCTCGCGCAGACGATCGCGCCCGCCTTCGTCTACCGCGTGCAGG  
ACGTCGCGCGGCTGCGCCGAACATCGTCCGGCTGATCGCGATGACGGCCGGCCTGATGACGGCCGGCGCGCTCGTGTGTCGACCGCCCGCCCGCTGATCGTTCGGC  
AAGGTGTTTCGGCCGCGGCTACGAGGCGTGGTTCGACATCTTCGCTGGGCCGCTGGCTGTCCGTGCCGGCCGGCATCGAGGCGATAGGCAATCTTATCGTTCCTCAA  
TATCAAGCAAATTCGTGTTGCTGTGAAATGGCTGCTCGCGCTGGCCATCGCCGCGATCGTCAACCTGTTTCGCGATCCCGCGGCTCGGCCTGTACGGCGCGCTCGTC  
GGGCTGGCGGCGGGTACCTGGCCGCGGCCCGCTCAACTTTTATTACATCCGTTTCAAGCTGCGAACATGA

**>IST439\_01736 glycosyl transferase family protein**

ATGACGTCCCCTGCTTGCCCGACCCCGCTCGACGACGTGGCCGCTGCTGATGCCGGCCTACAACGGGCACGACGACGTCGTCCGGACCCTCGCGTCGTTTCGCGAGGAC  
GCGCCCGTGCACGTGCTGATCGTCGACGACGGCAGCAGCCGCCGATCGTCGCGCCCGATCTGCCGGCCTGTGATCGACGTGCTGCGCATGCCCCGGAACGGCGG  
CATCGAACGCGCGCTGGCGGCCGGCATCGACGCGCTCGCGGGCGCGGGCTTCCGCTACGCGGCCCGTATCGACGCGGGCGATCTCGCCGCGCCGACGCGCTCGCGA  
AGGAGCGCGCCTATCTCGGCGCCACCCGCGCTAGCCTGCGTCGGCATGTGGACGCAGGTCGTGTCGCGCCGGCGAGCCGCGCTTCATGCTGACGCGCGCCGCC  
GATCCGCGCACGCTGCGCCGACGCGCTTCCTGCGCTCGCCGCTCGTGCATCCGTTGTGATGTTGCGCATCGACCCGTGCGCGAAGTGGGCAACTATCGGGCAAAG  
TACCGCGCGGCCGAGGATCTCGATCTTTTTTACGGTTAATGCAACGCTACGATTGCGCGAACCTGCCGAACTCGGCCTGTATTACGAGCTTAACGAGGGCGGGATC  
AGCGCGACCAAACGCCGGCGCCAGCTGGTGTGACGCTCACGCTGCTGCTGCGCCACTTCAACGTGCTGAACCCGTACGACTGGGCGGGCCTCGCCAAGAACCTGCT  
GCATTCGTGACGCCGTACCGTACGCTGCAGCGCATCAAGCAGACGCTGTTTCGCGCGCGGGCCGCTCCGCTGA

**>IST439\_01737 group 1 glycosyl transferase**

ATGTCGGCCACCTCCCGCTGCGCATCGCGCTCGTCTGCAACACGGCCTGGGCGATCCATACGTATCGCCACGGGCTGATCCGTGCGCTCGTCGCGCGGGCGCCGAG  
GTCGTGCTGATCGCGCCGACGATCGCACCGTGGCCTGCTGAGCAGATGGGCTGCCGCTACGTGGCGCTCGCGGTCGCGTGAAGGACGAGCCCGCGCGAGGA  
TCTCGGCACGCTCGCCGCGCTGGTGCGCCACTATCGCGCGCTGCGGCCGACCTCGTGTTCATTACACGATCAAGCCGAACATCTACGGCTCGGTGGCCGCTGGCT  
CGCGCGCTGCCGTCGATCGCGGTGACGACGGCCCTCGGCTACGTGTTTCATCCAGAAGAGCCGCGCGGCGAGCGTCGCCAAGCGCCTGTACCGGTTTCGCTTCGCT  
TCCCGCGTGAAGTGTGGTTTCCTGAACCGCGACGATCTCGCGACCTTACCCGACGAGCGCTTGTGTCGCGCACCCCGATCGCGCACGGCTGCTGCACGGCGAAGGCGTC  
GATCTCGAACAGTTCGCGCCCGTGGCCTGCCCCGGCGACGCGCCCGTGTTCATCTGATCGGCCGGCTGCTGTGGGACAAGGGCGTGCGCGAATACGTGGAGGC

CGCGCGTCTCGTGCAGCGCTTCCCGAACGCGCGCTTCCAGTTGCTCGGGCCGCTCGGCGTCGACAATCCGAGCGCGATCGGCCGCGCGGATGTCGACGCGTGGG  
 TCGACGAAGGCATCGTCAATATCTCGGCGAAGCGCACGACGTGCGACCCGACATCGCGGCGGCCGACTGCGTCTGTGCCGTCGTACCCGGAAGGCGTCCGCGC  
 ACGCTGATGGAAGCATCGGCGATGGGCCGCCGATCGTTCGCGACCGACGTGCCGGGTGCCGCGACGTCGTCGCCGACGGCGAAACAGGTTACCTGTGCCGTGTGG  
 CGACAGCGCGAGCCTCGCGGAGCAATTGACTCGCATGATCGAACTCGGAACGGCCGACGCGAAGCGATGGGCGCGCGGCCGCGAAGGTCACGGCGGAATT  
 GACGAACAGCAGGTCGTCGAACGCTACCGGCGTACCATCCACTCGTTAACCGGTATCACCTCTGA

**>IST439\_01738 UDP-glucose 4-epimerase**

ATGACCGCTAAAGGCACCATCTCTGTACCGGCGGAGCCGGCTATATCGGCTCGCACACGACCGTCTGAGCTGCTCGACAACGGCTACGACGTCGTGATCGTCGACAA  
 CCTCGTCAACAGCAAGGCCGAATCGGTGCGCCGGATCGAGCGGATCACGGGCAAGACGCCCGGTTCCACCAGGTGACGTGTGCGACGAAGCCGCGCTCGCGAAG  
 GTGTTTCGACGCGCATCCGATCACCGGCACGATTCACCTTCGCGGCGCTCAAGGCCGTCGGCGAATCGGTTCGAGAAGCCGCTCGAGTACTACCAGAACAACATCGGCGG  
 GCTGCTCGCCGTGCTCAAGGTCATGCGCGAGCGCAACGTCAGGCAGTTCGTGTTACGCTCGTCCGCGACCGTGTACGGCGTGCAGCGCTCGCCGATCGACGAAT  
 GTTCCCGCTGTCCGCGACGAACCCGTACGGCCAGTGAAGCTGATCGCCGAACAGATTCTGCGCGATCTCGAAGTCTCGGACCCGTCGTGGCGGATCGCGACGCTGC  
 GCTACTTCAACCCGGTCGGCGCGCACGCGAGCGGGTGTATCGGCGAGGATCCGGCCGGCATCCCGAACAACCTGATGCCGATGTGCGCGAAGTTCGCGGTTCGGAAG  
 CTGGAAGAGCTGCGCGTGTTCGGCTCGGACTACCCGACGCCGGACGGCACCCGGCGTGCAGTACTACATCCAGTCTGATCTCGCGAAGGCGCATCGCCGACT  
 CGACGCGCTCGCAAGCGCGACGCGAGCTTCCTGTAACCTCGGCACCGGGCAAGGCTACAGCGTGTTCGAAGTGGTTCGCGCGGTTTCGAGAAGGCGTCGGGTCTC  
 CGGTGCCGTACGAACCTCGTTCGCGCGCCGCCGGGCGACATCGCCGAGTGCTATGCGAACCCGACGGCCGCGGCCGACATCATCGGCTGGCGCGGACGCTCGGCATC  
 GAGGAAATGTGTGCCGACACTGGCGATGGCAGGAGGGGAACCCGCGCGGTTTTGTATAA

**>IST439\_01739 glycosyl transferase family protein**

ATGCTCAGTTTCGCTCCGGCTTATCGTCTCCCTGCTCGTACGCTGTTTCATCGTGCCTACGCGCACCTTACGAGAAATTCTCGATCGACAGCGACCTGGCCGGCG  
 TGCAGAAATTCATGTGCGGCCGGTGCCGCGGGTGGGCGGCATCGGGATCCTCGCCGGGGTTCGTCGTTGCCGCGCTGGTGTGTTCCGGCGCTATCCGACGATCTCCG  
 GCAGATTCTCGGGATCGTTGCCTGCGGGTCCCGGCTTCTGTCCGGCTCGTTCGAGGATCTGACCAAGCGGTTTCGCCGCGTGCAGCGGCTGCTCTGCACGAT  
 GCGCGCCGCGCTCGCATTCTGGTGTGATGCACATCGCGGTCACGCGTTCAGCTGCGTTCCTGCTCGGTCAGTTCATGATGTTTCGCGTTCGCTCGGATGTCGCGTT  
 GCTCGCGGTCGCCGCGCTCGCAATGCGATCAACATCATCGACGGCTTCAACGGGCTCGCGTTCGATGGTTCAGTTCATGATGTTTCGCGTTCGCTCGGATGTCGCGTT  
 CCACGTCAACGACCCGGTTCGTGATGTCCGCGTTCGATCATCATGATGGGCGCCGTGCTCGGCTTCTTCTGTGGAATTTCCCGGCCGGGCTGATCTTCTCGGCGACGG  
 CGGCGCGTATTTTCATCGGCTTTCATGCTCGCGGAACCTCGCGATCATGCTCGTGTGCGCAACCCGCAAGTGTCCGCGTGGTACCCGGTGTGCTGTTTCATGTATCCGAT  
 CTTCGAGACCTGCTTCTCGATCTACCGGAAGAAGTTCATTCGCGGGATGTGCCAGGGATTCGCGACGGCGTGCACCTGCACATGCTCGTGTACAAGCGGCTGATGCG  
 CTGGGCCGTCGGCACGAAGCATGCGCACGACCTGACGCGGCGCAATTCGCTGACTTACCTTATCTGTGGCTGCTGTGCCTCGTTCGCGGTGATACCCGGCGACGCTGTT  
 CTGGCGGCATACGGTGCACCTGTTTCGCGTTCGTCGTGCTGTTTCGCGGTGACTTATGTGTGGCTGTACATCAGTATCGTTCGGGTTCCGCGCACCCGCGATGGATGGTCTGA  
 CGCAAGCATCGGCACAGCCGCTGA

**>IST439\_01740 polysaccharide biosynthesis protein CapD**

ATGTTGCAATCCAGAGCATCTTGGCTGTCCCTGAGTGCTTTTCTGTTTCGACCTGACGGCGGTTGTCGCCGATGGTTGTTTCGCTATCTCGTTTCGTTTCAATGGCAGCGT  
 TCCGCATGATTTCTGAGCGGCGCGCTGATGGCGCTGACGTGGGTAAGTCCCGTCTACGCGTGTGTTCCACGGGTTTCGGCCTATATCGCGGGTGTGGGTGTTTCG  
 GAGCCTGCCGGACCTAATGCGTATCGCGAAAGCGGTGACCGGCGGCGGCATGATCGTGTGATCGGCGCCGTGATGTTCCAGCCGGCACCGATCATCCCCGCTCAG  
 TGCTGCTCGTGTCCGCGTGTGTTCTTGGCGATGGGTGGCGCGCGCGCGCTGTATCGCGCGACGAAGGAGTACTACCTGTACGGCGGACTCGTTCGCAAGGGC  
 AAACCGGTGCTGGTGTTCGCGCGCCGGCACGGCCGGCGCGAGCCTGGCACGCGAACTGTGCGGCTCCGGTGAATGGCGCCTCGTGGGCTGCTCGACGACGACGTCAC  
 GAAGCAGGGCCGTGAAATCTACGGCTACAAGGTGCTCGGCTCGTTCAACGACCTCAAGCACTGGACCGATGCGATGAAGGTGCAATACGCGATCATCGGATTCCTG  
 CGGATCGGTGCAACGCGAGCCGTCGCGACCTGTGCGTTCGCGGCGGCGTGAAGGCGATGGTGTGCTTTCGCTGACCGCGCTGATGCCGGGCGAGGGCTTC  
 CTGTGCGAGGTACGGAACATCGATCTCGAGGATCTGCTCGGCGCGAGGCCGTGACGATCGATACGCCCCAGTTCGAGGCGCTGCTGCGCGGCCGCGTGTGATGGT  
 GACGGGCGCGGGCGGCTCGATCGGCTCCGAGCTGTGCCGGCAGATCTTGCCTTCGCGCCTGCACAGCTCGTTGCGTTCGACCTCTCGGAATACGCGATGTACGGGCT  
 GACCGAGGAGTTGCGCGAGCGGTTTCCGGATTTGCCGTTGGTGGCGATCATCGGAGACGCGAAAGATTTCGCTGCTGCTCGATCAGGTGATGTGCGGCCATGCGCCGC  
 ACATCGTGTTCATGCGGCCGCTACAAGCACGTGCCGTTGATGGAAGAGCACAAACGATGGCAGGCGTGCGAACAACGTGCTCGGCACCTATCGCGTGGCGCGG  
 GCGGCGATCCGCCACGGCGTGCCTACTTCGTGCTGATCTCGACCGACAAGCGGTCAATCCGACCAACGTGATGGGCGCGAGCAAGCGTCTCGCCGAAATGGCCTG  
 TCAGGCGTTACAGCAGACGAGCGGGCGCACCCAGTTCGAGACGGTGCCTTCGGCAACGTGCTCGGCAGCGCGGGCAGCGTGATTCCGAAGTTTTCAGCAGCAGATCG  
 CGAAGGGCGGCCCGGTCACGGTACGCAACCCGGAGATCACGCGTTTCTTCATGACGATCCCGGAAGCGTTCGACGCTCGTGTGAGGCGTTCGAGCATGGGGCATGGC

GGCGAGATCTTCATTCTCGACATGGGCGAGCCGGTGAAGATCGTCGATCTCGCGTGGCAGCTGATCCGCCTGTACGGTTTCTCGGAAGAACAGATCCGGATCGAATTC  
 ACCGGGCTGCGGCCGGGGCGAGAAGCTTTACGAGGAAGTCTCGCGGACGACGAGGCGACGACGCGCACGCCGACCCGAAACTGCGGATCGCGCGGGCGCGCAAG  
 TGCCGGACAATTTCCCTCGACGAAGTCTGCCGTGGCTGATGCAGCATCGTGTGCTGCCGACGACGAAGTGCGGCGCGACCTGCGGGCGTGGGTGCCTGAGTACCAG  
 ACGGCCACCGTGCCAGCCCTGAAGAGTGTGAGAGTCGCATCGAACCCTGA

>IST439\_01741 glycosyl transferase family protein

ATGCATTTCCCGATCTCCACTTGGCCTGCCGCGTGGGGTGGCGCTGGTCCGCCCATCGCATCGACGGCCATACTGCGCATGCTGCTCGCGACCGGCTCGCATGG  
 CGGCTCGCTACCGATATCCGAACGACCGTTCCCTCCATACGTTGCCGACGCCGCGTGTCCGGCGATGGGGCATCGTGCCAGTGTCCGTTGCCACGCTGTTGATGCTG  
 GCACCGCAACTGTGGCTGGTCCGAGTCCGAGCAGCAGGGCTGGCCGCGATGTCGAGATTGACGACCGGCGCGGGCTGCCGGCGCGTGGGTTCTCGGCGCACCT  
 CGCGGCAGTCCGCTGATCGTTGATTTCCAGCCGACGCGCCCTGGTGGCTGCTCGCTGGCGTGGGTTTCGCGATGGTCTGGCTGACGAACCTGTACAACCTTCATG  
 GACGGTGCAGTGGGCTTGGCGGAGGATGGCGTTGTTCCGCTTCCGCGGGTATGCGGTCCGCGCGCTCGGCGGCGACAAGAGTCCCGGATCTCGTCCGTTCCGG  
 AGCGGCGGTGCCCGGCGCGGCTGGGCTTCTTCTGCTCAATTTCCACCCGGCGAGACTATTCCTTGGCGACGCCGGCTCGATTCCGCTCGGCTTCTGGCCGGTGC  
 GCTGGGCTACTGGGGTGGCGTACCGATGTCTGGCCGATCGGTTTCCGGCCATGGTGTCTCGCCCTTATTCTCGATGCATCTGTAACACTTCTGAGACGCTTGTA  
 CGCGGAGAAAAGTTCTGGCAGGCGCATCGCGAGCATTATTACAGAGGATGGTTCGATCGGGCGTACGCCACGGTCCGACCGCTCTTATTGGTACCTCATCATGCTC  
 GCCGCATAAATGTCCCGTGTGGGCAAAAGGCGCCCTGAATCGAGCAATGGCTGTGCTTCTCGCGTGGTATGGCGTCTGGCATGGTTCGGATGGTTGATCGAC  
 ATGCGTTGGCGCCGGGTTCAAGTGGCTGCCAAAACAATTCTTGA

>IST439\_01742 NAD-dependent epimerase/dehydratase

GTGAGTGATCTCGTCATCAGCGGCGGAACGGCTTCGTTGGCCGAGCGGTCTGCCGTCGCGCGCTGGAGGCCGGGCACACCGTACGGCGCTGGTCCGGCGACCGGG  
 CGGGTGCGTGCATGGCGTACGCGAATGGGTGCATGACGCAGCCGATTTCCGACGGCCTTGATGCGCGATGGCCGGCCGACCTGGCAGCCGATTGCGTGATCCATCTCG  
 CCGCTCGTGTGCATGTGATGCGCGACGAGTCCCGATCCCGATGCCGCGTTCGACGCTACCAATGTCGCCGGCACGCTGCGTCTCGCCGAGGCGTCCGCAATCACG  
 GCGTCCGCGCTTCGTTCCGCGAGCAGCATCAAGGATGTCCGCGAGGGCGATGGCGGTGTGCCGCTACGGAAGCCGTCGAACCCGTTCCGCGAGGACGCGTACCGG  
 CGACCAAGTACGCGGGAGCGGCAACTCCGCGAATTCGGCGCGTCCGCGGGTCCGACGTCGTCGTCGTCGCTCCGCGCTCGTGTACGGCCCGGCTCCGTCG  
 CAACTTCTGAGGATGATGGACGCGGTCCGCGCGGGGTGCCGTTGCCGTTCCGCGCGCTTCGCGCGCCGAGCGTCCGTCACGTCGACAATCTCGCCGATGCGCT  
 GCTGCGTTGCGCAATCGATCCACGCGCGCCGGCGAGTGTCTTCATATCGCCGACGACGACGCCCCGACGGTCCGCGGACCTACTGCGCCTCGTCCGCGACGCGCTT  
 GCCGCCCCGCGCGGCTAATTGCAGTACCTCCTGCACTACTGCGGGCGGCCGGGGCGCTGACAGGCCGCCGAGCGGTTATCGACCCGCTGACAGGGAGCTTGCAGCTC  
 GATACCGACCGGATAAGGCGCGTACTCGACTGGCATCCGCTTATACGCCCGGAAGGCCCTTGAAGCGACCGCCGCGTGGTATCGTTCGCGCGATAACAACAATA  
 G

>IST439\_01743 glycosyl transferase family protein

GTGCTTGACCGGATGAATGTAGCCTTCGACGTGATTACAGGCCACGGCAACGTCGGCTATGGCCGCGGCCATAATCTCGCGATCGAGACGGTATGCAGCCGTTACCA  
 CCTCGTGTGAATCCCGACATCGATCTCGCGCCGAATGCGCTTGCGAACGCGCTTGATTTTCTCGATGCCATCCAGAGGCCGGCCTGCTGACGCCCTACATTGGCGA  
 CGAAGCCGGCCGATCCAGTATCTGTGTCGTCGCTATCCCGCGATGCTCGATCTGTTCTGTCGAGGTTTTCTGCCGTCGGTTTTTACAAAATCCTTCGAGCGCCGGCTG  
 GCTCATTACGAAATGCGTGACGAGATCAACGACGCGGATATCGTCTGGGATCCACCGATCGTGAGCGGATGCTTTATGCTCTTCCGGACCGACCTGCTCAAGCGTCTG  
 AATGGTTTTCGATCCACGGTATTTTCTGTACTTCGAAGACTACGATCTCAGCTTCCGACCCGTGAGCTGGCGCGGATCGCATATGTGCCTGCGGTACGTGTGCTTACC  
 ACGGCGGTGGTGTGCCCGCAAGGATGGGCCACATCCGGATGTTCCGCGCGTCCGCTTCAAGTTCTACAACCGCTTCGGGTGGAGGCTGTGGTGA

>IST439\_01744 hypothetical protein

GTGGCAACCCCGCCCCGCTCACCTCGGTCCGAAGGCGGACTGGCACACGGGAAGATGCATGAAACTTGGCGGATGGCGCTCGCCGAGCAGGTGGATTGATTTGGAG  
 CATGAAAGAGGTGCGAGAGTGCGAACGGAAGTACGAGTACAGGACGACACAATGGACGTCTATAACGACACGCCCCAGGTCCGAGCGGCAACGGGCGGAATGG  
 TATGGAGGCGATCACGCAGCTCTGTGATACCGAACGAGCGACAGGACCGTTCCAGGGATCTGTCGCTTACGCTATATCTATTTACACGCCGACTTTATGACGCAGCACT  
 GAAGGACGGTGTAAGCGATCTGTTTTTCTTCGCCAGAGAAGGACTTTTGTGTAAGCAAATGTTCCGACCGGGTGAAGCCAAATGAGGACAAGACACCCATACGATCTC  
 ATTACCTAAAGGTGTCTCGTTCGACCTTTCTACTGTCACTGCGCCCGTTGGACCAGGAGCAGTTTGAAGTTCTCTTCCGACAGTACAGAAGGATCTCGATCGCCGA  
 GTTTCTGAAGAGCCTAGCGCTGGAAGAATATCTTCTGACGTTCCGCTACCGAGCTTAACGTCGAACCGTCCGCTGCTGGCCGAGCGCAGTGCAGATTTGCCGCACGATCC  
 ACTTTTCGTCGACTGCTGGACCTGCCGTCGTTGCGCAAGTGTACGAGAAGGAGCGACTTCCGCGCAGCAAGGCGTTCCGAAAATACGTGGGCAGTTTCTGCAACGG  
 CACCATCCCTGCGCGGTTGAGTGTGTTGACGTGGGATGGAAGGGGTCAATCCAGGACAACGATTTCCGGATGGATGCGTAATCTCCGGGGAGACGCGGCAAGCGTCA  
 AGGGCTACTACGTCGGCCTCATCGCAACCGCAATGCCGGCCCCCAACGAGAAGGAGGATTGTTATTTGAGAGCATCGGCGAGCGCACCGGGGCTACTACGTT

TTCAACGAAAACCGGTCCTCTTTGAAGTCTTGTGTCATGCCAACACGGGTCGGCACAGCGCTATATTCAGACATCCGCGAACGAGGTCATCGTCGTAGAAGACTCG  
TTTGAAGAAGGTCCGATGATCGAAGCGTTTGTAAACCGGTAGCCGATTGCATCATGGCCTCGTTCGTTTCGATCGCGGACGCGATGGCTACGCTTGATGTGTCCGAG  
CAAGACCTTCTGGCCTGCGCAGTCAAACGCCATTCACGCATGGTGTTCGGCCTTCATCAAGCGAAATCGAATGGATGCGCTCACTGTGCGACGTTGAAAATTCGGT  
GTCTTCGAGGCGTCGACGTTGACAATCGCCGGACCTCGCCTTCGTTGGGTTCCCGGCTAAGATTCTCGATCGATGATTCAGACACGGTCAACGCGACCTCGGTTTCT  
GGCCTTGGCTGACGATTGAAGAGCGCGCCCTCTGGGGAATGGCAGGTGTCTATCGGATCATTTCGCATGTGGCAGAACCAACGGACCATGCGTCTGCACTCGAGCCCA  
ACGGACAAAAAGCGAGTACCTAAATGA

**>IST439\_01745 group 1 glycosyl transferase**

ATGAAAAGTGTGTTTTCTTGTGGGGTCCATGGCGATCAGCGGTGGAACCTACGTCATTGTTACGACGCAAGCTATTTGCGTGACCACGGCTACGATGTCCTCGCC  
GTTCAAGAACCCTTCACGCCCCGAGACGCTGGCGTGGCACAACGAGGCGCCTCGGCTGCGTTGCATTCCCATTGCGACGGCCCGTCCGAGAAAATACGACCTCGTGATC  
GCCACATGGTGGAAAACGGCGCTCGAGCTTCATGAATTCGATGCTGCCCGTTACGCGTACTTTGTCCAGTCGATCGAGTCCCGATTCTATCCGGAAGCGGAGATCCCG  
TTACGCGCGCTCGTCGACGCGACCTACAAGCTCCAGTGACTTATGTGACGGAAGCGACGTGGATCAAGGATTACTTGCAAAGCCATTTCCGGCAGGAAGCGTCACT  
GGTCAAGAACGGTATCCGGAAGGATGTCTATACCCCGACCGGGCCAGCCGTCGAAGCAAGGAGTCCGACTCGTCTCCGCGTATTGGTAGAAGGACATTTCCGAGTAC  
CGTTCAAAAATACCGCGCTGGCAATTCGTCTCGCGAAGGGAAGCAGGTGCAAGCGAAATCTGGCTGCTGACTGGCTCACCCGTGCAATGGATTCCCGGTGTGTCCCGA  
GTATTTCCCGAGTACCTATGGTCAAGACGGCGGAAAATCTATCGGTCTGTGACGTGCTACTGAAGCTGAGCACGGTTCGAAGGAATGTTCCGACCCGCTCTTGAAAATG  
TTCCACTGTGGCGGAACAGCAGTCTTTCGACGTCACCGGGCAGGATGAATACATCGTCGACAACGAAAATTCGCGGGTAGTATCGACAAAACGATCTGGACGGAGT  
TGTGAAAACGATTCCGACGCTCCTGAACGACCGAAGCGAACTCGCTCGGCTTAAGGCAGGCGCCTTGCGCACAGCTCAAGCATGGCCCCAATGGGAAGACGCATCTG  
CGAAATTCAGACTGGATTCCGCGATTGCTTGCAAGGGCCGTCGTGAATCGAGATGAGTGTGACGACGTCATCGCAAAGGCCTGGAGCGACTATGACCGTGACGAA  
AAGCTGCGTCTCGCACAACTCGCCAGTCCGTCAGCCGATGAGGCTGCGTGCCTCGCCGAAAATGCCAGGCTCGTGACACGACGTCTCAAACAGCTCGAAGC  
AGTGAGCGAAGTGTGGTTGGCAAGCGTAAGGCCTATTGA

**>IST439\_01746 HAD superfamily hydrolase-like protein**

ATGTACAGACTGAAGACCATAGACGTGTGGAGACAGCTGCTCCGGCGCGATTGCCACCCGGAATTTGCAAACTCGCAACGGCGTTGCACGTGTTTCTCGATCCGTCG  
ATCGGTGTTTTCTGCCACTATGAAACGGCGTGGGACATCTATTGTGCGAGAGTCGAGCAAGAAATGCTGCTGGCCGATCGCGCGCGCGCAGCGGGAAAAGACAACGA  
GTACGAGATCGTCGAGGTAAGTACTGCGGCAATGGCTTACGCAGATCGTCGATCGACCCTACGACGACTCCCTGCCGTCCCGACTCGCCGATTTGAACTGGGCTTCGAACT  
TCGACACACCTATCCTGACCCAACGATCAGGGATGTATCCGAGCAGTATCCTGCGGAAACGACGCTGTTCTGTGCGACTTTTATATGTCCGCGGATCGCCTTTCCAC  
GCTGCTGCATCACCACGGAATCGACACGTTTGTGCCGGGAGGCTTGTGTTCTGCGACGTCGGTATCAACAAACGCTCGGGCAAGTTTTTACTACGTCGAAGATCT  
CTATCGAGTAATGCCGAATGAGCATGTTACATCGGCGACAATTTGACGTCCGATCGCGATGTGCCTCACGCGCTTGGTATCAATGCAGTCCATTTTCAACCGGAAGC  
AGAGCATGCGAAGCGACTGGCTCGCGAATCCGTCTTCGAGGATAGAGGCGAGTTGTTTCGCAAGATTCAAGGCAGAAGTAAAGGTAGCAGCAAGTGGCTACCAATTTG  
CCCATGAGCAGTCCGCTGAACCGGAAGCCGCTTACGCGCAGGGATTCAAGTCGAGCGTGGTGTGATTGGCTTCTGTCTCTACGTTGCCGAAAATGTGCTGTCTGGT  
GGCTCAAGAGTCTGTACTTCTTTTACGTGAAGGCGAATCTTTGTACGTGTGTGGAATGCACTATTTCCGAGCGGACAGCATGCCGGTCACACACTTCCGACGCGCTC  
TCTACTCGAAGTGAGTCGCTTTTCGACGTTCTGTGCGTCACTGCGGGAGGTCACCCCGTCAGAGATGCGCCGAATCTGGAATCTATATACAACACATTCTCTCGAAGC  
ACTCTTGAAGTGCCTGGGGCTTGAACCTCGCTCGTTCGCGGAACTGGCAGCGAAGCATCAGTTGCCGTTGACTGAAACGATCCAGCATCCATGGGAAGACGGCCGTG  
TCAAACGCCTCTTCGACGATCCGCAGTTTGTTCATCAAATCACGGAGAAGATTCACGCCGATCGTCGTGCATTGCTAGCTTACTTGGCGACACAAGGGTTACGTGATG  
ATGGATCGGTAGTGAGCGTCGTCGATATTGGTTGGCGCGGCACCATTCAGATAATCTTGATGGCTTCTGCCCTCCGTACAATCCAGGGGCAATATCTGGGATTGC  
AAATGTTCTGAACGAGCAGCCTGCGAATACGCACAAGCAAGCGTTCCGACCCAATGCAAAATGAAACACTGGAACACTGCCATCTTCTCGACGCCGTGTCTCCGCTC  
GAGATGTTGTTTCACTTCACTTCTCGGTTTCACTTCCGGTATCCGCATGAACGATGACGGCACCGGTGTCAGTTGCATTTTCGCTCGGAAAGCGATATGGAACACGCTTCC  
AGCAAAAATTTCTGCGAAAGCTTCCAGGATGGCGTGGTTGAAGCCGCGAAGGTGTGGGCCCGTACGTCGAAACATGTGCGATTGCTGCGTCAGAACTTCCGATCC  
GGCGTTGGCAACTGACAGGTTGATCGAAGCTCCACACGAAACCTTGGTCGACGCATACTCGGCACTCAATCAGAATGACGTGTTCCGGTCCCGGTCAAGTTCGTC  
GCAAACGTGCCACCCCGGGAGTAGGCAAGTTGTTGTGGGGAATCGTTTCGCGCGATGCGCGTCGGGAAGTGATCATGTACGTTTCGTCAAACCCAATGGACCGCCGGA  
ATCTGGGCGCGTCGCGACTTGACATTGGGAAAATCGGCTGGCGCTAGTGATGGTGTTCGATTGGCCAAAGGCTACAAGCGATTTATGCAGCGAAGAACCGCCAAAA  
TCGCACCCACAGTCTAGCCAAGGAGCGGCCAGGTCGTCGGACGGTCTGATCAATGA

**>IST439\_01748 type 11 methyltransferase**

ATGGAACGCATGCATTCCGATCATTTCAGATGCAGCACCTCGTATATGAGCATAACGCATCGCTATCAATTGCTTTCGCGTATCGCTTACGGGCATGTGTCGACTGCG  
CATGCCGATATCGGTTACGCATCTGAAATTATGGCGGCTGCCGGGTCCGGTCTACCACGGTTTCGATATTGATGAGCAGTCGGTTCAACTGGCGGTCCAACGATTTG

CGGCTGACGGACGTCAGTTTACGGTCGGCTCGATTCTGTGCTGCCGCTTGAGGATGCGTCGGTCGATACTTCTTGTGCTGGAAACACTTGAACATTTGCGGGAGC  
CTCAGCGCGGATTACAGGATTACCCGTGTGCTGAAGGCGGACGGGATCTTCATGGGCTCGGTGCCGACGCGTTCTTACGATGAGCGTTGTACGAGTTTACGGCC  
CGAACCCCTATCATGTGACGCGGTTTCGACGACGAGATCGTTCAACTGCTGCGAACCGGGTTTGTAGTACGATGGCTTGGGGTCATCAGCCTCGACGTAGTGTGCA  
TCGTCCGGCCGTGTCGGGCGCGCAATGCAGGTCCAGCGGAATGGGTGCGCAACAGCAAGTGGGACGCGACAAAGTATGGTTTCGTTTCATCTTGTGCTGCGTACATCAG  
CCTGTTGACGAACTGGTCCGCAAGCTGCTTCCGCAGCAATGCGCGGTGCTGCATGGGTTGCCTACTGTTGAAGTGGACGACAGAGCAGATATGCCGCTCTACAAGAC  
AATCACGGAATCGAGGGTCTTGTGAGGGATCGCGACGCCGATTGGATCGATCGGAAGGAATGATCCGGGACCGCGACGCTGCGTTGCACCGGATGGAAGAGATG  
ATTCCGAGCCGCGACGCGCAATTCGAAGCAATGAGCAGATGATCAATGACCGTGACGCAATTGATCGCGCAACTCGAACGAATGGTTCGCTGACCGGGACGCGGCGA  
TCGCATCGAATGAGAAGTTGATCAACGAGCGCGACGCGTTGCTGCGAGCTAAACGCACAGACTGA

>IST439\_01749 ABC transporter-like protein

ATGTCCTCTAATGCGGTCATTTGCGTTACCAGTCTTGGCAAGTGCTATCAGCTGTACGCGAAGCCGAGCGATCGCCTCAAGCAGATGCTCTTCCGGGCGGTTAAGCGC  
TATTACACGGAGTTTGGGCGCTGAGAGACGTGTGCTTACGTCCTGAAAGGTGAAACGATCGGCATCATCGGGCGGAATGGCGCCGGTAAGTCGACGCTCTTGCA  
GATGTTGTCAGGCGTGTCCAACAATCGACCGGGCAAAGCGAGGTGCGTGGTGCCTGTGTGCGCTGCTGGAACCTCGGCGCCGGTTCAACCCTGAATTCACCGGGC  
GCGAGAACGCGATTCTGACCGGCGCAATTTATGGTATTCCTGGCGAGCAAATGGAAGCGAAGTCCGGGAAATCGAAGCGTTCGCGGATATCGGTGATCATATCGAT  
CAGCCGGTCAAACATTATTCGAGCGGTATGTATGCGCGTCTTGCCTTCTCGGTCGCGATCCACGTTGATCCCGAGATTCTGATCGTCGACGAGATACTTGGGTTGGG  
GATGCATTGTTCCAGCGCAAGTGCATGAACAAGTTCTACGAATTTTCGCGATCGTGGCTGCACGGTGTCTTCTGACTCATGATGCGTACCAGATCAAGAGTACGTGT  
CAACGAGCGTTGTAAGAACGCGCCGACAGGTGCGGTCGCGCATGCCACGGAAGTCGTGGATCGTTATCTGCAGGATCTCGAACTCATTGAGAGTAAGCCGGC  
TGATGTGGTACAACGGGGTGCATCGCTGACGATGCGCCGACGTCACGGCTGTAGCGCAGCCGCAAACTGATTCGGATCACGGACGTGACGCTGACCGACGAAT  
TTGGCTCGCCATCACGAGGTGCGAAGCGGTGAGACAATCGACATTCGTTTCCACTATCACGCGCCGCACAAACAGTCGATTCCGAAGGTATCGTTTGTGCTCAATC  
TCTATCGCCACGATTCGTTTACATCTGTGGCAGCAGCAGATCATGGATGGTCTGGAGCCGTTTCGCAACGGAGGATGAGGGCGAGGTGATTGTGCGCTTCCCTGATC  
TCCAACCTTTCGCGGACGTTATATGTGGCGCGTCGCGATCAACGACGAGCGAGGGCTTGGTATCCTGGCAGAGGCAACGCCGGTCTGCGCGTTTTCAGGTGGTTCGACG  
ATTTCCAGGCGGTTGGTCTCGTCAACGTGCCGCGCAGTGGATCGTACACGGCCAGGAAGTAGAAAAGAGAGAATTATTTCAAAGGTTTGA

>IST439\_01750 ABC-2 type transporter

ATGAGGATAAATTTCTCCCATCTCCGGGGTGTCCGTTCCATCGTGTCTCATCGTGACTTGGTCCGCGCCTCGTGTGGCGGGACATCCAGGGGCGCTATCGTGGATCTG  
CAGGTGGGCTCGGCTGGGCGATCGTTACGCCGCTCGTGTGTTGAGCGTTTACGTTTCGTTTCAGCTTCGTGCTCAAGTCACGCTGGATCGGCGTGCCGAATCAGG  
ATTCGCACGCTACGTACGCGATCATGTGTTTTTCGGGGCTCATTGTGTACAGCGTATTTACAGAATGTATTTACGCGCGCCGACGTTGATGACCGGCAACGTGAAC  
ACGTGAAGAAGGTGCTGTTTCCGCTGGAGATTCTTCCAATCGTGTGCTCGGAAGTGCCTGTTCCATTTTCGCAATCGGCTTACCGTGTCTCCTGGTATTCATCGTGGT  
GTTCCGGCTCGGCGCTGCCACCGACCGCTTGGGCGTTCTGGCTCGTGTGCTGCCGCGCTCATGCTCTGGTGCATCGGTCTTGGATGGATCCTTTTCGTCACTCGGCGTGT  
CTGCGTGACCTGGGGCAACTCGTTACTGTTGCTGAGTGTGTTGATGTTTCGCTACGCCAATTTTCTATTCCGCGGATAGCCTGCCGGAGCAATATCGCATCTACATGG  
ACCTCAACCCGCTATCGTTTGTGATCGGGACATTGCGAGAAACCGTCACTTGGGTGATATCCCGATCCGTCGTCGATTGTCAGCATCTTCTCCGGGCATCCTGGT  
GGCAGTTTTTGGCATGTGGTGGTTCGAACGTACGCGTGGAGGGTTCGCGGATGTCTCTAA

>IST439\_01751 dTDP-4-dehydrorhamnose reductase

ATGAGTGATTGCGGTGTGAGCGTACCGGAACAGCGGACGATTCTGCTTACCGGAGCGAACGGGCAAGTGGGGTTTGAACCTTGCGCGAAGCCTGCAGGGATTGGGGCG  
CGTCATTGCCTGTGATCGCCGGCAGCTTGACCTGGCTGATCTTGACGGCATTTCGCGACGCCATGCGGGCACTGCGTCCGGCACTGGTTCGTAACGCGGCTGCGTACAC  
AGCCGTCGCCGATGCAGAGCGGGACGTGACCGGTAGCATGCGGATCAATGCCGACGCGCCGCGGGTGTGGCGGAAGAGGGCAAGCGTGTGCGTGGCGCTGGTC  
CATTATTCGACCGATTACGTATTTGACGGGCGAAAGATGGGCGCGTACGTCGAGGACGATGCGCCGCCCCGCTGAACGCATACGGCCGTAGCAAGCTCGCAGGCGA  
GGAGGCGATTTCGGGGACTGGCTGCGACCACCTGACTTTCCGTACGAGTTGGGTGTATGACCTGAGGTCCAACAATTTCTTACGACGATGCTGCGCATCGGGCCG  
AGCGTCCCGAGTTGCGTGTGGTACGCGATCAGTTTGGCGCGCCGACGTGGTCTTACGCGATCGCCAGCCTCACGGCGCAGGTGCTTGTCTAGCGGATCGCAGCCGGCC  
GCATGCCGAGCGATGGTACAGTGGTGGTCCACGCATACTGGTGTGTACCATCTGACCGCAATGGGCGAGACATCTTGGCACGGCTTCGCGGAAGCGATATTTGCT  
GCGTCCAATTGCACCAACAAGCCTTCCGTTATTCCAATCGACGCTGCGTGTACCCGGGTCCCGTGTCTCCGGCCGGCCAACTCGCTGCTGTGCAACGATAAGCTCGCA  
TGTGCGTTTGGTTTACGTGCGCCGACTGGCGTTTTTGTGTAACCTTTGCCTGGCCGACGCGCACTCGAGTTCGCGATGCCGAACGGAACCGGTCGATTGCCTAA

>IST439\_01752 dTDP-4-dehydrorhamnose 3,5-epimerase

ATGGCCATCCAGGTAACAGCAACGGCATTGCCGGAAGTCAAGATCATCGAGCCAAAGGTATTCGGCGACGCGCGGGCTACTTCTACGAGAGCTTCAATGCCAGGA  
GTTCCGCGAACAGGTGCGCGCCGGGCGTCGAGTTTCGTACAGGACAATCACTCGCGCTCGATGCGCGGCTGCTGCGCGGGCTGCACTACCAGATTACGATGCGCAGG

GAAAGCTGGTGC GCGTGGTTCGAGGGCGAGGTGTTTCGATGTCGCGGTTCGATGTACGTCGCAGCTCGCCGAACTTCGGCAAATGGGTTCGGCGTCCGTCTGTTCGGCCCCG  
GATCATCGGCAGCTATGGGTCCCGCCGGGCTTCGCGCACGGGTTCTGTGTCTGTCTGACATTGCCAGTTTCTGTACAAGACAAGCGACTATTGGTTTCCGCAGCAC  
GAACGGAGCCTGCTGTGGAATGATCCGGCGATCGGGATCGATTGGCAGATAGACACGTTGCCGACCCTGGCAGCCAAGGATGCGGCCGGGAAGCGGTTGTCCGAGG  
CGGAATGTTATGAGTGA

>IST439\_01753 glucose-1-phosphate thymidyltransferase

ATGGCACGTAAAGGCATTATCCTTGCCGGCGGTTTCGGGCACACGGCTGTATCCGATTACGCATGTCGTATCGAAGCAGCTGCTGCCGGTGTACGACAAACCGATGATT  
TACTACCCGCTGTTCGACGCTGATGGTGGCGGGCATCCGGGACGTGCTGATCATCTCGACGCCGCACGACACGCCGCGCTTCGGAGCGATGCTGGGCGACGGCAGTCA  
GTGGGGCATGAATATCCAGTTTCGCGGTGCAGCCGTCGCCGGACGGGCTCGCGCAGGCCTTCATCATCGGGAAAGTCGTTTCATCGGCAACGAGCCGTCGGCCCTGATTCT  
CGGCGACAACATCTTCTACGGACACGACCTCGCGAAGCAACTCGAACGGGCCGATGCGAATACGAATGGCGCAACGGGTGTTTCGCGTATCACGTTACAGGATCCCGAGC  
GGTACGGCGTGGTTCGAATTCGACCCGGGATTCCTGTCGTTGTTCGATCGAGGAAAAGCCGGTCAAGCCGCGTTCGCACTACGCGGTGACGGGCCTGTACTTCTACGAC  
AACC GCGTGTGCGATATCGCGTCCGACATCAAGCCGTCGCCGCGCGGGCAGCTCGAGATCACCGACGTGAACCTCGCGCTATCTGGCCGACGGTGC ACTCGATGTCGA  
GATCATGGGGCGGGGCTATGCGTGGCTCGATACGGGTACGCATGATTCGCTGATCGATGCGGCAATGTTTCATCGCGACGCTGCAGAAGCGCCAGGGGCTCGTCTGTTG  
CGTGCCCGGAAGAGATCGCCTACCGCAAGCGGTGGATCGACGCCGAGCAGTTGGGCAAGCTCGCCGCACCCTGTTCGAAAAACAGTTACGGGGCGATATCTCCAACAT  
ATTCTTCTGACCAAGTCGCATGGCCATCCAGGTAA

>IST439\_01754 dTDP-glucose 4,6-dehydratase

ATGATCCTGGTTACCGGCGGCGCGGGTTTTATCGGTGCCAATTTTCGTA CTGACTGACTGGATGGATGCTTCCGGCGAGGCCGTGCTCAATGTCGACAAGCTGACCTATGCG  
GGCAACCTGCGGACGCTGGTCCCGCTGGACGGGAATCCGGCCCATGTGTTCCGCCGCTGTCGACATCTGCGATCGCGCGGCACTCGATGCGCTGTTTGCCGAGCACAA  
GCCGCGCGCCGTGATCCACTTCGCCGCCGAAAGCCATGTCGACCGTTCGATCCACGACCGGCCGATTTCTGTCAGACTAACGTCGTCGGTACGTTTACGCTGCTCGA  
AGCGGCACGTACGCACTGGAATAGCCTGAACGATACGGACAAGGCGGCCCTCCGCTTCTGACAGTGTGACCGACGAGGTGTTTCGGATCGCTGTTCGGCCACCGACC  
CGCAATTCGCGAGACGACTCCGTACGCGCCGAACAGCCCGTATTCGGCGACGAAGGCCGGCTCTGATCATCTGGTGCCTGCGTATCACCATACTGACGGCCTGCCG  
ACTCTACGACGAATTGCTCGAACAACACTACGGCCCGTACCAAGTTCCCGAGAAACTGATTCGCTGATGATCGCGAACGCGCTCGCGGGCAAGCCGCTGCGGGTCTA  
CGGCACGGCCAGAACGTGCGCGACTGGCTGTACGTCGGCGACCACTGCAGCGGATCCGCGAAGTGCTCGCGCGCGGCGTGC CGGGCGAGACGTACAACATCGGT  
GGCTGGAACGAGAAGAAGAATCTCGACGTCGTGCATACGCTGTGCGATCTGCTCGACGCGGCGCGGCCGAAGGCCGGCCGGTTCGTACCGCGAGCAGATCACCTACGT  
GAAGGACCGTCCGGGCCACGACCGCCGCTACGCGATCGATGCGCGCAAGCTCGAGCGCGAACTCGGCTGGAAGCCCGCCGAGACGTTTCGAGACGGGGCTCGCGAAA  
ACCGTGCCTGGTATCTCGACAATCAGGCGTGGGTGGACGACGTCGCGTCCGGCGATTATCGCAAGTGGGTGCAAACCAATTACGCGCAACGCACGTGA

**Table S3.2** | Primer sequences used to amplify the desired regions with detailed description

Primers	Primer Sequences	Description
primer 824	5'-GCCCATTTTCCTGTCAGTAACGAGA-3'	Screening of the transformants
pSC rev	5'-GATGCCTGGCAGTTCCTACTCTCG-3'	
WbiI-flag-NdeI	5'-GCTACATATGGACTACAAGGACGACGACGACAAGTTGCAATCCAGAGCATCTTGGCT-3'	Cloning of <i>wbiI</i> with an N-terminal FLAG epitope incorporated to the expressed protein
WbiI_439_XbaI	5'-TAGCATCTAGATCAGCGGTTTCGATGCGACTCTCAC-3'	
wbiI 439 NdeI	5'-ACGCTCATATGTTGCAATCCAGAGCATCTTGGCTG-3'	Cloning of <i>wbiI</i> with a C-terminal FLAG epitope incorporated to the expressed protein
WbiI-flag-XbaI	5'-CGTTCTAGATCACTTGTCTGTCGTCGTCGTCCTTGTAGTCGCGGTTTCGATGCGACTCTCAC-3'	
Bmul_2510-439-NdeI	5'-AAGGCCTACATATGTACAGACTGAAGACCA-3'	Cloning of <i>bmul_2510</i> with a C-terminal FLAG epitope incorporated to the expressed protein
Bmul_2510-flag-XbaI	5'-CGTTCTAGATCACTTGTCTGTCGTCGTCGTCCTTGTAGTCTGTACGACCGTCCGACGACCC-3'	
Primer P1	5'-TGAAATTCAGCAGGATCACAACGCTCATATGTTGCAATC-3'	Cloning of both <i>wbiI</i> and <i>bmul_2510</i> with a C-terminal FLAG epitope incorporated to the expressed protein using the Gibson assembly method
Primer P2	5'-GCGACCTCCTCGTTCTAGATCACTTGTCTG-3'	
Primer P3	5'-ATCTAGAACGAGGAGGTCGCGCTCATATGTACAGACTGAAGACC-3'	
Primer P3	5'-GATCCCCGGGTACCATGGCACGTTCTAGATCACTTGTCTGTC-3'	

**Table S3.3** | Monosaccharide compositional analysis of the LPS/LOS isolated from *B. cenocepacia* isolates IST4113, IST4129, IST4134 and IST439.

Strains	Monosaccharide composition
<b>IST4113, IST4129, IST4134</b>	4-amino-4-deoxy-L-arabinose (L-Ara4N) D-glucose (D-Glc) D-galactose (D-Gal) D-glucosamine (D-GlcN)  L- <i>glycero</i> -D- <i>manno</i> -heptose (L,D-Hep) 3-deoxy-D- <i>manno</i> -oct-2-ulopyranosonic acid (D-Kdo) D- <i>glycero</i> -D- <i>talo</i> -oct-2-ulopyranosonic acid (D-Ko)
<b>IST439</b>	4-amino-4-deoxy-L-arabinose (L-Ara4N) D-glucose (D-Glc) D-galactose (D-Gal)  L- <i>glycero</i> -D- <i>manno</i> -heptose (L,D-Hep) 3-deoxy-D- <i>manno</i> -oct-2-ulopyranosonic acid (D-Kdo) D- <i>glycero</i> -D- <i>talo</i> -oct-2-ulopyranosonic acid (D-Ko)  4-amino-4-deoxy-L-arabinose (L-Ara4N) D-glucose (D-Glc) D-glucosamine (D-GlcN) D-ribose (D-Rib) D-galactosamine (D-GalN)

## 8.3.3. Data related to chapter IV

**Table S4.1** | Detailed description of the *Burkholderia* sequential isolates per patient examined in this study with their isolation dates. The presence or absence of the LPS OAg is presented by + or – symbols, respectively, and their corresponding profiles (Fig. 4.2) is provided. Information on the species identification and genotyping for the different strains tested are also presented: the genotyping data and the ribopatterns were obtained before (Cunha et al., 2003; Cunha et al., 2007; Coutinho et al., 2011b; Moreira et al., 2014; Coutinho et al., 2015; Moreira et al., 2017) and the RAPD profiles were obtained in this study (Fig. S4.1) together with the sequence type, when present in the MLST database. The different letter (A, B, C and D) of the isolate signature (as ISTnumberA, B, C or D) was used to indicate different clone at the same isolation date. ND – Not determined.

Patient	Isolate	Isolation date	species	O-antigen presence	OAg profiles (Fig. 4.2)	RAPD profiles (Fig. S4.1)	Isolate ripo-patterns	Sequence type	reference
<b>Patient B</b>	IST402	4/13/95	<i>B. stabilis</i>	-	G1	RAPD15	1	ND	(Cunha et al., 2003)
	IST409	4/24/95	<i>B. stabilis</i>	+	G2	RAPD14	1	ND	(Cunha et al., 2003)
	IST421	4/3/98	<i>B. stabilis</i>	-	G1	RAPD15	1	ND	(Cunha et al., 2003)
	IST428	9/28/98	<i>B. stabilis</i>	-	G1	RAPD15	1	ND	(Cunha et al., 2003)
	IST437	1/25/99	<i>B. stabilis</i>	-	G1	RAPD15	1	ND	(Cunha et al., 2003)
<b>Patient D</b>	IST407	12/6/95	<i>B. cenocepacia</i> IIIB	+	-	-	3	ND	(Cunha et al., 2003)
	IST4288A	1/11/07	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work
	IST4326A	4/27/07	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work
	IST4368A	6/20/07	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work

	IST4368B	6/20/07	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work
	IST4390A	3/12/08	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work
	IST4558A	7/13/10	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work
	IST4563A	10/26/10	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work
	IST4616A	2/8/11	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work
	IST4629A	3/29/11	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work
	IST4645A	5/9/11	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work
	IST4649A	6/1/11	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work
<b>Patient H</b>	IST412	1/21/97	<i>B. stabilis</i>	+	G3	RAPD14	1	ND	(Cunha et al., 2003)
	IST413	3/18/97	<i>B. stabilis</i>	+	G4	RAPD15	1	ND	(Cunha et al., 2003)
	IST414	5/13/97	<i>B. stabilis</i>	+	G3	RAPD14	1	ND	(Cunha et al., 2003)
	IST415	7/8/97	<i>B. stabilis</i>	+	G4	RAPD15	1	ND	(Cunha et al., 2003)
	IST420	3/11/98	<i>B. stabilis</i>	+	G3	RAPD14	1	ND	(Cunha et al., 2003)
	IST423	5/6/98	<i>B. stabilis</i>	+	G3	RAPD14	1	ND	(Cunha et al., 2003)
	IST425	7/7/98	<i>B. stabilis</i>	+	G3	RAPD14	1	ND	(Cunha et al., 2003)

	IST427A	8/13/98	<i>B. stabilis</i>	+	G3	RAPD14	1	ND	(Cunha et al., 2003)
	IST427B	8/13/98	<i>B. stabilis</i>	+	G4	RAPD15	1	ND	(Cunha et al., 2003)
	IST446	3/13/99	<i>B. stabilis</i>	+	G4	RAPD15	1	ND	(Cunha et al., 2003)
	IST448	6/4/99	<i>B. stabilis</i>	+	G3	RAPD14	1	ND	(Cunha et al., 2003)
	IST451	7/8/99	<i>B. stabilis</i>	+	G3	RAPD14	1	ND	(Cunha et al., 2003)
<b>Patient I</b>	IST416	9/30/97	<i>B. cenocepacia</i> IIIA	+	A1	RAPD01	8	ND	(Cunha et al., 2003)
	IST442	2/10/99	<i>B. cenocepacia</i> IIIA	+	A1	RAPD01	ND	ND	This work
	IST464	5/24/00	<i>B. cenocepacia</i> IIIA	+	A1	RAPD01	ND	ND	This work
Patient J	IST419	2/26/98	<i>B. multivorans</i>	+	E1	ND	9	ND	(Cunha et al., 2003)
	IST424	6/4/98	<i>B. multivorans</i>	+	E1	ND	9	ND	(Cunha et al., 2003)
	IST439	1/30/99	<i>B. cenocepacia</i> IIIA	+	A2	RAPD02	11	218	(Cunha et al., 2003), (Coutinho et al., 2011a)
	IST453	7/19/99	<i>B. multivorans</i>	-	E2	ND	9	375	(Cunha et al., 2003)
	IST455A	2/1/00	<i>B. multivorans</i>	-	E2	ND	ND	ND	This work
	IST455B	2/1/00	<i>B. multivorans</i>	-	E2	ND	ND	ND	This work

	IST461	4/4/00	<i>B. multivorans</i>	-	E2	ND	9	ND	(Cunha et al., 2003)
	IST495A	5/29/01	<i>B. multivorans</i>	-	E2	ND	ND	ND	This work
	IST595B	5/29/01	<i>B. multivorans</i>	-	E2	ND	ND	ND	This work
	IST4103	7/24/01	<i>B. cenocepacia</i> IIIA	-	A3	RAPD02	11	218	(Cunha et al., 2003), (Coutinho et al., 2011a)
	IST4110	9/25/01	<i>B. cenocepacia</i> IIIA	-	A3	RAPD02	11	614	(Cunha et al., 2003), (Coutinho et al., 2011a)
	IST4112	10/11/01	<i>B. cenocepacia</i> IIIA	-	A3	RAPD02	11	614	(Cunha et al., 2003), (Coutinho et al., 2011a)
	IST4113	11/6/01	<i>B. cenocepacia</i> IIIA	-	A3	RAPD02	11	614	(Cunha et al., 2003), (Coutinho et al., 2011a)
	IST4119	1/22/02	<i>B. multivorans</i>	+	E1	ND	9	ND	(Cunha et al., 2003)
	IST4116A	2/11/02	<i>B. cenocepacia</i> IIIA	-	A3	RAPD02	11	614	(Cunha et al., 2003), (Coutinho et al., 2011a)
	IST4116B	2/11/02	<i>B. cenocepacia</i> IIIA	-	A3	RAPD02	11	218	(Cunha et al., 2003), (Coutinho et al., 2011a)
	IST4131	2/26/02	<i>B. cenocepacia</i> IIIA	-	A3	RAPD02	11	218	(Cunha et al., 2003), (Coutinho et al., 2011a)
	IST4129	3/26/02	<i>B. cenocepacia</i> IIIA	-	A3	RAPD02	11	218	(Cunha et al., 2003), (Coutinho et al., 2011a)
	IST4130	5/14/02	<i>B. cenocepacia</i> IIIA	-	A3	RAPD02	11	218	(Cunha et al., 2003), (Coutinho et al., 2011a)
	IST4134	7/2/02	<i>B. cenocepacia</i> IIIA	-	A3	RAPD02	11	218	(Cunha et al., 2003), (Coutinho et al., 2011a)
<b>Patient N</b>	IST431A	8/30/98	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work

IST431B	8/30/98	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST443A	3/3/99	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST443B	3/3/99	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST443C	3/3/99	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST443D	3/3/99	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST444A	3/3/99	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST444B	3/3/99	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST444C	3/3/99	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST449A	6/9/99	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST449B	6/9/99	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST463A	5/23/00	<i>B. multivorans</i>	+	-	ND	ND	ND	This work
IST463B	5/23/00	<i>B. multivorans</i>	+	-	ND	ND	ND	This work
IST472A	10/18/00	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST472B	10/18/00	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST485A	4/1/01	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
IST485B	4/1/01	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work

	IST485C	4/1/01	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST491A	4/26/01	<i>B. cepacia</i>	+	C2	ND	12	ND	This work
	IST4104A	8/16/01	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4104B	8/16/01	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4105A	8/31/01	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4105B	8/31/01	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4106A	9/24/01	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4106B	9/24/01	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4106C	9/24/01	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4115A	11/8/01	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4115B	11/8/01	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4117A	1/3/02	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4117B	1/3/02	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
Patient O	IST430	9/30/98	<i>B. cenocepacia</i> IIIA	+	A4	RAPD03	11	ND	(Cunha et al., 2003)
	IST435	11/1/98	<i>B. cenocepacia</i> IIIB	+	B1	RAPD09	14	ND	(Cunha et al., 2003)

---

IST440	2/8/99	<i>B. cenocepacia</i> IIIB	+	B1	RAPD09	14	ND	(Cunha et al., 2003)
IST450	6/9/99	<i>B. cenocepacia</i> IIIB	+	B1	RAPD09	14	ND	(Cunha et al., 2003)
IST458	3/8/00	<i>B. cenocepacia</i> IIIB	+	B1	RAPD09	14	ND	(Cunha et al., 2003)
IST462	5/23/00	<i>B. cenocepacia</i> IIIA	+	A4	RAPD03	11	280	(Cunha et al., 2003)
IST463	5/23/00	<i>B. cenocepacia</i> IIIA	+	A4	RAPD03	11	ND	(Cunha et al., 2003)
IST467	7/19/00	<i>B. cenocepacia</i> IIIA	+	A4	RAPD03	11	ND	(Cunha et al., 2003)
IST470	10/18/00	<i>B. cenocepacia</i> IIIA	+	A4	RAPD03	11	ND	(Cunha et al., 2003)
IST471	10/18/00	<i>B. cenocepacia</i> IIIB	+	B1	RAPD09	14	ND	(Cunha et al., 2003)
IST478	2/15/01	<i>B. cenocepacia</i> IIIA	+	A4	RAPD03	11	ND	(Cunha et al., 2003)
IST486	4/3/01	<i>B. cenocepacia</i> IIIA	+	A4	RAPD03	11	ND	(Cunha et al., 2003)
IST490	4/26/01	<i>B. cenocepacia</i> IIIA	+	A4	RAPD03	11	ND	(Cunha et al., 2003)
IST497	5/31/01	<i>B. cenocepacia</i> IIIA	+	A4	RAPD03	11	ND	(Cunha et al., 2003)
IST4107	9/11/01	<i>B. cenocepacia</i> IIIA	+	A4	RAPD03	11	ND	(Cunha et al., 2003)
IST4114	11/8/01	<i>B. cenocepacia</i> IIIB	+	B1	RAPD09	14	ND	(Cunha et al., 2003)
IST4118	1/3/02	<i>B. cenocepacia</i> IIIB	+	B1	RAPD09	14	ND	(Cunha et al., 2003)
IST4123	2/28/02	<i>B. cenocepacia</i> IIIB	+	B1	RAPD09	14	ND	(Cunha et al., 2003)

IST4128	5/2/02	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2003)
IST4135	11/21/02	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4137	1/23/03	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4139	1/24/03	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4142	1/25/03	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4146	8/7/03	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4150	9/30/03	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4157	1/22/04	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4165	3/25/04	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4167	5/13/04	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4240	9/15/04	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4177A	11/17/04	<i>B. cenocepacia</i> IIIB	+	B2	RAPD10	ND	ND	This work
IST4180	1/26/05	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4183	3/16/05	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4189	5/25/05	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
IST4204A	7/27/05	<i>B. cenocepacia</i> IIIB	+	B2	RAPD10	ND	ND	This work

	IST4209	9/21/05	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
	IST4216	11/16/05	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
	IST4227	1/18/06	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
	IST4230	2/8/06	<i>B. cepacia</i>	+	C2	ND	12	ND	(Cunha et al., 2007)
	IST4297	3/21/07	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4327	4/10/07	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4315	4/20/07	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4355	7/4/07	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4369	7/12/07	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
	IST4383	10/31/07	<i>B. cepacia</i>	+	C2	ND	ND	ND	This work
<b>Patient P</b>	IST432	10/22/98	<i>B. cenocepacia</i> IIIA	+	A5	RAPD04	13	ND	(Cunha et al., 2003)
	IST468	10/18/00	<i>B. cenocepacia</i> IIIA	+	A5	RAPD04	13	ND	(Cunha et al., 2003)
	IST489	4/19/01	<i>B. cenocepacia</i> IIIA	+	A5	RAPD04	13	ND	(Cunha et al., 2003)
<b>Patient R</b>	IST438	1/25/99	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	43	(Cunha et al., 2003)
	IST445	3/8/99	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	43	(Cunha et al., 2003)

---

IST452	7/12/99	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	43	(Cunha et al., 2003)
IST454	12/22/99	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2003)
IST456	3/8/00	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	43	(Cunha et al., 2003)
IST465	5/31/00	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2003)
IST474	1/9/01	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2003)
IST475	1/16/01	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	43	(Cunha et al., 2003)
IST476	1/19/01	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2003)
IST484	3/29/01	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2003)
IST4144	7/10/03	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2007)
IST4149	9/25/03	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2007)
IST4155	11/27/03	<i>B. cenocepacia</i> IIIB	-	B4	RAPD10	15	ND	(Cunha et al., 2007)
IST4164	3/11/04	<i>B. cenocepacia</i> IIIB	-	B4	RAPD10	15	ND	(Cunha et al., 2007)
IST4178	11/17/04	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2007)
IST4203	7/28/05	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2007)
IST4205	8/10/05	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2007)
IST4210	9/21/05	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2007)

	IST4219	11/16/05	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2007)
	IST4228	1/18/06	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2007)
	IST4231	2/14/06	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2007)
	IST4232	2/21/06	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2007)
	IST4235	3/2/06	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	15	ND	(Cunha et al., 2007)
	IST4252	6/14/06	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	ND	ND	This work
	IST4255	7/12/06	<i>B. cenocepacia</i> IIIB	+	B3	RAPD10	ND	ND	This work
<b>Patient T</b>	IST466	6/12/00	<i>B. cenocepacia</i> IIIB	+	B5	RAPD11	16	ND	(Cunha et al., 2003)
	IST469	10/18/00	<i>B. cenocepacia</i> IIIB	+	B5	RAPD11	16	ND	(Cunha et al., 2003)
	IST473	12/20/00	<i>B. cenocepacia</i> IIIB	+	B5	RAPD11	16	ND	(Cunha et al., 2003)
	IST479	2/19/01	<i>B. cenocepacia</i> IIIB	+	B5	RAPD11	16	ND	(Cunha et al., 2003)
	IST4102	7/12/01	<i>B. cenocepacia</i> IIIB	+	B5	RAPD11	16	ND	(Cunha et al., 2003)
	IST4109	9/26/01	<i>B. cenocepacia</i> IIIB	+	B5	RAPD11	16	ND	(Cunha et al., 2003)
<b>Patient V</b>	IST481	2/21/01	<i>B. contaminans</i>	+	D1	ND	17	ND	(Cunha et al., 2003)
	IST4148	9/3/03	<i>B. contaminans</i>	+	D1	ND	17	ND	(Cunha et al., 2007)

IST4169	6/8/04	<i>B. contaminans</i>	+	D1	ND	17	ND	(Cunha et al., 2007)
IST4241	6/28/04	<i>B. contaminans</i>	+	D1	ND	17	ND	(Cunha et al., 2007)
IST4200	7/19/05	<i>B. contaminans</i>	+	D1	ND	17	ND	(Cunha et al., 2007)
IST4224	12/27/05	<i>B. contaminans</i>	+	D1	ND	17	ND	(Cunha et al., 2007)
IST4256	7/14/06	<i>B. contaminans</i>	+	D1	ND	ND	ND	This work
IST4323	4/24/07	<i>B. contaminans</i>	+	D1	ND	ND	ND	This work
IST4538	10/4/10	<i>B. contaminans</i>	+	D1	ND	ND	ND	This work
IST4657	6/14/11	<i>B. contaminans</i>	+	D1	ND	ND	ND	This work
IST4684	10/12/11	<i>B. contaminans</i>	+	D1	ND	ND	ND	This work
IST4751	10/17/12	<i>B. contaminans</i>	+	D1	ND	ND	ND	This work
IST4920	4/8/16	<i>B. contaminans</i>	+	D1	ND	ND	ND	This work

Patient AB	IST4121	1/31/02	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
	IST4136	1/23/03	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
	IST4140	5/15/03	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
	IST4141	5/15/03	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)

IST4151	10/28/03	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
IST4153	11/15/03	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
IST4154	11/15/03	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
IST4166	4/18/04	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
IST4170	6/24/04	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
IST4892B	6/24/04	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	ND	ND	(Moreira et al., 2017)
IST4173	8/5/04	<i>B. cepacia</i>	+	-	-	19	ND	(Cunha et al., 2007)
IST4179	1/3/05	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
IST4182	3/9/05	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
IST4187	4/20/05	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
IST4191	6/8/05	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
IST4202	7/27/05	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
IST4213	10/6/05	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
IST4234	3/2/06	<i>B. cenocepacia</i> IIIA	+	A6	RAPD05	7	ND	(Cunha et al., 2007)
IST4259	8/5/06	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
IST4893	8/5/06	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	(Moreira et al., 2017)

IST4894	8/5/06	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	(Moreira et al., 2017)
IST4346	6/28/07	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
IST4417	10/9/08	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	(Moreira et al., 2017)
IST4895	10/9/08	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	(Moreira et al., 2017)
IST4436	12/10/09	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
IST4450	3/4/10	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
IST4470	5/6/10	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
IST4559	7/22/10	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
IST4589	12/9/10	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
IST4624	2/17/11	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
IST4653	6/14/11	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
IST4676	10/13/11	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
IST4676S	10/13/11	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	(Moreira et al., 2017)
IST4676R	10/13/11	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	(Moreira et al., 2017)
IST4677S	10/13/11	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	(Moreira et al., 2017)
IST4725	3/29/12	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work

	IST4787	2/7/13	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
	IST4797	4/26/13	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
	IST4835	12/5/13	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
	IST4854	4/24/14	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
	IST4882	8/21/14	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
	IST4882A	8/21/14	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	(Moreira et al., 2017)
	IST4882B	8/21/14	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	(Moreira et al., 2017)
	IST4884A	8/21/14	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	(Moreira et al., 2017)
	IST4889	6/16/15	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
	IST4930	7/12/16	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
	IST4934	9/20/16	<i>B. cenocepacia</i> IIIA	-	A7	RAPD05	ND	ND	This work
<b>Patient AF</b>	IST4193	12/9/04	<i>B. contaminans</i>	+	D2	ND	2	96	(Coutinho et al., 2015)
	IST4194	2/10/05	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
	IST4186	4/14/05	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
	IST4188	5/4/05	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)

IST4192	6/30/05	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4206	8/10/05	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4207	8/10/05	<i>B. cenocepacia III B</i>	+	-	ND	ND	ND	(Coutinho et al., 2015)
IST4221	11/23/05	<i>B. cepacia</i>	+	-	ND	ND	ND	(Coutinho et al., 2015)
IST4238	3/14/06	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4237	3/20/06	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4240a	5/10/06	<i>B. cenocepacia III A</i>	+	A8	RAPD06	ND	ND	(Coutinho et al., 2015)
IST4240b	5/10/06	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4240c	5/10/06	<i>B. cepacia</i>	+	-	ND	ND	ND	(Coutinho et al., 2015)
IST4251	6/14/06	<i>B. cenocepacia III A</i>	+	A8	RAPD06	ND	ND	(Coutinho et al., 2015)
IST4253a	7/7/06	<i>B. cenocepacia III A</i>	-	A10	RAPD06	ND	ND	(Coutinho et al., 2015)
IST4253b	7/7/06	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4258a	8/2/06	<i>B. cenocepacia III A</i>	-	A10	RAPD06	ND	ND	(Coutinho et al., 2015)
IST4258b	8/2/06	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4272	12/11/06	<i>B. cenocepacia III A</i>	+	A9	RAPD07	ND	ND	(Coutinho et al., 2015)
IST4276	11/22/06	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)

IST4303	3/28/07	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4319	4/23/07	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4322	4/23/07	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4391	3/19/08	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4402a	7/9/08	<i>B. cenocepacia IIIA</i>	-	A10	RAPD06	ND	ND	(Coutinho et al., 2015)
IST4402b	7/9/08	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4408	8/13/08	<i>B. cenocepacia IIIA</i>	-	A10	RAPD06	ND	ND	(Coutinho et al., 2015)
IST4412	9/16/08	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4420a	10/17/08	<i>B. cenocepacia IIIA</i>	-	A10	RAPD06	ND	ND	(Coutinho et al., 2015)
IST4420d	10/17/08	<i>B. dolosa</i>	+	-	ND	ND	ND	(Coutinho et al., 2015)
IST4435	11/23/09	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4497	1/11/10	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4456	3/29/10	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4474	5/18/10	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4496	2/22/10	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
IST4485	5/10/10	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)

	IST4490	6/16/10	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
	IST4550	7/19/10	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
	IST4513	8/9/10	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
	IST4517	10/10/10	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
	IST4566	10/29/10	<i>B. contaminans</i>	+	D2	ND	ND	ND	(Coutinho et al., 2015)
<b>Patient AL</b>	IST4152	10/30/03	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)
	IST4158	2/12/04	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)
	IST4159	3/2/04	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)
	IST4160	4/2/04	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)
	IST4168	5/18/04	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)
	IST4175	10/6/04	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)
	IST4184	3/17/05	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)
	IST4220	11/16/05	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)
	IST4222	12/14/05	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)
	IST4226	1/18/06	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)

Patient AN	IST4197	5/25/05	<i>B. cenocepacia</i> IIIA	+	A11	RAPD08	21	ND	(Cunha et al., 2007)
	IST4896	5/25/05	<i>B. cenocepacia</i> IIIA	+	A11	RAPD08	ND	ND	(Moreira et al., 2017)
	IST4190	6/1/05	<i>B. cenocepacia</i> IIIA	+	A11	RAPD08	21	ND	(Cunha et al., 2007)
	IST4201	7/21/05	<i>B. cenocepacia</i> IIIA	+	A11	RAPD08	21	ND	(Cunha et al., 2007)
	IST4211	9/22/05	<i>B. cenocepacia</i> IIIA	+	A11	RAPD08	21	ND	(Cunha et al., 2007)
	IST4215A	10/27/05	<i>B. cenocepacia</i> IIIB	+	-	ND	ND	ND	This work
	IST4223A	12/15/05	<i>B. cenocepacia</i> IIIA	+	A11	RAPD08	ND	ND	This work
	IST4223B	12/15/05	<i>B. multivorans</i>	+	-	ND	ND	ND	This work
	IST4243	5/11/06	<i>B. cenocepacia</i> IIIA	+	A11	RAPD08	ND	ND	(Moreira et al., 2017)
	IST4897	5/11/06	<i>B. cenocepacia</i> IIIA	+	A11	RAPD08	ND	ND	(Moreira et al., 2017)
	IST4304A	3/29/07	<i>B. cenocepacia</i> IIIA	+	A11	RAPD08	ND	ND	This work
	IST4304B	3/29/07	<i>B. multivorans</i>	+	-	ND	ND	ND	This work
	IST4364	6/21/07	<i>B. cenocepacia</i> IIIA	+	A11	RAPD08	ND	ND	This work
	IST4386	11/22/07	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	(Moreira et al., 2017)
	IST4388	1/10/08	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	This work
	IST4410	8/21/08	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	(Moreira et al., 2017)

IST4898B	8/21/08	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	(Moreira et al., 2017)
IST4442	10/8/09	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	This work
IST4446	1/21/10	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	This work
IST4466	5/6/10	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	This work
IST4466S	5/6/10	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	(Moreira et al., 2017)
IST4466R	5/6/10	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	(Moreira et al., 2017)
IST4469S	5/6/10	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	(Moreira et al., 2017)
IST4531	8/13/10	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	This work
IST4586	12/9/10	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	This work
IST4628	3/14/11	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	This work
IST4632	4/28/11	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	This work
IST4680	10/13/11	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	This work
IST4691	11/3/11	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	This work
IST4721A	3/29/12	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	This work
IST4721B	3/29/12	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	This work
IST4721	3/29/12	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	(Moreira et al., 2017)

	IST4722	3/29/12	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	(Moreira et al., 2017)
	IST4788	2/7/13	<i>B. cenocepacia</i> IIIA	-	A12	RAPD08	ND	ND	(Moreira et al., 2017)
<b>Patient AP</b>	IST4199	7/14/05	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)
	IST4218	11/17/05	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)
	IST4236	3/8/06	<i>B. cepacia</i>	+	C1	RAPD12	12	ND	(Cunha et al., 2007)
	IST4183	5/3/06	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work
	IST4406	5/6/08	<i>B. cepacia</i>	+	C1	RAPD12	ND	ND	This work
<b>Patient AQ</b>	IST4198	7/7/05	<i>B. cepacia</i>	+	C5	ND	24	ND	(Cunha et al., 2007)
	IST4214	10/13/05	<i>B. cepacia</i>	+	C5	ND	24	ND	(Cunha et al., 2007)
	IST4225	12/26/05	<i>B. cepacia</i>	+	C5	ND	24	ND	(Cunha et al., 2007)
	IST4229	1/31/06	<i>B. cepacia</i>	+	C5	ND	24	ND	(Cunha et al., 2007)
	IST4233	2/22/06	<i>B. cepacia</i>	+	C5	ND	24	ND	(Cunha et al., 2007)
	IST4254	7/12/06	<i>B. cepacia</i>	+	C5	ND	ND	ND	This work
	IST4294	1/29/07	<i>B. cepacia</i>	+	C5	ND	ND	ND	This work
	IST4308	4/4/07	<i>B. cepacia</i>	+	C5	ND	ND	ND	This work

	IST4399	1/8/08	<i>B. cepacia</i>	+	C5	ND	ND	ND	This work
	IST4395	4/16/08	<i>B. cenocepacia</i> IIIB	-	-	ND	ND	ND	This work
	IST4407	6/18/08	<i>B. cepacia</i>	+	C5	ND	ND	ND	This work
<b>Patient AR</b>	IST4208	8/25/05	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
	IST4370	7/12/07	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
	IST4377	8/13/07	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
	IST4385	11/22/07	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
	IST4387	1/10/08	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
	IST4389	3/13/08	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
	IST4409	8/4/08	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
	IST4616	8/5/09	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
	IST4481 IIIB	5/19/10	<i>B. cenocepacia</i> IIIB	+	B6	RAPD10	ND	43	(Moreira et al., 2014)
	IST4481 VI	5/19/10	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
	IST4484 IIIB	4/30/10	<i>B. cenocepacia</i> IIIB	+	B6	RAPD10	ND	43	(Moreira et al., 2014)
	IST4484 VI	4/30/10	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)

	IST4499	7/27/10	<i>B. cenocepacia</i> IIIB	+	B6	RAPD10	ND	43	(Moreira et al., 2014)
	IST4503	7/22/10	<i>B. cenocepacia</i> IIIB	+	B6	RAPD10	ND	43	(Moreira et al., 2014)
	IST4507	10/7/10	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
	IST4583	12/9/10	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
	IST4601S	1/12/10	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
	IST4601R	1/12/10	<i>B. dolosa</i>	-	F	ND	ND	668	(Moreira et al., 2014)
<b>Patient AU</b>	IST4283	2/7/07	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
	IST4299	3/21/07	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
	IST4371	7/11/07	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
	IST4378	9/12/07	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
	IST4384	11/7/07	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
	IST4392	4/8/08	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
	IST4401	6/18/08	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
	IST4413	9/17/08	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
	IST4422	10/15/08	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work

---

IST4424	11/12/08	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4433	7/6/09	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4527	10/18/10	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4541	6/14/10	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4546	7/19/10	<i>B. cepacia</i>	+	C4	RAPD13	ND	ND	This work
IST4597	12/20/10	<i>B. cepacia</i>	+	C4	RAPD13	ND	ND	This work
IST4604	1/24/10	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4623	2/7/11	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4640	5/2/11	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4665	6/27/11	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4673	10/10/11	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4709	12/9/11	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4729	4/16/12	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4746	9/10/12	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4755	10/29/12	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work

---

---

IST4768	11/26/12	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4776	1/21/13	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4793	3/18/13	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4815	6/24/13	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4832	11/18/13	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4844	1/13/14	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4919	4/2/16	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4923	5/9/16	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4928	6/6/16	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work
IST4933	9/12/16	<i>B. cepacia</i>	+	C3	RAPD12	ND	ND	This work

---