Insights into the regulatory networks of Pseudomonas aeruginosa

Julian Trouillon

▶ **To cite this version:**

Julian Trouillon. Insights into the regulatory networks of Pseudomonas aeruginosa. Virology. Université Grenoble Alpes [2020-..], 2020. English. NNT : 2020GRALV027 . tel-03331635

HAL Id: tel-03331635
https://theses.hal.science/tel-03331635v1

Submitted on 2 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE GRENOBLE ALPES

Spécialité : **Virologie - Microbiologie - Immunologie**

Arrêté ministériel : 25 mai 2016

Présentée par

Julian Trouillon

Thèse dirigée par **Sylvie Elsen**

préparée au sein du **Laboratoire BCI – Biologie du Cancer et de l'Infection**
dans **l'École Doctorale Chimie et Sciences du Vivant**

Insights into the regulatory networks of *Pseudomonas aeruginosa*

Etude des réseaux de régulation chez *Pseudomonas aeruginosa*

Thèse soutenue publiquement le **13 novembre 2020**,
devant le jury composé de :

Monsieur François Parcy
Directeur de recherche, CNRS, Président
Madame Carmen Buchrieser
Professeur, Institut Pasteur, Rapportrice
Monsieur Patrick Viollier
Professeur, Université de Genève, Rapporteur
Monsieur Thomas Hindré
Maître de conférences, Université Grenoble Alpes, Examinateur
Monsieur Stephen Lory
Professeur, Harvard Medical School, Examinateur
Madame Sylvie Elsen
Chargée de recherche, CNRS, Directrice de thèse

Acknowledgements

First, I would like to thank my jury. Especially Carmen Buchrieser and Patrick Viollier who are reviewers of my thesis, as well as Thomas Hindré, Stephen Lory and François Parcy. I really appreciate that you are taking the time to read and evaluate my work.

I am also grateful to the members of my thesis advisory committee, Stéphan Lacour, François Parcy and Bertrand Toussaint, for their guidance and our interesting discussions. Especially, I thank François and members of his lab for their experimental and computational advices.

I thank the Doctoral School of Chemistry and Life Sciences of the Grenoble Alpes University and the French Ministry of Education and Research for my thesis funding.

Sylvie, un grand merci pour ces trois ans passés avec toi. J'ai vraiment apprécié ce duo que nous avons formé, aussi bien au bureau qu'à la paillasse. Tu m'as laissé la liberté de faire ce que je voulais tout en m'apportant un soutien immense constant, ce qui a permis de faire de cette thèse ce qu'elle est, et pour ça je te remercie pleinement. Durant ces trois ans, j'ai énormément appris de toi, ce sont ton implication et ton excitation pour la recherche qui m'ont permis d'arriver où je suis aujourd'hui. Merci pour tout.

Je voudrais remercier Ina Attrée pour son accueil dans son équipe. Ina, tu as su créer un environnement de travail exceptionnel à tous points de vue et j'ai vraiment adoré mes années dans ce laboratoire. Je souhaite également te remercier pour ton soutien et ta confiance durant ces trois ans, qui ont aussi été très importants pour moi.

I would like to thank Stephen Lory for welcoming me in his lab for three months. Steve, I sincerely enjoyed my time in your lab in Boston, I only wish it could have lasted longer. It was a truly insightful experience, on both scientific and human sides. Thank you for your trust and the numerous inspiring interactions.

I also want to thank the people that I've met in Boston, especially the entire Lory lab and notably Kook Han, who helped me a lot and with whom I had a lot of fun in the lab. Kook, my bench neighbor and partner in many interesting experiments and discussions, I really enjoyed working with you. I would also like to thank the amazing people that I've met in the lab and other neighboring labs in the building, with whom I spent most of my days and nights and who made my stay in Boston unforgettable.

J'aimerais remercier toute l'équipe PBRC, j'ai vraiment apprécié ces trois ans passés auprès de vous. D'un point de vue scientifique, je souhaite tous vous remercier

pour nos nombreuses interactions qui ont aidé à façonner cette thèse. Je remercie particulièrement Michel Ragno pour son aide et son implication précieuses qui m'ont permis d'aller plus vite et plus loin. D'un point de vue humain, je n'oublierai jamais l'ambiance extraordinaire dans ce labo et toutes nos sorties qui ont rendu ces trois ans très appréciables.

Particulièrement, je remercie Stéphane et Vincent, avec qui tout a commencé il y a trois ans et avec qui on s'apprête maintenant à finir ensemble, ainsi que Manon qui nous à rejoins un an après. Vous trois avez été au centre de ces trois années, tous les jours, au labo comme au bar, et les avez rendus inoubliables. Nous partons maintenant chacun de notre côté mais j'ose espérer vous revoir, et souvent.

Je remercie du fond du cœur mes parents, qui ont réussi à rendre toute chose possible pour nous. Merci aussi à ma sœur Léa et mon frère Théo qui m'accompagnent au quotidien depuis toujours.

Merci à Anaëlle, ta présence et ton soutien m'ont été indispensables durant ces années. J'espère te retrouver bien vite, moins loin.

Enfin, je voudrais remercier tous mes amis et mes colocataires, votre soutien inconditionnel durant toutes ces années est ce qui m'a fait arriver jusqu'ici.

Contents

List of Figures

List of Tables

List of Abbreviations

ANI	**A**verage **N**ucleotide **I**dentity
AcP	**Ac**etyl **P**hosphate
DBD	**D**NA-**B**inding **D**omain
ECF	**E**xtra**C**ytoplasmic **F**unction
HGT	**H**orizontal **G**ene **T**ransfer
HK	**H**istidine **K**inase
ISR	**I**nduced **S**ystemic **R**esistance
NAP	**N**ucleoid-**A**ssociated **P**rotein
OCS	**O**ne-**C**omponent **S**ystem
QS	**Q**uorum **S**ensing
RBP	**R**NA-**B**inding **P**rotein
RNAP	**RNA P**olymerase
RR	**R**esponse **R**egulator
sRNA	**S**mall non-coding regulatory **RNA**
T3SS	**T**ype **3** **S**ecretion **S**ystem
TCS	**T**wo-**C**omponent **S**ystem
TF	**T**ranscription **F**actor
TPS	**T**wo-**P**artner **S**ecretion
TU	**T**ranscriptional **U**nit

Forewords

Pseudomonas aeruginosa is a human opportunistic pathogen and a leading cause of nosocomial infections. This bacterium is present ubiquitously in nature and is capable of adapting to a wide array of different conditions. In order to sense and respond to so many environments, *P. aeruginosa* possesses one of the most complex bacterial regulatory network. However, most of this network is still uncharacterized. The major objective of this thesis was to study several aspects of regulation in *P. aeruginosa* and related species in order to better understand the physiology and adaptive capacity of this pathogen.

When I arrived in the lab of Ina Attrée as a Master student in 2017, I joined Sylvie Elsen's subgroup, working on genetics and gene regulation in *P. aeruginosa*. At that time, a new virulence factor - the ExlBA two-partner secretion system - had been discovered in the lab three years before and Sylvie and her PhD student, Alice Berry, were studying its regulation. Alice discovered the first regulator of *exlBA*, Vfr, which activates this operon in a cAMP-dependent manner. I participated in this work which led to a publication in 2018 (Section 4.1.1).

For my own project, I aimed at characterizing additional regulators of *exlBA* and used a genome-wide screening approach to find a second transcription factor involved in *exlBA* regulation, ErfA. After its identification at the end of my internship, I continued with the characterization of ErfA during my PhD. ErfA was found to be a strong inhibitor of *exlBA*, counteracting Vfr activation of the operon. I also studied ErfA and ExlBA in other related *Pseudomonas* species and found that both *exlBA* regulatory pathways that we identified are specific to *P. aeruginosa*. Additionally, each species seemed to have different regulation for this operon, revealing a mechanism of evolution of regulatory circuits through promoter regulatory sequence diversification. This project led to a publication in 2020 (Section 4.1.2). To bring that new knowledge into context and compare with other systems and bacteria, we also reviewed the common and specific mechanisms of regulations of two-partner secretion systems across bacterial species, for which a manuscript is in preparation (Section 4.1.4).

After thoroughly studying the ErfA-*exlBA* regulatory interaction, there still was unanswered questions about ErfA main function and other regulatory targets. To better understand these and also bring knowledge on so far uncharacterized regulators, I then studied the family of 8 transcription factors that shared ErfA protein architecture. I found that these regulators dictate the expression of small numbers of genes, usually located around their own genes and mainly involved in metabolism. They thus act as neighbor specialized regulators that probably sense and respond to signals specific to the regulated metabolic pathway. Transcription factors of this family were found across the entire *Pseudomonas* genus with different numbers of such regulators per species, potentially reflecting different metabolic capacities. This work led to the redaction of a manuscript currently in revision for publication (Section 4.3).

In an attempt to globally increase our knowledge on *P. aeruginosa* regulatory network, we developed a project on the characterization of 51 transcription factors of the two-component system family in collaboration with two platforms at IBS in Grenoble, France. Indeed, in my previous projects I set up and used a DAP-seq approach for the genome-wide characterization of transcription factors. In this work we expressed and purified all 51 proteins that I then used to determine their genome-wide binding sites in three different *P. aeruginosa* strains representing the three main phylogenetic lineages. So far, I have performed 312 DAP-seq for this project for which I am currently analyzing the results (Section 4.4).

In several of my projects, I have looked at conservation of regulatory interactions between strains or species, with the notable example of *exlBA* divergent regulation between *Pseudomonas* species due to differences in its promoter sequences. In order to assess how common is this mechanism of promoter diversity, I studied conservation of all promoters across the *Pseudomonas* genus. I found that a large proportion of genes exhibit different contents of both putative and known transcription factor binding sites in their promoters between closely related strains or species (Section 4.5). This result highlights the need for further experimental characterization of this mechanism.

While most of my work focused on transcriptional regulation, I also studied post-transcriptional regulation, notably during my 3-months stay in the lab of Stephen Lory at Harvard Medical School in Boston, USA in the summer of 2019. For

this project I studied the major regulatory RNA-binding protein Hfq and assessed its role during growth in the three major *P. aeruginosa* phylogenetic lineages. This allowed the identification of several strain-specific and common regulatory pathways in important processes such as metabolism and virulence and highlighted once again major regulatory differences between strains of the same species. The corresponding manuscript is currently in preparation (Section 5.1).

Introduction

Chapter 1

Pseudomonas aeruginosa

Contents

Pseudomonas aeruginosa was first discovered in 1882 by the french pharmacist Carle Gessard while he was trying to understand the origin of the blue color sometimes observed on infected war wounds (Gessard, 1882). During the 20th century, this bacterium increasingly drew the interest of the scientific community as it began to be better characterized and found to be responsible for many infections (Haynes, 1951; Bodey et al., 1983). More recently, the genome of *P. aeruginosa* main reference strain, PAO1, was among the first bacterial genomes to be sequenced, which dramatically boosted the research on this pathogen (Stover et al., 2000). Today, *P. aeruginosa* is considered as a global health issue, notably due to a often high antibiotic resistance (Strateva and Yordanov, 2009; Poole, 2011), and a major cause of nosocomial infections (Obritsch et al., 2005; Driscoll et al., 2007).

P. aeruginosa is a gram-negative bacterium found ubiquitously in nature. It is famous for its adaptability and ability to thrive in a wide range of environments (Moradali et al., 2017; Diggle and Whiteley, 2020). To be able to do so, *P. aeruginosa* strains possess relatively large genomes (5.5-7 Mbp) with one of the highest number of regulatory genes (Cases et al., 2003; Winsor et al., 2016) (http://www.pseudomonas.com). This complex regulatory network orchestrates the bacterium's considerable metabolic capacity; it can use numerous different carbon sources and encompasses many pathways for secondary metabolites and polymers (Mathee et al., 2008; Frimmersdorf et al., 2010). While widely found in the environment, *P. aeruginosa* also excels at being a pathogen to a variety of different hosts, including humans, due to its versatility and an important array of potent virulence factors (Moradali et al., 2017).

Today, *P. aeruginosa* is one of the most studied bacterial species, with more than 5,000 sequenced genomes (Winsor et al., 2016). This remarkable organism has been the subject of thousands of studies worldwide and continues to amaze and trouble the scientific and medical communities.

1.1 The *Pseudomonas* genus

The *Pseudomonas* genus is one of the most diverse bacterial genera, encompassing more than 200 species that have been isolated worldwide (Madigan et al.,

2017). These gram-negative gammaproteobacteria are defined by their impressive metabolic diversity and adaptability which allow them to thrive and colonize all sorts of environments, from crops, soil, water - both fresh and salt -, cities to even clouds (Romero and Karp, 2003; Batrich et al., 2019; Uğur et al., 2012; Joly et al., 2013). Among their specific characteristics, most of the *Pseudomonas* species secrete pyoverdine, a siderophore making these bacteria fluorescent (Cézard et al., 2015). Due to their large physiologic diversity, *Pseudomonas* species are able to infect a wide variety of plants, invertebrates and animals, including humans, which leads to many economical and health repercussions. On the other hand, some species are also known to protect and promote plant growth (Preston, 2004). The clear phylogenetic definition of the genus proved challenging over the last century but is now more established (Peix et al., 2009; Gomila et al., 2015). Several lineages encompassing groups of strains represented by few major species can be delineated (Figure 1.1). Among these are the well characterized *P. aeruginosa*, *P. stutzeri*, *P.*

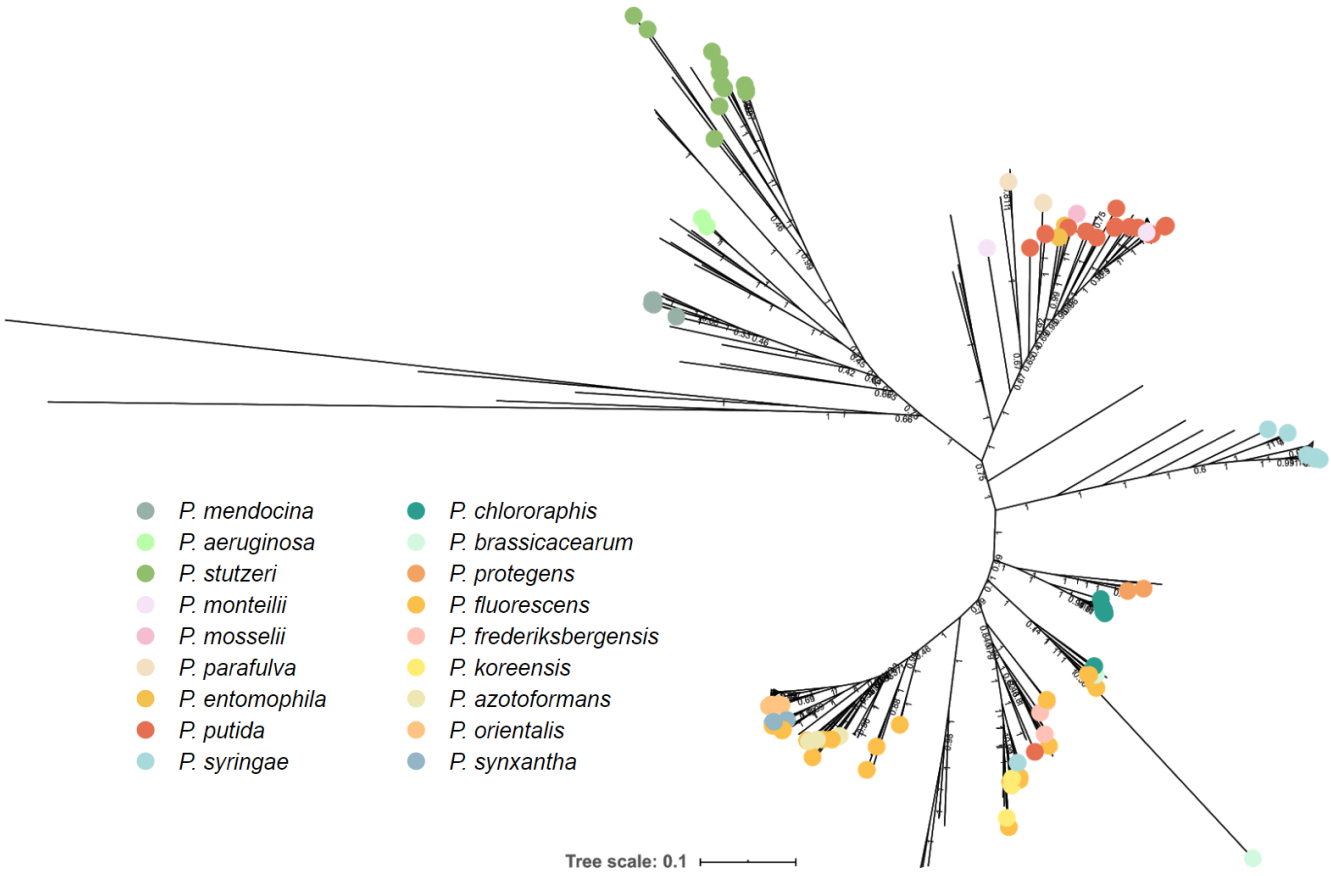


Figure 1.1: Phylogenetic tree of the *Pseudomonas* genus. Maximum-Likelihood phylogenetic tree of 503 *Pseudomonas* complete genomes. The tree was generated from the multiple alignment of the concatenated sequences of 66 core genes for each strain with 100 bootstraps.

putida, *P. syringae*, *P. chlororaphis*, *P. protegens* and *P. fluorescens*, each known for their specific features described below. *Pseudomonas* genomes are relatively large, with an average size of 5.63 Mbp, which reflects their important versatility (Figure 1.2) (Hesse et al., 2018; Cases et al., 2003). They also exhibit high GC contents, with a genus average of 61.2% (66.6% on average for *P. aeruginosa*), which has been linked to larger genome sizes and specific types of DNA polymerase III α subunits (Wu et al., 2012). *Pseudomonas* species are genetically highly diverse, with a genus core genome recently determined as less than 2% of the pan-genome (Yi and Dalpke, 2020). Among the only few genes functionally conserved in all species are genes involved in resource usage with broad substrate specificity and bacterial motility regulation, which are both instrumental in environmental adaptation.

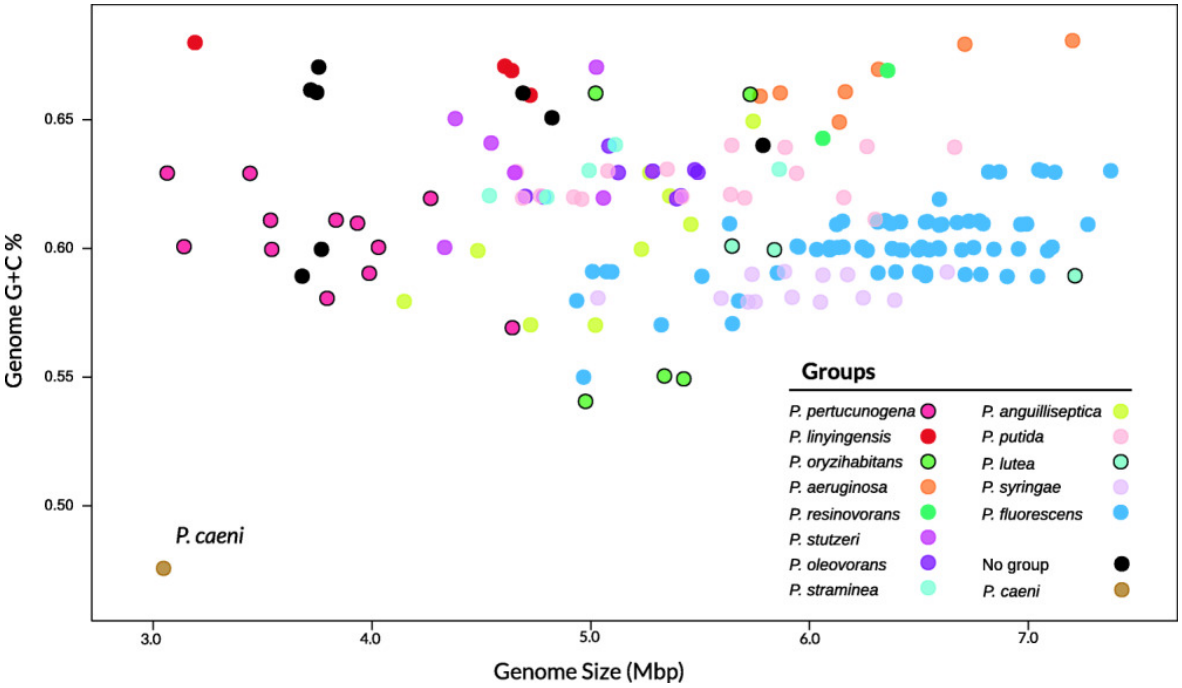


Figure 1.2: Relationship of genome size and GC content across 166 strains of the *Pseudomonas* genus. Figure from (Hesse et al., 2018).

1.1.1 Pathogenicity of *Pseudomonas* species

Pathogenicity is one of the most investigated traits of some *Pseudomonas* species, due to its health and economical impact. In this regard, *P. aeruginosa* stems as an excellent pathogen, being able to infect a wide array of hosts, including amoeba, plants, animals and humans (Pukatzki et al., 2002; Rahme et al., 2000;

De Bentzmann and Plésiat, 2011). This opportunistic pathogen is able to colonize most tissues in humans and is a leading cause of nosocomial infections, notably in the context of cystic fibrosis (Lyczak et al., 2000); its clinical importance will be further addressed in section 1.3. Another important *Pseudomonas* pathogen is *P. syringae* (Xin et al., 2018). This highly complex species regroups more than 60 phytopathogenic variants, being able to infect as many different plant hosts, including nearly all economically important crop species (Bull et al., 2010). Incidentally, it has been the cause of major crop disease outbreaks leading to large economic spin-offs in the past (Vanneste, 2017; Scortichini and Troplano, 1994). To be such a successful pathogen, *P. syringae* relies on several virulence factors, including the Type III Secretion System (T3SS) which allows the secretion of toxin effectors directly into host cells (Jin et al., 2003). There have been more than 50 families of T3SS effectors identified in *P. syringae* (Lindeberg et al., 2012). This huge diversity is thought to come from molecular co-evolution of plant host immune systems and bacterial effector proteins and it illustrates the wide array of plants infected by *P. syringae*. Other major virulence factors include the IaaL protein involved in the inactivation of auxin, a plant growth hormone, ice-nucleation proteins and extracellular polysaccharides (Ichinose et al., 2013). This large arsenal is believed to have been acquired by pathoadaptation of a nonpathogenic *P. fluorescens*-like ancestor to plant environments over time, leading to the development of a highly versatile and thriving plant pathogen (Xin et al., 2018). Several other less studied species are infectious, such as the insect- or fish-pathogens *P. entomophila* and *P. plecoglossicida*, respectively (Vodovar et al., 2006; Park et al., 2000).

1.1.2 Biocontrol, bioremediation and biotechnological applications

While some species might be highly pathogenic, many *Pseudomonas* species are harmless or even contribute to plant health. Bacteria from the *Pseudomonas* genus can either live inside plants as endophytes or in the rhizosphere and influence plant development by either directly promoting plant growth or by protecting plants from pathogenic microorganisms (Haas and Défago, 2005; Eljounaidi et al., 2016). This latter biocontrol phenomenon often involves the secretion by the bacteria of secondary metabolites with antimicrobial properties against phytopathogens (Mishra and Arora, 2018). These antibiotic compounds help clearing the rhizosphere from pathogens and consequently protect plants from root infection. It

has also been proposed that biocontrol agents can promote plant growth by a mechanism called induced systemic resistance (ISR) (Pieterse et al., 2014). ISR induced by beneficial bacteria often involves regulation of plant hormones such as salicilic acid and leads to the activation of immune defenses that protect plants from future infections. Other mechanisms beneficial to plants include bacterial-mediated nitrogen fixation, phosphate solubilization and production of plant hormones such as the growth-inducing auxin (Glick, 2012). *Pseudomonas* species have been reported to be involved in all of the above mechanisms, notably *P. fluorescens*, *P. protegens*, *P. putida*, *P. chlororaphis* and *P. stutzeri* (David et al., 2018; Patten and Glick, 2002; Anderson and Kim, 2018; Pham et al., 2017). Some of these species are even used as biocontrol agents in agriculture to improve crop yields (Carlier et al., 2008; Panpatte et al., 2016).

Another interesting feature of some *Pseudomonas* species is their potential in bioremediation, or the degradation or detoxification of hazardous compounds such as heavy metals, pesticides or phenolic contamination. Several species have been found suitable for use in bioremediation, including *P. aeruginosa*, *P. fluorescens* and *P. putida* (Wasi et al., 2013).

Bacteria from the *Pseudomonas* genus exhibit remarkable adaptive abilities, as illustrated by all of the above-mentioned niches they can colonize. These versatility and complexity make them great chassis for the biotechnological development of synthetic biology tools (Nikel et al., 2014). Additionally, the large range of molecular tools available for the genetic manipulation of *Pseudomonas* strains also directed the choice of the scientific community towards these bacteria. The *P. putida* KT2440 strain stemmed as a particularly interesting candidate due to its high physiological robustness and metabolic versatility and is now largely used for different types of biotechnological applications, including chemical synthesis (Martínez-García and Lorenzo, 2019).

1.2 The genomes of *P. aeruginosa* strains

The genome of the *P. aeruginosa* strain PAO1, which encompasses 5,572 genes, was the first *Pseudomonas* genome to be sequenced in 2000 (Stover et al., 2000). This milestone brought the first insights on the species vast genomic complexity, which

is now well characterized due to the thousands of sequenced genomes available (Winsor et al., 2016). *P. aeruginosa* strains possess large genomes with sizes ranging between 5.5 and 7 Mbp, reflecting a large intra-species diversity.

1.2.1 The core and pan-genomes

In the last two decades, many studies have compared and analyzed *P. aeruginosa* genomes in order to delineate the core set of genes shared by all members of the species (Valot et al., 2015; Ozer et al., 2019). Recently, the core genome of *P. aeruginosa* was defined as a group of 665 genes, representing 1.2% of the entire pan-genome (Freschi et al., 2019). This set of genes spans mostly housekeeping functions, including RNA processing, chromatin structure, cell division and metabolism (Figure 1.3). On the other hand, the accessory genome is enriched in mobile elements, secondary metabolic pathways and secretion systems. Moreover, about 30% of dispensable genes have unknown functions, while these represent only 10% of the core genome. Indeed, this accessory genome is known to be constantly reshaped by several evolutionary mechanisms, including horizontal gene transfer (HGT) or selective genome reduction (Kung et al., 2010), leading to important genetic mixing. HGT plays a key role in *P. aeruginosa* physiological diversification, especially regarding accessory functions such as virulence and antibiotic resistance (Qiu et al., 2009; Botelho et al., 2019). This often involves transferable mobile genetic elements such as transposons, genomic islands, phages and plasmids. It was found that 7.8 and 18% of *P. aeruginosa* pan-genome come from phages and plasmids, respectively (Freschi et al., 2019), illustrating the major contribution of these two mobile elements to accessory genome diversification.

1.2.2 The different *P. aeruginosa* lineages

The delineation of the core and pan-genomes comes concomitantly with phylogenetic analyses allowing the identification of several groups of *P. aeruginosa* strains (Freschi et al., 2015; Freschi et al., 2019). Three major groups are identified among *P. aeruginosa* strains, each represented by their reference strain, PAO1, PA14 and PA7 (Figure 1.4). The PAO1- and PA14-like are the two major phylogenetic groups, with their corresponding reference strains being the two first strains of *P. aeruginosa*

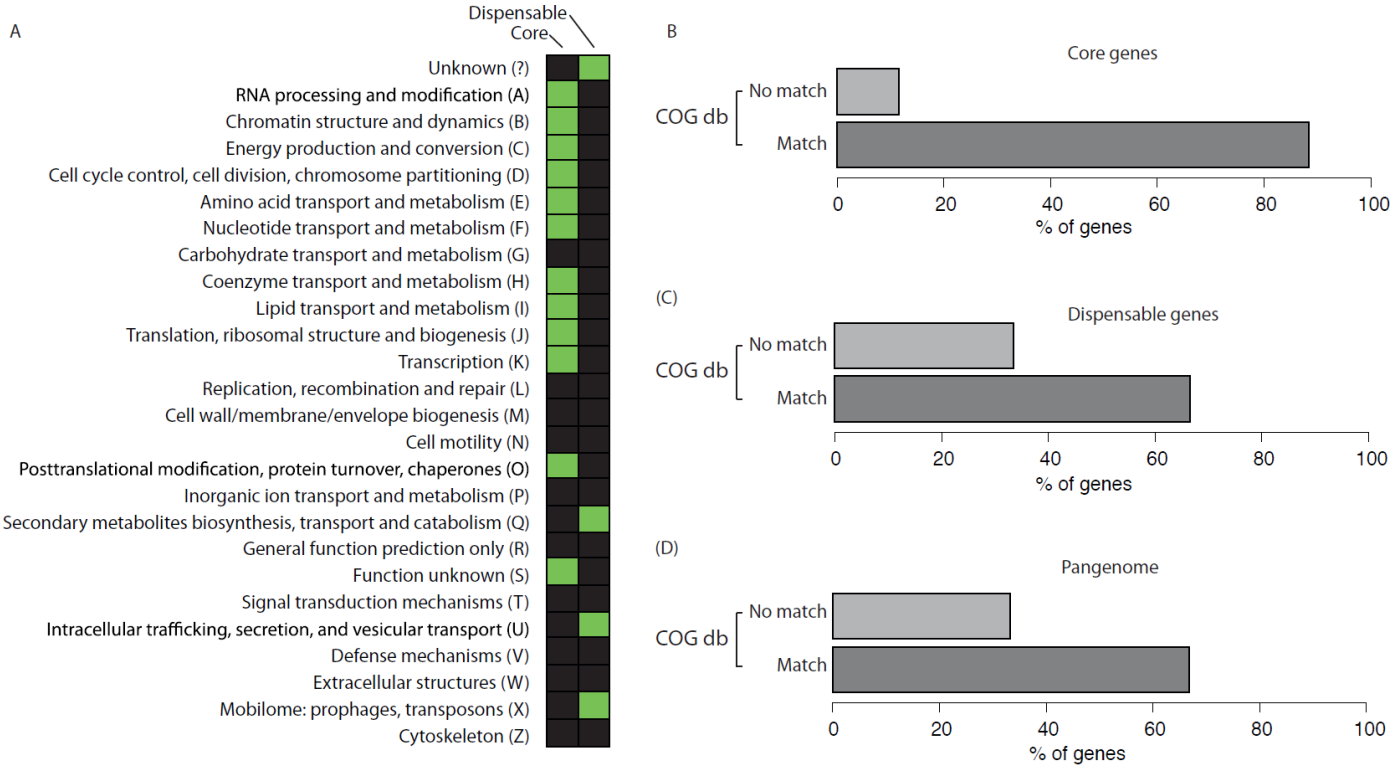


Figure 1.3: Functional analysis of core and pan-genomes of *P. aeruginosa*. (A) Functional enrichment analysis for core and dispensable genes. The green color indicates significant enrichment. (B-D) Percentage of genes that have a match in the COG database for core genes (B), dispensable genes (C) and all genes (D). Figure from (Freschi et al., 2019).

to be fully sequenced (Stover et al., 2000; Lee et al., 2006). While representing two distinct lineages, more than 95% of PAO1 genes are also found in PA14. Major differences between the two include a 1.7 Mbp inversion in the PAO1 genome and differences in virulence factors. Particularly, both PAO1- and PA14-like lineages possess the T3SS and the associated ExoT and ExoY toxins, but except for a few exceptions, they also each harbor lineage-specific toxins, namely ExoS and ExoU, respectively (Sawa et al., 2014b; Ozer et al., 2019). The molecular function of T3SS and T3SS-related toxins will be further discussed in section 1.4.1. The *exoU* gene is part of the PAPI-2 genomic pathogenicity island, along with the *spcU* gene which encodes the ExoU chaperone, and it is thought to have been obtained in the PA14-like lineage by plasmid-mediated HGT (Kulasekara et al., 2006). On the other hand, the *exoS* gene is always found between *spcS*, encoding ExoS and ExoT chaperone, and *PA3840* in PAO1-like strains. In PA14-like strains, *exoS* is thought to have been lost due to a targeted deletion event (Kulasekara et al., 2006).

The third phylogenetic lineage, which encompasses the PA7-like strains, is much

more distant from the other two (Figure 1.4). Indeed, while the other two lineages share average nucleotide identity (ANI) values of more than 98%, reflecting high similarity, the PA7-like strains are rather distantly related to the rest of the species, with ANI values of 93-94%, and are thus considered as taxonomic outliers (Freschi et al., 2019). The first identified and representative member of this group, the PA7 strain, has been identified and studied in 2010 for its antibiotics multiresistance phenotype (Roy et al., 2010). Since then, dozens of other PA7-like strains have been identified but they still represent only around 2% of all sequenced *P. aeruginosa* strains. However, there is still a bias in the strains that are sequenced, as they often come from cystic fibrosis patients, which might select for a specific type of strains. For instance, several PA7-like strains were isolated in the environment (Reboud et al., 2016) and environmental strains represent only a small proportion of

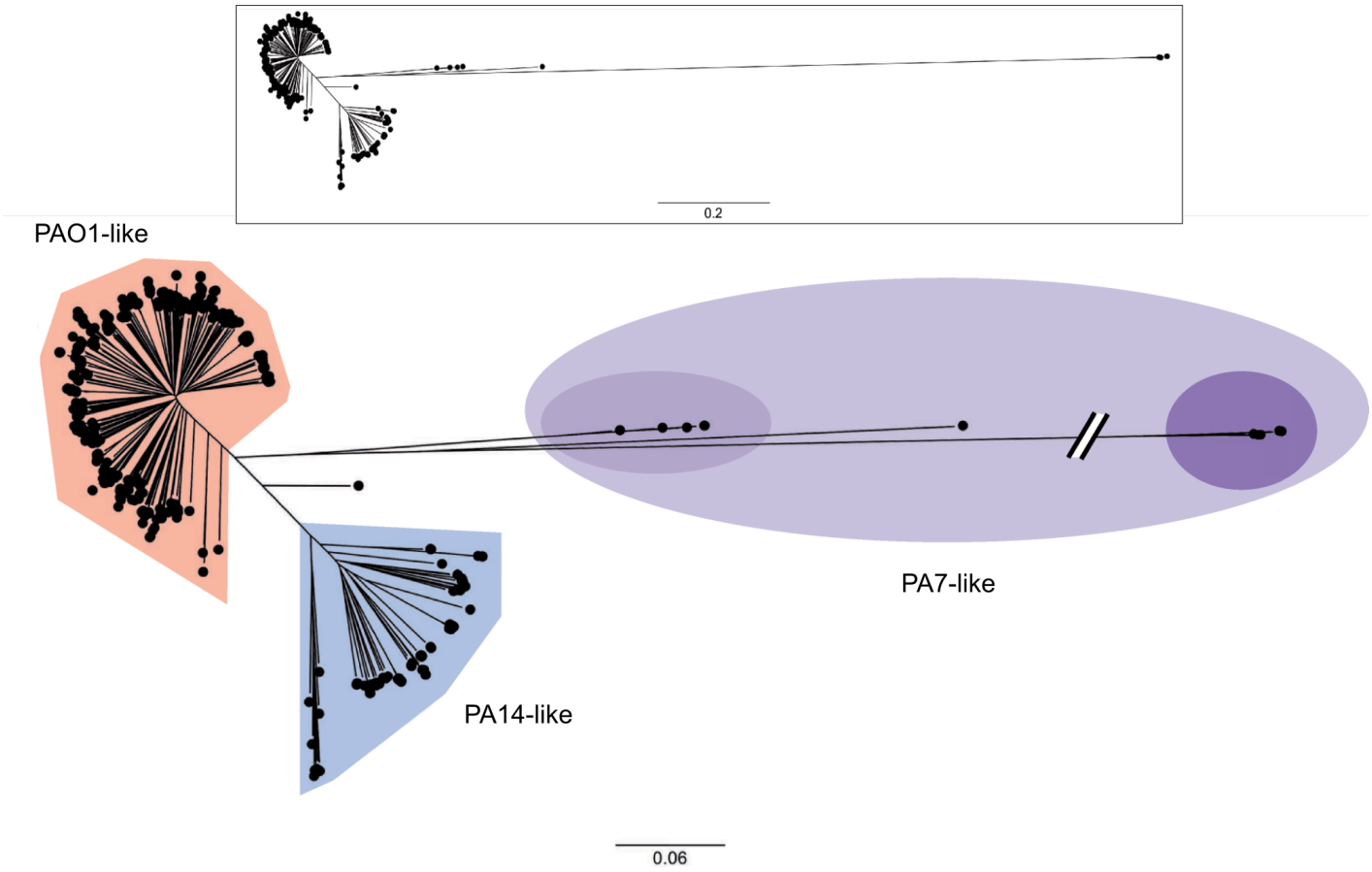


Figure 1.4: Phylogenetic tree of *P. aeruginosa* isolates. The maximum likelihood phylogenetic tree was generated from core genome SNPs in 739 isolates. The different phylogenetic groups of isolates are highlighted by colored areas. PA7-like strains are composed of two subgroups, one phylogenetically closer to and one very distant from the PAO1- and PA14-like strains. The smaller tree on the top of the panel shows the actual distances between groups. Figure adapted from (Ozer et al., 2019).

sequenced genomes, which means that the frequency of PA7-like strains might be overlooked. Among the major differences of this group is the absence of the entire set of 36 genes encoding the T3SS, as well as all genes coding for the associated toxins. Strikingly, all T3SS-related genes were excised from PA7-like genomes, even though they are present at different genomic locations. Two subgroups can be delineated among PA7-like strains (Figure 1.4) (Reboud et al., 2016), and notably differ by their specific *T3SS* deletion events. Indeed, while the deletion of the *T3SS* main locus is total for the more distant subgroup, PA7-like strains phylogenetically closer to the two other lineages harbor an apparent genetic scar of the deletion, containing fragments of some deleted genes. On the other hand, PA7 exhibits 18 novel genomic islands not present in PAO1 or PA14, including four putative prophages. Among their lineage-specific genes are the Type II and V Secretion Systems Txc and ExlBA, respectively (Cadoret et al., 2014; Elsen et al., 2014). ExlBA has since stemmed as a major virulence factor of *P. aeruginosa* PA7-like strains and even other *Pseudomonas* species (Huber et al., 2016; Reboud et al., 2017b; Basso et al., 2017a), and will be further discussed in section 1.4.2.

1.3 Clinical importance of *P. aeruginosa*

Pseudomonas aeruginosa is a renowned opportunistic pathogen and has been extensively studied because of its apparent clinical importance. It is a leading cause of nosocomial infections (Driscoll et al., 2007), often exhibiting high antibiotic resistance (Bassetti et al., 2018), and has been listed among the three priority bacterial pathogens for which the development of new treatments is urgently needed (along with *Acinetobacter baumannii* and bacteria of the *Enterobacteriaceae* family, such as *Klebsiella pneumoniae* and *Escherichia coli*) by the World Health Organization in 2017.

1.3.1 Infections caused by *P. aeruginosa*

P. aeruginosa is known for respiratory infections, notably in the lungs of cystic fibrosis or chronic obstructive pulmonary disease patients, but it also induces many other types of infection such as infections of wounds, burns or urinary tracts, bacteremia, otitis and keratitis (Gellatly and Hancock, 2013). This opportunistic pathogen also often induces infections in hospitalized or immunocompromised patients. Notably, *P. aeruginosa* infections in clinical settings are frequent following

surgery or the use of catheters or mechanical ventilation (Streeter and Katouli, 2016).

This pathogen is often described as having two different lifestyles (Valentini et al., 2018). In this perhaps-oversimplified view, *P. aeruginosa* is found either as planktonic, motile bacteria, or attached to a surface, forming communities called biofilms (Figure 1.5). In planktonic form, *P. aeruginosa* cells use a single flagellum in order to be motile, and also express many virulence factors such as the T3SS, inducing host cell death. In biofilm however, most of these factors are repressed, making *P. aeruginosa* cells much less aggressive. Cell attachment is helped by Type IV pili (T4P) which serve as adhesins, leading to aggregation associated with exopolysaccharides (EPS) production and biofilm formation. Bacteria are protected in biofilms, wich are matrices composed of exopolysaccharides, proteins and extracellular DNA, and are thus much less sensitive to the host immune system or antibiotics treatments (Høiby et al., 2010).

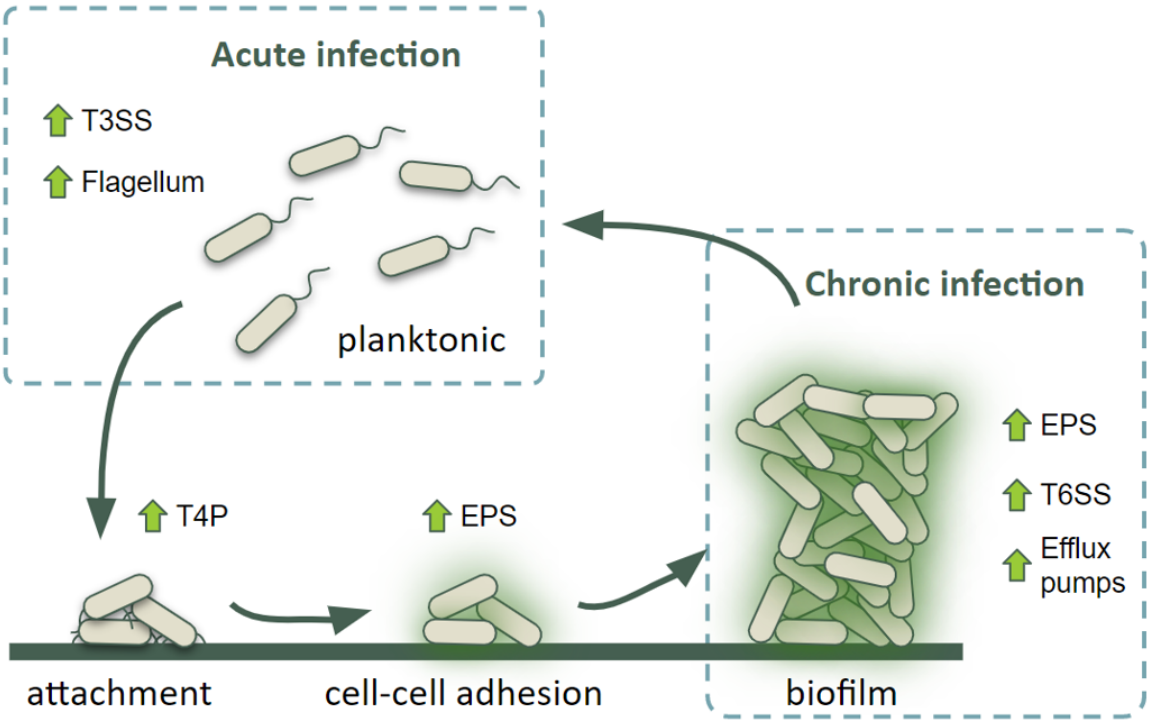


Figure 1.5: Schematic representation of the two different *P. aeruginosa* lifestyles. Upon detection of a surface, planktonic bacteria that are motile can attach to the surface via Type IV pili (T4P) and further aggregate through secretion of exopolysaccharides (EPS), leading to formation of biofilm structures.

Depending on its physiological state, *P. aeruginosa* will induce different types of infections. In its planktonic lifestyle, *P. aeruginosa* will cause acute infections, usually happening more dramatically and at shorter time scales, while the biofilm lifestyle is associated with chronic infections (Valentini et al., 2018). Chronic infections can last months, years or even decades, such as in case of cystic fibrosis patients where the vast majority of patients carry infections for their entire life (Gaspar et al., 2013). Cystic fibrosis is a recessive genetic disorder caused by mutations in the *CFTR* gene, leading to the inactivation, misfolding or absence of the CFTR anion channel (Davis, 2006). This causes the buildup of thick mucus in the lungs, which represents a thriving environment for bacterial pathogens, and especially *P. aeruginosa* which is the leading cause of death in cystic fibrosis patients. Consequently, cystic fibrosis patients suffer from lung infections even from early stages of their lives, reaching around 80% of adult patients with a *P. aeruginosa* lung infection, which explains their dramatically shortened life expectancy of around 40 years (Gaspar et al., 2013; MacKenzie et al., 2014). Many studies have focused on the adaptation of *P. aeruginosa* to the lungs of cystic fibrosis patients over years of infection and found interesting evolutionary features involved in the rewiring of regulatory networks and acquisition of antibiotics resistance (Yang et al., 2011; Folkesson et al., 2012; Winstanley et al., 2016). One major example of such evolution involves mutations in the quorum sensing regulator LasR, which has been notably linked to increased β-lactamase activity (Hoffman et al., 2009). Other such examples include mutations in the σ factor AlgU, an activator of alginate production, and its cognate anti-σ factor MucA, frequently found in an evolutionary study of cystic fibrosis clinical isolates (Marvig et al., 2015). These two mutations were shown to happen sequentially, with *mucA* mutations happening first and later acquisition of *algU* mutations, reflecting an evolutionary trajectory in which high mucoidity might serve at a precise time during infection and later needs to be abrogated. Other transcription factors were found mutated in clinical isolates in this study, such as the multidrug efflux pump regulator MexZ or the global regulator GacA (Matsuo et al., 2004; Brencic et al., 2009). Another interesting example of regulatory rewiring in clinical strains was recently published and involves mutations in intergenic regions (Khademi et al., 2019). Several of the selected mutations were found within regulatory elements inside of promoters. Some of these intergenic mutations were tested and found to induce change in expression of the neighbor genes, such as with *PA4837* which is involved in nickel and zinc uptake (Lhospice et al., 2017), revealing an overlooked mechanism of clinical evolution through promoter modification. All of which once again illustrates the impressive adaptability of *P. aeruginosa* to the

above-mentionned clinical contexts.

1.3.2 Antimicrobial resistance

Pseudomonas aeruginosa has emerged as an important cause of infections worldwide, partly due to its versatility and large panel of virulence factors, as already mentioned. Another major feature of this bacterium that makes it such a successful pathogen is its high antibiotic resistance (Lambert, 2002). In this regard, it strikes as particularly efficient due to (i) a large panel of intrinsic resistance mechanisms, (ii) its ability to acquire mutations improving existing mechanisms or new genetic material containing antibiotic resistance genes and (iii) adaptive resistance mechanisms leading to unstable, transient resistance phenotypes (Figure 1.6) (Blair et al., 2015; Arzanlou et al., 2017). These different mechanisms participate in the continuously increasing apparition of multi-drug resistant and sometimes incurable infections, which illustrates the urgent need for the development of new treatments (Medina and Pieper, 2016). While most projects focus on antibiotics development, a more recent approach involves the development of anti-virulence molecules which do not kill bacteria, and thus induces less resistance development, but improve clinical outcomes through inhibition of different virulence factors (Veesenmeyer et al., 2009; O'Loughlin et al., 2013; Sawa et al., 2014a). Another now more investigated option for the treatment of multiresistant *P. aeruginosa* is the use of bacteriophages for the specific killing of these bacteria, which proved effective in the few first clinical tests (Waters et al., 2017).

Among *P. aeruginosa* intrinsic antiobiotic resistance mechanisms are its ability to prevent entry of specific antiobitics due to the nature of its gram-negative cell structure and the absence of non-specific porins and transporters that could facilitate antibiotics import (Tamber and Hancock, 2003; Chevalier et al., 2017). Inducible β-lactamases such as AmpC are also widely found in *P. aeruginosa* and provide intrinsic low levels of resistance to β-lactam antibiotics (Upadhyay et al., 2010). This bacterium also possesses a large capacity of antibiotic efflux which considerably participates in its multi-resistance (Figure 1.6). The presence of four different actively expressed multi-drug efflux pumps allows the constant expelling of a wide range of antimicrobial compounds (Poole, 2001; Du et al., 2018).

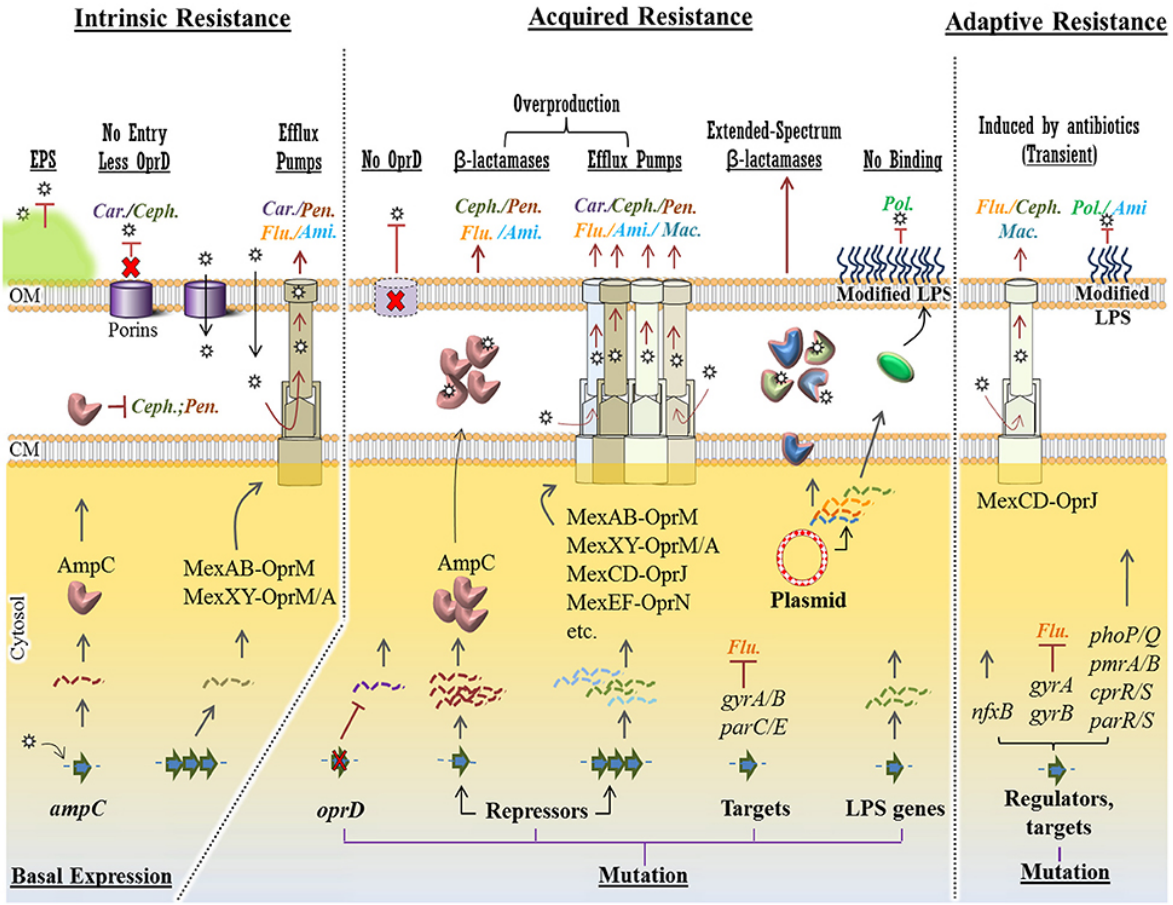


Figure 1.6: Intrinsic, acquired and adaptive antibiotic resistance mechanisms in *P. aeruginosa*. For each mechanism, various examples of strategies conferring antiobiotic resistance are presented. (Car., Carbapenems; Ceph., Cephalosporins; Pen., Penicillins; Ami., Aminoglycosides; Flu., Fluoroquinolones; Mac., Macrolides; Pol., Polymyxins; EPS, exopolysaccharides; CM, cytoplasmic membrane; OM, outer membrane). Figure from (Moradali et al., 2017).

While under the selective pressure of ongoing antibiotics treatment, *P. aeruginosa* is able to mutate or acquire new genetic material involved in the apparition of new antibiotic resistances (Figure 1.6) (Paramythiotou et al., 2004; Breidenstein et al., 2011). Once acquired, these traits are stable and transmitted during cell division, leading to the development of resistant bacterial populations. Many of such acquired mechanisms have been reported (Figure 1.6), including mutations in regulatory networks leading to over-expression of resistance genes (Marvig et al., 2015; Xavier et al., 2010; Cabot et al., 2011), inactivation of genes involved in antibiotics target or import (Pirnay et al., 2002; Gorgani et al., 2009), and acquisition of new genes such as with the exchange of genes encoding extended-spectrum β-lactams via plasmid-driven HGT (Poirel et al., 2001). Plasmids, and more generally

HGT, play a major role in antibiotic resistance acquisition and the prevalence of multi-resistance keeps increasing worldwide due to this mechanism (Hong et al., 2015). There already are plasmids carrying resistance genes for even the most recently developed antibiotics such as colistin, used as a last resort in multi-drug resistant infections (Liu et al., 2016).

During adaptive resistance, bacteria temporarily react to specific environmental conditions or antibiotics via different regulatory responses or reverting mutations, leading to a transient form of resistance, which fades away after the end of the stress (Khaledi et al., 2016; Lee et al., 2016). However, this mechanism is still poorly understood in comparison with the two others (Moradali et al., 2017).

Some *P. aeruginosa* infections are also believed to resist treatment through an antibiotic persistence mechanism. In these cases, persister bacteria that are transiently resistant to antibiotics appear inside of sensible populations (Mulcahy et al., 2008). During antibiotic treatment, this phenotypic heterogeneity phenomenon leads to the clearance of the vast majority of the population but not of the persisters. These persisters are thought to be metabolically inactive, rendering drug targets absent or temporarily dispensable, leading to this transient resistance phenotype (Maisonneuve and Gerdes, 2014). Several explanations have been proposed for this phenomenon, such as the activation of toxin-antitoxin systems (Page and Peti, 2016), ATP depletion (Conlon et al., 2016) or (p)ppGpp-mediated stringent response mechanisms (Germain et al., 2015).

1.4 *P. aeruginosa* virulence factors

P. aeruginosa possesses a vast array of virulence factors contributing to its high pathogenic potential. Lipopolysaccharides are major components of the outer membrane that activate host immune system (Raoust et al., 2009). The flagellum allows motility and is associated with the planktonic lifestyle, while type IV pili allow twitching as well as adhesion and thus permit the transition to biofilm formation (Barken et al., 2008). The production of alginate induces the formation of a mucoid exopolysaccharide capsule which helps cell adherence and resistance to host defenses (Leid et al., 2005). Several proteases, such as the LasB elastase, are also secreted and are involved in host tissue degradation (Casilag et al., 2016). Many major

virulence factors are secretion systems, and *P. aeruginosa* possesses a large number of them (Bleves et al., 2010). Among these is the Type II secretion system, notably involved in the secretion of the Exotoxin A, which induces host cell apoptosis through inhbition of protein synthesis (Wolf and Elsässer-Beile, 2009). The type III secretion system allows the direct secretion of toxins into host cells. Type V secretion systems includes autotransporters, such as the EstA esterase required for biofilm formation (Wilhelm et al., 2007), and two-partner secretion systems such as ExlBA (Elsen et al., 2014). In this section I will focus on two major ones, more related to the work presented below, as the entire arsenal has been previously extensively reviewed (Moradali et al., 2017; Conrad et al., 2011; Gellatly and Hancock, 2013; Wang et al., 2013; Filloux, 2011).

1.4.1 Type III Secretion System

The Type III Secretion System (T3SS) is a key bacterial secretion system involved in the virulence of many pathogens, including *P. aeruginosa*, *Salmonella typhimurium*, *E. coli*, or *Yersinia pestis* among others (Waterman and Holden, 2003; Mills et al., 2008; Brodsky et al., 2010). Probably being the most investigated bacterial secretion system, it has been the topics of thousands of studies and has even been leveraged as a molecular delivery tool for a large array of biotechnology and biomedical applications (Widmaier et al., 2009; Bai et al., 2018).

In *P. aeruginosa*, the T3SS has long been considered one of the major virulence factors and it is associated with highly virulent acute infections (Le Berre et al., 2011). This large needle-shaped machinery allows the injection of toxins directly into host cells (Hauser et al., 2002). It is present in the two main lineages of the species, which encompass the PAO1- and PA14-like strains (Freschi et al., 2019). The T3SS is composed of more than 30 proteins encoded at one major locus, except for the secreted toxins (Figure 1.7). The regulation of this large complex involves a specific regulation through the activator ExsA and the different proteins regulating its activity, ExsCDE, providing a secretion-synthesis coupling mechanism that activates transcription upon secretion activity (Yahr and Wolfgang, 2006). More global regulatory mechanisms impact T3SS expression and will be addressed in chapter 2. The T3SS is anchored at the bacterial membrane and forms a needle-like filament composed of repeated subunits of the protein PscF (Pastor et al., 2005). The translocation apparatus at the tip of the needle then inserts itself into the membrane of the host cell, forming a pore. This process involves the oligomerization of the

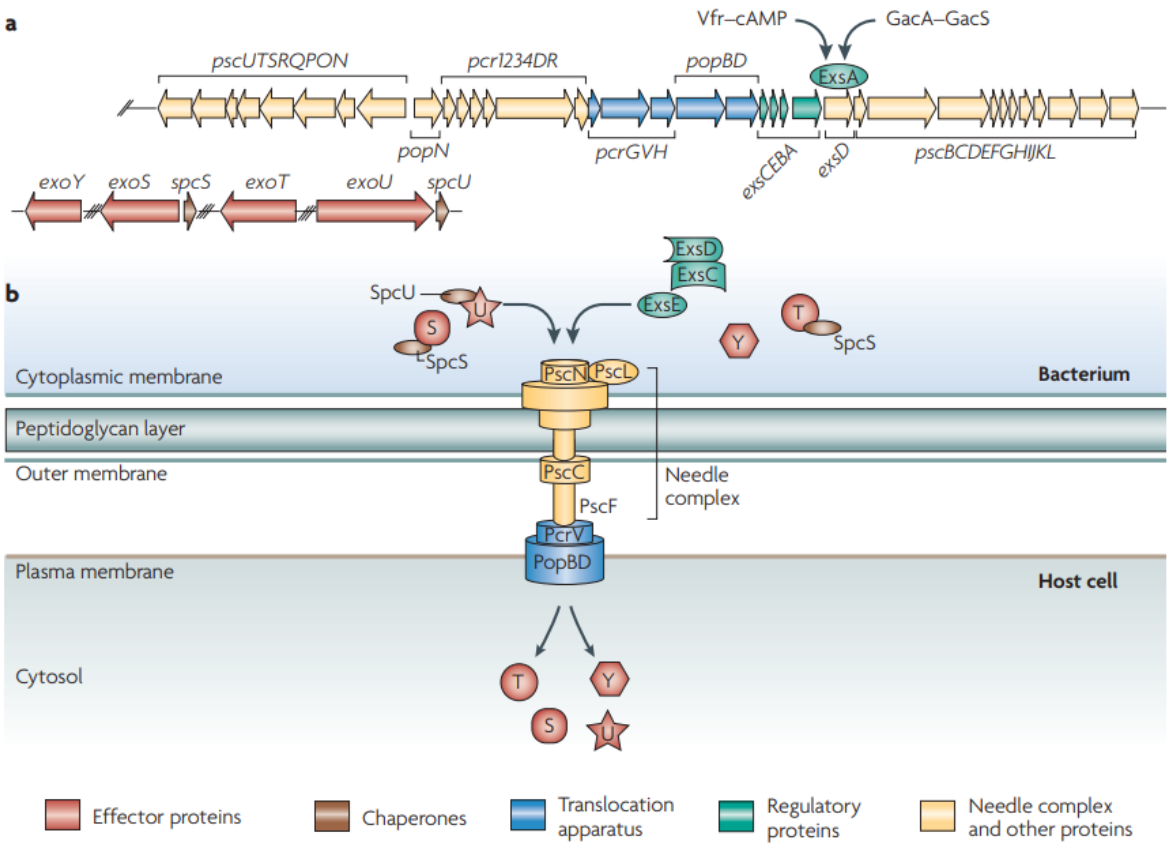


Figure 1.7: The Type III Secretion System in *P. aeruginosa*. (A) Genetic organization of the genes involved in T3SS. (B) Schematic view of the T3SS during injection into a cell. Figure from (Hauser, 2009).

translocator proteins PopB and PopD (Figure 1.7) (Schoehn et al., 2003; Sundin et al., 2004). Effectors are then secreted inside of the target cell. Four effectors have been identified in *P. aeruginosa*, the conserved ExoT and ExoY, and the two lineage-specific ExoS and ExoU, mainly found in PAO1- and PA14-like isolates, respectively (Hauser et al., 2002). Each of these effectors has a specific molecular function and participates to host cell death in different ways. ExoS and ExoT share 76% amino acid identity and are both bifunctional toxins with GTPase-activating and ADP ribosyl transferase domains. Once secreted, ExoS is addressed to the plasma membrane of the host cell, from where it goes to the endoplasmic reticulum and Golgi apparatus through endosome trafficking (Zhang et al., 2007). The toxin then inactivates the GTPases Rho, Rac and CDC42, leading to host cell actin cytoskeleton disruption (Würtele et al., 2001). While ExoS has been extensively studied, ExoT is less characterized but seems to have a similar mechanism of action (Shafikhani and Engel, 2006). However, its contribution to pathogenicity seems less important (Lee et al., 2005). ExoY is an nucleotidyl cyclase that induces an increase

of intracellular cGMP and cUMP concentrations after injection (Beckert et al., 2014). This then induces the disruption of cNMP-dependent cell signalling which alters cell physiology and eventually leads to cell death (Belyy et al., 2018). Finally, ExoU is a very potent phospholipase targeted to the plasma membrane where it cleaves membrane phospholipids, leading to rapid host cell death (Sato et al., 2003; Pazos et al., 2017).

1.4.2 ExlBA

The two-partner secretion (TPS) system (or Type Vb Secretion System) ExlBA has been recently identified in the PA7-like clinical isolate CLJ1 (Elsen et al., 2014). This TPS system is now known to be conserved across all PA7-like strains and absent in all others (Huber et al., 2016), as no strain possessing both the T3SS and ExlBA has been identified so far. It is composed of an outer membrane transporter, ExlB, which allows the specific secretion of a pore-forming toxin, ExlA (Huber et al., 2016). Both ExlB and ExlA are first transported into the periplasm through the general Sec machinery before integration of ExlB into the outer membrane (Figure 1.8A). Once secreted through ExlB, ExlA inserts into host cell membrane, leading to the formation of pores and eventually cell death. This pore-forming activity requires the cooperation of Type IV pili, which probably serve to adhere and promote close contact between bacteria and host cell, allowing the secretion of ExlA in the vicinity of the host cell membrane (Basso et al., 2017b). The ExlA-dependent formation of pores has been described in several cell types and induced different host cell responses (Figure 1.8B). Briefly, ExlA has been shown to induce pyroptosis in macrophages through activation of the inflammasome complex (Basso et al., 2017a). Indeed, the pore formation in the plasma membrane induces potassium ions efflux leading to the activation of the NLRP3/ASC inflammasome, which in turn activates caspase-1, eventually leading to cell death by pyroptosis. Another study showed that the formation of pores by ExlA induces a massive entry of calcium ions in epithelial and endothelial cells, leading to ADAM10 activation which in turn cleaves the cadherins involved in cell-cell junction, thus leading to epithelial or endothelial barrier disruption (Figure 1.8B) (Reboud et al., 2017b).

While being the major virulence factor of the *P. aeruginosa* PA7-like strains, ExlBA has also been found in several other *Pseudomonas* species (Basso et al., 2017a; Job et al., 2019). Among these are also non-pathogenic species, or plant or insect

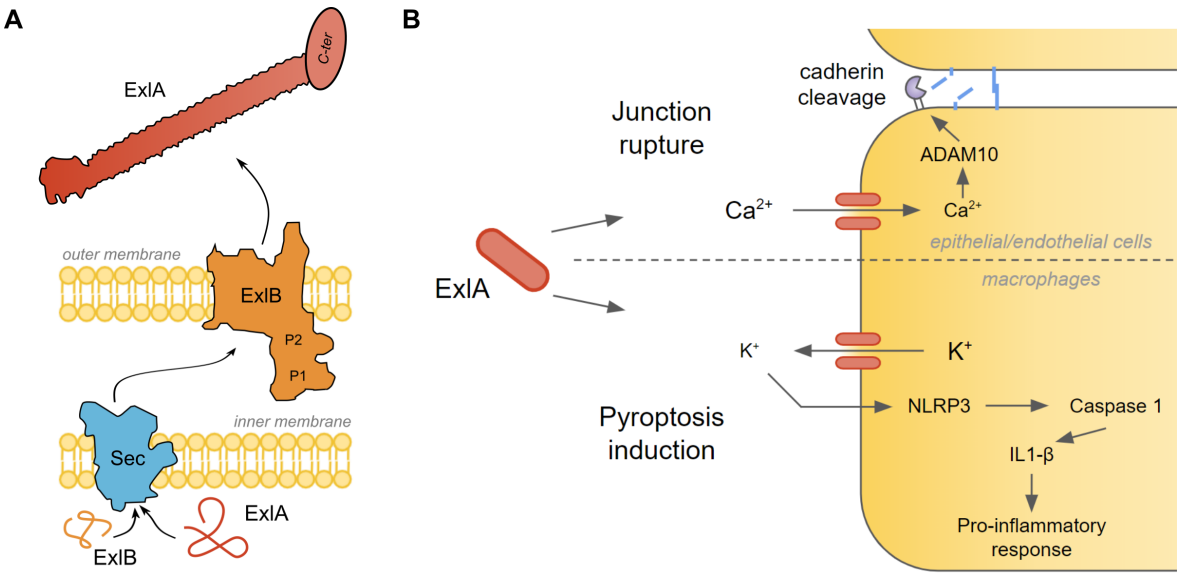


Figure 1.8: The ExlA secretion system and its effects on host cells. (A) Schematic representation of the ExlBA TPS system. Both ExlB and ExlA proteins are translocated into the periplasm by the Sec machinery before ExlB inserts itself into the outer membrane. P1 and P2: the two periplasmic POTRA domains of ExlB, which are necessary for ExlA secretion (Basso et al., 2017b). C-ter: Pore-forming domain of ExlA, which has no predicted structure (Reboud et al., 2017a). (B) Two modes of action of ExlA on host cells. ExlA-dependent pore formation induces junction disruption in epithelial and endothelial cells (top). Later on, Ca^{2+} entry also induces necrotic cell death through unknown mechanisms. ExlA-dependent activation of inflammasome leads to pyroptosis in macrophages (bottom).

pathogens, raising the question of ExlBA origin and function. The fact that the *exlBA* two-gene operon is scattered across the *Pseudomonas* genus where it is found in very different genetic environments implies its exchange by HGT which might explain its presence in very different bacteria.

Chapter 2

Transcriptional regulatory networks

Contents

In all kingdoms of life, transcription regulation has evolved as a fundamental mechanism allowing diversification and adaptation (Wray et al., 2003). While pinpointed in the early stages of molecular biology (Britten and Davidson, 1969), the importance of gene regulation has often been overlooked compared to the function or conservation of genes themselves when trying to understand physiological mechanisms and evolution. However, it is now well established that transcriptional regulation constitutes a pillar of phenotypic evolution (Carroll, 2008). The recent technological advances in large-scale methodologies now allow a better characterization of this mechanism and the different molecular partners involved (Di Iulio et al., 2018; Wittkopp and Kalay, 2012; Touzain et al., 2011; Shen-Orr et al., 2002).

2.1 Bacterial transcription regulation

While eukaryotes have much larger genomes, more genes, a highly structured chromatin, three RNA polymerases, and complex networks of *cis*-regulatory elements able to impact the transcription of genes even located tens of thousands of base pairs away (Castellanos et al., 2020; Sosinsky et al., 2007), bacteria tend to have simpler regulatory mechanisms. In this section, the key concepts of bacterial transcription regulation will be briefly reviewed.

2.1.1 Core promoters and sigma factors

Bacteria possess a single RNA polymerase (RNAP) responsible for the transcription of all classes of RNAs (Darst, 2001). To initiate transcription, the RNAP interacts with a σ factor, recognizing and binding to the core promoter elements, which eventually leads to transcription initiation (Figure 2.1) (Browning and Busby, 2004; Mejia-Almonte et al., 2020). The major σ factor RpoD drives the expression of housekeeping genes and recognizes the most characterized consensus core promoter sequence (TTGACA-N17-TATAAT) (Feklistov et al., 2014). However, a wide variety of σ factors can be found in bacteria, each having a different sequence specificity (Paget and Helmann, 2003; Helmann, 2019), allowing for large-scale regulatory changes when switching between them. Indeed, many σ factors are expressed or active only in specific conditions and thus serve as major response tools to different environmental stresses by reprogramming the expression of often hundreds or thousands of genes. In *P. aeruginosa* for instance, there are 24 different σ factors (Potvin et al., 2008). Most of them (19) are extracytoplasmic function

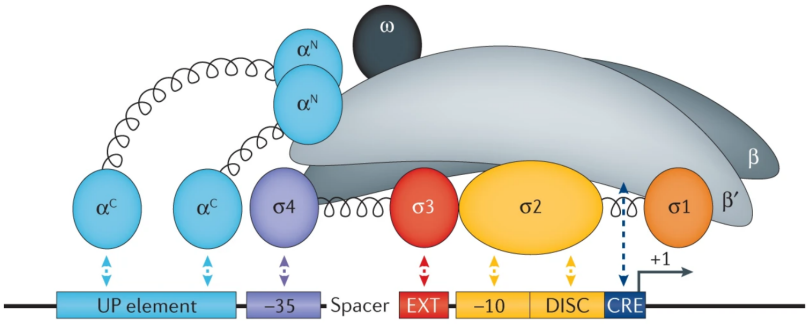


Figure 2.1: Schematic view of RNAP holoenzyme binding to promoter regions. The core RNAP is composed of five subunits: 2 α, β, β' and ω. The RNAP α subunits interact with the upstream promoter (UP) element located upstream of the -35 element. The RNAP core enzyme interacts in a sequence-specific manner with the template-strand positions -4 to +2, which constitute the core recognition element (CRE). The σ^{70}-related factors contain up to four functional domains (σ1-4). The σ2 domain recognizes and interacts with the -10 element, and the σ4 domain interacts with the -35 element. The extended -10 element interacts with σ3. Some promoters have an element called a discriminator (DISC), which is recognized by and interacts with the σ2 domain. Figure from (Mejia-Almonte et al., 2020).

(ECF) σ factors, which control the expression of subsets of genes in response to various extracellular changes. ECF σ factors are usually found associated with an outer-membrane receptor protein and an anti-σ factor (Heimann, 2002; Sineva et al., 2017). Upon sensing of an extracellular signal, the receptor protein induces the release of the σ factor from the anti-σ factor, allowing it to exert its regulatory function.

2.1.2 Transcription factors

Transcription factors can tweak RNAP recruitment by binding to different locations on or around the core promoter, leading to either repression or activation of transcription (Figure 2.2) (Browning and Busby, 2016; Balleza et al., 2008). Activators usually bind upstream of the core promoter to facilitate RNAP recruitment, while repressors usually bind inside of core promoter, onto the transcription start site or just downstream of it to prevent RNAP binding or transcription initiation. The pool of different TFs found in the cell at a given time thus defines the transcription rates of all genes and, consequently, the current physiological state. Numerous different types of TFs have been identified, each with their specific feature. Most TFs are considered as one-component systems (OCSs), as opposed to two-component systems (TCSs), that are able to directly sense a signal. OCSs possess DNA-binding and

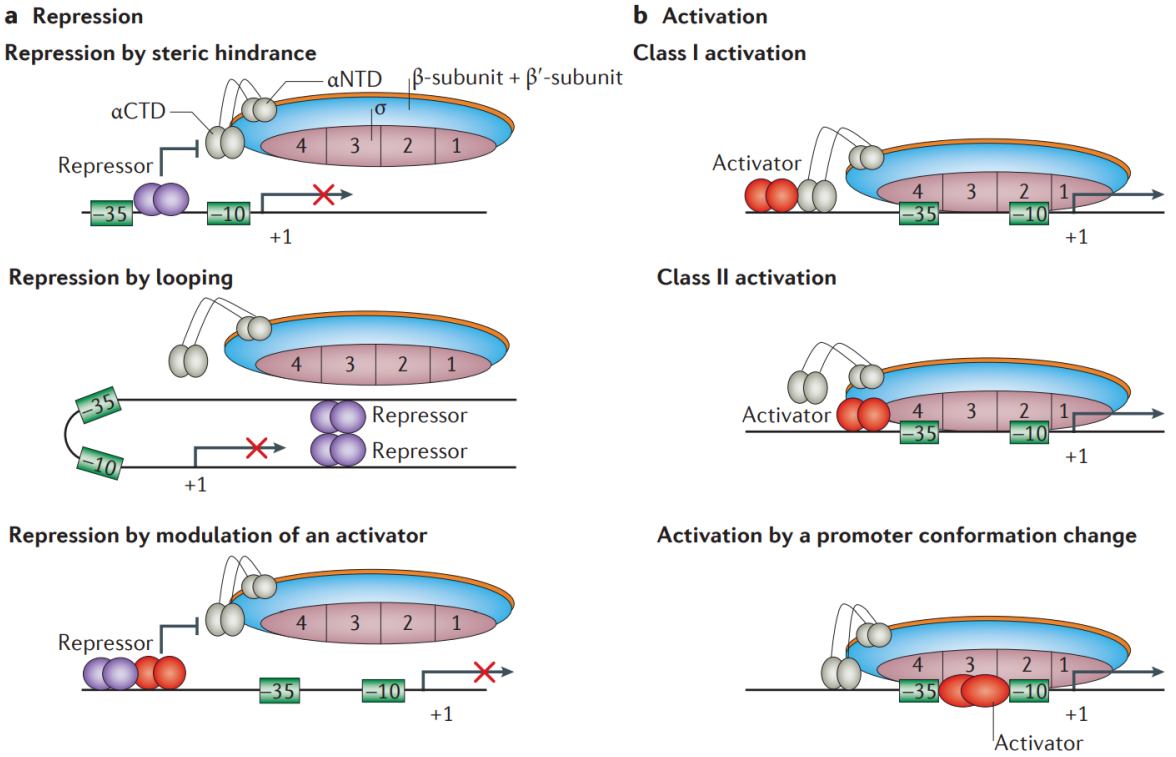


Figure 2.2: Regulation of transcription by transcription factors. (A) Mechanisms of transcription repression. TF can inhibit transcription by (i) steric hindrance by binding directly on core promoter elements, (ii) DNA looping induced by interactions between two repressors up- and downstream of promoter, or (iii) modulation of an activator to prevent RNAP recruitment. (B) Activation of transcription mechanisms. Direct recruitment of RNAP can occur through interaction of a DNA-bound activator with the RNAP α-subunit (i) or the σ factor (ii). Other mechanisms of activation involves a TF-driven DNA conformational change to realign promoter elements (iii). Figure from (Browning and Busby, 2016).

signal-sensing domains and their DNA-binding activity can consequently by modulated upon direct sensing of a signal, which allows appropriate regulatory responses to changing environments (Ulrich et al., 2005). TFs can also be sorted by their type of DNA-binding domain (DBD). There are a few major superfamilies that each regroups several families of DBDs based on their structural features (Table 2.1) (Perez-Rueda and Martinez-Nuñez, 2012). Notably, winged helix and homeodomain-like domains represent about 70% of all bacterial DBDs and encompass all major TF families such as LysR-like, AcrR-like, GntR-like or AraC-like. Winged helix DBDs are composed of three alpha helices and three beta sheets and bind the major groove of DNA in a sequence-specific manner (Gajiwala and Burley, 2000). Several DBD families are considered as homeodomain-like as they have sequence and structural similarities with homeodomains. Homeodomains represent another subgroup of helix-turn-helix DNA-binding domains and possess three alpha helices (Mannervik,

Table 2.1: List of DBD superfamilies in prokaryotes. Data from (Perez-Rueda and Martinez-Nuñez, 2012).

Superfamily	Proportion (%)	Families
Winged helix DNA-binding domain	44.18	CRP, LysR, GntR, AsnC, MarR
Homeodomain-like	26.3	AraC/XyIS, TetR/AcrR, Fis
Lambda repressor-like DNA-binding domains	10.55	GalR/LacI
C-ter effector domain of response regulators	9.68	OmpR, Spo0A, NarL
Putative DNA-binding domain	3.02	MerR
Nucleic acid-binding proteins	1.86	Cold shock DNA-binding domain-like
Putative TM1602, C-terminal domain	1.75	HTH_11
IHF-like DNA-binding proteins	0.72	IHF
AbrB/MazE/MraZ-like	0.71	AbrB N-terminal domain-like
Ribbon-helix-helix	0.19	Met repressor, MetJ (MetR), Arc/Mnt-ike phage repressors, CopG
TrpR-like	0.19	Trp repressor, TrpR
KorB DNA-binding domain-like	0.1	KorB DNA-binding domain-like
Flagellar transcriptional activator FlhD	0.08	Flagellar transcriptional activator FlhD
H-NS histone-like proteins	0.67	HN-S

1999). While only a few types of DBDs represent the majority of TFs, there are a large number of different TF families due to the numerous possible DBD associations with other domains (Ortet et al., 2012; Rivera-Gómez et al., 2017). Among these, several families are most represented in bacteria, including LysR-, AcrR-, GntR-, AraC- and Cro-like TFs (Figure 2.3). Out of these 5 families, 4 are families of OCSs that all possess a signal-sensing domain in addition to their DBD and the last one represents the monodomain family of Cro-like repressors. Notably, the LysR-like family of TF is the largest family of bacterial TFs and encompass OCSs that can act as both activators and repressors of genes involved in numerous cellular processes (Maddocks and Oyston, 2008). Many other families, that are smaller, exist and encompass as many different types of domain architectures.

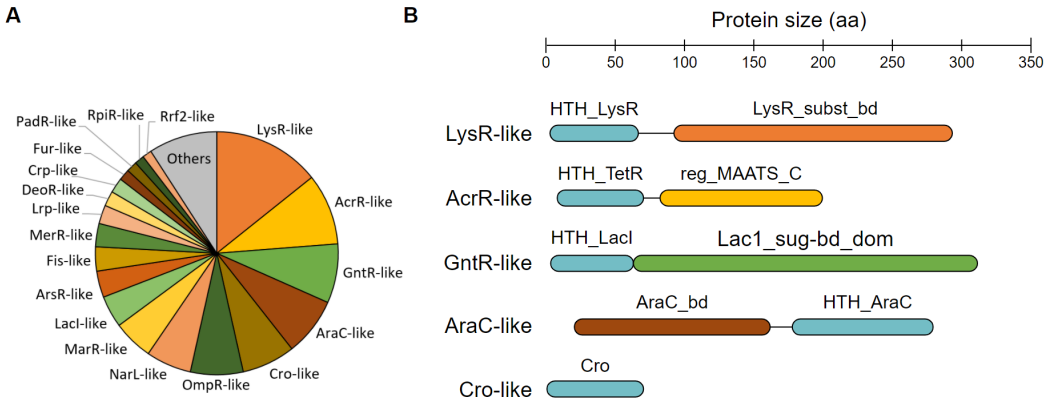


Figure 2.3: The most represented TF families. (**A**) Proportion of each TF family found in bacteria. Only families that represent at least 1% of all TFs are labeled. Data from (Minezaki et al., 2005). (**B**) Schematic view of protein domain architectures of the 5 most represented TF families. DBDs are shown in blue. Domains are named after their corresponding Interpro domain family IDs.

As opposed to OCSs, the input and output signals are transmitted between two different proteins in TCSs (Jacob-Dubuisson et al., 2018). In the majority of TCSs, a membrane-bound histidine kinase (HK) can activate its cognate response regulator (RR) by phosphoryl transfer upon sensing of an extracellular signal. The signal sensing by the HK induces autophosphorylation at a conserved Histidine residue in its transmitter domain. The phosphoryl group is then transferred to a conserved Aspartic acid residue in the receiver domain of the RR, which then induces a conformation change leading to RR activation. In classical TCSs, this signal transduction mechanism happens directly between the two proteins (Figure 2.4). More complex TCSs exist and involve phosphorelay mechanisms. Indeed, hybrid and unorthodox HKs possess a receiver domain in addition to their transmitter domain. A Hpt phosphotransfer module is thus necessary for RR phosphorylation and it is present either on an independent protein or as an integral part of unorthodox HKs (Figure 2.4) (Casino et al., 2010). Many different types of RRs exist, most of them are TFs

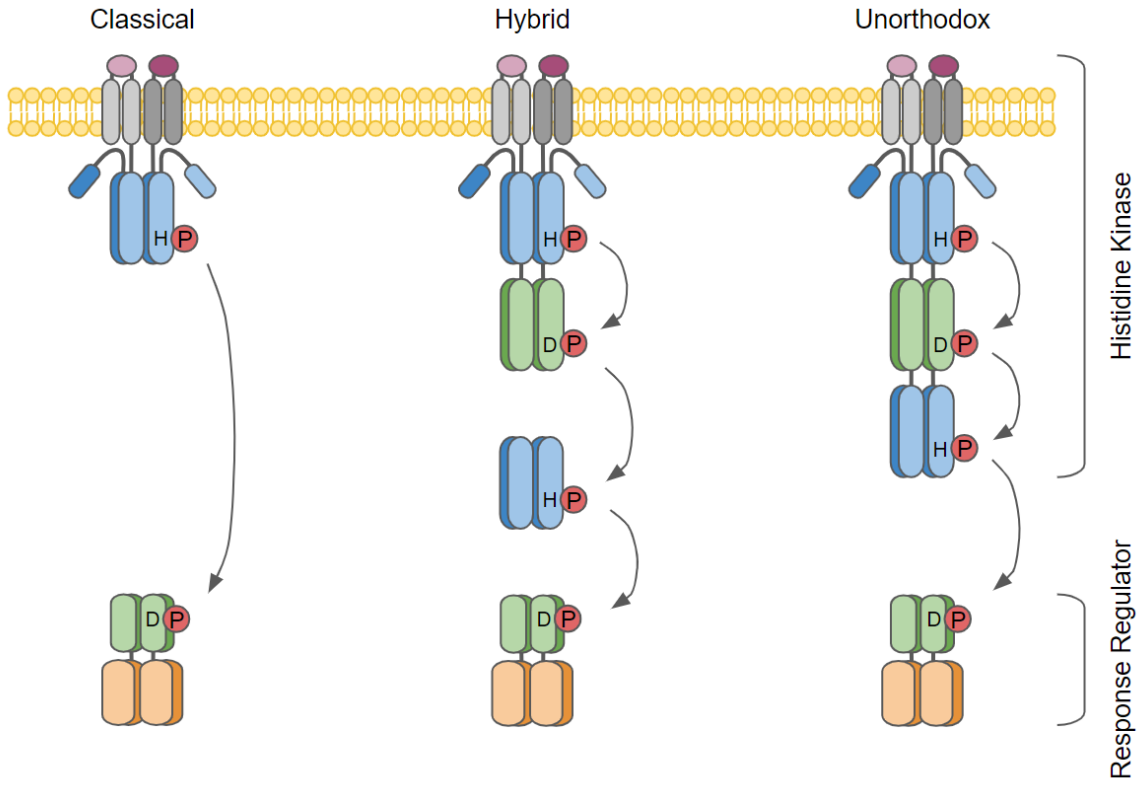


Figure 2.4: Structural organization of two-component systems. Histidine kinases are usually membrane-bound proteins and possess a signal-sensing domain (purple). Phosphorylation transmitter and receiver domains are shown in blue and green, respectively. The output response domain is shown in orange, it is a DNA-binding domain in most response regulators. H: Histidine, D: Aspartic acid.

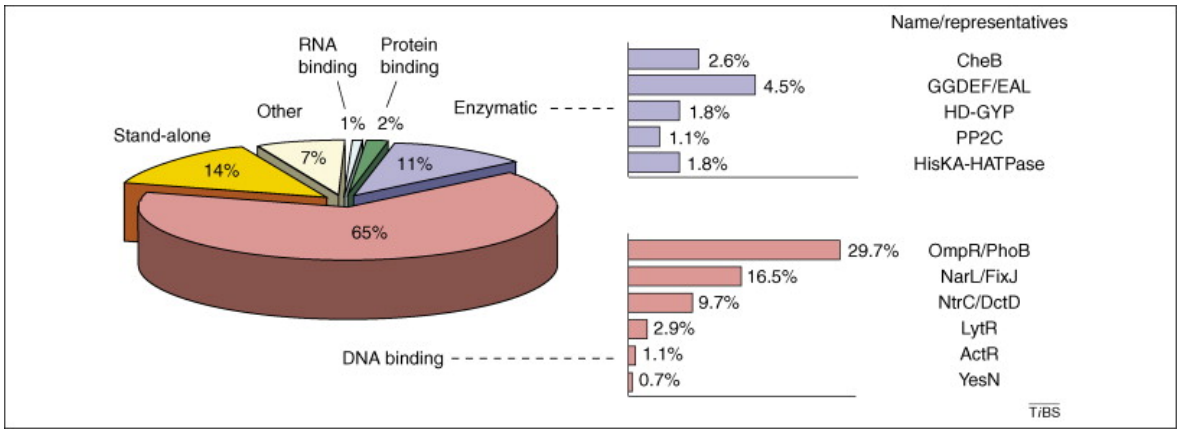


Figure 2.5: The different families of response regulators. Bacterial RRs are categorized by their effector domains and are divided into subfamilies based on functions and structures. The set of about 9,000 RRs analyzed in (Galperin, 2006) was used to generate this functional distribution. Figure from (Gao et al., 2007).

but some are not and carry either the receiver domain alone or in association with RNA- or protein-binding domains or domains with enzymatic functions (Figure 2.5) (Gao et al., 2007). Notably, among the RRs with enzymatic activities is the GGDEF/EAL subfamily of RRs, which contain domains responsible for c-di-GMP synthesis or degradation. An example of such RR is WspR in *P. aeruginosa*, which is involved in biofilm formation (De et al., 2009). The major DNA-binding RR subfamilies will be discussed in section 2.2.1.

The huge diversity of input and output signals of TFs contributes to their pivotal role in bacterial response and adaptation to changing environments (Ulrich et al., 2005). Many different types of input signals can be integrated into transcriptional regulatory networks. However, even though most TFs are predicted to bind to a metabolite (Chubukov et al., 2014), in most cases we still lack information on their metabolite input signal and how it modulates their activity. Many types of input signals can be integrated either directly by TFs or through TCS systems and modify TF activity. In addition to metabolite-sensing, numerous input signals can be sensed through different mechanisms, such as interactions with other proteins (Browning et al., 2019), direct oxygen or redox sensing (Crack et al., 2004; Zheng and Storz, 2000), metal sensing (Chandrangsu et al., 2017), indirect light-sensing through photosensory modules (Elías-Arnanz et al., 2011), temperature, pH or mechanic (contact) cues (Shivaji and Prakash, 2010; Chakraborty et al., 2017; O'Toole and Wong, 2016).

2.1.3 Nucleoid-associated proteins

Although bacteria do not possess histones and their genomic DNA is thus not organized in nucleosomes as it is in eukaryotes (Khorasanizadeh, 2004), they have DNA-binding proteins called nucleoid-associated proteins (NAPs) that are able to modulate DNA compaction (Browning et al., 2010). Major NAPs include HU, H-NS, IHF and Fis, all inducing different kinds of DNA organization due to their specific way of interacting with DNA (Dillon and Dorman, 2010). Binding of NAPs have been found to mainly occur in intergenic regions (Grainger et al., 2006), and to profoundly impact gene transcription by modification of promoter accessibility (Browning et al., 2010). Notably, NAPs have been studied for their role in stress response or growth regulation such as during switch from exponential to stationary phases (Hołówka and Zakrzewska-Czerwńska, 2020). In this context, NAPs, which exhibit growth-dependent expression patterns (Verma et al., 2019), play a central role on DNA compaction allowing growth rates modification through regulation of transcription or DNA replication.

2.1.4 Diversity and evolution of transcriptional regulatory networks

The investigation of molecular mechanisms is often restricted to a small number of bacterial species, and among each of them, very few reference strains. This bias often leads to the false assumption that regulatory networks function in the same way between species and strains, i.e. that if a TF regulates a gene in one strain and that they are both conserved in another strain, the regulatory relationship is too. While proved wrong early on (Lozada-Chavez et al., 2006), this postulate remains hard to debunk due to the still low number of experimental inter-species or inter-strains comparative studies. Moreover, the very nature of bacterial genetic evolution, driven by high gene exchange rates notably due to HGT, makes diversification of regulatory interaction a pivotal and frequent adaptation mechanism in prokaryotes (Balleza et al., 2008; Price et al., 2008).

Different types of regulatory diversification have been reported so far (Perez and Groisman, 2009a), the most straightforward being the loss or inactivation of a regulating TF. Target genes can also be lost or gained, along with the corresponding regulatory binding site on their promoter, leading to the reshaping of TF

regulons. These two mechanisms are the most studied ones because they rely on gene presence/absence, which is often studied in phylogenetic analyses. A third mechanism occurs by tweaking TF-DNA interactions, often by modification of binding sites recognized by TFs on target gene promoters, thus affecting promoter structure (Perez and Groisman, 2009a; Gelfand, 2006). It is less characterized as it requires to have strong previous knowledge on TF DNA consensus motif and genome-wide binding site profiles and is often studied only between distant species in order to apprehend large-scale evolutionary mechanisms (Perez and Groisman, 2009b). Recent studies (Connolly et al., 2020), as well as results presented in this thesis, have shown that this mechanism also occur between closely related species or even between strains of the same species, revealing the major importance of this mechanism in bacterial adaptation. These discoveries highlight again how biased is our way of mostly studying reference strains and gene content to explain phenotypic observations.

2.2 Transcription regulation in *P. aeruginosa*

2.2.1 *P. aeruginosa* transcription factors

With more than 9% of genes encoding TFs, *P. aeruginosa* has proportionally more TFs than any other bacterial species (Cases et al., 2003), more than any major "higher organisms", including humans, and thus potentially more than any other living cell. In total, there is around 500 predicted TFs in PAO1, spanning many different families (Figure 2.6). Among these, more than two third possess at least two domains (Ortet et al., 2012), many of which are predicted as OCSs, including LysR-like OCSs which are the most abundant TFs in *P. aeruginosa* (Figure 2.6) (Reen et al., 2013). Additionally, *P. aeruginosa* carry one of the highest numbers of TCSs in bacteria (Francis et al., 2017), with 64 HKs and 72 RRs (including 51 predicted as TFs). DNA-binding RRs exhibit a N-terminal receiver domain (REC) and a C-terminal DNA-binding domain that can be of different families of helix-turn-helix (HTH) domains. In PAO1, five different types of DNA-binding RRs are found, including the three major OmpR-, NarL and NtrC-like families and two RRs belonging to rare subfamilies, AlgR (LytTR-like) and RoxR (ActR-like) (Figure 2.7). As in most bacteria (Gao et al., 2007), OmpR-like RRs are the most abundant in *P. aeruginosa*. While the other families have only two domains, NtrC-like RRs also possess a σ^{54}-binding domain in between their REC and Fis-like domains.

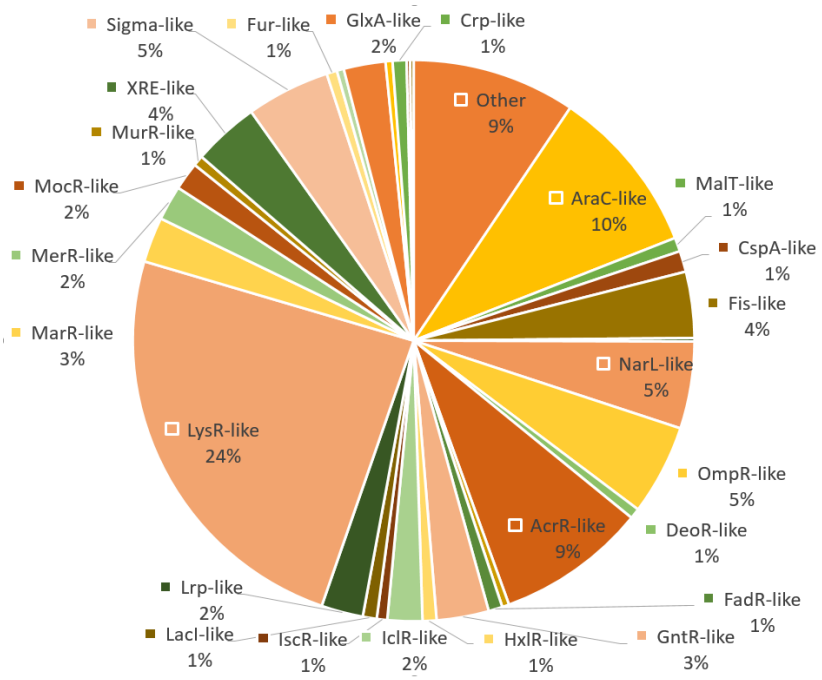


Figure 2.6: The families of *P. aeruginosa* transcription factors. Functional annotation of 500 predicted TFs were obtained from PseudoCAP (Brinkman et al., 2000), the P2TF database (Ortet et al., 2012) and the Pseudomonas Genome database (Winsor et al., 2016), and manually pooled and curated.

This domain is a AAA+ ATPase domain that induces the RNA-polymerase open complex formation in a ATP-dependent manner when interacting with σ^{54} (Gao et al., 2007).

Overall, many of *P. aeruginosa* TCSs have been studied and showed to be involved in the regulation of major cellular processes. Among them is the production of surface adhesins required for the initiation of biofilm formation, called Cup fimbriae (Vallet et al., 2001). Their expression is regulated by the Roc pathway, composed of two HKs, RocS1 and RocS2, which are each able to activate three RRs, RocR, RocA1 and RocA2 (Francis et al., 2017; Kulasekara et al., 2005). Upon activation, the RRs induce the downregulation of the MexAB-OprM efflux pump and either activation (RocA1) or repression (RocR) of Cup fimbriae, which fits with their role in biofilm regulation (Vettoretti et al., 2009; Mikkelsen et al., 2011b). Another example of the importance of TCSs in *P. aeruginosa* is the regulation of response to antimicrobial peptides which involves the *arn* operon responsible for LPS modification leading to increased resistance. Although the effect of this response mechanism has recently been shown to be strain-specific (Sciuto et al., 2020), it involves 5 independent TCSs - PmrAB, PhoPQ, ColRS, ParRS and CprRS - which respond either to cationic peptides or to magnesium or zinc ions, which are increased in presence of antimicrobial peptides (Gooderham and Hancock, 2009; Fernández et al., 2010; Lee and Ko, 2014). The FimS-AlgR is another important TCS

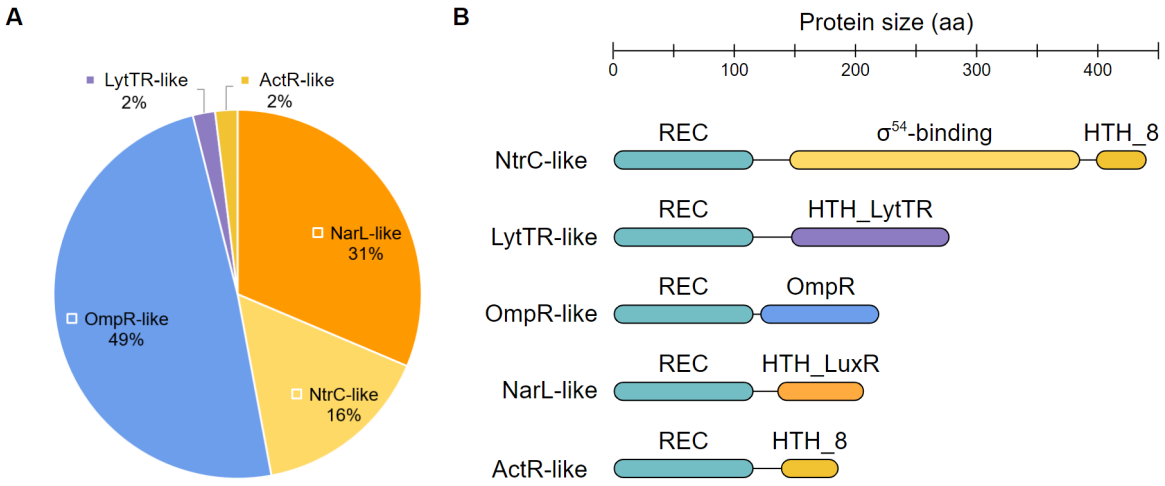


Figure 2.7: The families of *P. aeruginosa* response regulators. (**A**) Functional classification of 51 predicted RRs were obtained from PseudoCAP (Brinkman et al., 2000), the P2TF database (Ortet et al., 2012) and the Pseudomonas Genome database (Winsor et al., 2016), and manually pooled and curated. (B) Schematic views of domain architectures of the different RR subfamilies.

in *P. aeruginosa*. Indeed, it regulates alginate synthesis as well as twitching motility and biofilm formation, notably by impacting c-di-GMP synthesis (Kong et al., 2015). Other TCSs play major roles in the regulation of *P. aeruginosa* lifestyle transitions and will be discussed in further details in section 2.2.3.

Although this major human pathogen possesses this massive regulatory network (Galán-Vásquez et al., 2011), which is the main reason behind its high versatility, most of it is still uncharacterized. Indeed, I estimate that about a third of *P. aeruginosa* TFs have been characterized by at least one study, while the remaining two thirds are composed of uncharacterized, unnamed, predicted TFs. Additionally, only about 10% of *P. aeruginosa* TFs have been studied using high-throughput methods such as RNA-seq or ChIP-seq, which means that the rest of the "characterized" TFs have been studied through targeted methods and thus probably have incompletely determined regulons. Overall, this knowledge gap represents a limitation for both the development of potential new treatments against this pathogen and our global understanding of bacterial adaptation mechanisms.

As mentioned above, *P. aeruginosa* possesses 24 σ factors, more than 70 TCSs and hundreds of OCSs, all of which participate in its ability to adapt to new conditions. Although most if not all of them could probably contribute to pathogenicity in some infection conditions, including those involved in regulation of metabolism (Turner

et al., 2014), most of the scientific effort has focused on studying the regulation of processes with more direct effect on virulence such as toxin secretion systems or biofilm formation so far. A good example of this phenomenon is the fact that the CbrAB TCS is among the best characterized regulatory systems involved in central carbon metabolism, which is due to the fact that it also regulates directly virulence-related processes such as T3SS or biofilm formation (Sonnleitner et al., 2012; Linares et al., 2010). This bias is mostly due to the common use of very simple infection models such as cell cytotoxicity assays which often reflect only one phenotype - in most cases toxin secretions -, as opposed to more complex models where numerous cellular processes are needed for survival and colonization (Skurnik et al., 2013). Some untargeted genome-wide approaches have used Tn-seq (Van Opijnen et al., 2009) to identify genes important in different infection-related conditions (Potvin et al., 2003; Skurnik et al., 2013; Turner et al., 2014). Everytime, numerous TCSs were found important for bacterial survival during infection. The regulation of some individual metabolic pathways have been studied (Dolan et al., 2020; Schweizer and Po, 1996; Hoffmann and Rehm, 2004; Indurthi et al., 2016), or the regulation of major shifts such as the aerobiosis/anaerobiosis switch dictated by the Anr regulator (Trunk et al., 2010), or central carbon metabolism (Sonnleitner et al., 2012). But most of our knowledge concerns virulence-related pathways, which will be described below.

2.2.2 Quorum sensing

Quorum sensing (QS) is one of the major regulatory mechanisms in bacteria (Waters and Bassler, 2005). It is often referred to as a communication mechanism that allows individual bacteria to coordinate their behavior by secreting and sensing chemical signals called autoinducers. Usually, QS takes place in high cell density communities where the concentration of autoinducers is increased. When a threshold concentration of autoinducers is reached, autoinducer-TF interactions lead to the activation of specific autoinducer synthases which constitutes a positive feedback loop, further increasing the signal. These TFs also regulate many other genes upon autoinducer sensing, often involved in virulence, antibiotics resistance or biofilm formation (Barr et al., 2015).

In *P. aeruginosa*, three main QS pathways have been thoroughly described, the Las, Rhl and Pqs pathways, which regulate around 10% of all genes (Figure 2.8)

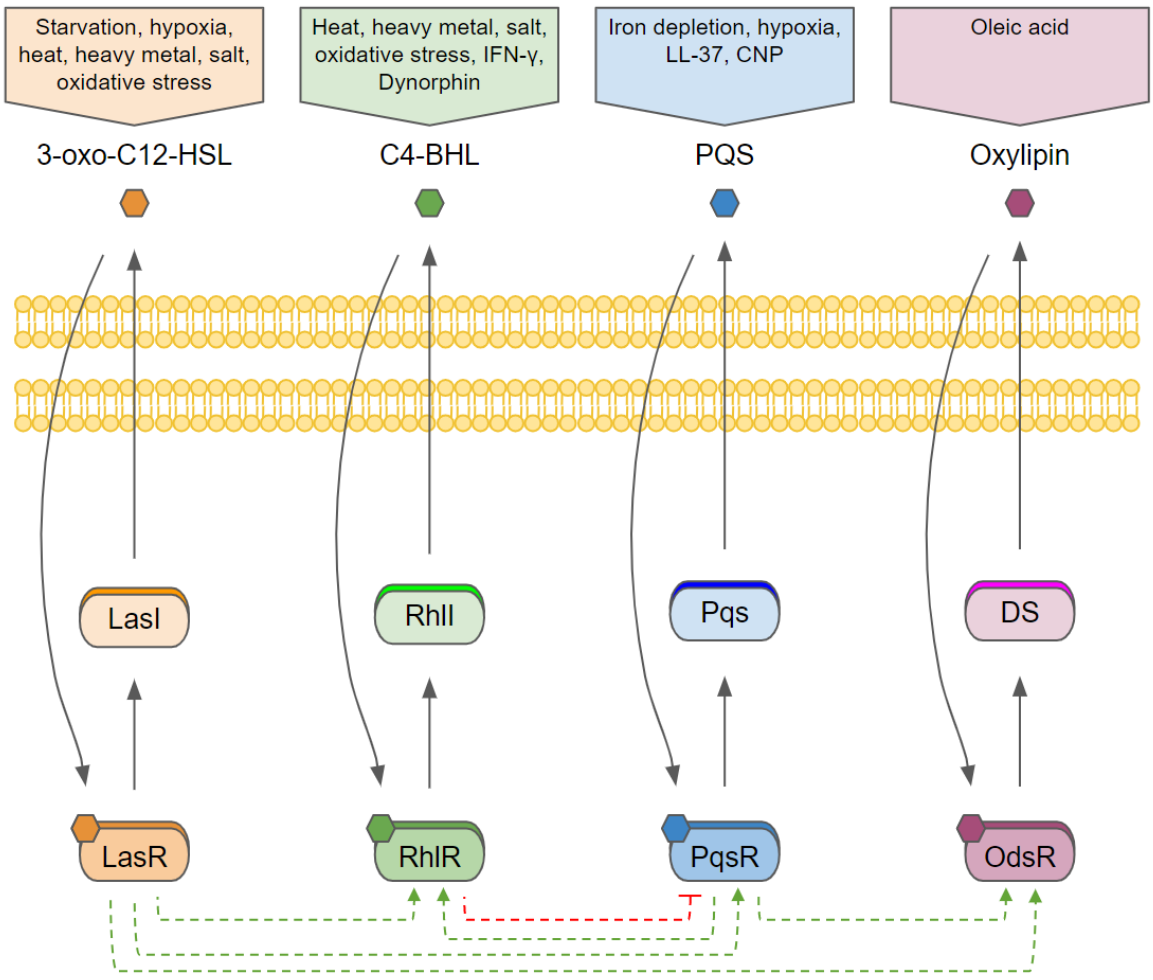


Figure 2.8: The *P. aeruginosa* quorum sensing network. The four different QS pathways (Las, Rhl, Pqs and Oxylipin) are depicted. Top to bottom: signals triggering the different QS systems, autoinducers, autoinducer synthesis pathways (Pqs: PqsABCDH, DS: PA0277-PA2078, diol synthase), autoinducer-sensing TFs. The regulatory interplays between the different pathways are shown with dashed arrows.

(Moradali et al., 2017; Schuster and Greenberg, 2006). Recently, a fourth QS pathway has been identified and involves oxylipin - a common eukaryotic communication molecule (Pohl and Kock, 2014) - and the OdsR TF (Martínez et al., 2019). Each of these regulatory pathways responds to specific external stimuli, such as hypoxia, oxidative stress or starvation, and involves the major TFs LasR, RhlR, PqsR and OdsR and their respective target autoinducer synthases (Figure 2.8) (Moradali et al., 2017; García-Contreras, 2016). Some cross-regulation has been demonstrated between the different QS systems, as LasR activates both Rhl and Pqs systems, RhlR inhibits Pqs and PqsR activates Rhl (Figure 2.8) (Lee and Zhang, 2015). Additionally, the production of oxylipin was shown to be partially impaired in *lasR*

and *PqsR* mutants (Martínez et al., 2019). This interplay allows a better control on global regulatory networks, notably for physiological adaptation to different growth phases (Choi et al., 2011). Upon QS activation, many virulence factors are produced, such as pyocyanin or Exotoxin A (Schalk and Guillon, 2013; Michalska and Wolf, 2015), making QS a central mechanism of *P. aeruginosa* pathogenicity and infective potential (Feng et al., 2016). The recently discovered oxylipin QS system seems to have a smaller regulon, containing genes mostly involved in metabolism (Martínez et al., 2019). Another pathway called IQS has been thought to be involved in QS (Moradali et al., 2017), but recent work proved that IQS is in fact a pyochelin byproduct and that its supposed biosynthesis pathway is not involved in QS (Cornelis, 2020). Other TFs, called orphans QS regulators, were found to regulate QS, such as QscR and VqsR which regulate the Las and Rhl pathways (Lee and Zhang, 2015).

2.2.3 Planktonic-biofilm switch

The transition between *P. aeruginosa* two lifestyles (described in Chapter 1) is an active process involving complex regulatory pathways, the main one being the Gac/Rsm pathway (Figure 2.9) (Moradali et al., 2017). This pathway is composed of the major GacSA TCS, which activity is also influenced by other HKs such as LadS, RetS, SagS, BfiS and PA1611, which either activate or inhibit this system in response to different mostly unknown stimuli (Francis et al., 2018; Kong et al., 2013; Bhagirath et al., 2017; Broder et al., 2016). The major targets of the GacSA TCS are the two small non-coding regulatory RNAs (sRNA) RsmY and RsmZ which in turn regulate by sequestration the regulatory RNA-binding protein (RBP) RsmA. By modulating translation initiation of target mRNAs, RsmA inhibits biofilm formation and induces virulence factors such as T3SS and pili, linked to the planktonic lifestyle (Brencic and Lory, 2009; Gebhardt et al., 2020). This regulatory pathway is also impacted by the HptB pathway (Figure 2.9). Three hybrid histidine kinases - PA1611, ErcS' and SagS - drive the phosphorylation of the HsbR RR through the HptB protein (Bouillet et al., 2019; Petrova et al., 2017; Kong et al., 2013). In turn, HsbR controls the phosphorylation of HsbA, which activity depends on its phosphorylation state (Bordi et al., 2010). When unphosphorylated, HsbA sequestrates the anti-σ factor FlgM, which allows the σ factor FliA to be active and to induce genes involved in the planktonic lifestyle. When phosphorylated, HsbA favors HsbD activity, a diguanylate cyclase inducing c-di-GMP synthesis and thus

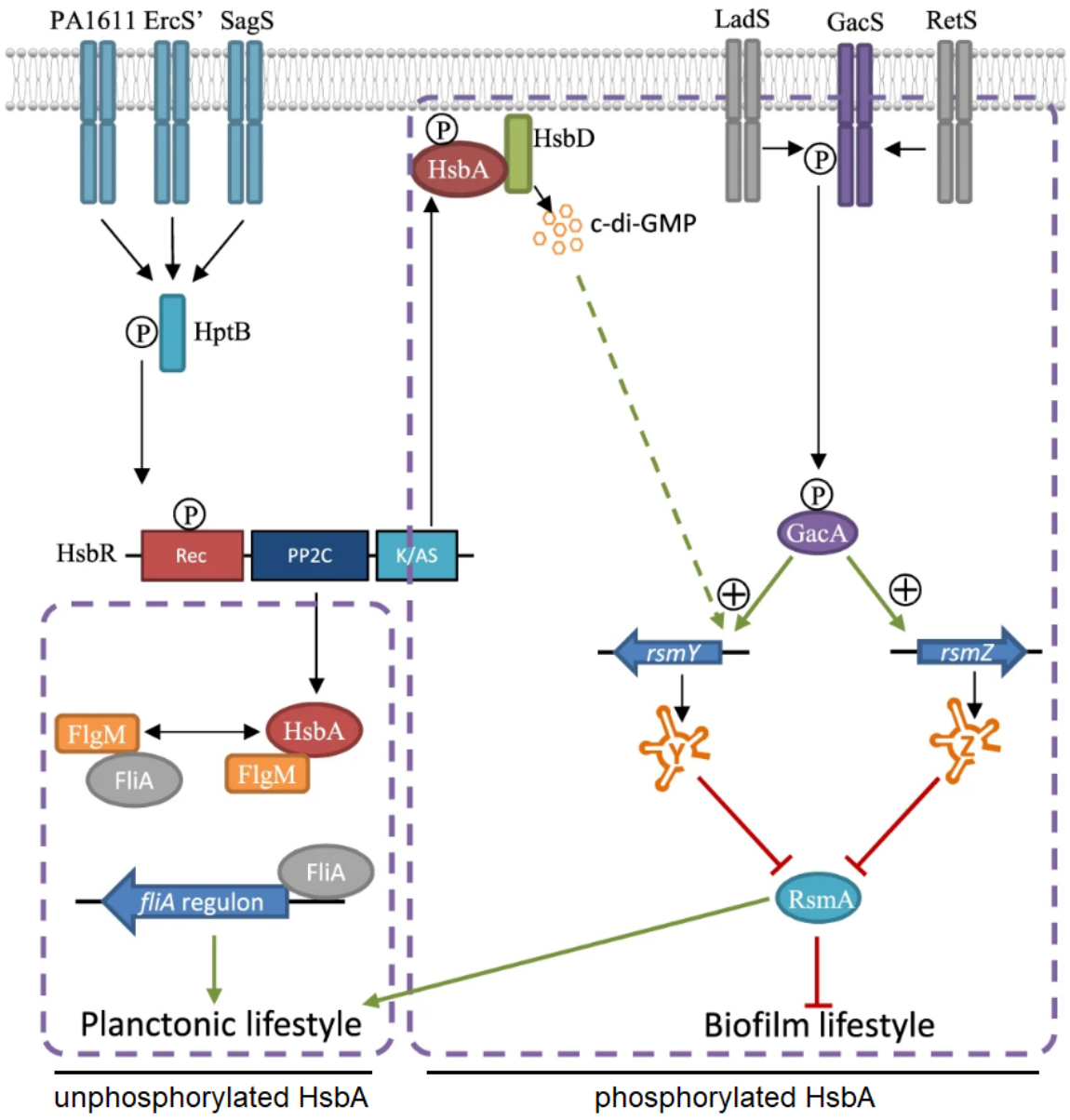


Figure 2.9: The regulatory networks driving *P. aeruginosa* lifestyle switch. The GacSA TCS regulate the expression of the two sRNAs RsmY and RsmZ which sequestrate the RsmA regulator. Free, active RsmA leads to expression of virulence factors and induces planktonic lifestyle. The HptB pathway also influences this switch, involving several HKs and the anti-σ factor antagonist HsbA which either induces FliA release from FlgM or HsbD-mediated c-di-GMP synthesis, depending on its phosphorylation status. Adapted from (Bouillet et al., 2019).

promoting biofilm formation.

The two secondary messengers c-di-GMP and cAMP, which synthesis and degradation are tightly regulated, influence this lifestyle switch (Pesavento and

Hengge, 2009). Notably, the regulation of c-di-GMP levels involves dozens of proteins including TCSs and various other sensing mechanisms responding to nearly as many signals and stimuli (Valentini and Filloux, 2016). These proteins possess GGDEF or EAL domains which correspondingly have diguanylate cyclase or phosphodiesterase activities, inducing the synthesis or degradation of c-di-GMP, respectively (Hengge, 2009). Among these is the Wsp TCS, responsible for c-di-GMP synthesis upon sensing of surface attachment (Hickman et al., 2005). The concentration of these secondary messengers regulates in turn various major cellular processes such as motility, virulence and biofilm formation through protein sensors including global TFs such as the c-di-GMP-sensing FleQ or the cAMP-sensing Vfr (Hickman and Harwood, 2008; Kanack et al., 2006). Vfr is a CRP homolog that has been shown to activate the transcription of several major virulence factors including ToxA, Type IV pili and the T3SS (Fuchs et al., 2010). Vfr also activates the Las QS pathway and thus acts on the expression of hundreds of genes (Schuster et al., 2003). cAMP and c-di-GMP play a key role in *P. aeruginosa* lifestyle switch as they are either associated with planktonic or biofilm lifestyle, respectively. Moreover, high concentrations of each one of the two were shown to induce a decrease in cellular content of the other, revealing a balance mechanism instrumental for the phenotypic switch (Almblad et al., 2015; Almblad et al., 2019).

Chapter 3

Post-transcriptional regulatory networks

Contents

Post-transcriptional regulation is a key mechanism allowing cells to control protein synthesis via modification of mRNAs stability or accessibility. In all kingdoms of life, post-transcriptional regulation has been shown to involve different molecular mechanisms and being instrumental in rapid cellular adaptation. Often involving noncoding RNAs (Filipowicz et al., 2008), several key processes have been revealed and the major bacterial ones are reviewed in this chapter.

3.1 Bacterial post-transcriptional regulation

3.1.1 Non-coding sRNAs

Bacterial sRNAs are 50-400bp-long untranslated RNAs that regulate translation through base pairing with target mRNAs (Waters and Storz, 2009). sRNAs interactions with mRNAs are generally mediated by RBPs such as Hfq (Holmqvist and Vogel, 2018). sRNAs recognize mRNAs through their seed sequence, which initiates base-pairing (Wagner and Romby, 2015). The seed sequence is often small or noncontiguous, rendering sRNA target prediction challenging (Li et al., 2012). Most sRNAs are involved in translation regulation through direct binding to mRNAs but some can act on transcription such as the well known 6S RNA, which inhibits transcription and mediates growth phase transition by specific binding to RNA polymerase (Wassarman and Storz, 2000; Wassarman, 2018). The discovery of regulatory sRNAs and of how common they are in bacteria is relatively recent (Toffano-Nioche et al., 2012; Brantl and Brückner, 2014), and this growing understanding only became possible with the technological advances made in high-throughput sequencing and new approaches based on RNA sequencing (Saliba et al., 2017).

Most sRNAs act on translation by direct binding to ribosome binding site, thus preventing ribosome recruitment and consequently inhibiting mRNA translation (Bouvier et al., 2008). sRNAs can also activate translation by the relieving of inhibitory mRNA structure occluding the ribosome binding site (Wagner and Romby, 2015; Mandin and Gottesman, 2010). Other regulatory mechanisms include the induction of mRNA degradation by binding inside of target mRNA coding sequence and recruitment of RNase (Bandyra et al., 2012). However, some sRNAs do not bind to mRNAs. Instead, this class of sRNAs acts by binding to and sequestrating regulatory RBPs, preventing them to exert their regulatory function.

An example of such mechanism is the sequestration of RsmA by RsmY and RsmZ in *P. aeruginosa* (Moradali et al., 2017).

Different types of sRNAs were identified, (i) those with their own promoter, either in intergenic regions or directly inside of genes both in the same or opposite orientation of genes, and (ii) sRNAs processed from mRNAs, generated from either 5'- or 3'-end of mRNAs (Wagner and Romby, 2015; Vogel et al., 2003). Most sRNAs are not constitutively expressed but are usually induced in stress conditions (Hoe et al., 2013) and play a role of regulation of stress response. One well documented example concerns iron homeostasis. Several sRNAs have been identified as inhibiting nonessential iron-binding proteins and activating iron storage mechanisms under iron limitation (Holmqvist and Wagner, 2017). The synthesis of these well studied sRNAs - RuhB in *E. coli* and PrrF1 and Prrf2 in *P. aeruginosa* - is activated in response to iron starvation, due to release of inhibition by the Fur TF, which is instrumental in bacterial adaptation to this condition (Massé et al., 2005; Wilderman et al., 2004; Han et al., 2016).

3.1.2 RNA-binding proteins

RNA-binding proteins (RBP) play a pivotal role in post-transcriptional regulation by acting on mRNAs translation or stability (Holmqvist and Vogel, 2018). They can induce both activation or inhibition of translation through different mechanisms (Figure 3.1). All of which involve mRNA-RBP interactions and, in some instances, a tripartite interaction with a regulatory sRNA. In such cases, the sRNA specifically recognizes the mRNA target by base pairing, and the RBP facilitates and stabilizes the sRNA-mRNA interaction.

Several major RBP families are known and usually found across numerous bacterial phyla. The Csr/Rsm family of RBPs is among the most conserved ones and notably encompasses the two well-characterized *E. coli* CsrA and *P. aeruginosa* RsmA (Romeo et al., 2013). In both species, this RBP activity is directly controlled by several sRNAs that prevent mRNA binding by sequestrating the RBP (Babitzke and Romeo, 2007). Csr/Rsm RBPs were found to regulate numerous key bacterial processes such as quorum sensing, virulence, motility and biofilm formation (as briefly described for *P. aeruginosa* in the previous chapter). These RBPs usually act by competing with ribosomes and thus inhibiting translation but they have

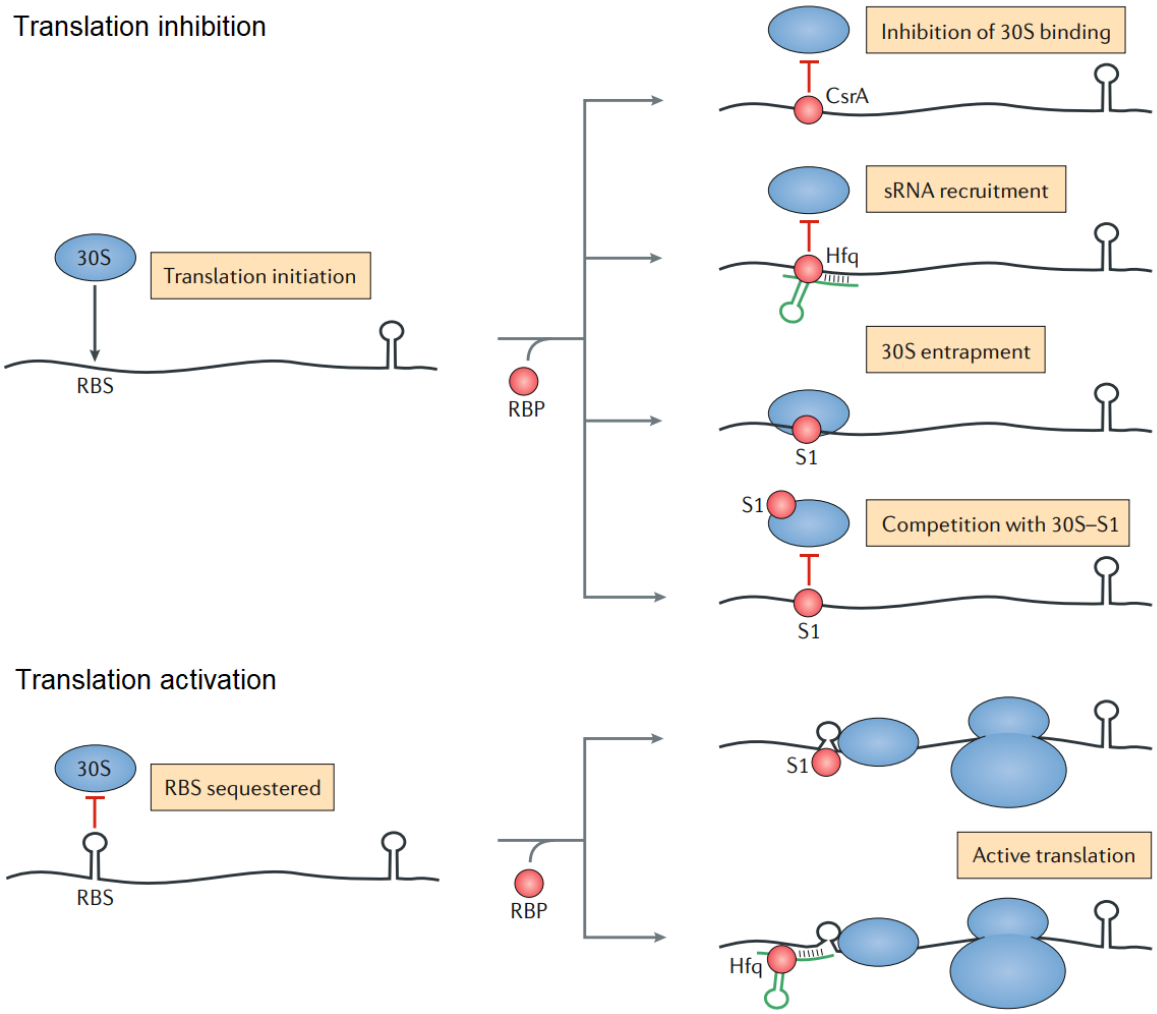


Figure 3.1: The different RBPs mechanisms of action. RBPs can inhibit or activate mRNA translation through different molecular processes. Inhibition mechanisms include preventing of the binding of the 30S ribosomal subunit to mRNAs in different manners or sequestration of 30S. Activation mechanisms promote 30S recruitment. Several examples are shown with three major RBPs, Hfq, CsrA and S1. Figure from (Holmqvist and Vogel, 2018).

also been shown to act in several other ways, including by promoting translation through prevention of RNaseE-mediated mRNA degradation (Yakhnin et al., 2013).

Hfq represents another major regulatory RBP through its chaperone activity, mediating hundreds of sRNA-mRNA interactions. Hfq is conserved across the bacterial kingdom and has been shown to be involved in the regulation of translation of thousands of proteins, notably in all major bacterial pathogens (Chao and Vogel, 2010). It was among the first bacterial RBP to be thoroughly studied, especially in pathogens where it was shown to regulate the synthesis of numerous virulence factors, such as the T3SS, and many other key processes such as quorum sensing

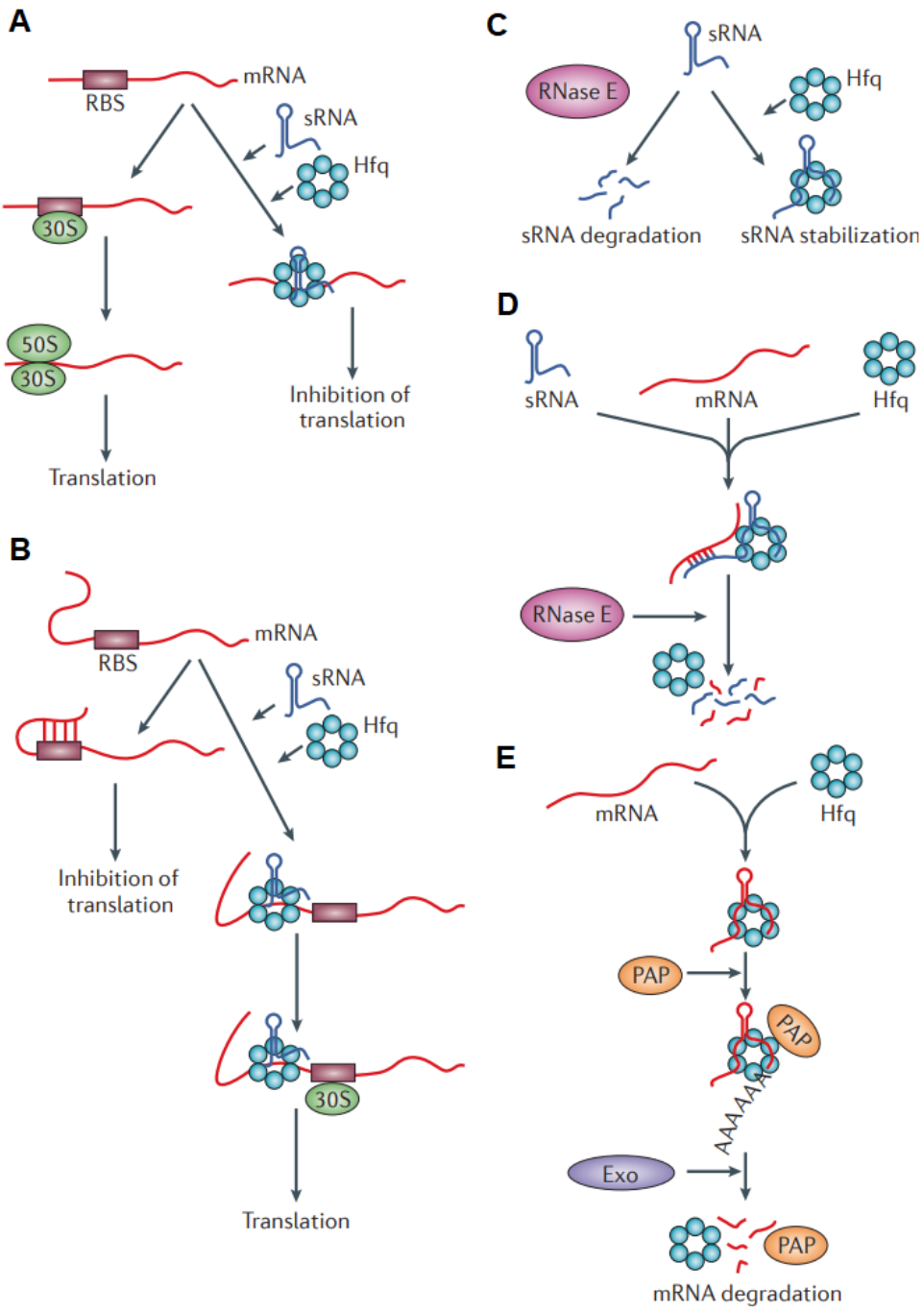


Figure 3.2: The modes of Hfq action. (A) Hfq, in association with a sRNA, can sequester the ribosome binding site, leading to translation inhibition. (B) Hfq, in association with a sRNA, can promote ribosome binding and thus activate translation. (C) Hfq can stabilize and protect sRNA from RNase degradation. (D) Hfq, in association with a sRNA, can promote RNase-mediated mRNA degradation. (E) Hfq can stimulate mRNA polyadenylation by poly(A) polymerase (PAP), leading to mRNA degradation by exoribonucleases (Exo). Figure from (Vogel and Luisi, 2011).

or biofilm formation. Hfq adopts a hexameric ring-like architecture, typical of the Hfq-Sm-LSm RBP family found across both prokaryotes and eukaryotes, which harbors two RNA-binding interfaces, one involved in sRNA interactions and the other one in mRNA interactions (Vogel and Luisi, 2011). Indeed, Hfq mediates sRNA-mRNA interactions through specific binding to each partner involved, leading to either activation or inhibition of mRNA translation, often by promoting or preventing ribosome recruitment (Figure 3.2). Similarly to CsrA, Hfq was also found to be involved in RNaseE-mediated mRNA degradation (Massé et al., 2003; Morita et al., 2005), by either recruiting or protecting mRNAs from RNasE.

Other important RBP families include the more recently discovered FinO-like RBPs (Glover et al., 2015). This family is widely conserved in proteobacteria (Smirnov et al., 2016) and encompasses several members that have been shown to regulate major virulence processes in bacterial pathogens such as *Salmonella* and *Legionella* (Smirnov et al., 2017; Attaiech et al., 2016; Smirnov et al., 2016). These RBPs, ProQ and RocC, were found to mediate sRNA-mRNA interactions in a structure-specific manner (Holmqvist et al., 2018), as opposed to the more classical sequence-specificity found with Hfq or CsrA.

3.1.3 Riboswitches and RNA thermometers

Riboswitches represent another mechanism of regulation in bacteria. Riboswitches are RNA structures found in 5′ untranslated regions of mRNAs which are able to sense different ligands, usually metabolites, and regulate the target mRNA translation, or transcription in some cases, accordingly (Winkler and Breaker, 2005). Riboswitches can both activate or inhibit translation and have been classified into more than 40 different families (McCown et al., 2017). They are composed of an aptamer region, responsible for ligand binding, and of an expression platform, which drives gene regulation. Upon binding of the signal in the aptamer region, a conformational change occurs in the expression platform which leads to translation or transcription regulation. Translation regulation often works by modification of ribosome binding site accessibility while transcription regulation involves the formation of a RNA hairpin inducing transcription termination. Riboswitches were reported to sense a wide variety of metabolites, usually linked to the function of the regulated mRNA (Nudler and Mironov, 2004). Notable examples include glycine-, adenine- and thiamine pyrophosphate-sensing riboswtiches, all regulating

the corresponding different metabolic pathways (Mandal et al., 2004; Lemay et al., 2006; Serganov et al., 2006). In *P. aeruginosa*, there have been several riboswitches identified, including the thiamine pyrophosphate-, lysine or guanidine-binding ones (Pavlova and Penchovsky, 2019; Reiss and Strobel, 2017).

Similarly to riboswitches, RNA thermometers are also regulatory RNA structures found in mRNA 5′ untranslated regions that regulate translation (Narberhaus et al., 2006). However, these structures do not sense small molecules, as opposed to riboswitches, but sense changes in temperature and modulate ribosome recruitment and translation accordingly. Different families of RNA thermometers exist and function by one of two mechanisms: zipper-like RNA thermometers induce changes between open or closed conformations in RNA structure while switches work by shifting between two different RNA structures (Kortmann and Narberhaus, 2012). One major RNA thermometer is the one controlling expression of the heat-shock responsive σ factor RpoH in *E. coli* (Storz, 1999). When switching from 30°C to 42°C, *rpoH* RNA thermometer unfolds, leading to translation activation and consequent regulatory adaptation through expression of the σ factor. Another example is the ROSE-like (repression of heat-shock gene expression) RNA thermometer found in the 5′ untranslated region of *P. aeruginosa rhlAB* mRNA. This RNA thermometer induces RhlR expression at 37°C, leading to activation of numerous QS-dependent virulence genes (Grosso-Becerra et al., 2014). Additionally, another ROSE-like RNA thermometer is found in the *lasI* mRNA, also encoding a master QS regulator (Grosso-Becera et al., 2015).

3.1.4 Diversity and evolution of post-transcriptional regulatory networks

Similarly to transcription regulation, post-transcriptional regulatory interactions are not necessarily conserved across strains or species. However, only very few studies have focused on investigating that. Most of our knowledge derives from the few works that analyzed sRNA conservation across groups of species, or from comparison of RBPs regulons between species reported one by one in individual studies. I believe this represents yet another bias in our way of seeing regulation and, for that reason, one project of this thesis presents a first example of experimental characterization of intra-species functional diversification of RBPs (in this case

Hfq) (see section 5.1).

The identification of sRNAs and of their targets has long been a limiting factor in our understanding of their function (Livny and Waldor, 2007). With the recent advances in bioinformatics and experimental methods in the field (Tjaden, 2012; Saliba et al., 2017; Georg et al., 2020), more and more studies aim at the genome-wide identification and characterization of sRNAs (Wurtzel et al., 2010; Tree et al., 2014; Saliba et al., 2017; Hör et al., 2020). Few studies have looked at sRNA conservation between species or strains, but they always found differences in sRNA content (Backofen and Hess, 2010; Zhang et al., 2004; Wurtzel et al., 2012). One example has been the phylogenetic analysis of the conservation of the MmgR sRNA across α-proteobacteria. By looking at MmgR distribution across 70 species, this study showed that different copies of MmgR homologs could be found in varying numbers (0 to 6) between species or strains (Lagares Jr et al., 2016). Another study showed that even closely related sRNA homologs can exhibit differences in their regulatory function. Indeed, the Pxr sRNA was shown to have divergent roles in developmental gene regulation across *Myxococcus* species (Chen et al., 2017). These sparse examples highlight the diversity found in sRNA conservation and function and the need of such comparative studies.

Some studies of RBPs also led to insights into functional conservation between species. Hfq has been studied for decades and genome-wide interactomes have been individually determined for this RBP in numerous species, including *E. coli*, *P. aeruginosa* and *S. typhimurium* (Melamed et al., 2020; Holmqvist et al., 2016; Kambara et al., 2018). However, it has recently been shown that Hfq exhibit species-specific binding profiles, notably between *Listeria monocytogenes, Caulobacter crescentus and E. coli*, which was explained by differences in Hfq protein sequence between these species (Tran et al., 2020). A recent study compared ProQ RNA targets between *E. coli* and *Salmonella* and found interesting features (Holmqvist et al., 2018). If was found that numerous targets were conserved between the two species, even with low target RNA sequence conservation.

3.2 Post-transcriptional regulation in *P. aeruginosa*

With hundreds of regulatory sRNAs recently identified and one of the richest bacterial regulatory network, *P. aeruginosa* exhibits a complex post-transcriptional regulatory network.

3.2.1 sRNAs

In *P. aeruginosa*, several studies have tried to identify sRNAs at genome-wide scales. The number of identified sRNAs varies around 200 to 500, depending on the sequencing and analysis methods used (Gómez-Lozano et al., 2012; Wurtzel et al., 2012). However, a much smaller number has been functionally characterized. The most studied sRNAs often are highly expressed or impact bacterial physiology globally. Indeed, most of the well characterized sRNAs play key roles in *P. aeruginosa* virulence regulatory network (Figure 3.3). These include the above-mentionned RsmY and RsmZ that act as RBP sequestrators and a few others (Pita et al., 2018). Among them, the PhrS sRNA is known to activate PqsR - a major QS regulator - synthesis in a Hfq-dependent manner (Sonnleitner et al., 2011). More recently, PhrS was shown to repress transcription termination of CRISPR RNAs in a Hfq-independent way, promoting CRISPR adaptive immunity against bacteriophages (Lin et al., 2019). PrrF1 and PrrF2 are two sRNAs known for their Hfq-dependent role in response to iron starvation (Wilderman et al., 2004), as discussed above. The ErsA sRNA was shown to inhibit *algC* and activate *amrZ*, demonstrating its central role in alginate production and biofilm formation (Ferrara et al., 2015; Falcone et al., 2018). Several sRNAs are known for their role in stress response, notably through regulation of RpoS synthesis, such as Real and RgsA (Thi Bach Nguyen et al., 2018; Lu et al., 2018). Some sRNAs strongly impact virulence, including sr0161 which directly interacts with *exsA* mRNA to regulate T3SS expression (Zhang et al., 2017), or sRNA 179 which indirectly regulates Vfr (Janssen et al., 2020).

While the identification of potential sRNAs from transcriptomics data has now become common, there is still an important lack of knowledge regarding their functions. Several recent methodological developments and improvements now allow for more large scale functional investigation such as with genome-wide sRNA-mRNA interaction mapping (Zhang et al., 2017; Saliba et al., 2017). However,

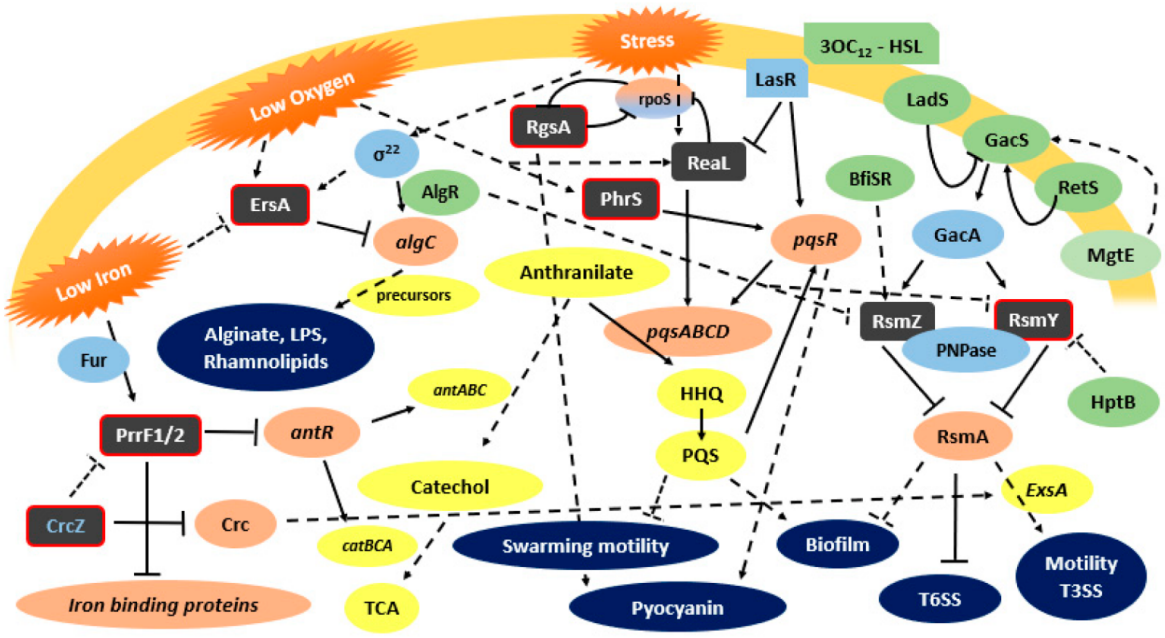


Figure 3.3: *P. aeruginosa* major sRNAs and their involvement in the regulation of virulence. sRNAs are shown by dark grey boxes (with red border if interacting with Hfq), green and blue boxes show indirect and direct regulators, respectively, orange and yellow boxes show direct and indirect targets, respectively. Indirect regulation is shown by dashed lines. Figure from (Pita et al., 2018).

there is still a lot to do in order to fully understand sRNA regulatory networks.

3.2.2 Rsm proteins

As mentionned above, RsmA is part of the Gac/Rsm pathway. Its activity is regulated by the RsmY and RsmZ sRNAs and the two more recently discovered RsmV and RsmW (Janssen et al., 2018; Miller et al., 2016). RsmA is a global post-transcriptional regulator of the Csr/Rsm family of RBPs that have been shown to regulate over 500 mRNAs in *P. aeruginosa* and to play a pivotal role in the planktonic/biofilm switch (Brencic and Lory, 2009; Moradali et al., 2017). When not sequestered, RsmA inhibits biofilm formation and activates functions necessary for the planktonic lifestyle such as toxin secretion and motility. Interestingly, RsmA has an overlapping regulon with another Csr/Rsm RBP, RsmF (Marden et al., 2013). RsmF was found more abundant in strains lacking RsmA, partially complementing the effect of its deletion. Recently, it was shown that RsmA associates with nascent transcripts (Gebhardt et al., 2020), during transcription-translation coupling mechanisms (Kohler et al., 2017). This study revealed a new regulatory process where the

translation of mRNAs is regulated as soon as they emerge from RNA polymerase, even before they are fully synthesized. Through this mechanism, RsmA regulates key processes, notably through regulation of AmrZ, a TF controlling the expression of numerous genes involved in motility, virulence and biofilm formation (Gebhardt et al., 2020; Jones et al., 2014). Other important targets include several genes involved in Type VI secretion system, the anti-σ factor MucA, numerous TFs including several TCSs, the global regulator Anr, the σ-factor RpoS and genes involved in LPS, pili or flagella production.

3.2.3 Hfq

While Hfq was first studied in *E. coli*, it was shown that *P. aeruginosa* Hfq could partially complement a *E. coli hfq* mutant, revealing some level of functional conservation (Sonnleitner et al., 2002). In *P. aeruginosa*, Hfq regulates hundreds of genes, notably involved in the synthesis of virulence factors (Sorger-Domenigg et al., 2007), carbon metabolism (Sonnleitner et al., 2003), antibiotic resistance (Zhang et al., 2017), as well as the switch between planktonic and biofilm lifestyle (Chihara et al., 2019). Hfq was notably found to regulate carbon catabolite repression, which allows the ordered use of preferred carbon sources before less preferred ones (Deutscher, 2008). To do so, Hfq forms a complex with the Crc protein to interact with target mRNAs in a sRNA-independent manner (Sonnleitner et al., 2018; Pei et al., 2019). When preferred carbon sources are exhausted, the CrcZ sRNA becomes more abundant and sequesters Hfq, abrogating its inhibitory effect on catabolic proteins involved in less preferred carbon sources processing (Figure 3.3) (Sonnleitner and Bläsi, 2014). Hfq was recently shown to target nascent transcripts (Kambara et al., 2018), similarly to RsmA with which it also shares common regulatory targets (Gebhardt et al., 2020).

Results

Chapter 4

Transcriptional regulation

Contents

4.1 Regulation of ExlBA and TPSs

4.1.1 The Vfr-cAMP regulation of ExlBA

Alice Berry, Kook Han, Julian Trouillon, Mylène Robert-Genthon, Michel Ragno, Stephen Lory, Ina Attrée & Sylvie Elsen✉
Published in *Journal of Bacteriology* in 2018

This work represents Alice Berry's PhD project where she discovered the first regulator of *exlBA*. As a starting point for this completely new operon, Alice looked at the major known regulatory pathways and virulence-related TFs to see if they had a role in *exlBA* regulation (Berry, 2019). My project shared the same goal - characterizing *exlBA* regulation - but using an untargeted approach, as described in the next section. Alice had just found out that *exlBA* is under control of the global regulator Vfr. I participated in this project by creating and using an *exlBA* transcriptional fusion to show the cAMP-Vfr-dependent regulation of *exlBA* and the implication of different adenylate cyclases in this pathway. In this work, we identified the first regulator of *exlBA*: the global TF Vfr. While Vfr regulates the T3SS in PAO1- and PA14-like strains, this result reveals a new function of Vfr in PA7-like strains which is there again to activate a major virulence factor.
For this project, I performed the transcriptional fusion activity measures shown in Figure 5A.

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

RESEARCH ARTICLE

cAMP and Vfr Control Exolysin Expression and Cytotoxicity of *Pseudomonas aeruginosa* Taxonomic Outliers

Alice Berry,[a] Kook Han,[b] Julian Trouillon,[a] Mylène Robert-Genthon,[a] Michel Ragno,[a] Stephen Lory,[b] Ina Attrée,[a] Sylvie Elsen[a]

[a]Biology of Cancer and Infection, U1036 INSERM, CEA, University of Grenoble Alpes, ERL5261 CNRS, Grenoble, France

[b]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, USA

ABSTRACT The two-partner secretion system ExlBA, expressed by strains of *Pseudomonas aeruginosa* belonging to the PA7 group, induces hemorrhage in lungs due to disruption of host cellular membranes. Here we demonstrate that the *exlBA* genes are controlled by a pathway consisting of cAMP and the virulence factor regulator (Vfr). Upon interaction with cAMP, Vfr binds directly to the *exlBA* promoter with high affinity (equilibrium binding constant [K_{eq}] of ≈2.5 nM). The *exlB* and *exlA* expression was diminished in the Vfr-negative mutant and upregulated with increased intracellular cAMP levels. The Vfr binding sequence in the *exlBA* promoter was mutated *in situ*, resulting in reduced cytotoxicity of the mutant, showing that Vfr is required for the *exlBA* expression during intoxication of epithelial cells. Vfr also regulates function of type 4 pili previously shown to facilitate ExlA activity on epithelial cells, which indicates that the cAMP/Vfr pathway coordinates these two factors needed for full cytotoxicity. As in most *P. aeruginosa* strains, the adenylate cyclase CyaB is the main provider of cAMP for Vfr regulation during both *in vitro* growth and eukaryotic cell infection. We discovered that the absence of functional Vfr in the reference strain PA7 is caused by a frameshift in the gene and accounts for its reduced cytotoxicity, revealing the conservation of ExlBA control by the CyaB-cAMP/Vfr pathway in *P. aeruginosa* taxonomic outliers.

IMPORTANCE The human opportunistic pathogen *Pseudomonas aeruginosa* provokes severe acute and chronic human infections associated with defined sets of virulence factors. The main virulence determinant of *P. aeruginosa* taxonomic outliers is exolysin, a membrane-disrupting pore-forming toxin belonging to the two-partner secretion system ExlBA. In this work, we demonstrate that the conserved CyaB-cAMP/Vfr pathway controls cytotoxicity of outlier clinical strains through direct transcriptional activation of the *exlBA* operon. Therefore, despite the fact that the type III secretion system and exolysin are mutually exclusive in classical and outlier strains, respectively, these two major virulence determinants share similarities in their mechanisms of regulation.

KEYWORDS *Pseudomonas aeruginosa*, Vfr, cAMP, gene regulation, TPS, exolysin, PA7 group, virulence, cyclic AMP

Bacterial virulence factors are tightly regulated by complex regulatory networks integrating multiple stimuli to ensure their synthesis only when required (1). *Pseudomonas aeruginosa*, a major inhabitant of various environmental reservoirs, is also an important human opportunistic pathogen and, consequently, a frequently used model for studying virulence gene regulation. Its exceptional regulatory complexity was suggested from the 6.3-Mb genome of the reference strain PAO1, containing 5,570 genes with roughly 12% of them devoted to regulation (2–4). Additionally, more than

Received 6 March 2018 **Accepted** 4 April 2018

Accepted manuscript posted online 9 April 2018

Citation Berry A, Han K, Trouillon J, Robert-Genthon M, Ragno M, Lory S, Attrée I, Elsen S. 2018. cAMP and Vfr control exolysin expression and cytotoxicity of *Pseudomonas aeruginosa* taxonomic outliers. J Bacteriol 200:e00135-18. https://doi.org/10.1128/JB.00135-18.

Editor George O'Toole, Geisel School of Medicine at Dartmouth

Copyright © 2018 American Society for Microbiology. All Rights Reserved.

Address correspondence to Sylvie Elsen, sylvie.elsen@cea.fr.

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

500 putative noncoding RNAs carrying potential regulatory functions were found in different strains (5–7) further increasing the complexity of regulatory networks used by this versatile microorganism. Quorum sensing (QS [with four interconnected systems: Las, Rhl, PQS, and IQS]), the two-component regulatory systems (comprising around 60 of both histidine kinases and response regulators and 3 Hpt proteins), the Gac/Rsm system (with regulatory RNAs RsmY and RsmZ negatively controlling the activity of the posttranscriptional regulator RsmA), and the 2 second messengers (cyclic AMP [cAMP] and cyclic di-GMP [c-di-GMP]) (1, 8–10) are among the well-characterized networks controlling virulence. These regulatory pathways govern the reciprocal expression of the virulence factors, depending on the *P. aeruginosa* "lifestyle." During planktonic growth, the bacteria express factors involved in acute virulence, such as pili, flagella, and toxins, and most of these factors are under the positive control of QS and cAMP, an allosteric activator of the virulence factor regulator, Vfr (11). In contrast, other factors like exopolysaccharides and adhesins are synthesized under growth in a structured and multicellular community called "biofilm." A high c-di-GMP level is commonly associated with this mode of life in several pathogenic bacteria, even if only a few direct targets have been identified up to now (12, 13). These regulatory circuits were mainly studied in most common laboratory strains (PAO1, PAK, and PA14), and they are encoded within their core genome. However, there are some strain variations that have been reported, such as strains with point mutations or deletions in key regulatory genes affecting their virulence. For instance, the defective genes *mexS* in PAK, *ladS* in PA14, and *gacS* in cystic fibrosis (CF) isolate CHA are responsible for overexpression of genes encoding the type 3 secretion system (T3SS), the major virulence factor of *P. aeruginosa* (14–16).

The existence of *P. aeruginosa* taxonomic outliers with the fully sequenced PA7 strain as the representative was recently reported (17–22). They constitute one of the three major groups of *P. aeruginosa* strains defined from phylogenetic analysis of core genome sequences (23). The PA7 strain exhibits low sequence identity of its core genome to the "classical" strains PAO1 (group 1) and PA14 (group 2) (95% instead of the >99% usually found) and possesses numerous specific regions of genome plasticity (RGPs) encoding notably a third specific type 2 secretion system, Txc (17, 24). An unexpected but important feature of PA7 is the absence of the *toxA* gene, as well as the absence of the determinants for the entire T3SS apparatus and all T3SS-exported toxins (17). While characterizing the PA7-like strain CLJ1, isolated from a patient suffering from chronic obstructive pulmonary disease (COPD) and hemorrhagic pneumonia, we discovered a new type 5 secretion system (T5SS) responsible for this strain's pathogenicity. This T5SS is composed of two proteins—ExlB and ExlA. ExlB is the outer membrane protein that facilitates the secretion of the pore-forming protein exolysin (ExlA) responsible for membrane permeabilization and cell death (19, 21, 25). Since 2010, approximately 30 additional PA7-like strains have been reported in publications, and some of these were studied in more detail. We analyzed the phenotypes of the majority of *exlA*$^+$ PA7-like strains and their ability to produce ExlA. Interestingly, different levels of ExlA secretion were observed among these strains leading to variations in the extent of cytotoxicity and pathogenicity in mice (21). This observation strongly suggested that ExlB and ExlA synthesis is controlled by as yet unidentified regulatory pathways.

With this objective in mind, we undertook to analyze the regulation of *exlBA* expression in strain IHMA879472 (IHMA87), a PA7-like strain isolated recently from a urinary infection. This strain is able to secrete ExlA, exhibits high ExlA-dependent cytotoxicity on different eukaryotic cell lines, and can be genetically manipulated (21, 25). We found that *exlBA* is under the control of cAMP and its cognate receptor Vfr, which regulate the expression of acute virulence factors in the classical *P. aeruginosa* strains. We further demonstrate that the absence of a functional Vfr in PA7 accounts for its reduced cytotoxicity.

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

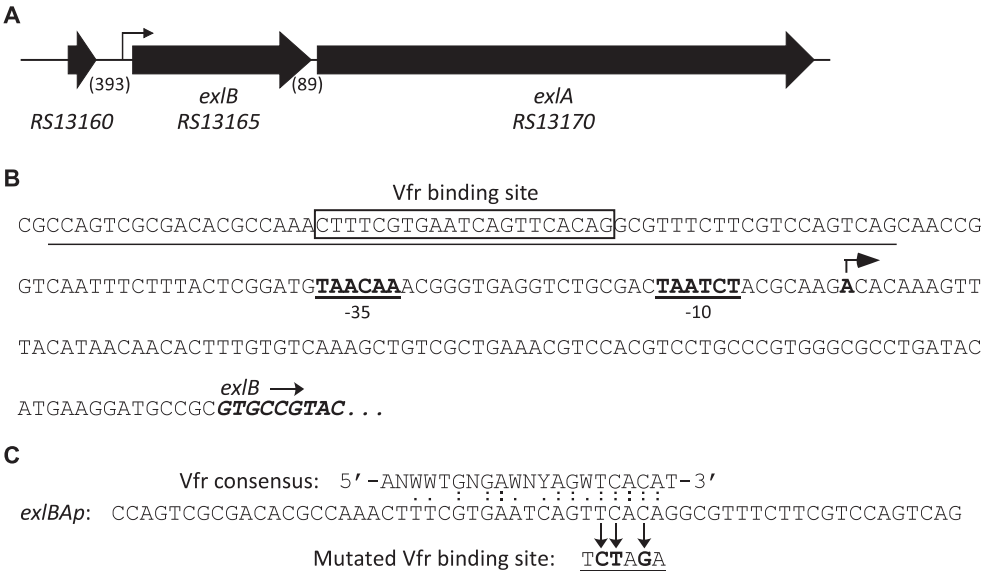


FIG 1 Features of the *exlBA* chromosomal region. (A) The *exlBA* genes as well as the upstream gene are depicted by thick arrows with the nomenclature from Pseudomonas Genome Database website (28; www.pseudomonas.com). The lengths of intergenic sequences (in base pairs) are in parentheses. The *exlBA* promoter (*exlBAp*) region identified in this study is represented by a thin arrow. (B) Sequence of the *exlBA* promoter. The three first codons of *exlB* are in italics and boldface. The nucleotide identified as potential *exlBA* TSS found by circularization assay is in boldface and pinpointed by an arrow, with the corresponding putative −10 and −35 boxes indicated. The Vfr binding site (VBS) is boxed. Underlined is the 60-mer sequence used in EMSA and provided in panel C. (C) Comparison of the Vfr consensus sequence to the sequence found in *exlBAp*. The mutated VBS used in EMSA and created in the chromosome of the VBM mutant is also indicated.

RESULTS

The *exlBA* promoter region contains a putative Vfr binding site. ExlBA constitutes a T5SS belonging to the "b" subgroup, also called a two-partner secretion (TPS) system, whose determinants are commonly encoded in a single transcriptional unit (26, 27). As depicted in Fig. 1A, the *exlB* and *exlA* genes are in the same orientation, separated by 89 bp, and are predicted in PA7 to form an operon without the presence of an intergenic transcriptional terminator (28; www.pseudomonas.com). We sought the transcription start site (TSS) of *exlBA* mRNA using a circularization approach and identified one nucleotide located 91 bp upstream from the *exlB* start codon as the TSS candidate, and predicted putative "−10/−35" sequences using BPROM (29; www.softberry.com) (Fig. 1B). Because the −10 and −35 sequences were poorly conserved, we validated the existence of the promoter by site-directed mutagenesis. First, we introduced a reporter *lacZ* gene directly into the *exlA* locus, creating a transcriptional reporter in the wild-type (WT) strain. Then we mutated the −10 and −35 boxes directly on the chromosome, which affected the *exlBA* expression *in vivo* (see Fig. S1 in the supplemental material). By examining the promoter sequence using Regulatory Sequence Analysis Tools (RSAT) (30; http://embnet.ccg.unam.mx/rsa-tools/) and the Vfr consensus binding sequence (31), we identified a putative binding site for the transcription factor Vfr, CTTTCGTGAATCAGTTCACA centered at around 95 bp upstream from the TSS of *exlBA* (Fig. 1B and C).

Expression of both *exlB* and *exlA* requires Vfr. To examine the possible role of Vfr in *exlBA* regulation, we deleted the *vfr* gene in the ExlA-secreting IHMA87 strain and assessed the impact of the loss of Vfr on *exlBA* expression. We tested secretion of ExlA *in vitro* and observed a clear reduction of its secretion in the Δ*vfr* mutant that was restored upon complementation of the strain with the wild-type copy of the gene (Fig. 2A). Then we measured the impact of *vfr* deletion on *exlB* and *exlA* mRNA levels by reverse transcription-quantitative PCR (RT-qPCR) and observed a significant reduction of both genes' expression in strains lacking Vfr. The reduced expression was partially

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

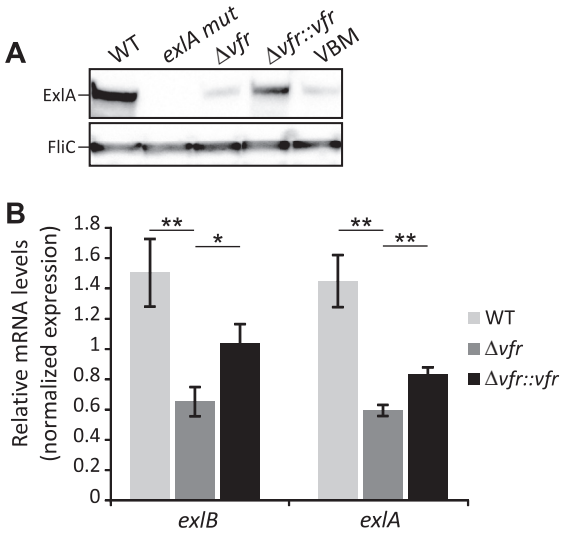


FIG 2 Vfr regulates ExlA secretion and *exlBA* expression. (A) The secretion of ExlA by the indicated strains grown in LB medium. The VBM strain carries a mutation in the VBS present in the *exlBA* promoter. Immunoblots of 100-fold-concentrated secretomes revealing the ExlA protein and FliC, used as a loading control. (B) RT-qPCR analysis of *exlB* and *exlA* relative expression in the wild-type (WT), the *vfr* mutant, and complemented strains grown in LB medium. The *rpoD* gene was used as a reference. The experiments were performed in triplicate, and the error bars indicate the standard errors of the means. The *P* value was determined using the Mann-Whitney U test and is indicated by * ($P < 0.05$) or ** ($P < 0.01$) when the difference between the mutant strain and the wild-type or complemented strains is statistically supported.

restored upon complementation (Fig. 2B; see Fig. S2 in the supplemental material). These results verify that Vfr is required for *exlBA* expression.

Vfr binds *in vitro* directly to *exlBA* promoter. To test the ability of Vfr to directly interact with the *exlBA* promoter (*exlBAp*), we performed electrophoretic mobility shift assays (EMSAs) using recombinant His-tagged Vfr protein. Figure 3A illustrates that Vfr binds with a high affinity to the 60-mer Cy5-labeled *exlBAp* fragment encompassing the putative binding site of Vfr, as revealed by the apparent equilibrium binding constant

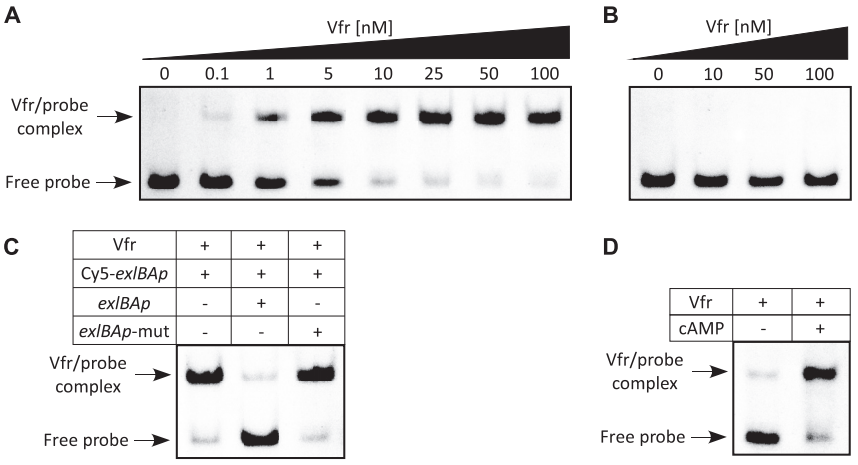


FIG 3 Vfr binds specifically to *exlBA* promoter in a cAMP-dependent manner. (A and B) The recombinant His_6-Vfr protein was incubated with 0.5 nM Cy5-labeled *exlBAp* probe (A) or the mutated *exlBAp*-mut probe (B) for 15 min before electrophoresis. Arrows indicate the positions of unbound free probes and Vfr-promoter probe complexes. The concentrations of Vfr used in the assay are also shown above each gel. (C) His_6-Vfr (50 nM) was incubated with 0.5 nM Cy5-*exlBAp* probe, and where indicated, a 200-fold excess of unlabeled wild-type or mutated probe was added to the reaction mixture prior to incubation. (D) His_6-Vfr (50 nM) was incubated with Cy5-*exlBAp* probe (0.5 nM) in the presence or absence of 20 μM cAMP in the binding buffer, as indicated. Note that in panels A, B, and C, cAMP (20 μM) was provided in the binding assay.

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

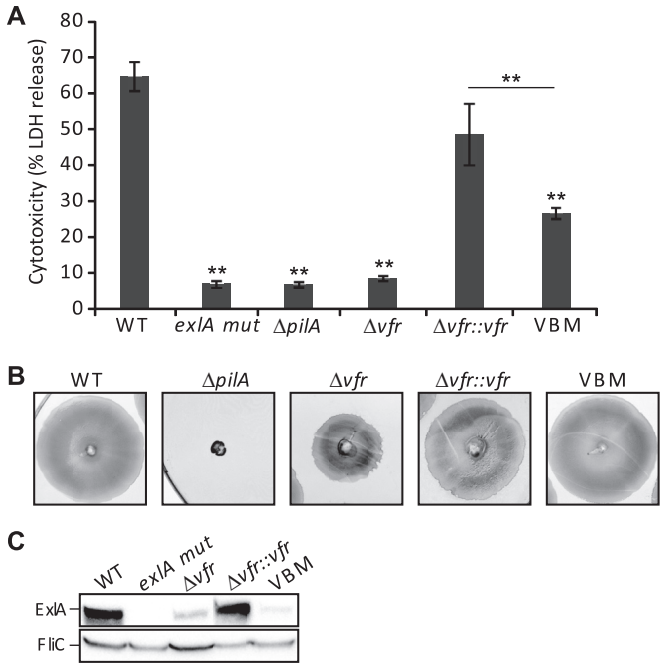


FIG 4 Vfr controls synthesis of TFP and ExlBA, and both are required for cytotoxicity. (A) ExlA-dependent cytotoxicity on A549 epithelial cells of the wild type (WT), various mutants, and the complemented strains. Cytotoxicity was measured after 4.5 h of incubation with the bacteria added at an MOI of 10. The experiments were performed in triplicate, and the bars indicate the standard deviations. The *P* value ($P < 0.01$) was determined using a Mann-Whitney U test and is indicated by ** when the difference between the WT strain and each mutant is statistically supported. (B) Twitching ability of the indicated strains after 48 h at 37°C. The twitching area was visualized following staining of the plates with Coomassie blue. (C) Analysis of ExlA secreted by the different strains after a 2-h infection of A549 cell monolayers. Proteins from cell culture medium were TCA precipitated and used in an immunoblot analysis with anti-ExlA and anti-FliC antibody probes. FliC was used as a loading control.

(K_{eq}) of approximately 2.5 nM. The binding specificity was demonstrated by observing that the electrophoretic mobility of a fragment (Cy5-*exlBAp* mut) with mutation (in boldface) in three conserved bases in the consensus (CTTTCGTGAATCAGT**CT**A**G**A) (Fig. 1C) was not altered by Vfr, even at a high protein concentration (Fig. 3B). Moreover, Vfr binding to Cy5-*exlBAp* could be outcompeted with unlabeled probe containing the wild-type binding site but not with probe containing the mutated sequence (Fig. 3C). This clearly shows that the sequence present in *exlBA* promoter is a true Vfr binding site (VBS).

Inactivation of *vfr* affects both type 4 pilus function and ExlBA-dependent cytotoxicity. As ExlA leads to permeabilization of epithelial cell membranes and cell death (19, 25), we directly measured the impact of the *vfr* deletion on the cytotoxicity of the ExlA-producing strain IHMA87. The absence of Vfr strongly reduced the cytotoxicity of this strain toward A549 epithelial cells to a level similar to that observed with an ExlA-deficient strain (IHMA87 *exlA* mut). Additionally, the phenotype was restored to an almost wild-type level in the complemented strain (Fig. 4A).

Vfr is a global regulator and was shown to modulate the expression of several virulence factors in the PAO1- and PA14-like strains. For instance, Vfr regulates directly the transcription of the *fimS-algR* locus, with AlgR activating the expression of *fimU pilVWXY$_1$Y$_2$E*, whose products are required for type 4 pilus (TFP) synthesis (32–34). The VBS identified in the *fimS* promoter of PAO1 (31) is also present in the one of IHMA87 (28; www.pseudomonas.com), suggesting that the TFP regulation is conserved. The twitching motility that relies on TFP was impaired in IHMA87 Δ*vfr*, although to a lesser extent than when the major TFP subunit PilA was absent (IHMA87 Δ*pilA*). Introduction of a copy of *vfr* corrected the motility defect of the mutant strain (Fig. 4B).

The Vfr-dependent regulation of TFP is significant, because ExlA cytotoxicity is facilitated by functional TFP (25) (Fig. 4A). To distinguish the effect of Vfr on TFP from

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

the one on ExlBA during eukaryotic cell infection, we introduced the mutations in the VBS directly on the *exlBA* promoter in the chromosome by replacing the original VBS with the 3-base-mutated sequence (Fig. 1C) shown not to interact with Vfr in the EMSA (Fig. 3B and C). In this VBS-mutated strain (VBM), the ExlA secretion was reduced to a level similar to that found in the *vfr* mutant in both liquid medium (LB) (Fig. 2A) and during eukaryotic cell infection (Fig. 4C). This further confirms the direct regulation of *exlBA* expression by the transcriptional regulator. The VBM strain displayed a cytotoxic phenotype on epithelial cells intermediate between those observed for the wild-type strain and the Δ*vfr* strain (Fig. 4A). Therefore, both the TFP and ExlBA are regulated by Vfr, and Vfr-dependent activation of *exlBA* expression during cell infection is required for full cytotoxicity.

***exlBA* expression requires cAMP.** Vfr controls gene expression in response to levels of the second messenger cAMP, as binding of the allosteric cAMP molecule increases its DNA-binding activity to target promoters (31, 35–37). However, Vfr regulation of the QS regulator-encoding *lasR* gene is cAMP independent (36). Using the EMSA, we have demonstrated that cAMP is required for *in vitro* binding of Vfr to the *exlBA* promoter, as the recombinant His_6-Vfr shifted the fluorescent probe only in the presence of the cyclic nucleotide (Fig. 3D). In *P. aeruginosa*, the cAMP levels can be modulated by calcium concentrations (38) and EGTA-induced calcium depletion of the medium has been shown to increase intracellular cAMP levels (33). We assessed the effect of different cAMP levels on *exlBA* expression *in vivo* by introducing the *lacZ* gene into *exlA* in the *vfr* mutant, its complemented version, and the strain with mutated VBS (VBM), creating transcriptional reporters, and then comparing their activities to that of the IHMA87 *exlA*::*lacZ* strain. When the calcium chelator EGTA was added to the medium, we observed increased β-galactosidase activity in the wild-type strain (Fig. 5A). This effect was abolished in the *vfr* mutant, in which *exlBA* expression was decreased and became "blind" to EGTA addition, but was restored in the complemented strain. As expected, the β-galactosidase activities in the latter strain were lower because of the reduced amount of Vfr in the complemented strain compared to the wild-type strain (Fig. S2). Moreover, the VBM mutant carrying mutation of the VBS in *exlBA* promoter exhibited *exlBA* expression similar to that seen in the *vfr* mutant (Fig. 5A). These data further confirmed that the expression of *exlBA in vivo* requires both Vfr and cAMP since cAMP is needed to promote Vfr binding.

CyaB is the adenylate cyclase controlling ExlA activity. Production and degradation of intracellular cAMP are tightly controlled by adenylate cyclases (ACs) and phosphodiesterases (PDEs) to ensure homeostasis. In PAO1, three ACs have been identified: CyaA, the cytosolic class I AC; CyaB, the membrane-bound class III AC; and ExoY, the exotoxin secreted by the T3SS and active only in eukaryotic cells (reviewed in reference 39). Examination of the IHMA87 genome identified *cyaA* and *cyaB* genes, while it lacked *exoY*. This is in accordance with the previously noted absence of the genes for the T3SS and the effectors—a distinctive feature of the PA7-like strains (17, 22).

To address which AC provides to Vfr the cAMP molecule required for *exlBA* expression, we created individual deletions in *cyaA* and *cyaB* in the IHMA87 background and combined the two mutations in a single strain. We then created the *exlA-lacZ* transcriptional fusion in the mutants and assessed the effect of the *cya* mutations on *exlA-lacZ* expression. Only the absence of *cyaB* impaired the activation of *exlBA* expression under the conditions of calcium limitation (Fig. 5A). Moreover, inactivation of both *cya* genes did not reduce further the promoter activity and reintroduction of the *cyaB* copy restored the wild-type phenotype of IHMA87 Δ*cyaB*.

We also examined the ExlA-dependent cytotoxicity on epithelial A549 cells of the different mutants to determine which AC might affect the cAMP/Vfr pathway during host cell infection. As shown in Fig. 5B, solely the absence of *cyaB* affected strongly the cytotoxicity of IHMA87 to levels similar to that observed with IHMA87 Δ*vfr*. Inactivation of both genes did not further decrease cytotoxicity, and introduction of a copy of the

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

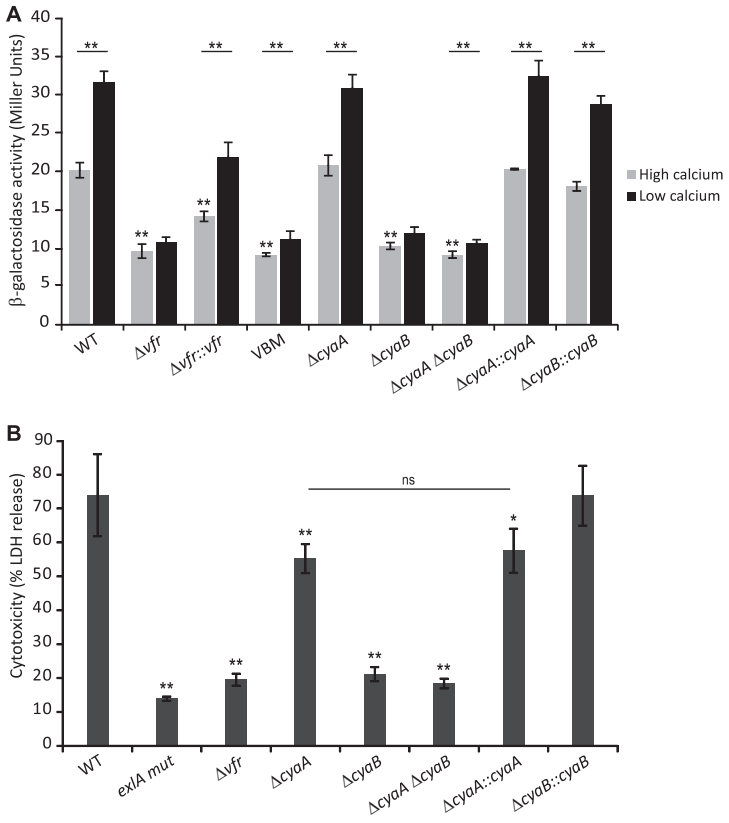


FIG 5 Vfr controls *exlBA* expression in a cAMP-dependent manner. (A) β-Galactosidase activities of the indicated wild-type (WT), mutant, and complemented strains carrying a chromosomal *exlA-lacZ* transcriptional fusion. Strains were grown in LB medium supplemented with $CaCl_2$ (high-calcium condition) or EGTA-$MgCl_2$ (low-calcium condition) at 37°C to an OD_{600} of 1.5. (B) ExlA-dependent cytotoxicity on A549 epithelial cells of the indicated wild-type (WT), *vfr*, *cyaA*, and/or *cyaB* mutant, and corresponding complemented strains measured after 4.5 h of infection (MOI of 10). The experiments were performed in triplicate, and the error bars indicate the standard deviations. The *P* value was determined using Mann-Whitney U test and is indicated above the error bars by * ($P < 0.05$) or ** ($P < 0.01$) when the difference from the wild-type strain is statistically supported. In panel A, the asterisks above the horizontal line compare each strain under the two different growth conditions. Please note that in panel B, while the Δ*cyaA* mutant seems slightly less cytotoxic than the wild-type strain, its phenotype is not complemented (ns, not supported by statistics) and is probably due to genetic manipulation.

cyaB gene into the IHMA87 Δ*cyaB* strain reversed the cytotoxicity to wild-type levels (Fig. 5B). Therefore, in IHMA87 the expression of *exlBA* relies on the secondary messenger cAMP produced by CyaB, when grown under laboratory conditions and during the interaction with mammalian cells.

The absence of ExlBA-dependent cytotoxicity in PA7 is due to a frameshift in the *vfr* gene. We recently analyzed several phenotypes of PA7-like strains, including cytotoxicity on various cell types and different types of motility (21). Compared to the ExlA-producing strains IHMA87 and CLJ1 (19), the reference PA7 strain secreted negligible amounts of ExlA *in vitro*, was poorly cytotoxic toward eukaryotic cells, and was devoid of TFP-dependent twitching motility (21). All these traits being consistent with the phenotypes of the IHMA87 Δ*vfr* mutant, we wondered whether Vfr synthesis or function was affected in PA7. We first assessed the amount of Vfr in the PA7 strain by immunoblotting and then extended our analysis to all the PA7-like strains from our collection (21). We observed different Vfr levels in cells grown in calcium limiting medium. For example, we detected large amounts in CLJ1 and IHMA87 and smaller amounts in PA70 and IHMA567230, and the protein was absent from PA7 (Fig. 6A).

The comparison of the *vfr* gene in PA7 (28; www.pseudomonas.com) to that of IHMA87 revealed the absence of the adenine 293, resulting in a frameshift that creates two putative overlapping ORFs. The first one encodes a 161-amino-acid (aa)-long

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

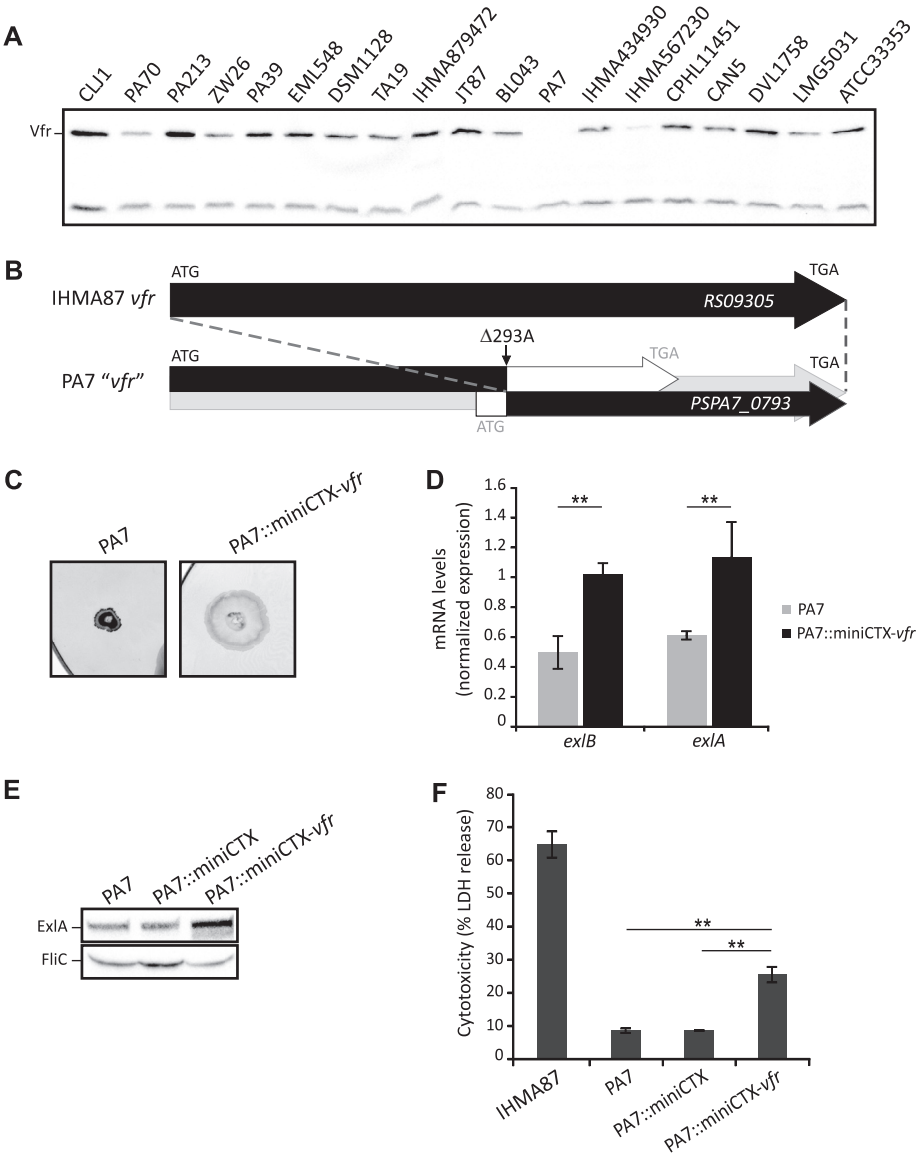


FIG 6 Natural mutation in *vfr* of *P. aeruginosa* PA7 impacts TFP synthesis and ExlA-dependent virulence. (A) Immunoblot of the cytosolic fraction of the different PA7-like strains analyzed in reference 21. The upper band corresponds to Vfr, while the lower band is a contaminant used as a loading control. (B) The *PSPA7_0793* gene is shorter than the corresponding *RS09305* gene of IHMA87 (AZPAE15042) (28; www.pseudomonas.com). Deletion of the adenine at 293 creates a frameshift that leads to two putative ORFs, as shown in the box. The 3′ ORF corresponds to the annotation "*vfr*." (C) Twitching abilities of PA7 and PA7 containing the miniCTX-*vfr* plasmid integrated into the *att* chromosomal site strains. To visualize the twitching area, after 48 h of incubation at 37°C, the plates were stained with Coomassie blue. (D) Complementation of PA7 with *vfr* impacts mRNA levels of *exlB* and *exlA* genes, as assessed by RT-qPCR. The relative mRNA quantity of the PA7 strain compared to the PA7::miniCTX-*vfr* strain is shown. The experiments were performed in triplicate, and the error bars indicate the standard errors of the means. The Mann-Whitney U test was used to calculate the *P* values (**, $P < 0.01$). (E) Immunoblot analysis of ExlA secreted by the different PA7 strains grown in LB. FliC is used as a loading control. (F) ExlA-dependent cytotoxicity on A549 epithelial cells of IHMA87, PA7, and PA7 containing in the chromosome either the empty miniCTX1 plasmid or the miniCTX-*vfr* plasmid. Cytotoxicity was measured after 4.5 h of infection at an MOI of 10. The experiments were performed in triplicate, and the error bars indicate the standard deviations. The *P* value ($P < 0.01$), determined using Mann-Whitney U test, is indicated by ** when the difference between the PA7 strains is statistically supported.

protein consisting of N-terminal 98 aa identical to Vfr, and the other, annotated as *vfr* (*PSPA7_0793*), encodes a putative 118-aa-long protein translated from a wrong-start codon corresponding to the 116-aa C-terminal portion of Vfr (Fig. 6B). The base deletion in the PA7 *vfr* gene was confirmed by PCR amplification and resequencing.

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

In order to examine whether the phenotypes of PA7 were due to *vfr* inactivation or additional mutations, we introduced one copy of intact *vfr* into the PA7 chromosome and verified the production of Vfr in this complemented strain (Fig. S2). The PA7::miniCTX-*vfr* strain exhibited an increased ability to twitch (Fig. 6C) and higher *exlA* and *exlB* mRNA levels compared to the wild-type PA7 (Fig. 6D), thus producing higher ExlA secretion (Fig. 6E). The introduction of functional *vfr* in PA7 also led to the creation of a strain with increased cytotoxicity toward the A549 cells (Fig. 6F), demonstrating the effect of *vfr* expression on PA7 virulence. These results clearly show that the presence of a functional Vfr regulator is required for full cytotoxicity of PA7-like strains as seen in the other strains of this clade.

DISCUSSION

Since its discovery in 1994 (37), Vfr has emerged as an important global regulator of *P. aeruginosa* virulence. This transcriptional factor, which belongs to the cAMP receptor protein (CRP) family, affects directly or indirectly the expression of more than 200 genes, controlled either positively or negatively (33, 40). The virulence traits positively regulated by Vfr include the TFP, QS, exotoxin A, T2SS, and T3SS, among others. However, when overproduced, the Vfr protein can also impact negatively flagellar biosynthesis by inhibiting the synthesis of the major flagellar regulator FleQ. In this work, we extended the regulon of Vfr by demonstrating that it controls directly and positively the expression of the main virulence factor of the PA7-like strains, namely, the T5SS ExlBA. Therefore, the two major virulence determinants in *P. aeruginosa*, T3SS and ExlBA, are encoded by genes that are mutually excluded from the genomes of different *P. aeruginosa* groups (PAO1/PA14 versus PA7), but they share similarities in their mechanisms of action and regulation. Effectively, both systems require close host cell contact for translocation and action of their toxins, and the interaction is mediated by extracellular appendages, the *P. aeruginosa* TFP (25, 41). Furthermore, for full T3SS- and ExlA-dependent cytotoxicity, they both require CyaB, the main cAMP-producing enzyme, and the cAMP receptor Vfr (33; this work). While Vfr was known to control T3SS synthesis for over a decade (33), the underlying mechanism was deciphered only recently (42). Here, we demonstrated the cAMP-dependent binding of Vfr to the VBS identified in the *exlBA* promoter of PA7-like strains and its requirement for cytotoxicity toward eukaryotic cells. Deletions of *cyaB* and/or *vfr* in a classical T3SS-positive strain led to an attenuated phenotype with reduced colonization and dissemination in a mouse model of lung infection, mostly through downregulation of T3SS (43). Thus, we can infer the same cAMP/Vfr regulatory pathway requirement *in vivo* for the virulence of the PA7-like strains that relies on exolysin ExlA, supported by our previous observation that PA7 is avirulent in mice (19).

The *P. aeruginosa* T5bSS ExlBA is a new TPS found in strains of the PA7 group. TPSs are typically encoded by contiguous genes organized in a single operon, with the gene encoding the transporter preceding the sequence encoding the toxin (26). However, alternate gene organizations for other TPS-encoding genes have been described. For example, the *P. aeruginosa* PdtBA system is encoded by two noncontiguous genes transcribed in different transcriptional units, but they are coregulated (27). We showed that the *exlBA* genes are transcribed as a single transcriptional unit from the promoter controlled by Vfr. The *exlBA* expression is low when the IHMA87 strain is cultivated in rich liquid medium, as suggested by very low activity of a reporter *lacZ* fusion in the wild-type IHMA87 strain. This could reflect the poorly conserved −35 (TAACAA) and −10 (TAATCT) sequences that deviate from the consensus sequences (−35 TTGACA/−10 TATAAT) but govern *exlBA* transcription (Fig. S1), suggesting that other factors besides Vfr might be required for an efficient transcription.

Vfr binding to its target promoters relies on VBS whose location relative to the TSSs determines the mechanism of transcriptional regulation. Vfr downregulates expression of few genes where the VBS overlaps the −10 sequence or is located just downstream of the TSS, as reported, respectively, for *fleQ* and one of the four *rhlR* promoters (40, 44). For them, Vfr binding probably reduces transcription initiation by impairing the RNA

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

polymerase recruitment. However, most genes are positively controlled by Vfr, the VBSs being located upstream of the TSS at variable distance. For instance, in *algD* far upstream (FUS) promoter, VBS is centered at bp −362.5 (31), while it is at bp −58 from the TSS at the *vfr* promoter (36). The VBSs can also overlap the −35 sequence, as reported for the *ptxR* T2 and *exsA* promoters (42, 45). Consequently, Vfr employs several mechanisms for RNA polymerase recruitment to promote gene expression, similar to what was observed for the CRP of *Escherichia coli*, with which it shares 67% sequence identity (37, 46, 47). With its VBS centered at around bp −97, the *exlBA* promoter should belong to a class I CRP promoter type, whereby Vfr recruits RNA polymerase by interacting with the C-terminal domain of its α-subunit (α-CTD). The *in vitro* binding of Vfr on *exlBA* promoter occurs with a high affinity (K_{eq} of ≈2.5 nM), placing it in the top range of those previously reported for other target sequences of this regulator (36, 42). As for most genes, binding activity of Vfr depends on the cyclic nucleotide cAMP and factors affecting its cellular levels.

In the PA7 reference strain, two key regulatory genes, *vfr* and *pqsR* (*mvfR*), carry mutations and are inactivated. Although the base deletion within *vfr* was not identified in the original report (17), it is reminiscent of the annotated *pqsR* pseudogene resulting from the deletion of two consecutive bases, the guanidine and cytosine at positions 698/699 (28; www.pseudomonas.com). This *pqsR* mutation accounts for some phenotypic differences between PA7 and PAO1 (17), as the PQS system controls the expression of several virulence factors such as elastase, pyocyanin, and rhamnolipids (48). We demonstrated that the absence of Vfr also impacts PA7 virulence, particularly by reducing TFP and ExlBA synthesis, required for full cytotoxicity. The importance of the Vfr regulator in *P. aeruginosa* physiology and adaptation during chronic infection is highlighted by mutations in its gene commonly found among CF isolates (49). Loss of regulatory genes affecting virulence is a well-known means of reducing the aggressive phenotype of CF isolates and promotes chronic long-term infections (49, 50). Interestingly, spontaneous mutations in *vfr* were also observed in PAO1 after several cycles of static growth, leading to a secretion-defective and impaired-motility phenotype that could potentially benefit the bacterium under these conditions (51).

Recently we found that other *Pseudomonas* species, such as *Pseudomonas putida*, *Pseudomonas protegens* and *Pseudomonas entomophila*, express ExlA-like toxins which are responsible for pyroptotic death of macrophages (52). Orthologs of *vfr* and *cyaB* genes are present in these species, inferring a common regulatory mechanism. Identification of the signaling events triggering cAMP synthesis upon host cell contact, as well as other regulatory circuits controlling ExlBA synthesis, will be the next steps toward understanding of the virulence exerted by *P. aeruginosa* taxonomic outliers.

MATERIALS AND METHODS

Bacterial strains and culture conditions. The *P. aeruginosa* strains used in this study are described in Table 1. Bacteria were grown in lysogeny broth (LB) at 37°C with agitation. To assess the effect of calcium on *exlBA* expression, overnight cultures were diluted in LB to an optical density at 600 nm (OD_{600}) of 0.1. When cultures reached an OD_{600} of 0.5, either 5 mM calcium (high-calcium condition) or 20 mM $MgCl_2$–5 mM EGTA at pH 8.0 (low-calcium condition) was added, and growth was continued until the OD_{600} reached 1.0 to 1.5. For the introduction of plasmids into *P. aeruginosa*, strains were selected following mating with *E. coli* donors on LB plates supplemented with 25 μg/ml irgasan. The following antibiotics were added when needed: 75 μg/ml gentamicin (Gm) and 75 μg/ml tetracycline (Tc).

Plasmids and genetic manipulation. The plasmids utilized in this study and the primers used for PCR are listed in Table 1 (and see Table S1 in the supplemental material). To generate *P. aeruginosa* deletion mutants, upstream and downstream flanking regions of *vfr*, *cyaA*, and *cyaB* were fused by the splicing by overlap extension PCR (SOE-PCR) procedure using the appropriate primer pairs (53). The resulting fragments of 873, 884, and 850 bp were cloned into the pCR-Blunt II-TOPO vector and then subcloned into the PstI-BamHI (for *vfr*) or HindIII-XhoI (for *cya*) sites of the pEXG2 plasmid to obtain the pEXG2-Mut-vfr, pEXG2-Mut-cyaA, and pEXG2-Mut-cyaB plasmids, respectively. Each wild-type gene with its own promoter region was amplified by PCR and sequenced. To complement the mutants, the DNA fragments containing the *cyaA* (3,385-bp) and *cyaB* (1,998-bp) genes were inserted into the SmaI site of the integrative vector miniCTX1 by sequence- and ligation-independent cloning (SLIC) (54), leading to miniCTX-*cyaA* and miniCTX-*cyaB*. The *vfr* fragment (1,220 bp) was cloned in the BamHI-HindIII sites of the plasmid, leading to miniCTX-comp-*vfr*. To create the transcriptional *exlA-lacZ* fusion, the upstream region of *exlA*, the *lacZ* gene, and the downstream region of *exlA* were amplified using the

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

TABLE 1 Bacterial strains and plasmids used in this work

Strain or plasmid	Genotype or relevant properties	Reference or source
Strains		
P. aeruginosa		
IHMA879472 (IHMA87)	Wild-type strain (urinary infection)[a]	58
IHMA87 *exlA mut*	IHMA87 with pEXG2 inserted into *exlA* gene (Gm^r)	25
IHMA87 Δ*pilA*	IHMA87 with nonpolar *pilA* deletion	25
IHMA87 Δ*vfr*	IHMA87 with nonpolar *vfr* deletion	This work
IHMA87 Δ*vfr*::*vfr*	IHMA87 Δ*vfr* with miniCTX-comp-*vfr* (Tc^r)	This work
IHMA87 VBM	IHMA87 carrying a mutation in the Vfr binding site upstream of *exlB* gene, in *exlBA* promoter	This work
IHMA87 Δ*cyaA*	IHMA87 with nonpolar *cyaA* deletion	This work
IHMA87 Δ*cyaA*::*cyaA*	IHMA87 Δ*cyaA* with miniCTX-*cyaA* (Tc^r)	This work
IHMA87 Δ*cyaB*	IHMA87 with nonpolar *cyaB* deletion	This work
IHMA87 Δ*cyaB*::*cyaB*	IHMA87 Δ*cyaB* with miniCTX-*cyaB* (Tc^r)	This work
IHMA87 Δ*cyaA* Δ*cyaB*	IHMA87 with nonpolar *cyaA* and *cyaB* deletion	This work
IHMA87 *exlA*::*lacZ*	IHMA87 with promoterless *lacZ* in *exlA*	This work
IHMA87 Δ*vfr exlA*::*lacZ*	IHMA87 Δ*vfr* with promoterless *lacZ* in *exlA*	This work
IHMA87 Δ*vfr*::*vfr exlA*::*lacZ*	IHMA87 Δ*vfr*::*vfr* with promoterless *lacZ* in *exlA*	This work
IHMA87 VBM *exlA*::*lacZ*	IHMA87 VBM with promoterless *lacZ* in *exlA*	This work
IHMA87 Δ*cyaA exlA*::*lacZ*	IHMA87 Δ*cyaA* with promoterless *lacZ* in *exlA*	This work
IHMA87 Δ*cyaA*::*cyaA exlA*::*lacZ*	IHMA87 Δ*cyaA*::*cyaA* with promoterless *lacZ* in *exlA*	This work
IHMA87 Δ*cyaB exlA*::*lacZ*	HMA87 Δ*cyaB* with promoterless *lacZ* in *exlA*	This work
IHMA87 Δ*cyaB*::*cyaB exlA*::*lacZ*	IHMA87 Δ*cyaB*::*cyaB* with promoterless *lacZ* in *exlA*	This work
IHMA87 Δ*cyaA* Δ*cyaB exlA*::*lacZ*	IHMA87 Δ*cyaA* Δ*cyaB* with promoterless *lacZ* in *exlA*	This work
PA7	Wild-type strain (nonrespiratory infection)	17
PA7::miniCTX	PA7 with miniCTX1 empty vector (Tc^r)	This work
PA7::miniCTX-*vfr*	PA7 with miniCTX-comp-*vfr* (Tc^r)	This work
E. coli		
Top10	Chemically competent cells	Invitrogen
BL21 Star(DE3)	F$^-$ *ompT hsdSB*(r_B^- m_B^-) *gal dcm rne131*(DE3)	Invitrogen
Plasmids		
pCR-Blunt II-TOPO	Commercial cloning vector (Kn^r)	Invitrogen
pRK2013	Helper plasmid conjugative properties (Kn^r)	59
pEXG2	Allelic exchange vector (Gm^r)	60
pEXG2-Mut-vfr	pEXG2 carrying SOE-PCR fragment for deletion of *vfr*	This work
pEXG2-VBM	pEXG2 carrying SOE-PCR fragment for mutation of Vfr binding site in *pexlBA*	This work
pEXG2-Mut-cyaA	pEXG2 carrying SOE-PCR fragment for deletion of *cyaA*	This work
pEXG2-Mut-cyaB	pEXG2 carrying SOE-PCR fragment for deletion of *cyaB*	This work
pEXG2-exlBA-lacZ	pEXG2 carrying *lacZ* integrated within *exlBA* fragment for integration into chromosome	This work
miniCTX1	Site-specific integrative plasmid (*att B* site, Tc^r)	61
miniCTX-comp-*vfr*	miniCTX carrying *vfr* gene	This work
miniCTX-*cyaA*	miniCTX carrying *cyaA* gene	This work
miniCTX-*cyaB*	miniCTX carrying *cyaB* gene	This work
pET15BVP	Expression vector for inducible production of N-terminally His_6-tagged proteins (Ap^r)	62
pET15BVP-purif-Vfr	Expression vector pET15BVP carrying PCR fragment for Vfr overproduction	This work

[a]Referred to as AZPAE15042 at www.pseudomonas.com.

SLIC_pEXG2_exlBA′_F/SLIC_exlBA′_lacZ_R, SLIC_lacZ_F/SLIC_lacZ_R and SLIC_lacZ_exlA′_F/SLIC_exlA′_pEXG2_R primer pairs. The resulting fragments of 700, 3,123, and 700 bp were then cloned into SmaI-cut pEXG2 by SLIC. Flanking *exlA* regions were designed so that the *lacZ* coding sequence with its own ribosome binding site can be inserted in frame after the third codon of *exlA*. The miniCTX and pEXG2-derived vectors were transferred into *P. aeruginosa* IHMA87 or PA7 strains by triparental mating using pRK2013 as a helper plasmid. For allelic exchange using the pEXG2 plasmids, cointegration events were selected on LB plates containing Irgasan and gentamicin. Single colonies were then plated on NaCl-free LB agar plates containing 10% (wt/vol) sucrose to select for the loss of plasmid, and the resulting sucrose-resistant strains were checked for gentamicin sensitivity and mutant genotype.

For overproduction of Vfr in *E. coli*, the *vfr* gene was PCR amplified using IHMA87 genomic DNA as the template, and the amplicon was inserted into NdeI-BamHI sites of the pET15b plasmid. All the constructions were verified by sequencing.

Mapping of the 5′ end using circularization of *exlB* cDNA. The procedure for 5′-end mapping was carried out as previously described (55) with some modification. Total RNA was isolated from IHMA87 at an OD_{600} of 2.0 using hot phenol-chloroform extraction. The residual DNAs in 30 μg of total RNA were removed with Turbo DNase I (4 U; Thermo Fisher Scientific). Following phenol-chloroform extraction and ethanol precipitation, 15 μg of total RNA was reverse transcribed using 5′-phosphorylated *exlB* gene-specific primer (R_exlB+60; 2 pmol) and SuperScript III reverse transcriptase (400 U, 40 μl of reaction

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

volume; Thermo Fisher Scientific). After reverse transcription, the residual RNAs were removed using RNase H (5 U) and RNase I_f (50 U). The cDNA was purified using the Oligo Clean & Concentrator kit (Zymo Research) with a reduced ethanol volume (200 μl) and recovered with 19 μl of nuclease-free water. The 8.75 μl of cDNA was used for the circularization reaction (15 μl of reaction mixture volume, 100 U of CircLigase II; Epicentre). The control cDNA was incubated under the same condition in the absence of CircLigase II. After incubation at 60°C for 2 h, the reaction was terminated by further incubating at 80°C for 10 min. The circularized cDNA (4 μl) was used as the template in an amplification of the junction sequence using GoTaq green master mix (40 μl of reaction mixture volume; Promega) and *exlB* primer pairs (F_exlB+4 and R_exlB+3). The cycling conditions were 94°C for 3 min, followed by 36 cycles of 94°C for 20 s, 55°C for 45 s, 72°C for 60 s, and 72°C for 2 min. PCR products were separated and eluted by 2% agarose and cloned into pJET2.1 vector (Thermo Fisher Scientific). Positive clones obtained from colony PCR were sequenced with the pJET reverse primer. All the primers used for circularization are listed in Table S1.

Cell culture and cytotoxicity assay. The epithelial lung carcinoma cell line A549 (ATCC CCL-185) was cultured in Roswell Park Memorial Institute (RPMI) medium supplemented with 10% fetal bovine serum (Gibco), at 37°C, in 5% CO_2. The day before the infection, the A549 cells were seeded in a 48-well tissue culture plate (Falcon) at 8×10^4 cells/well in fresh RPMI medium. For the infection, RPMI medium was replaced by nonsupplemented endothelial basal medium 2 (EBM2; Lonza). Monolayers were infected with bacteria at an OD_{600} of ≈1.0 and at a multiplicity of infection (MOI) of 10. The bacterial cytotoxicity was determined by measuring the release of lactate dehydrogenase (LDH) in the supernatant 4.5 h postinfection using the Roche cytotoxicity LDH detection kit. The negative controls were noninfected cells, and the positive controls were cells lysed by the addition of Triton X-100. The degree of cytotoxicity was calculated by determining the percentage of cytotoxicity compared to that of the positive control. The experiments were performed at least in triplicate.

ExlA secretion analysis. ExlA proteins were precipitated and concentrated using the sodium deoxycholate-trichloroacetic acid (DOC-TCA) method, either in LB cultures or after cell infection, as previously described (25). Briefly, proteins secreted by the bacteria were collected when the bacterial cultures reached an OD_{600} of ≈1.0, or after 2 h of A549 infection at an MOI of 10. The secreted proteins were precipitated by adding DOC at a final concentration of 0.02% for 30 min at 4°C. TCA was then added to 10%, and the solution was incubated overnight at 4°C. After centrifugation for 15 min at 15,000 $\times$ *g* at 4°C, the pellet was resuspended in Laemmli loading buffer corresponding to 1/100 of the original volume. Samples were separated on 8% SDS-PAGE and transferred onto 0.2-μm-pore polyvinyl difluoride (PVDF) membranes (Amersham) for Western immunoblotting. Rabbit polyclonal primary antibodies raised against the recombinant C-terminal part of ExlA and against the purified ExlA protein with its C-terminal part deleted (52) were mixed at a 1:1,000 dilution. The anti-FliC antibodies, used at a 1:1,000 dilution, were raised in rabbits (Biotem) with the recombinant FliC4 protein (residues 43 to 443 of the FliC b-type). His_6-FliC4 was expressed from pET28a.TEV-FliC4 and purified by affinity chromatography using the HIS60 Ni SuperFlow cartridge (Clontech Laboratories). Before immunization of rabbits, the tag was cleaved off by the tobacco etch virus (TEV) protease and the FliC4 protein was further purified by size exclusion chromatography using HiLoad 16/60 Superdex 75 Prep grade column (GE Healthcare). The immunoblots were developed using horseradish peroxidase (HRP [Sigma])-conjugated anti-rabbit secondary antibodies at a 1:40,000 dilution. The detection was performed with the Luminata Classico Western HRP substrate kit (Millipore) and analyzed using a Chemidoc (Bio-Rad).

Vfr analysis in PA7-like strains. Bacterial cultures (30 ml at an OD_{600} of ≈1.0) were centrifuged for 10 min at 6,000 $\times$ *g* and resuspended in 600 μl of a mixture of 25 mM Tris-HCl, 500 mM NaCl, and 10 mM imidazole (pH 8.0), supplemented with protease inhibitor cocktail (Roche). After sonication (3$\times$ 150 J), the broken cells were removed by centrifugation at 22,000 $\times$ *g* for 30 min at 4°C. A 7.5-μl concentration of each sample was subjected to SDS-PAGE (15% gel) and then transferred onto a PVDF membrane. The primary anti-Vfr antibody was used at a 1:25,000 dilution and the secondary HRP-conjugated anti-rabbit antibody at 1:50,000.

Twitching motility assay. For the twitching motility assay, bacteria were inoculated at the plastic-agar interface of agar plates containing 10 g/liter tryptone, 5 g/liter yeast extract, 1% agar, and 10 g/liter NaCl and kept at 37°C for 48 h. The agar medium was then removed, and the twitching diameter was observed after staining with 0.1% Coomassie blue (21).

***β*-Galactosidase assays.** β-Galactosidase activity was assayed when the bacterial cultures reached an OD_{600} of ≈1.5 as described previously (56), with technical details reported in reference 57.

RT-qPCR. Total RNA from 2.0 ml of cultures (OD_{600} of ≈1.0) was extracted either with hot phenol followed by ethanol precipitation or using the TRIzol Plus RNA purification kit (Invitrogen) and then treated with DNase I (amplification grade; Invitrogen). The yield, purity, and integrity of RNA were further evaluated on a NanoDrop spectrophotometer and by agarose gel migration. cDNA synthesis was carried out using 3 μg of RNA with the SuperScript III first-strand synthesis system (Invitrogen) in the presence or absence of the SuperScript III RT enzyme to assess the absence of genomic DNA. The CFX96 real-time system (Bio-Rad) was used to PCR amplify the cDNA, and the quantification was based on use of SYBR green fluorescent molecules. Two microliters of cDNA was incubated with 5 μl of Gotaq master mix (2$\times$) (Promega) and reverse and forward specific primers at a final concentration of 125 nM in a total volume of 10 μl. The cycling parameters of the real-time PCR were 95°C for 2 min (for activation of the *Taq* polymerase), 40 cycles of 95°C for 15 s and 60°C for 45 s, and finally a melting curve from 65°C to 95°C by increments of 0.5°C for 5 s to assess the specificity of the amplification. To generate standard curves, serial dilutions of the cDNA pool of the CLJ strains were used. The experiments were performed in duplicate with 3 biological replicates for each strain, and the results were analyzed with the CFX Manager

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

software (Bio-Rad). The relative expression of mRNAs was calculated using the quantification cycle ($\Delta\Delta C_q$) method relative to *rpoD* reference C_q values. The sequences of primers were designed using Primer3Plus and are given in Table S1.

Vfr purification. Overproduction of His_6-Vfr was performed in *E. coli* BL21(DE3)Star harboring pET15BVP grown at 37°C in ampicillin-containing LB medium. Overnight culture was diluted to an OD_{600} of 0.05, and expression was induced at an OD_{600} of 0.6 with 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside). After 3 h of growth, bacteria were harvested by centrifugation (4,000 $\times$ *g*, 10 min, 4°C) and resuspended in a mixture of 25 mM Tris-HCl, 500 mM NaCl, and 10 mM imidazole (pH 8.0) supplemented with protease inhibitor cocktail (Roche). The bacteria were then broken using an M110-P Microfluidizer (Microfluidics). After centrifugation at 4°C and 30,000 rpm for 30 min, the soluble fraction was directly loaded onto a 1-ml nickel column (Protino Ni-nitrilotriacetic acid [NTA]; Macherey-Nagel). The column was washed with buffer containing increasing imidazole concentrations (20, 40, and 60 mM) using an ÄKTA purifier system (GE Healthcare), and the proteins were eluted with 200 mM imidazole. Aliquots from the peak protein fractions were analyzed by SDS-PAGE, and the fractions containing Vfr were pooled and dialyzed against buffer (50 mM Tris-HCl, 100 mM KCl, 50 mM NaCl, 2 mM dithiothreitol [DTT], 2 mM EDTA, 10% glycerol, 0.5% Tween 20 [pH 7.0]) as previously described (36).

EMSA. The 60-mer DNA probes (5′Cy5-p*exlBA*_EMSA_F/p*exlBA*_EMSA_R, p*exlBA*_EMSA_F/p*exlBA*_EMSA_R, 5′Cy5-p*exlBA_mut*_EMSA_F/p*exlBA_mut*_EMSA_R, and p*exlBA_mut*_EMSA_F/p*exlBA_mut*_EMSA_R) were generated by annealing complementary pairs of oligonucleotides as described in Table S1. The probes (0.5 nM) were incubated for 5 min at 25°C in binding buffer (10 mM Tris-HCl, 50 mM KCl, 1 mM DTT, 5% glycerol, 100 μg/ml bovine serum albumin, 1 mM EDTA [pH 7.5]) containing 25 ng/μl poly(dI-dC) ($\times$1,330) and, unless indicated otherwise, 20 μM cAMP (Sigma-Aldrich). Vfr protein was added at the indicated concentrations in a final reaction volume of 20 μl and incubated for an additional 15 min at 25°C. Samples were immediately loaded on a native 5% Tris-borate-EDTA (TBE) polyacrylamide gel and run at 100 V and 4°C with cold 0.5$\times$ TBE buffer. Fluorescence imaging was performed using the Chemidoc MP apparatus. EMSAs were repeated a minimum of two times, and the representative gels are presented.

Statistical analysis. Statistical analyses were carried out using a nonparametric Mann-Whitney U test. *P* values of <0.05 and <0.01 were considered significant.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/JB.00135-18.

SUPPLEMENTAL FILE 1, PDF file, 0.6 MB.

ACKNOWLEDGMENTS

We are grateful to Katrina Forest for the Vfr antiserum, Annabelle Varrot for the pET28a.TEV-FliC4 plasmid, and Peter Panchev for proofreading the manuscript. International Health Management Association (IHMA, USA) kindly provided the *P. aeruginosa* IHMA879472 strain.

This work was supported by grants from Agence Nationale de la Recherche (ANR-15-CE11-0018-01), the Laboratory of Excellence GRAL (ANR-10-LABX-49-01), and the Fondation pour la Recherche Médicale (Team FRM 2017, DEQ20170336705). We further acknowledge support from CNRS, INSERM, CEA, and Grenoble Alps University. Alice Berry received a Ph.D. fellowship from the CEA Irtelis program.

REFERENCES

1. Moradali MF, Ghods S, Rehm BH. 2017. Pseudomonas aeruginosa lifestyle: a paradigm for adaptation, survival, and persistence. Front Cell Infect Microbiol 7:39. https://doi.org/10.3389/fcimb.2017.00039.
2. Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ, Brinkman FS, Hufnagle WO, Kowalik DJ, Lagrou M, Garber RL, Goltry L, Tolentino E, Westbrock-Wadman S, Yuan Y, Brody LL, Coulter SN, Folger KR, Kas A, Larbig K, Lim R, Smith K, Spencer D, Wong GK, Wu Z, Paulsen IT, Reizer J, Saier MH, Hancock RE, Lory S, Olson MV. 2000. Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen. Nature 406:959–964. https://doi.org/10.1038/35023079.
3. Rodrigue A, Quentin Y, Lazdunski A, Mejean V, Foglino M. 2000. Two-component systems in Pseudomonas aeruginosa: why so many? Trends Microbiol 8:498–504. https://doi.org/10.1016/S0966-842X(00)01833-3.
4. Galan-Vasquez E, Luna B, Martinez-Antonio A. 2011. The regulatory network of Pseudomonas aeruginosa. Microb Inform Exp 1:3. https://doi.org/10.1186/2042-5783-1-3.
5. Wurtzel O, Yoder-Himes DR, Han K, Dandekar AA, Edelheit S, Greenberg EP, Sorek R, Lory S. 2012. The single-nucleotide resolution transcriptome of Pseudomonas aeruginosa grown in body temperature. PLoS Pathog 8:e1002945. https://doi.org/10.1371/journal.ppat.1002945.
6. Cattoir V, Narasimhan G, Skurnik D, Aschard H, Roux D, Ramphal R, Jyot J, Lory S. 2013. Transcriptional response of mucoid Pseudomonas aeruginosa to human respiratory mucus. mBio 3:e00410-12. https://doi.org/10.1128/mBio.00410-12.
7. Gomez-Lozano M, Marvig RL, Tulstrup MV, Molin S. 2014. Expression of antisense small RNAs in response to stress in Pseudomonas aeruginosa. BMC Genomics 15:783. https://doi.org/10.1186/1471-2164-15-783.
8. Coggan KA, Wolfgang MC. 2012. Global regulatory pathways and cross-talk control pseudomonas aeruginosa environmental lifestyle and virulence phenotype. Curr Issues Mol Biol 14:47–70.
9. Jimenez PN, Koch G, Thompson JA, Xavier KB, Cool RH, Quax WJ. 2012. The multiple signaling systems regulating virulence in Pseudomonas aeruginosa. Microbiol Mol Biol Rev 76:46–65. https://doi.org/10.1128/MMBR.05007-11.
10. Balasubramanian D, Schneper L, Kumari H, Mathee K. 2013. A dy-

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

namic and intricate regulatory network determines Pseudomonas aeruginosa virulence. Nucleic Acids Res 41:1–20. https://doi.org/10.1093/nar/gks1039.
11. Suh SJ, Runyen-Janecky LJ, Maleniak TC, Hager P, MacGregor CH, Zielinski-Mozny NA, Phibbs PV, Jr, West SE. 2002. Effect of vfr mutation on global gene expression and catabolite repression control of Pseudomonas aeruginosa. Microbiology 148:1561–1569. https://doi.org/10.1099/00221287-148-5-1561.
12. Hengge R. 2009. Principles of c-di-GMP signalling in bacteria. Nat Rev Microbiol 7:263–273. https://doi.org/10.1038/nrmicro2109.
13. Valentini M, Filloux A. 2016. Biofilms and cyclic di-GMP (c-di-GMP) signaling: lessons from Pseudomonas aeruginosa and other bacteria. J Biol Chem 291:12547–12555. https://doi.org/10.1074/jbc.R115.711507.
14. Jin Y, Yang H, Qiao M, Jin S. 2011. MexT regulates the type III secretion system through MexS and PtrC in Pseudomonas aeruginosa. J Bacteriol 193:399–410. https://doi.org/10.1128/JB.01079-10.
15. Mikkelsen H, McMullan R, Filloux A. 2011. The Pseudomonas aeruginosa reference strain PA14 displays increased virulence due to a mutation in ladS. PLoS One 6:e29113. https://doi.org/10.1371/journal.pone.0029113.
16. Sall KM, Casabona MG, Bordi C, Huber P, de Bentzmann S, Attree I, Elsen S. 2014. A gacS deletion in Pseudomonas aeruginosa cystic fibrosis isolate CHA shapes its virulence. PLoS One 9:e95936. https://doi.org/10.1371/journal.pone.0095936.
17. Roy PH, Tetu SG, Larouche A, Elbourne L, Tremblay S, Ren Q, Dodson R, Harkins D, Shay R, Watkins K, Mahamoud Y, Paulsen IT. 2010. Complete genome sequence of the multiresistant taxonomic outlier Pseudomonas aeruginosa PA7. PLoS One 5:e8842. https://doi.org/10.1371/journal.pone.0008842.
18. Dingemans J, Ye L, Hildebrand F, Tontodonati F, Craggs M, Bilocq F, De Vos D, Crabbe A, Van Houdt R, Malfroot A, Cornelis P. 2014. The deletion of TonB-dependent receptor genes is part of the genome reduction process that occurs during adaptation of Pseudomonas aeruginosa to the cystic fibrosis lung. Pathog Dis 71:26–38. https://doi.org/10.1111/2049-632X.12170.
19. Elsen S, Huber P, Bouillot S, Coute Y, Fournier P, Dubois Y, Timsit JF, Maurin M, Attree I. 2014. A type III secretion negative clinical strain of Pseudomonas aeruginosa employs a two-partner secreted exolysin to induce hemorrhagic pneumonia. Cell Host Microbe 15:164–176. https://doi.org/10.1016/j.chom.2014.01.003.
20. Boukerb AM, Marti R, Cournoyer B. 2015. Genome sequences of three strains of the Pseudomonas aeruginosa PA7 clade. Genome Announc 3:e01366-15. https://doi.org/10.1128/genomeA.01366-15.
21. Reboud E, Elsen S, Bouillot S, Golovkine G, Basso P, Jeannot K, Attree I, Huber P. 2016. Phenotype and toxicity of the recently discovered exlA-positive Pseudomonas aeruginosa strains collected worldwide. Environ Microbiol 18:3425–3439. https://doi.org/10.1111/1462-2920.13262.
22. Huber P, Basso P, Reboud E, Attree I. 18 July 2016. Pseudomonas aeruginosa renews its virulence factors. Environ Microbiol Rep https://doi.org/10.1111/1758-2229.12443.
23. Freschi L, Jeukens J, Kukavica-Ibrulj I, Boyle B, Dupont MJ, Laroche J, Larose S, Maaroufi H, Fothergill JL, Moore M, Winsor GL, Aaron SD, Barbeau J, Bell SC, Burns JL, Camara M, Cantin A, Charette SJ, Dewar K, Deziel E, Grimwood K, Hancock RE, Harrison JJ, Heeb S, Jelsbak L, Jia B, Kenna DT, Kidd TJ, Klockgether J, Lam JS, Lamont IL, Lewenza S, Loman N, Malouin F, Manos J, McArthur AG, McKeown J, Milot J, Naghra H, Nguyen D, Pereira SK, Perron GG, Pirnay JP, Rainey PB, Rousseau S, Santos PM, Stephenson A, Taylor V, Turton JF, Waglechner N, Williams P, Thrane SW, Wright GD, Brinkman FS, Tucker NP, Tummler B, Winstanley C, Levesque RC. 2015. Clinical utilization of genomics data produced by the International Pseudomonas aeruginosa Consortium. Front Microbiol 6:1036. https://doi.org/10.3389/fmicb.2015.01036.
24. Cadoret F, Ball G, Douzi B, Voulhoux R. 2014. Txc, a new type II secretion system of Pseudomonas aeruginosa strain PA7, is regulated by the TtsS/TtsR two-component system and directs specific secretion of the CbpE chitin-binding protein. J Bacteriol 196:2376–2386. https://doi.org/10.1128/JB.01563-14.
25. Basso P, Ragno M, Elsen S, Reboud E, Golovkine G, Bouillot S, Huber P, Lory S, Faudry E, Attree I. 2017. Pseudomonas aeruginosa pore-forming exolysin and type IV pili cooperate to induce host cell lysis. mBio 8:e02250-16. https://doi.org/10.1128/mBio.02250-16.
26. Leo JC, Grin I, Linke D. 2012. Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. Philos Trans R Soc Lond B Biol Sci 367:1088–1101. https://doi.org/10.1098/rstb.2011.0208.
27. Faure LM, Llamas MA, Bastiaansen KC, de Bentzmann S, Bigot S. 2013. Phosphate starvation relayed by PhoB activates the expression of the Pseudomonas aeruginosa sigmavrel ECF factor and its target genes. Microbiology 159:1315–1327. https://doi.org/10.1099/mic.0.067645-0.
28. Winsor GL, Griffiths EJ, Lo R, Dhillon BK, Shay JA, Brinkman FS. 2016. Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas Genome Database. Nucleic Acids Res 44:D646–D653. https://doi.org/10.1093/nar/gkv1227.
29. Solovyev V, Salamov A, Seledtsov I, Vorobyev D, Bachinsky A. 2011. Automatic annotation of bacterial community sequences and application to infections diagnostic, p 346–353. *In* Bioinformatics 2011. Proceedings of the International Conference on Bioinformatics Models, Methods, and Algorithms, Rome, Italy, 26 to 29 January 2011.
30. Medina-Rivera A, Defrance M, Sand O, Herrmann C, Castro-Mondragon JA, Delerce J, Jaeger S, Blanchet C, Vincens P, Caron C, Staines DM, Contreras-Moreira B, Artufel M, Charbonnier-Khamvongsa L, Hernandez C, Thieffry D, Thomas-Chollier M, van Helden J. 2015. RSAT 2015: Regulatory Sequence Analysis Tools. Nucleic Acids Res 43:W50–W56. https://doi.org/10.1093/nar/gkv362.
31. Kanack KJ, Runyen-Janecky LJ, Ferrell EP, Suh SJ, West SE. 2006. Characterization of DNA-binding specificity and analysis of binding sites of the Pseudomonas aeruginosa global regulator, Vfr, a homologue of the Escherichia coli cAMP receptor protein. Microbiology 152:3485–3496. https://doi.org/10.1099/mic.0.29008-0.
32. Beatson SA, Whitchurch CB, Sargent JL, Levesque RC, Mattick JS. 2002. Differential regulation of twitching motility and elastase production by Vfr in Pseudomonas aeruginosa. J Bacteriol 184:3605–3613. https://doi.org/10.1128/JB.184.13.3605-3613.2002.
33. Wolfgang MC, Lee VT, Gilmore ME, Lory S. 2003. Coordinate regulation of bacterial virulence genes by a novel adenylate cyclase-dependent signaling pathway. Dev Cell 4:253–263. https://doi.org/10.1016/S1534-5807(03)00019-4.
34. Luo Y, Zhao K, Baker AE, Kuchma SL, Coggan KA, Wolfgang MC, Wong GC, O'Toole GA. 2015. A hierarchical cascade of second messengers regulates Pseudomonas aeruginosa surface behaviors. mBio 6:e02456-14. https://doi.org/10.1128/mBio.02456-14.
35. Serate J, Roberts GP, Berg O, Youn H. 2011. Ligand responses of Vfr, the virulence factor regulator from Pseudomonas aeruginosa. J Bacteriol 193:4859–4868. https://doi.org/10.1128/JB.00352-11.
36. Fuchs EL, Brutinel ED, Jones AK, Fulcher NB, Urbanowski ML, Yahr TL, Wolfgang MC. 2010. The Pseudomonas aeruginosa Vfr regulator controls global virulence factor expression through cyclic AMP-dependent and -independent mechanisms. J Bacteriol 192:3553–3564. https://doi.org/10.1128/JB.00363-10.
37. West SE, Sample AK, Runyen-Janecky LJ. 1994. The vfr gene product, required for Pseudomonas aeruginosa exotoxin A and protease production, belongs to the cyclic AMP receptor protein family. J Bacteriol 176:7532–7542. https://doi.org/10.1128/jb.176.24.7532-7542.1994.
38. Inclan YF, Huseby MJ, Engel JN. 2011. FimL regulates cAMP synthesis in Pseudomonas aeruginosa. PLoS One 6:e15867. https://doi.org/10.1371/journal.pone.0015867.
39. Lory S, Wolfgang M, Lee V, Smith R. 2004. The multi-talented bacterial adenylate cyclases. Int J Med Microbiol 293:479–482. https://doi.org/10.1078/1438-4221-00297.
40. Dasgupta N, Ferrell EP, Kanack KJ, West SE, Ramphal R. 2002. fleQ, the gene encoding the major flagellar regulator of Pseudomonas aeruginosa, is sigma70 dependent and is downregulated by Vfr, a homolog of Escherichia coli cyclic AMP receptor protein. J Bacteriol 184:5240–5250. https://doi.org/10.1128/JB.184.19.5240-5250.2002.
41. Sundin C, Wolfgang MC, Lory S, Forsberg A, Frithz-Lindsten E. 2002. Type IV pili are not specifically required for contact dependent translocation of exoenzymes by Pseudomonas aeruginosa. Microb Pathog 33:265–277. https://doi.org/10.1006/mpat.2002.0534.
42. Marsden AE, Intile PJ, Schulmeyer KH, Simmons-Patterson ER, Urbanowski ML, Wolfgang MC, Yahr TL. 2016. Vfr directly activates exsA transcription to regulate expression of the Pseudomonas aeruginosa type III secretion system. J Bacteriol 198:1442–1450. https://doi.org/10.1128/JB.00049-16.
43. Smith RS, Wolfgang MC, Lory S. 2004. An adenylate cyclase-controlled signaling network regulates Pseudomonas aeruginosa virulence in a mouse model of acute pneumonia. Infect Immun 72:1677–1684. https://doi.org/10.1128/IAI.72.3.1677-1684.2004.
44. Croda-Garcia G, Grosso-Becerra V, Gonzalez-Valdez A, Servin-Gonzalez L, Soberon-Chavez G. 2011. Transcriptional regulation of Pseudomonas aeruginosa rhlR: role of the CRP orthologue Vfr (virulence factor regu-

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

lator) and quorum-sensing regulators LasR and RhlR. Microbiology 157:2545–2555. https://doi.org/10.1099/mic.0.050161-0.
45. Ferrell E, Carty NL, Colmer-Hamood JA, Hamood AN, West SE. 2008. Regulation of Pseudomonas aeruginosa ptxR by Vfr. Microbiology 154:431–439. https://doi.org/10.1099/mic.0.2007/011577-0.
46. Busby S, Ebright RH. 1999. Transcription activation by catabolite activator protein (CAP). J Mol Biol 293:199–213. https://doi.org/10.1006/jmbi.1999.3161.
47. Browning DF, Busby SJ. 2016. Local and global regulation of transcription initiation in bacteria. Nat Rev Microbiol 14:638–650. https://doi.org/10.1038/nrmicro.2016.103.
48. Diggle SP, Cornelis P, Williams P, Camara M. 2006. 4-Quinolone signalling in Pseudomonas aeruginosa: old molecules, new perspectives. Int J Med Microbiol 296:83–91. https://doi.org/10.1016/j.ijmm.2006.01.038.
49. Smith EE, Buckley DG, Wu Z, Saenphimmachak C, Hoffman LR, D'Argenio DA, Miller SI, Ramsey BW, Speert DP, Moskowitz SM, Burns JL, Kaul R, Olson MV. 2006. Genetic adaptation by Pseudomonas aeruginosa to the airways of cystic fibrosis patients. Proc Natl Acad Sci U S A 103:8487–8492. https://doi.org/10.1073/pnas.0602138103.
50. Jeukens J, Boyle B, Kukavica-Ibrulj I, Ouellet MM, Aaron SD, Charette SJ, Fothergill JL, Tucker NP, Winstanley C, Levesque RC. 2014. Comparative genomics of isolates of a Pseudomonas aeruginosa epidemic strain associated with chronic lung infections of cystic fibrosis patients. PLoS One 9:e87611. https://doi.org/10.1371/journal.pone.0087611.
51. Fox A, Haas D, Reimmann C, Heeb S, Filloux A, Voulhoux R. 2008. Emergence of secretion-defective sublines of Pseudomonas aeruginosa PAO1 resulting from spontaneous mutations in the vfr global regulatory gene. Appl Environ Microbiol 74:1902–1908. https://doi.org/10.1128/AEM.02539-07.
52. Basso P, Wallet P, Elsen S, Soleilhac E, Henry T, Faudry E, Attree I. 2017. Multiple Pseudomonas species secrete exolysin-like toxins and provoke caspase-1-dependent macrophage death. Environ Microbiol 19:4045–4064. https://doi.org/10.1111/1462-2920.13841.
53. Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR. 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene 77:61–68. https://doi.org/10.1016/0378-1119(89)90359-4.
54. Li MZ, Elledge SJ. 2007. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. Nat Methods 4:251–256. https://doi.org/10.1038/nmeth1010.
55. Polidoros AN, Pasentsis K, Tsaftaris AS. 2006. Rolling circle amplification-RACE: a method for simultaneous isolation of 5′ and 3′ cDNA ends from amplified cDNA templates. BioTechniques 41:35–36, 38, 40. https://doi.org/10.2144/000112205.
56. Miller JH. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
57. Thibault J, Faudry E, Ebel C, Attree I, Elsen S. 2009. Anti-activator ExsD forms a 1:1 complex with ExsA to inhibit transcription of type III secretion operons. J Biol Chem 284:15762–15770. https://doi.org/10.1074/jbc.M109.003533.
58. Kos VN, Deraspe M, McLaughlin RE, Whiteaker JD, Roy PH, Alm RA, Corbeil J, Gardner H. 2015. The resistome of Pseudomonas aeruginosa in relationship to phenotypic susceptibility. Antimicrob Agents Chemother 59:427–436. https://doi.org/10.1128/AAC.03954-14.
59. Figurski DH, Helinski DR. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc Natl Acad Sci U S A 76:1648–1652. https://doi.org/10.1073/pnas.76.4.1648.
60. Rietsch A, Vallet-Gely I, Dove SL, Mekalanos JJ. 2005. ExsE, a secreted regulator of type III secretion genes in Pseudomonas aeruginosa. Proc Natl Acad Sci U S A 102:8006–8011. https://doi.org/10.1073/pnas.0503005102.
61. Hoang TT, Kutchma AJ, Becher A, Schweizer HP. 2000. Integration-proficient plasmids for Pseudomonas aeruginosa: site-specific integration and use for engineering of reporter and expression strains. Plasmid 43:59–72. https://doi.org/10.1006/plas.1999.1441.
62. Arora SK, Ritchings BW, Almira EC, Lory S, Ramphal R. 1997. A transcriptional activator, FleQ, regulates mucin adhesion and flagellar gene expression in Pseudomonas aeruginosa in a cascade manner. J Bacteriol 179:5574–5581. https://doi.org/10.1128/jb.179.17.5574-5581.1997.

Downloaded from http://jb.asm.org/ on August 24, 2020 at CEA/GRENOBLE

cAMP and Vfr control Exolysin expression and cytotoxicity of *Pseudomonas aeruginosa* taxonomic outliers

Alice Berry,[a] Kook Han,[b] Julian Trouillon,[a] Mylène Robert-Genthon,[a] Michel Ragno,[a] Stephen Lory,[b] Ina Attrée[a] and Sylvie Elsen[a,#]

Biology of Cancer and Infection, U1036 INSERM, CEA, University of Grenoble Alpes, ERL5261 CNRS, Grenoble, France[a]; Department of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, USA[b]

[#]Address correspondence to Sylvie Elsen, sylvie.elsen@cea.fr.

Supplemental Material and Methods

Mutagenesis of *exlBA* promoter. Three mutations were introduced into the promoter of *exlBA* directly in the chromosome of the strain IHMA87 *exlA::lacZ*, two abolishing completely the consensus boxes "-10" ("Box -10 mut") and "-35" ("Box -35 mut"), and the other optimizing the predicted "-35" box to become closer to the consensus one ("Box -35 opt"). To create the mutations, the upstream region and the downstream region flanking the targeted box were amplified using for: 1) "Box -10 mut" : Mut-pexlBA-sF1 / Mut-pexlBA1-new10-sR1 (fragment of 539 pb) and Mut-pexlBA1-new10-sF2 / Mut-pexlBA-sR2 (fragment of 559 pb) ; 2) "Box -35 mut" : Mut-pexlBA-sF1 / Mut-pexlBA1-new35-sR1 (fragment of 515 pb) and Mut-pexlBA1-new35-sF2 / Mut-pexlBA-sR2 (fragment of 583 pb) ; 3) "Box -35 opt" : Mut-pexlBA-sF1 / Mut-pexlBA1-35-sR1 (fragment of 514 pb) and Mut-pexlBA1-35-sF2 / Mut-pexlBA-sR2 (584 pb). The resulting fragments were then cloned into *Sma*I-cut pEXG2 by

Sequence- and Ligation-Independent Cloning (SLIC) (1) and the mutated sequences were introduced into the chromosome by allelic exchange. Each mutation created restriction site (*Sal*I or *Sph*I) allowing selection of mutated genotype. The primers used for PCR are listed in the Table S1.

1. **Li MZ, Elledge SJ.** 2007. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. Nat Methods **4:**251-256.

Supplemental Tables

TABLE S1. Primers used in this work

Name	Sequence (5'->3')	Use
Mut-pexlBA-sF1	GGTCGACTCTAGAGGATCCCCTGGTGATGG CGGCTGGCACC	*pexlBA* mutagenesis
Mut-pexlBA-sR2	ACCGAATTCGAGCTCGAGCCCCGACGTAGC CGG CATCGACAT	*pexlBA* mutagenesis
Mut-pexlBA1-new10-sR1	GTCGACGTCGCAGACCTCACCCGTTTG	*pexlBA* mutagenesis
Mut-pexlBA1-new10-sF2	GGTGAGGTCTGCGACGTCGACACGCAAGAC ACAAAGTTTACATAAC	*pexlBA* mutagenesis
Mut-pexlBA1-new35-sR1	GCATGCCATCCGAGTAAAGAAATTGACCG	*pexlBA* mutagenesis
Mut-pexlBA1-new35-sF2	TTCTTTACTCGGATGGCATGCACGGGTGAG GTCTGCGACTAA	*pexlBA* mutagenesis
Mut-pexlBA1-35-sR1	GTCGACATCCGAGTAAAGAAATTGACCG	*pexlBA* mutagenesis
Mut-pexlBA1-35-sF2	TTTCTTTACTCGGATGTCGACAACGGGTGAG GTCTGCGACTAA	*pexlBA* mutagenesis
Mut-vfr-F1	CTGCAGGCGCCAGCTTAGCACAGGGC	*vfr* deletion
Mut-vfr-R1	GGCGACGCGTCCGGTGACTCAGCGGCCGTC GTCGTCCTCG	*vfr* deletion
Mut-vfr-F2	GTCACCGGACGCGTCGCCC	*vfr* deletion
Mut-vfr-R2	GGATCCTCTCAACCGGGCCGACGTGG	*vfr* deletion
IHMA-Mut-cyaA-F1	CCTCTAGACCGACGCTGTTCCCTAGTCC	*cyaA* deletion
IHMA-Mut-cyaA-R1	CTGTCGGTTCATGGGCGTCC	*cyaA* deletion

IHMA-Mut-cyaA-F2	GGACGCCCATGAACCGACAGCAGGCGCTGCTGGAGCAATGA	*cyaA* deletion
IHMA-Mut-cyaA-R2	GGCTCGAGATGCCTTCCTGTGCCTGCTG	*cyaA* deletion
IHMA-Mut-cyaB-F1	CCTCTAGAGTGCTCTTCCACGCGCTGGC	*cyaB* deletion
IHMA-Mut-cyaB-R1	GAGGGTGGGCTTCATGCGCT	*cyaB* deletion
IHMA-Mut-cyaB-F2	AGCGCATGAAGCCCACCCTCTACGTCGAGCACGAACTGCCC	*cyaB* deletion
IHMA-Mut-cyaB-R2	GGCTCGAGCCTGGTGATGCTCGAAGCC	*cyaB* deletion
Vfr-binding-F1	GGATCCTACGGCATGCATCTCGATGTC	p*exlBA* Vfr binding site mutation
Vfr-binding-R1	GCCTCTAGACTGATTCACGAAAGTTTGGCG	p*exlBA* Vfr binding site mutation
Vfr-binding-F2	CTTTCGTGAATCAGTCTAGAGGCGTTTCTTCGTCCAGTCAGCAAC	p*exlBA* Vfr binding site mutation
Vfr-binding-R2	CTCGAGGCAATTGGCGTTGGCGTTCCT	p*exlBA* Vfr binding site mutation
IHMA-Comp-vfr-F	CCGGATCCCGGCCTCGAGGAAGGCCTCGCAGC	*vfr* complementation
IHMA-Comp-vfr-R	CCAAGCTTCTGACTGATCCGCGCTGTCGA	*vfr* complementation
Comp-cyaA-sF	GATATCGAATTCCTGCAGCCCGGCCGAACACCTGCTCGAGC	*cyaA* complementation
Comp-cyaA-sR	CTCTAGAACTAGTGGATCCCCCGATAGCCATGGATTACGTCCCT	*cyaA* complementation
IHMA-Comp-cyaB-sF	GATATCGAATTCCTGCAGCCCTTTCGCCGAGTTCTACCCCTAT	*cyaB* complementation
IHMA-Comp-cyaB-sR	CTCTAGAACTAGTGGATCCCCCGAGCAATCCTGGCGGGCCTC	*cyaB* complementation
SLIC_pEXG2_exlBA'_F	TGCAGGTCGACTCTAGAGGATCCCCAGCGCGGTCTGGAGTCTCG	*exlA-lacZ* fusion
SLIC_exlBA'_lacZ_R	AAGTTAAAATGCCGCGCCCCTACCTTCTATGCATGAGAACCTCTTCG	*exlA-lacZ* fusion
SLIC_lacZ_F	AGGTAGGGGCGCGGCATTTT	*exlA-lacZ* fusion
SLIC_lacZ_R	TTATTTTTGACACCAGACCAACTG	*exlA-lacZ* fusion
SLIC_lacZ_exlA'_F	CCAGTTGGTCTGGTGTCAAAAATAAGACAATCCTGTCTTCCACCTC	*exlA-lacZ* fusion
SLIC_exlA'_pEXG2_R	AGGTACCGAATTCGAGCTCGAGCCCTCGAGCTCGGTCGACGCCT	*exlA-lacZ* fusion
Purif-vfr-NdeI-F	GGCATATGGTAGCTATTACCCACACACCCA	Vfr over-production
Purif-vfr-BamHI-R	GGGGATCCTTCAGCGGGTGCCGAAGACCA	Vfr over-production
R_exlB+60	CAGCAGCAGGGCTCGGCAG	circularization

F_exlB+4	CGTACCGCTCTACCGAATCATC	circularization
R_exlB+3	CACGCGGCATCCTTCATGTATC	circularization
rpoD-F3	CTGTTCATGCCGATCAAGCTG	RTqPCR Reference gene
rpoD-R3	AACGCTGTCGACCCACTTCTC	RTqPCR Reference gene
exlB-F	CCTATGGCTACTGGACCTACA	RTqPCR
exlB-R	AGGTAGCTGTCGACATCCTTG	RTqPCR
exlA-F	CGCTGAAGGACAAGCTGGAA	RTqPCR
exlA –R	CATTACGGTCGATGCCGTTC	RTqPCR
5’Cy5-p*exlBA*_EMSA_F	CCAGTCGCGACACGCCAAACTTTCGTGAATC AGTTCACAGGCGTTTCTTCGTCCAGTAG	EMSA
p*exlBA*_EMSA_F	CCAGTCGCGACACGCCAAACTTTCGTGAATC AGTTCACAGGCGTTTCTTCGTCCAGTCAG	EMSA
p*exlBA*_EMSA-R	CTGACTGGACGAAGAAACGCCTGTGAACTG ATTCACGAAAGTTTGGCGTGTCGCGACTGG	EMSA
5’Cy5-p*exlBA*_mut_EMSA_F	CTGACTGGACGAAGAAACGCCTGTGAACTG ATTCACGAAAGTTTGGCGTGTCGCGACTGG	EMSA
p*exlBA*_mut_EMSA_F	CCAGTCGCGACACGCCAAACTTTCGTGAATC AGTCTAGAGGCGTTTCTTCGTCCAGTCAG	EMSA
p*exlBA*_mut_EMSA-R	CTGACTGGACGAAGAAACGCCTCTAGACTG ATTCACGAAAGTTTGGCGTGTCGCGACTGG	EMSA

A

GTCAATTTCTTTACTCGGATGTAACAAACGGGTGAGGTCTGCGACTAATCTACGCAAGACACAAAGTT

-35 -10

consensus :	TTGACA $N_{16/18}$ TATAAT $N_{5/8}$ +1 (A/G)
exlBAp :	TAACAA N_{18} TAATCT N_7 +1 (A)
Box -10 mut :	TAACAA N_{18} **gtcgac**
Box -35 mut :	**gcAtgc** N_{18} TAATCT
Box -35 opt :	T**cgac**A N_{18} TAATCT

B

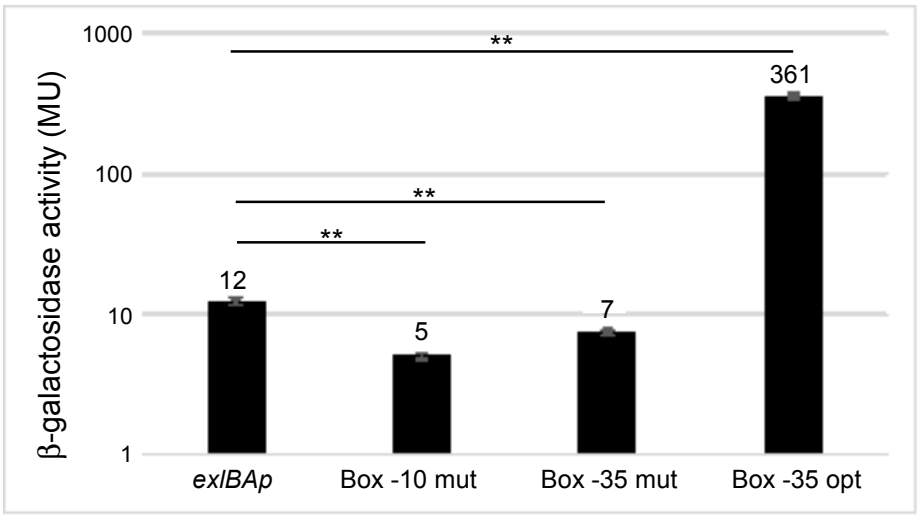


FIG S1 Definition of the *exlBA* promoter. (A) DNA sequence of the 5' region of exlB, with the nucleotide "A" identified as the TSS in bold and pinpointed by an arrow. The putative "-10" and "-35" boxes are underlined and compared to the consensus "-10" and "-35" sequences, with the conserved nucleotides underlined. The mutations introduced in the chromosome of the IHMA87 *exlA::lacZ* strain are indicated with the mutated bases in bold lower cases. The selected changes either reduce (Box -10 mut and Box -35 mut) or increase (Box -35 opt, "opt" for "optimal") the homology to the consensus. (B) β-galactosidase activities of IHMA87 *exlA::lacZ* strain with the indicated *exlBA* sequences. Strains were grown in LB medium at 37°C to an OD_{600} of 1.5 then β-galactosidase activity was measured. The experiments were performed in triplicate and the mean values of the β-galactosidase activities are indicated on the top of the histograms (in Miller Units) while the error bars indicate the standard deviations. The p-value ($p<0.01$) is determined using Mann-Whitney U test and indicated by two stars when the difference with the wild-type *exlBA-lacZ* fusion is statistically supported.

Note that the implication of another putative -35 box (underlined with dashed line, panel (A)) closer to the consensus was ruled out as its modification from GTAACA to GTCGAC ("box -35 opt") led to a higher activity of *exlBAp*. The poor matches of the identified "-10" and "-35" boxes to consensus might explain the low *exlBA* transcriptional activity measured in the IHMA87 strain harboring the native promoter fusion and the mild impact of introduced mutations. On the other hand, the change of the putative "-35" box to fit better the consensus led to strong increase (30 fold) of *exlBAp* activity.

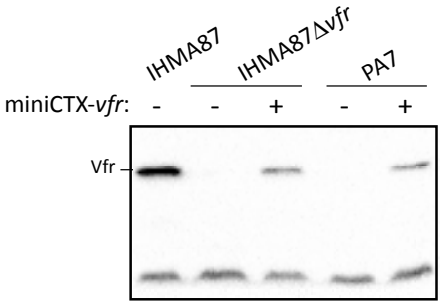


FIG S2 Vfr synthesis in wild-type and complemented strains. Western blot of the cytosolic fractions of the indicated strains grown in low calcium condition (EGTA/$MgCl_2$). The upper band corresponds to Vfr, while the lower band is a non-specific cross-reacting protein used as a loading control. The amount of Vfr observed in the complemented strains is lower compared to the parental strain IHMA87, probably due to the ectopic location (*att* site) of the own-promoter driven *vfr* gene.

4.1.2 The species-specific ErfA-dependent regulation of ExlBA

Julian Trouillon✉, Erwin Sentausa, Michel Ragno, Mylène Robert-Genthon, Stephen Lory, Ina Attrée & Sylvie Elsen✉

Published in *Nucleic Acids Research* in 2020

As stated above, I initiated this project during a 6-months internship as a Master student, and followed it at the beginning of my PhD. I searched for new regulators of *exlBA* in one of our lab's reference strains, IHMA879472 (IHMA87) (Kos et al., 2015), for which the genome was also fully sequenced and assembled in this work. Using transposon-mediated random mutagenesis in a strain carrying an *exlBA::lacZ* transcriptional fusion, I identified a second regulator of *exlBA*, which we called ErfA (for ExlBA regulatory Factor A). In this work, we show that ErfA is a direct inhibitor of *exlBA* in *P. aeruginosa*, counteracting the Vfr-dependent activation of the operon. Additionally, ErfA and Vfr regulations were shown to be species-specific as *exlBA* was found to be regulated differently in other *Pseudomonas* species. This phenomenon was explained by differences found at the promoter level, where each species was found to have specific TF binding sites, illustrating an evolutionary mechanism allowing virulence regulatory adaptation.

For this project, I performed all experiments and analyses except for IHMA87 genome assembly (ES), protein purifications (MR), RT-qPCRs (SE & MRG), some of the clonings (SE) and cytotoxicity tests for PA70, AL198 and PA23 strains (MRG).

2388–2400 *Nucleic Acids Research, 2020, Vol. 48, No. 5* *Published online 11 January 2020*
doi: 10.1093/nar/gkz1232

Species-specific recruitment of transcription factors dictates toxin expression

Julian Trouillon[1,*], Erwin Sentausa[1], Michel Ragno[1], Mylène Robert-Genthon[1], Stephen Lory[2], Ina Attrée[1] and Sylvie Elsen[1,*]

[1]Université Grenoble Alpes, CNRS ERL5261, CEA-IRIG-BCI, INSERM UMR1036, Grenoble 38000, France and [2]Department of Microbiology, Harvard Medical School, Boston, Massachusetts 02115, USA

Received November 20, 2019; Revised December 17, 2019; Editorial Decision December 19, 2019; Accepted December 21, 2019

ABSTRACT

Tight and coordinate regulation of virulence determinants is essential for bacterial biology and involves dynamic shaping of transcriptional regulatory networks during evolution. The horizontally transferred two-partner secretion system ExlB–ExlA is instrumental in the virulence of different *Pseudomonas* species, ranging from soil- and plant-dwelling biocontrol agents to the major human pathogen *Pseudomonas aeruginosa*. Here, we identify a Cro/CI-like repressor, named ErfA, which together with Vfr, a CRP-like activator, controls *exlBA* expression in *P. aeruginosa*. The characterization of ErfA regulon across *P. aeruginosa* subfamilies revealed a second conserved target, the *ergAB* operon, with functions unrelated to virulence. To gain insights into this functional dichotomy, we defined the pan-regulon of ErfA in several *Pseudomonas* species and found *ergAB* as the sole conserved target of ErfA. The analysis of 446 *exlBA* promoter sequences from all *exlBA*$^+$ genomes revealed a wide variety of regulatory sequences, as ErfA- and Vfr-binding sites were found to have evolved specifically in *P. aeruginosa* and nearly each species carries different regulatory sequences for this operon. We propose that the emergence of different regulatory *cis*-elements in the promoters of horizontally transferred genes is an example of plasticity of regulatory networks evolving to provide an adapted response in each individual niche.

INTRODUCTION

Horizontal gene transfer (HGT) is a major mechanism for the evolution of Prokaryotes facilitating their ability to adapt to specific environmental niches. The majority of genes expressed by bacteria are regulated at the transcriptional level through the action of transcription factors (TFs). In many instances, the genes encoding TFs and their target regulated genes are genetically linked; consequently, they are acquired as a single unit during HGT (1). However, the environment of the recipient, whether they are distinct species or even different strains of the same species, may be sufficiently different from that of the donor to make it necessary to rewire the regulation of the acquired structural genes and integrate them into existing regulatory networks. Mutations in the regulatory sequences near the promoters of horizontally acquired genes readily alter the specificity of recognition by the TFs and represent a simple way of placing them under the control of new regulatory elements that respond to input signals of the specific environment.

A new taxonomic group of the human pathogen *P. aeruginosa* has recently emerged, characterized by major features in their virulence factor repertoire; namely the absence of several important toxins, including the Type III Secretion System (T3SS) effectors and the associated secretion and regulatory machinery (2). Instead, they express ExlB–ExlA, a Two-Partner Secretion (TPS) system secreting a potent cytotoxin (3). ExlB (PSPA7_4641) is the cognate outer membrane transporter of the 172 kDa pore-forming cytotoxin, the Exolysin ExlA (PSPA7_4642) (4). In strains harboring the *exlBA* operon, apparent genetic scars at the T3SS-encoding locus can be identified, suggesting an unfavorable functional incompatibility between the two secretion systems or their respective exported toxins resulting in the evolutionary selection of a single secretion system (5). Whole-genome-based population studies demonstrated that the *exlBA* operon is present in two distinct phylogenetic groups, one sharing an average nucleotide identity (ANI) of ∼98% with the *T3SS*+ major group, and another representing clonal outliers with an ANI of ∼93% (2,5–8). The current cohort of strains with the *exlBA* operon and lacking the T3SS-encoding genes comprises isolates found in the environment or recovered from both acute and chronic human infections (5,6,9,10).

The presence of the *exlBA* operon in specific *P. aeruginosa* phylogenetic groups, as well as in some other *Pseudomonas*

*To whom correspondence should be addressed. Tel: +33 0438783074; Email: julian.trouillon@cea.fr; Email: sylvie.elsen@cea.fr
Present address: Erwin Sentausa, Evotec ID (Lyon) SAS, Marcy l'Étoile, France.

© The Author(s) 2020. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

species, implies its acquisition by HGT and therefore its expression might be controlled by TFs of the recipients, found at other locations on the chromosome. We recently investigated the *exlBA* regulation in the human urinary tract isolate *P. aeruginosa* IHMA879472 (IHMA87 (11)). We showed that the operon is under direct control of the global regulator Vfr, a member of the cyclic AMP receptor (CRP) family, which together with the co-activator cAMP stimulates *exlBA* expression (12). The consensus recognition sites for the CRP proteins in different bacterial species, including the *P. aeruginosa* Vfr (13), are well conserved and can be identified immediately upstream of the *exlBA* core promoter. This sequence is required for the expression of *exlBA* and was shown to specifically bind Vfr (12). Therefore, after the acquisition of the *exlBA* operon by HGT, it became part of the global cAMP/Vfr regulatory network that controls the expression of a number of virulence factors and biofilm determinants in *P. aeruginosa*. Here, we further probed the regulatory mechanisms controlling *exlBA* expression by attempting to identify additional regulators, assess their distribution and function in several *P. aeruginosa* groups and compare these to other *Pseudomonas* species.

MATERIALS AND METHODS

Bacterial strains

The bacterial strains used in this study are listed in Supplementary Table S5. *P. aeruginosa* and *E. coli* strains were grown in Lysogeny Broth (LB) at 37°C under agitation. *P. chlororaphis*, *P. protegens* and *P. putida* were cultivated at 28°C. *P. aeruginosa* strains were selected on LB plates supplemented with 25 μg/ml irgasan. Antibiotics for *P. aeruginosa* were added when needed at the following concentrations: 75 μg/ml gentamicin and 75 μg/ml tetracycline. For *P. chlororaphis*, 25 μg/ml rifampicin and 25 μg/ml gentamicin were used.

Genome sequencing and assembly

The genome of *P. aeruginosa* IHMA87 was sequenced using Illumina HiSeq (11) and completed with PacBio (Base Clear, Leiden, Netherlands) technology. Reads from both platforms were assembled using the hybrid assembler Unicycler version 0.4.0 (14) in normal mode to obtain two circular contigs with an average read depth of 136.5X. Genome annotation was carried out using Prokka version 1.12 (15) and annotation was manually curated to include or correct known gene names. The average nucleotide identity (ANI) between the chromosomes of PA7 and IHMA87 was calculated as the OrthoANIu value (16), while the synteny between the two genomes was identified and visualized using Mauve version snapshot_2015-02-13 by aligning them using the progressive Mauve algorithm with default parameters (17).

Transposon mutagenesis

A transposon mutant library was constructed in *P. aeruginosa* IHMA87 *exlBA::lacZ* using the Himar-1 mariner transposon on pBTK24 plasmid, which carries an outward-directed P*tac* promoter, making it able to either disrupt or overexpress adjacent genes. The library was generated by triparental mating the *P. aeruginosa* strain with the *E. coli* donor strain carrying pBTK24 and a pRK2013-containing helper strain. After overnight culture on LB agar plates with appropriate antibiotics, *P. aeruginosa* and *E. coli* were resuspended in LB at $OD_{600} = 1$. After incubation of *P. aeruginosa* at 42°C without agitation for 2 h, the three strains were then combined at a 1:2:2 recipient-to-donor/helper ratio, concentrated 30× and spotted for a total of 16 50-μl puddles on LB agar plates. After 4 h of incubation at 37°C (allowing one bacterial doubling, as checked by CFU counting), the puddles were scraped off, pooled and stored at −80°C. The mutant library size was estimated at 100 000 mutants by CFU counting, allowing a complete coverage of the genome (transposon insertions every 65 bp on average). Blue colonies were isolated after plating on LB agar plate containing 25 μg/ml irgasan, 75 μg/ml gentamicin and 40 μg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and overnight incubation at 37°C followed by 48 h incubation at 4°C. Among 236 mutants selected on plates and re-tested in an ONPG-based β-galactosidase activity assay, 36 displayed at least a 2-fold increase of the reporter activity in comparison to the parental strain. Transposon insertion regions in these candidate mutants were amplified by semi-random PCR and sequenced, leading to the identification of 13 genes potentially involved in *exlBA* regulation (Supplementary Table S1). All primers are listed in Supplementary Table S6.

Cell culture and cytotoxicity assay

A549 epithelial cells (ATCC CCL-185) were grown in Roswell Park Memorial Institute medium (RPMI) 1× supplemented with 10% fetal bovine serum (FBS) at 37°C, 5% CO_2. Cells were seeded in 96-well plates (50 000 cells/well, 200 μl/well). One hour before infection, the medium was replaced with non-supplemented RPMI medium. Ten minutes before infection, it was supplemented with 0.25 μM Syto24 (Thermofisher) and 0.5 μg/ml Propidium Iodide (PI). In parallel, bacteria were diluted to $OD_{600} = 0.1$ from 16 h-grown liquid cultures and further incubated for 3 h. For the infection, bacterial cultures at $OD_{600} = 1.0$ were diluted and added to monolayers to fit the multiplicity of infection (MOI) of 10. Live-cell microscopy and image analysis were done using an Incucyte S3 Live-cell Imaging System (Essen Bioscience). Total cell population was automatically counted at $T = 0$ in each well by monitoring the Syto24 labeling and further used for the normalization of the PI incorporation data. The number of cells with incorporated PI was counted every 45 min. Areas under the curves were estimated from PI incorporation curves by linear trapezoidal integration.

Routine cytotoxicity assays were performed by measuring only the PI fluorescence (excitation 544 nm/emission 590 nm) every 10 min with Fluoroskan Ascent FL2.5 (Thermo Corporation), during 8 h at 37°C. Cell shrinkage assays were performed as described previously (18). Briefly, A549 cells were stained with Dii dye (Life Technologies) for 1 h then washed and placed in RPMI containing 10% FBS before infection. Dii staining detection was done by image acquisition every hour on an Incucyte S3 Live-cell Imag-

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

ing System. Percent Red Object Confluence in RED channel was measured to quantify cell surface and normalized to t_0 to obtain cell shrinkage values.

Galleria mellonella infection assay

Larvae of *G. mellonella* were obtained from Sud-Est Appâts (Queige, France) and kept in a dark container at room temperature. White larvae of 2.5- to 3-cm size were selected and kept until infection the next day. *Pseudomonas* strains were diluted to $OD_{600} = 0.1$ from overnight cultures and grown in LB under agitation until $OD_{600} = 1$ was reached. Bacteria were then pelleted and resuspended in sterile PBS and diluted to $\sim 5.10^{-1}$ bacteria/μl or $\sim 6.10^{3}$ bacterial/μl for *P. aeruginosa* and *P. chlororaphis*, respectively. An insulin pen (HumaPen Luxura, Lilly Nederland) was used to inject 10 μl of bacterial suspension to the last proleg of the larvae. Animals injected with sterile PBS served as a control for physical trauma. CFUs for each dilution were systematically counted from the insulin pen to ensure proper bacterial loads. Forty larvae were injected per condition. Infection development was followed for 24 h at 37 or 30°C for *P. aeruginosa* or *P. chlororaphis*, respectively, and the animals were considered dead when they failed to react to touch. Strains were independently randomized and counting was blinded and done by another person, ensuring no bias in spotting and counting of CFU, as well as in counting dead larvae. Statistical significance was assessed using a Log-rank test.

RNA isolation

Total RNA was isolated from liquid cultures at $OD_{600} = 1$ using the hot phenol–chloroform extraction method. Briefly, after cell lysis in hot Lysis-phenol solution (40 mM sodium acetate, 1% SDS, 2 mM EDTA in acid phenol solution), RNA was isolated by sequential phenol–chloroform extractions and ethanol precipitation. Residual DNA was removed by treatment with DNAse following manufacturer's instructions (Invitrogen). Quantification of RNA was done using a QuBit 3.0 Fluorimeter. RNA sample quality was then assessed on an Agilent Bioanalyzer, yielding RINs of 9 or higher.

Construction of libraries for RNA-seq, sequencing and data analysis

After RNA isolation and DNAse treatment, ribosomal RNAs were depleted using the Ribo-Zero rRNA Removal Kit (Illumina) following manufacturer's instructions. The cDNA libraries were constructed from 50 ng of depleted RNA using the NEBNext Ultra II Directional RNA library prep kit following manufacturer's instructions (NEB). Libraries were size-selected to 200–700 bp fragments using SPRIselect beads, and quality was assessed on the Agilent Bioanalyzer using High Sensitivity DNA chips. Sequencing was performed at the Biopolymers Facility at Harvard Medical School on an Illumina NextSeq500. Approximately 18 million single-end 50 bp reads per sample were generated on average. More than 90% of reads were uniquely aligned for each sample to the IHMA87 or PAO1 genomes using Bowtie2 (19). Read counts per feature were then obtained with htseq-count (20). Differential gene expression between mutant and wild-type strains was assessed using DESeq2 (21). RNA-seq was performed in biological duplicates for each strain.

Chromatin immuno-precipitation

A VSV-G tag-encoding sequence was inserted in the *erfA* gene on the chromosomes of IHMA87 and PAO1 by two-step allelic exchange with the pEXG2-erfA-VSVG and pEXG2-PA0225-VSVG plasmids, respectively. The correct production of C-terminal VSV-G-tagged ErfA proteins in the corresponding IHMA87 ErfA-VSVG and PAO1 ErfA-VSVG strains was assessed by immunoblotting, and the correct activity of the tagged protein was verified by measuring the *exlA::lacZ* transcriptional fusion activity of a strain producing ErfA-VSV-G (Supplementary Figure S3). Each strain was grown in 30 ml LB medium from overnight cultures diluted to $OD_{600} = 0.1$ at 37°C under agitation. At $OD_{600} = 1$, 600 μl of 37% formaldehyde solution was added and samples were incubated for 10 min at 22°C on a rotating wheel. Crosslinking was stopped by the addition of 3 ml of sterile 1.5 M glycine and further incubation for 10 min at 22°C on a rotating wheel. Cells were collected by centrifugation and washed two times with sterile PBS before resuspension in 600 μl of Lysis Buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, pH 7.4, containing EDTA-free Roche protease inhibitor cocktail). Cell lysis and DNA fragmentation to 100–800 bp were done by sonication in a Qsonica bath sonicator for 8 min at 70% of amplitude. Lysates were cleared by centrifugation. Anti-VSV-G antibodies (Sigma) were bound on magnetic Protein A beads according to manufacturer's instructions (Dynabeads Protein A, Invitrogen) and further washed in Lysis Buffer. Cleared lysates were first incubated with bare beads only for 30 min at 4°C on a rotating wheel to prevent excess of non-specific binding to the beads in the subsequent steps. After removal of the bare beads, the lysates were incubated overnight at 4°C on a rotating wheel with antibodies bound to the beads. Beads were then washed three times using the provided Washing Buffer (Invitrogen) and transferred into a new tube after the last wash. Supernatants were removed and 70 μl of Elution Buffer (50 mM Tris-HCl, 10 mM EDTA, 1% SDS, pH 7.5) added to the beads and incubated for 10 min at 70°C under vigorous agitation (600 rpm). About 1 μl of 100 mg/ml RNase A was added to the supernatants followed by incubation for 30 min at 65°C, after which 5 μl of 20 mg/ml Proteinase K were added for another 1 h incubation at 50°C. The crosslinks were reversed by incubating the samples for 12 h at 65°C. Sample volumes were then adjusted to 100 μl with nuclease-free water before addition of 3 μl of 3 M sodium acetate (pH 5). DNA was purified on Monarch DNA Clean Up columns (NEB). ChIP was done in biological duplicates for each strain with wild-type strains serving as controls. Before constructing the libraries, each sample was tested for proper ErfA expression by western immunoblotting and DNA fragmentation by purification of DNA after lysis and agarose gel electrophoresis.

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

ChIP-seq library construction, sequencing and data analysis

After ChIP, DNA was quantified on a QuBit 3.0 Fluorimeter and 3.5 ng of DNA was used to construct libraries using the NEBNext Ultra II DNA library prep kit (NEB). The quality of DNA libraries was assessed on an Agilent Bioanalyzer. Sequencing was performed at the Biopolymers Facility at Harvard Medical School on an Illumina NextSeq500, generating an average of 18 million single-end 50 bp reads per sample with >90% of reads from each sample uniquely aligning to the IHMA87 or PAO1 genomes using Bowtie2 (19). Peak calling was done using MACS2 (22) for each duplicate against the two corresponding negative control samples. Peaks found in both replicates were selected using the Intersect tool from BEDTools (23) with a minimum overlap of 50% on each compared peaks. The closest gene and distance from it were identified for each peak using the ClosestBed tool from BEDTools (23). MEME-ChIP (24) was used to find motifs, using the 100 bp DNA regions centered on the summits of all detected peaks.

Electrophoretic mobility shift assay

Target DNA regions were amplified by PCR using Cy5-labeled primers and purified on DNA Clean up columns (NEB). The resulting 80-bp DNA probes were incubated at 0.5 nM for 5 min at 37°C in EMSA Buffer (10 mM Tris-HCl, 50 mM KCl, 10 mM $MgCl_2$, 10% glycerol, 0.1 mg/ml BSA, pH 8) containing 25 ng/μl poly(dI-dC). For competition assays, 100 nM unlabeled DNA probes (200-fold excess) were incubated with the labeled probes. ErfA protein was added at the indicated concentrations in a final reaction volume of 20 μl and incubated for an additional 15 min at 37°C. Samples were then loaded on a native 5% Tris-borate (TB) polyacrylamide gel and run at 100 V and 4°C in cold 0.5× TB Buffer. Fluorescence imaging was performed using a Chemidoc MP.

RT-qPCR

After total RNA isolation and DNase treatment, cDNA synthesis was carried out using 2 μg of RNA with the SuperScript IV first-strand synthesis system (Invitrogen) in the presence or absence of reverse transcriptase to assess the absence of genomic DNA. The CFX96 real-time system (Bio-Rad) was used to PCR amplify the cDNA, and the quantification was based on use of SYBR green fluorescent molecules. About 2 μl of cDNA were incubated with 10 μl of Luna Universal qPCR 2X Master Mix (NEB) and forward and reverse specific primers at final concentrations of 250 nM in a total volume of 20 μl. The real-time PCR was done according to manufacturer's instructions. To generate standard curves for each pair of primers, serial dilutions of the cDNA were used. The experiments were performed with three biological replicates for each strain, and the relative expression of mRNAs was analyzed with the CFX Manager software (Bio-Rad) using the Pfaffl method relative to *rpoD* reference *C*q values. Statistical analyses were performed by *T*-test. The sequences of primers are listed in Supplementary Table S6.

DAP-seq

To construct DNA libraries, genomic DNA (gDNA) was extracted from overnight cultures of *P. aeruginosa* IHMA87 and PAO1, *P. chlororaphis* PA23, *P. putida* KT2440 and *P. protegens* CHA0 using the GenElute Bacterial Genomic DNA kit (Sigma). Purified gDNAs were fragmented to 100–500 bp by sonication in a Qsonica bath sonicator for 6 min at 70% of amplitude. DNA end-repair was performed on 5 μg fragmented DNA using the NEBNext End-Repair Module (NEB). The dA-tailing was then performed using the NEBNext dA-tailing Module (NEB). Truncated Y adaptors were annealed by mixing adaptors A and B (Supplementary Table S6) in sterile water to a final concentrations of 30 μM each and incubating at 96°C for 2 min before allowing the samples to cool down to room temperature. Ligation of Y adaptors to the dA-tailed libraries was performed using T4 DNA Ligase (NEB) following manufacturer's instructions. DNA was purified on Monarch Clean Up columns between each of these steps. The quality of the libraries was checked on High Sensitivity DNA chips on an Agilent Bioanalyzer and by PCR tests with primers specific to adaptors A and B before and after adaptors ligation.

DAP-seq experiments were conducted in triplicates, and negative controls were performed by not adding any protein to the beads. About 20 μl of Dynabeads His-Tag Isolation and Pulldown magnetic beads (Invitrogen) were washed three times in 500 μl of Binding Buffer (sterile PBS containing 0.01% Tween20). About 500 ng of His-tagged ErfA protein was diluted in 500 μl of Binding Buffer and incubated for 20 min with the washed beads on a rotating wheel at room temperature. After binding, the bead–protein complexes were washed six times in 500 μl of Binding Buffer and then resuspended in 80 μl of Binding Buffer containing 50 ng of adaptor-ligated gDNA libraries and further incubated on a rotating wheel at room temperature for 1 h. The bead–protein–DNA complexes were washed six times in 500 μl of Binding Buffer to eliminate unbound DNA before transferring into a new tube following the final wash. The beads were then resuspended in 25 μl of sterile 10 mM Tris-HCl pH 8.5, and incubated for 10 min at 98°C for elution. After incubation, the samples were placed on ice for 5 min, beads were magnetically removed and the released DNA was used for PCR amplification as previously described (25), using a different indexed pair of primers in each sample to allow pooling for sequencing. PCR products were purified using SPRIselect beads at a 1:1 ratio. Quality of each library was assessed using High Sensitivity DNA chips on an Agilent Bioanalyzer.

DAP-seq sequencing and data analysis

Sequencing was performed by the high-throughput sequencing core facility of I2BC (Centre de Recherche de Gif – http://www.i2bc.paris-saclay.fr) using an Illumina NextSeq500 instrument. Approximately 5 million single-end 75 bp reads per sample were generated on average with >90% of reads uniquely aligning to the IHMA87, PAO1, PA23, CHA0 or KT2440 genomes using Bowtie2 (19). Peak calling was done using MACS2 (22) for each triplicate against the three corresponding negative control sam-

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

ples. Peaks found in the three replicates were selected using the Intersect tool from BEDTools (23) with a minimum overlap of 50% on each compared peaks. The closest gene and distance from it were identified for each peak using the ClosestBed tool from BEDTools (23).

exlBA promoter analysis

The ExlA sequence from *P. aeruginosa* PA7 was searched against the 4846 *Pseudomonas* genomes from the *Pseudomonas* database version 18.1 (http://pseudomonas.com, (26)) using a best bi-directional blast hits approach. A total of 446 *Pseudomonas* strains were found to possess an *exlA* homolog in their genomes, with BLASTp coverage >95% and sequence identity >30%. In these strains, the presence of an *exlB*-5′ adjacent gene was confirmed and the 200 bp sequence upstream of *exlB* start codon was retrieved using Python scripts. A multiple alignment was done between the 446 obtained sequences using Clustal Omega (27) with 5 guide-tree and HMM iterations. The corresponding phylogenetic tree was visualized and annotated with iTOL (28). The presence of *erfA*, *ergAB* and *vfr* homologs in the corresponding strains was assessed using *P. aeruginosa* PA7 protein sequences as a query, with coverage >95% and sequence identity >30%. The presence of putative transcription factor binding sites in the 200-bp sequences upstream from the *exlB* start codon was determined using RSAT matrix-scan (29) with default parameters using either Vfr core consensus binding motif (5′-TGNGANNNAGNTCACAT-3′) or ErfA binding motif from ChIP-seq results. Core promoter predictions were performed using BPROM (30). Conservation rates of P*exlBA*, *exlB* and *exlA* nucleotide sequences were obtained using pairwise alignments on all 446 corresponding sequences with Clustal Omega.

See Supplementary Data for Materials and Methods section on genetic manipulations, β-galactosidase activity assays, proteins purifications and western blot analysis.

RESULTS

exlBA promoter sequence shows differences in several *Pseudomonas* species

A number of *Pseudomonas* species carry the *exlBA* operon; this provides an opportunity to examine the evolution of regulatory pathways following acquisition of genes by HGT. Toward this goal, we analyzed the 5′ regulatory sequences at the *exlBA* promoters in four representative species carrying *exlBA*. A search of these sequences using Regulatory Sequence Analysis Tools (RSAT) ((31); http://embnet.ccg.unam.mx/rsa-tools/) identified the Vfr-binding site only in the *exlBA*-carrying lineage of *P. aeruginosa* and not in other *Pseudomonas* (Figure 1), although orthologues encoding Vfr (CRP) are found in all species. Binding sites for various TFs have been identified in all sequences; however, none is shared by any two species. Therefore it appears that, in these four species, the *exlBA* operon might be regulated differently, reflecting a potential evolution of different regulatory sequences in each of them.

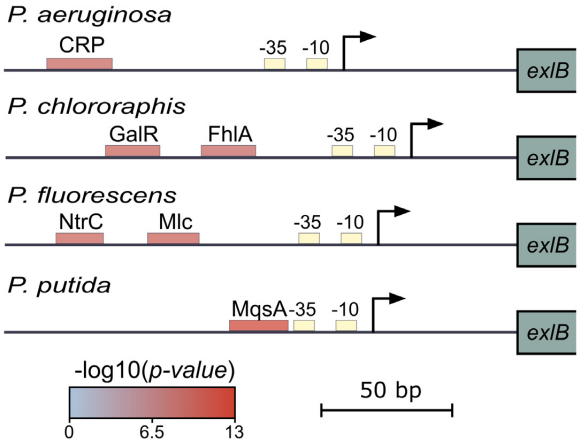


Figure 1. *exlBA* promoters display diverse predicted *cis*-regulatory elements across *Pseudomonas* species. Schematic representation of predicted *cis*-regulatory elements in the *exlBA* promoter regions of *P. aeruginosa* PA7, *P. chlororaphis* PA23, *P. fluorescens* Pt14 and *P. putida* KT2440. Core promoter sequences indicated as '-10' and '-35' boxes were identified by BPROM (30). Predicted binding sites for TFs (CRP, GalR, FhlA, NtrC, Mlc and MqsA) are indicated as boxes colored according to the *P*-value for each predicted site using RSAT matrix-scan (29).

Screen for regulators of *exlBA* identifies a novel repressor

To further characterize *exlBA* regulation, we built and screened a comprehensive transposon mutant library in *P. aeruginosa* IHMA87 carrying a *lacZ* gene incorporated into the *exlA* gene, providing a transcriptional readout for mutants with altered expression of the reporter (see 'Materials and Methods' section for details). The location of transposon insertions in 13 candidate genes that resulted in an increased expression of the *lacZ* reporter was determined (Supplementary Table S1). The insertion with the highest effect on expression was in the gene *IHMA87_00215* (*PSPA7_0311*), encoding an uncharacterized transcriptional regulator containing a Cro/CI-type DNA binding domain, which we named ErfA, for Exolysin Regulatory Factor A. We deleted this gene from the chromosome of *P. aeruginosa* IHMA87 and confirmed that the loss of *erfA* gene product leads to a significant increase in the transcription of the *exlBA* operon and that the various phenotypic changes expected from the rise in ExlA production are correspondingly enhanced. In the Δ*erfA* strain, we observed an approximately 40-fold increase in β-galactosidase activity of the *exlA::lacZ* transcriptional fusion, which was restored to the wild-type levels by trans-complementation (Figure 2A). The increase in transcription of *exlA* in the *erfA* mutant was accompanied by an increase of both ExlA synthesis and secretion (Figure 2B and C), and consequently of bacterial cytotoxicity during infection of epithelial cells (Figure 2D and E). To assess the impact of de-repression of Exolysin expression *in vivo*, we compared the infectivity of strains in a wax moth *Galleria mellonella* model of infection. Here again, the *erfA* mutant displayed faster killing kinetics of the larvae than the wild-type and complemented strains (Figure 2F) indicating that ErfA negatively regulates ExlA-dependent virulence of

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

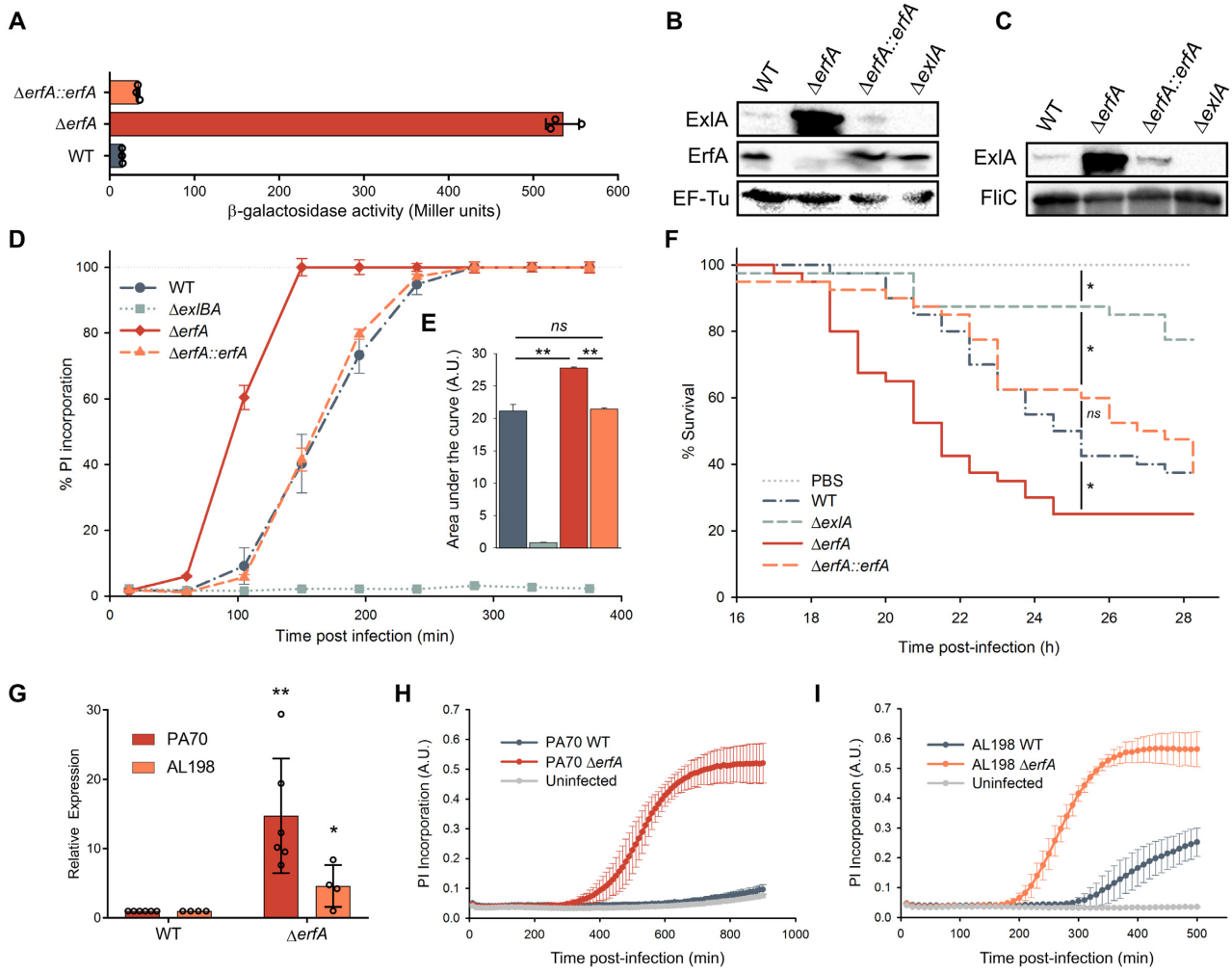


Figure 2. ErfA regulates ExlA-dependent virulence through inhibition of *exlBA* expression in *P. aeruginosa*. (**A**) β-Galactosidase activities of IHMA87 strains harboring *exlBA::lacZ*. The strains were grown in LB, and activities of the transcriptional fusion were measured at $OD_{600} = 1$. The enzyme activities are represented as mean ± standard deviation (SD) from three independent experiments. (**B** and **C**) Immunodetection of ExlA in bacterial cell extracts (B) and supernatants (C). Whole bacteria and culture supernatants were sampled and analyzed by SDS-PAGE for ExlA, ErfA, EF-Tu and FliC content using appropriate antibodies. (**D**) Kinetics of bacterial cytotoxicity on epithelial cells. A549 epithelial cells grown in 96-well plates were infected with indicated strains at a multiplicity of infection (MOI) of 10 in the presence of propidium iodide (PI). PI incorporation, which reflects membrane permeabilization, was monitored every 45 min by automated live-cell microscopy and normalized to the total number of cells measured at the start of the experiment by Syto24 staining. Data are represented as mean ± s.e.m. from three independent experiments. (**E**) The areas under each curve presented in (D) were calculated from the values obtained after 380 min of infection. ns: not significant. (**F**) Survival curves of *Galleria mellonella* larvae infected with an average of 5 bacteria per larva. Forty larvae were infected per strain. Significance testing was performed using log-rank test ($P < 0.05$). (**G**) RT-qPCR analysis of *exlA* expression in PA70 and AL-198 strains and *erfA* isogenic mutants. All strains were grown in LB medium to $OD_{600} = 1$. Expressions were normalized to the abundance of *rpoD* mRNA. Error bars indicate the SD. The *P*-value was determined using one-tailed *T*-test and is indicated by * ($P < 0.05$) or ** ($P < 0.01$). (**H** and **I**) Kinetics of PI incorporation (A.U.: Arbitrary Unit) into A549 epithelial cells infected at a MOI of 10 with PA70 (H) and AL-198 (I) strains and isogenic *erfA* mutants.

IHMA87 *in vivo*. Finally, to test whether ErfA plays the same inhibitory role in the second group of *P. aeruginosa* strains phylogenetically closer to the *T3SS+* groups (5,8), we deleted the *erfA* gene in PA70, a strain isolated from expectoration of a non-CF bronchiectasis patient, and in AL-198, a CF isolate. Both mutants exhibited higher level of *exlB* and *exlA* expression, and even the completely innocuous PA70 strain became cytotoxic when ErfA was absent (Figure 2G–I), showing the conserved function of ErfA on *exlBA* repression across two subgroups of *P. aeruginosa* strains.

The binding repertoire of ErfA in *P. aeruginosa*

An ortholog of *erfA* can be identified in the genomes of *P. aeruginosa* strains lacking *exlBA*, including the reference strain PAO1 (PA0225; www.pseudomonas.com (26)). We therefore assumed that ErfA might have other regulatory targets in these strains. In order to apply genome-wide approaches toward the identification of genes under ErfA control, we sequenced and assembled the genome of IHMA87, which represents the third closed genome of the PA7-like clade after PA7 and CR1 (6,32). In addition to a 6.53 Mb chromosome, IHMA87 harbors a 185 kb mega

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

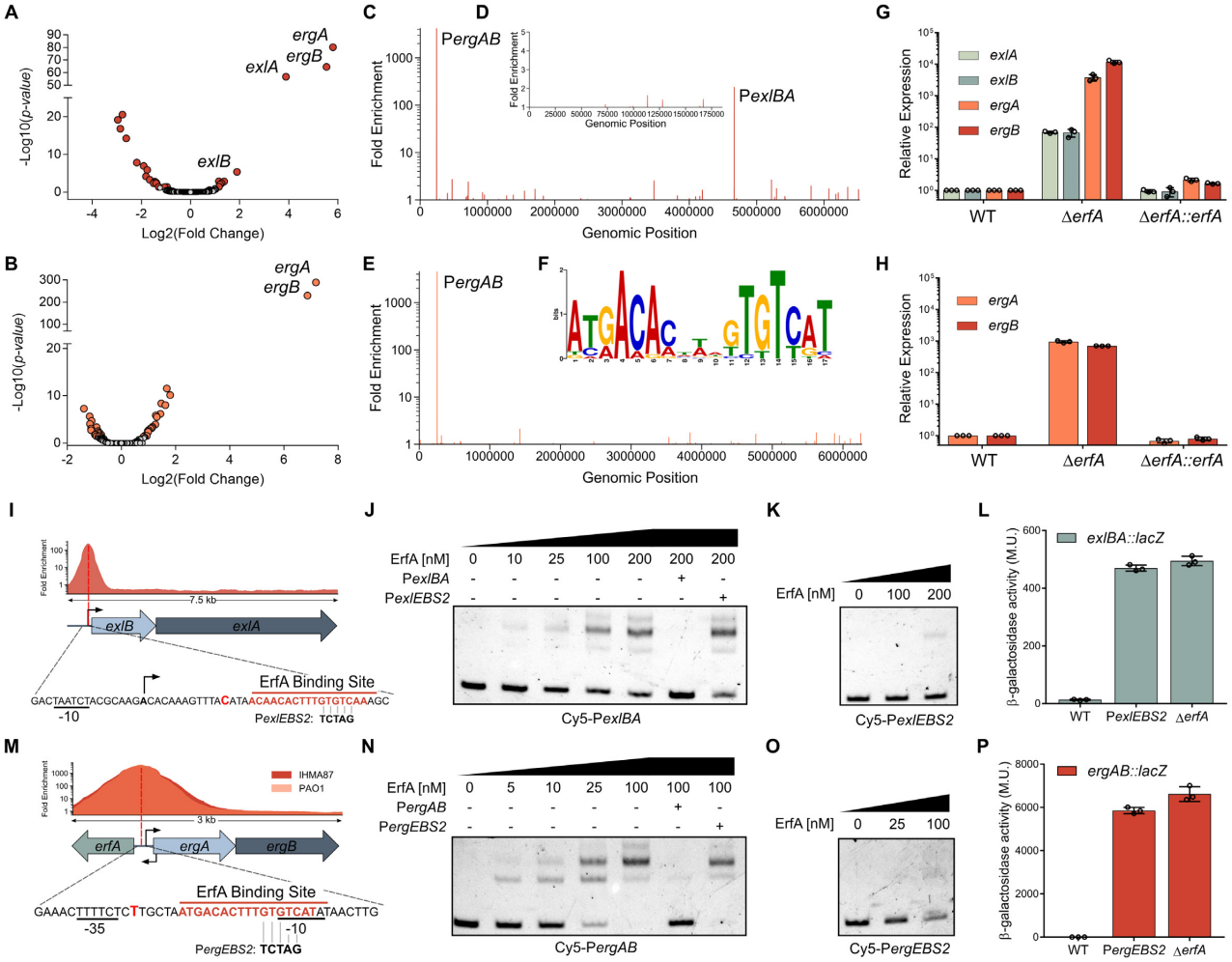


Figure 3. ErfA directly regulates a second operon in addition to *exlBA*. (**A** and **B**) Volcano plots displaying the RNA-seq results of the genes differentially expressed in respective Δ*erfA* mutants versus IHMA87 (A) and PAO1 (B) wild-type strains. Genes with q-value < 0.05 are depicted in red (IHMA87) or orange (PAO1). (**C**–**E**) Enrichment of normalized mapped reads after ChIP-seq on the whole IHMA87 chromosome (C) and IHMA87 plasmid (D), and the PAO1 genome (E). (**F**) Enriched DNA motif obtained with MEME-ChIP on relevant ChIP-seq peaks corresponding to near-summit regions. (**G** and **H**) RT-qPCR analysis of *ergA*, *ergB*, *exlA* and *exlB* mRNA levels in wild-type, *erfA* mutant and complemented strains in IHMA87 (G) and PAO1 (H). Experiments were performed in triplicates, with RNA extracted from bacteria at OD_{600} = 1 in LB and normalized to the *rpoD* transcript. Error bars indicate the SD. (**I**) Enrichment of normalized mapped reads after ChIP-seq and location of ErfA binding site on the 7.5 kb region encompassing *exlBA*. Black arrows indicate transcription start sites. The position of the summit of the ChIP-seq peak is denoted as a bold bright red letter. Bases changed in the P*exlEBS2* mutation of the binding site are shown. (**J** and **K**) Electrophoretic mobility shift assay of ErfA on *exlBA* promoter (P*exlBA*). Recombinant ErfA-His10 protein (0–200 nM) was incubated with 0.5 nM Cy5-labeled P*exlBA* 80-mer probe (J) or the mutated Cy5-P*exlEBS2* probe (K) for 15 min before electrophoresis. For competition assays, excess of unlabeled P*exlBA* or P*exlEBS2* probes (100 nM) are denoted '+' for the corresponding probe. (**L**) β-Galactosidase activities of the wild-type (WT) and Δ*erfA* strains harbouring *exlBA::lacZ* transcriptional fusion. The strain IHMA87*exlBA::lacZ* P*exlBA-EBS2* carries the P*exlEBS2* mutation indicated in I on ErfA binding site. Experiments were performed in triplicates, on bacteria in LB at OD_{600} = 1. Error bars indicate the SD. (**M**–**P**) These panels display the same experiments described in I,J,K,L, but focused on P*ergAB*.

plasmid (Supplementary Figure S1). With an average nucleotide identity of 98.95% between them, the genomes of PA7 and IHMA87 are also generally syntenic (Supplementary Figure S2).

To examine the global role of ErfA in IHMA87 (*exlBA*+) and PAO1 (*T3SS*+), we combined RNA-seq and ChIP-seq approaches. A comparison of the transcriptomes of wild-type and *erfA* mutants identified 8 and 2 genes that were differentially expressed (log2(fold change)>2) in IHMA87 and PAO1, respectively (Figure 3A and B; Supplementary Table S2). ChIP-seq performed with a functional VSV-G-tagged ErfA (Supplementary Figure S3) led to the identification of 2 and 1 regions exhibiting significant enrichment in IHMA87 and PAO1, respectively (Figure 3C–E and Supplementary Table S2). The analysis of all detected peaks allowed us to identify a conserved 17-bp palindromic motif (5′-ATGACACntnGTGTCAT-3′) as a likely ErfA DNA-binding site (EBS) (Figure 3F). In IHMA87, the second highest peak observed by ChIP-seq was centered on the *exlBA* promoter, showing that ErfA directly exerts its negative control on the operon. In addition to *exlBA*, one additional region was enriched in ChIP-seq in

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

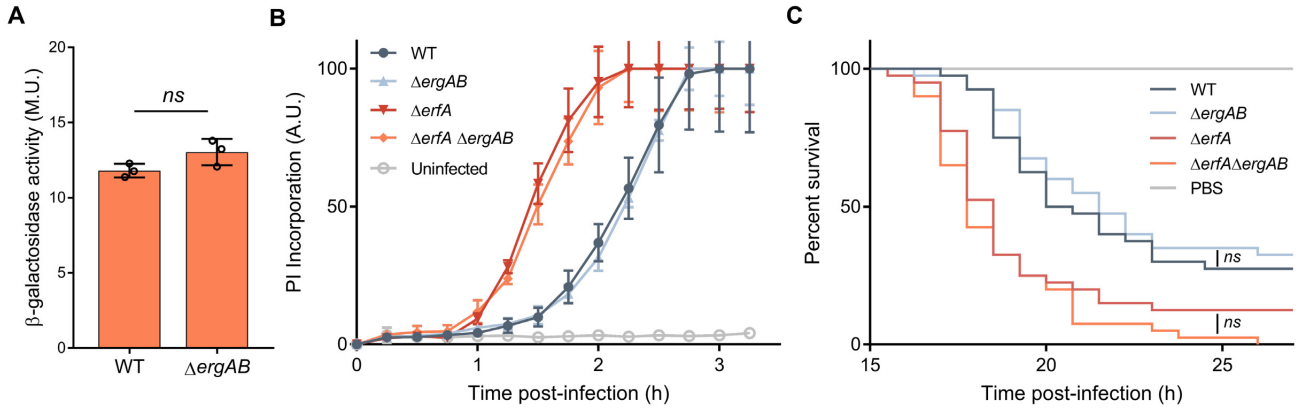


Figure 4. ErgA and ErgB are functionally unrelated to ExlBA. (**A**) β-Galactosidase activities of IHMA87 wild-type (WT) and *ergAB* mutant (Δ*ergAB*) carrying *exlBA::lacZ*. Activities were measured after growth in LB at $OD_{600} = 1$ and are represented as mean (in Miller Units) ± SD from three independent experiments. Statistical significance was assessed using two-tailed *T*-test. (**B**) Kinetics of bacterial cytotoxicity on A549 epithelial cells. Cells were infected at a MOI of 10 with the indicated strains in presence of propidium iodide (PI). PI incorporation is represented as mean ± s.e.m. from three independent experiments. (**C**) Survival curves of *Galleria mellonella* larvae infected with an average of 5 bacteria per larva. Forty larvae were used for each strain. Statistical testing was performed using log-rank test ($P < 0.05$).

both strains, corresponding to the intergenic region between *erfA* and a two-gene operon transcribed in the opposite direction. These two genes (*PA0224/IHMA87_00214* and *PA0223/IHMA87_00213*) were also the most upregulated genes in the *erfA* mutants in RNA-seq, as confirmed by RTqPCR (Figure 3G,H), and we propose to name them *ergA* and *ergB* for ErfA regulated gene A and B. Mutagenesis experiments showed that ErfA binding on this EBS only slightly diminished *erfA* expression (Supplementary Figure S4), excluding any auto-regulatory mechanism, as suggested by the location of ErfA binding. Overall, *exlBA* and *ergAB* were the only two ErfA targets found in both RNA-seq and ChIP-seq experiments. In addition, genes encoding proteins of the T3SS, including PopB and PopD translocators and exotoxin ExoT, were found slightly downregulated in PAO1 *erfA* mutant. The downregulation was confirmed for *popB* by RTqPCR and analysis of a P*popN-lacZ* transcriptional fusion, although it has no impact on cytotoxicity toward epithelial cells (Supplementary Figure S5). This slight regulatory effect seems to be indirect and would need further investigation. Therefore, we conclude that ErfA directly regulates two operons: *exlBA* and *ergAB*.

The EBS identified by ChIP-seq was centered at +25 nucleotides downstream of the transcription start site of *exlBA* (12) (Figure 3I). The direct interaction between ErfA and *exlBA* promoter (P*exlBA*) as well as specificity to the EBS were confirmed *in vitro* by EMSA, using either wild-type or mutated probes (Figure 3J,K). The same mutation that abolished ErfA binding *in vitro*, when introduced in the promoter of the IHMA87 *exlBA::lacZ* strain, led to increased β-galactosidase activity similar to that measured in the *erfA* mutant (Figure 3L). This confirmed that the inhibitory effect of ErfA on *exlBA* expression is the consequence of its direct and specific binding to a palindromic DNA motif found on *exlBA* promoter. In addition, ErfA binding to this regulatory site was found to counteract the positive effect of the cAMP-responsive Vfr activator on *exlBA* transcription (Supplementary Figure S6). The configuration of the Vfr- and ErfA-binding sites (VBS and EBS) would allow independent bindings, and thus independent signal integration of both the activator and the repressor, which is reminiscent of what is seen with CRP and LacI regulation of the *lac* operon (33). The EBS situated upstream of *ergAB* overlaps the '-10' box of the putative promoter and is conserved in IHMA87 and PAO1 strains (Figure 3M). EMSA revealed a much higher affinity of ErfA for the *ergAB* promoter than for the *exlBA* promoter (Figure 3N), fitting well with the higher effect seen on transcription, and the mutation of the site also prevented the binding of the protein both *in vitro* and *in vivo* (Figure 3O,P), confirming the direct control of these genes by ErfA.

ErfA regulates two functionally unrelated operons

The product of the *ergA* gene is predicted to be a class II aldolase while *ergB* encodes a putative dihydrodipicolinate synthase. However, two independent studies tested ErgB activity as a dihydrodipicolinate synthase without any conclusive results (34,35), leaving both gene products with unknown functions. We investigated whether *ergA* and *ergB* play a role in regulation or function of ExlBA by analyzing the transcription of *exlBA* (Figure 4A), the cytotoxicity on epithelial cells and the virulence in the *Galleria* model of infection (Figure 4B,C) of *ergAB* mutant. Collectively, these experiments indicated that ErgA and ErgB have no effect on ExlBA function, at least in the conditions tested herein. Therefore, we conclude that ErfA directly regulates two operons with unrelated functions.

ErfA targets uniquely *ergAB* in other *Pseudomonas* species

Orthologues of the *exlBA* operon have been identified in several soil and plant dwelling *Pseudomonas* such as *P. fluorescens*, *P. protegens*, *P. putida* and *P. chlororaphis* ((36,37), see later this work). In order to investigate the role of ErfA in these species, we determined the ErfA pan-regulon using

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

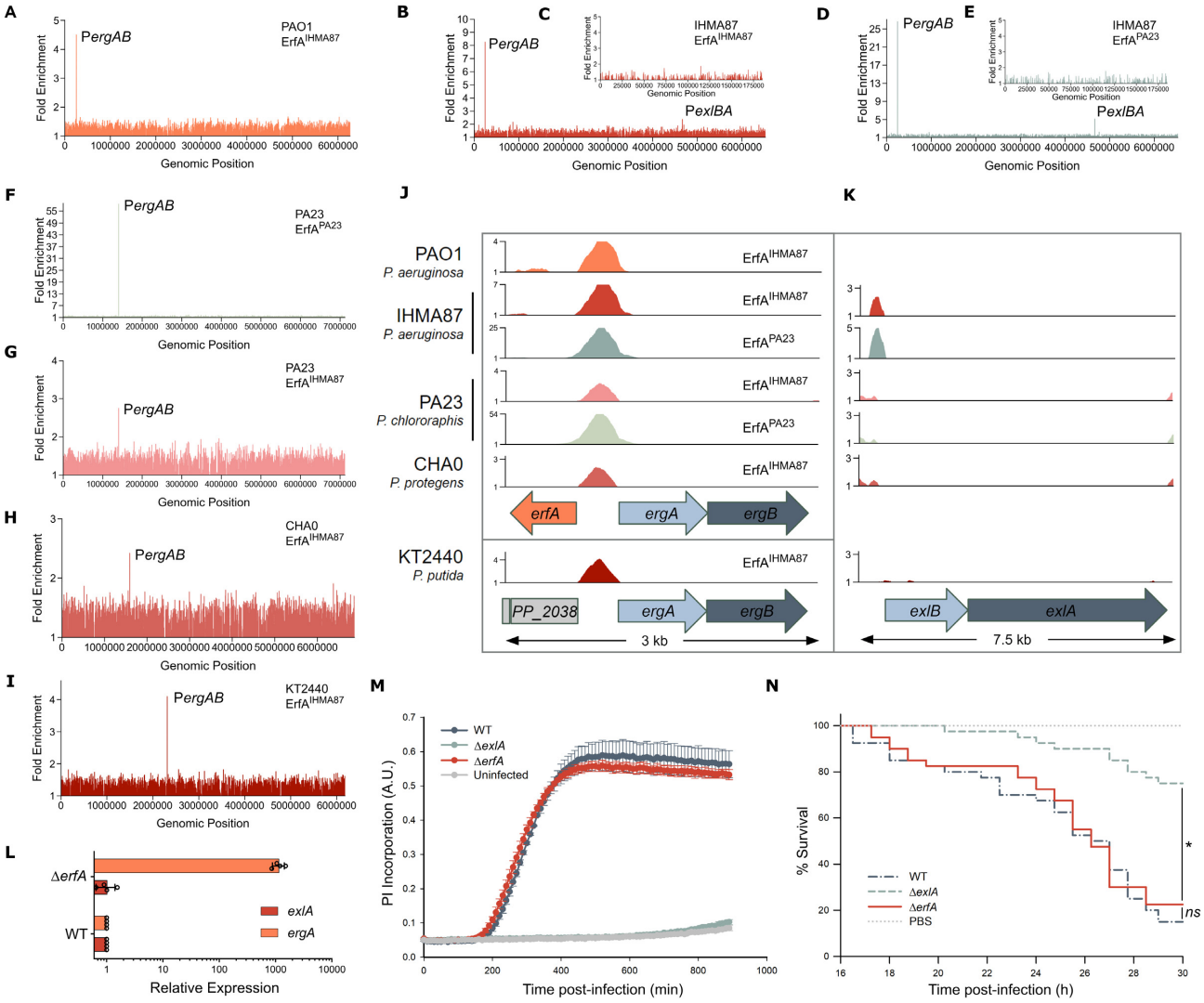


Figure 5. *exlBA* is regulated by ErfA only in *P. aeruginosa*. (**A**–**C**) Enrichment of normalized mapped reads after DAP-seq with purified ErfA from IHMA87 (ErfA^{IHMA87}) on genomes of PAO1 (A), IHMA87 chromosome (B) and plasmid (C). (**D**–**F**) Enrichment of normalized mapped reads after DAP-seq with purified ErfA from *P. chlororaphis* PA23 (ErfA^{PA23}) on whole IHMA87 chromosome (D), and plasmid (E), and PA23 genome (F). (**G**–**I**) Enrichment of normalized mapped reads after DAP-seq performed with ErfA^{IHMA87} on genomic DNA of *P. chlororaphis* PA23 (G), *P. protegens* CHA0 (H) and *P. putida* KT2440 (I). (**J** and **K**) Alignment of enrichments of normalized mapped reads for all DAP-seq experiments on the 3 and 7.5 kb region encompassing *erfA-ergAB* (J) and *exlBA* (K), respectively. Colors correspond to the ones from the genome-wide read density maps with either ErfA^{IHMA87} or ErfA^{PA23}. (**L**) RT-qPCR analysis of *ergA* and *exlA* mRNA levels in wild-type and isogenic *erfA* mutant in *P. chlororaphis* PA23. Experiments were performed in triplicates, with RNA extracted from bacteria at $OD_{600} = 1$ in LB and normalized to *rpoD* mRNA. Error bars indicate the SD. (**M**) Kinetics of PI incorporation in A549 epithelial cells infected at a MOI of 10 with *P. chlororaphis* PA23 strains at 30°C. (**N**) Survival curves of *Galleria mellonella* larvae infected with *P. chlororaphis* PA23 strains. Larvae were infected with an average of ~6.10^4 bacteria and incubated at 30°C. Significance testing was performed using log-rank test ($P < 0.05$).

an *in vitro* genome-wide binding assay called DAP-seq (25), utilizing a purified recombinant ErfA^{IHMA87} protein from *P. aeruginosa* IHMA87. We first confirmed that DAP-seq yielded the same profiles of DNA binding as ChIP-seq in both PAO1 (Figure 5A and Supplementary Table S3) and IHMA87 (Figure 5B and C; Supplementary Table S3). Indeed, only *exlBA* and *ergAB* promoters were found significantly enriched. Notably, the difference in fold enrichment between *ergAB* and *exlBA* correlated well with *in vitro* affinity for the two sites observed by EMSA and the *in vivo* impact on levels of transcripts measured by RTqPCR, making DAP-seq a convenient and reliable tool to assess ErfA binding. As ErfA proteins from the selected *Pseudomonas* species share between 72 and 76% amino acid identity, we also purified the recombinant ErfA^{PA23} from *P. chlororaphis* PA23, which is 75% identical to ErfA^{IHMA87}, and carried out DAP-seq with both proteins on *P. aeruginosa* IHMA87 and *P. chlororaphis* PA23 chromosomes to further validate our approach. Both proteins revealed the same binding patterns in the genomes of both of these organisms (Figure 5D–G and Supplementary Table S3), i.e. two promoter regions (*exlBA* and *ergAB*) in the *P. aeruginosa* genome and exclu-

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

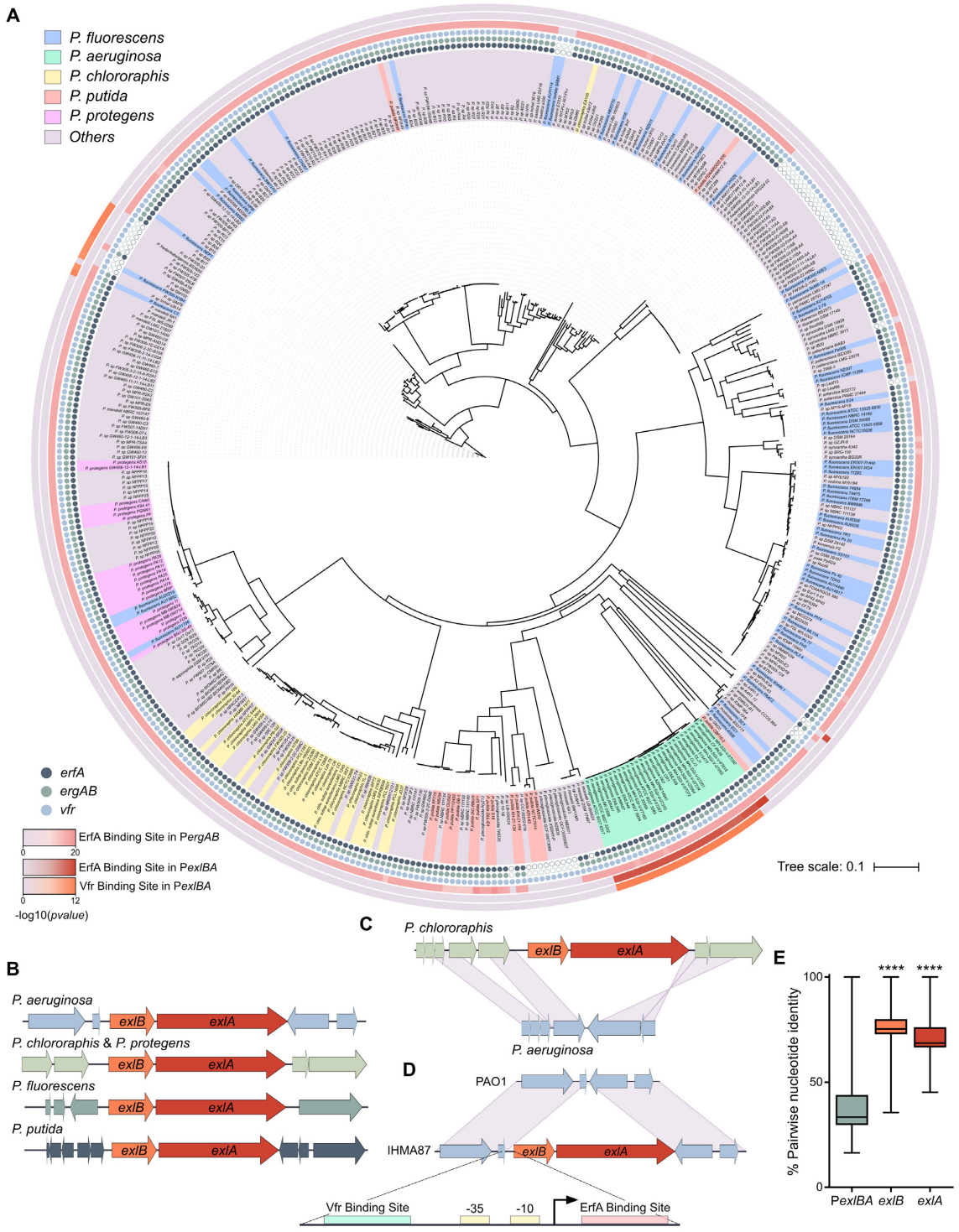


Figure 6. *exlBA* promoters evolved divergent sequences across *Pseudomonas* species. (**A**) Phylogenetic tree of P*exlBA* sequences in 446 *Pseudomonas* harboring *exlBA*-like genes. The background color for the strain name is given for the top five species. Presence of the genes *erfA*, *ergAB* and *vfr* are denoted by solid circles at each species name, empty circles indicate the absence of the gene. The *P*-values, computed with RSAT matrix-scan (29), supporting the presence of ErfA- and Vfr-binding sites on P*ergAB* and P*exlBA*, are displayed as heat maps on the external rings of the tree, the light gray color indicating the absence of tested binding sites. (**B**) Synteny of the *exlBA* genetic environment in *P. aeruginosa* IHMA87, *P. chlororaphis* PA23, *P. protegens* CHA0, *P. putida* KT2440 and *P. fluorescens* Pt14. (**C**) Schematic comparison of the *P. chlororaphis exlBA* location to that in *P. aeruginosa*. (**D**) Schematic comparison of *exlBA* locus between *T3SS*$^{+}$ (PAO1) and *exlBA*$^{+}$ (IHMA87) with a zoom on the 296 bp intergenic region encompassing Vfr-binding site, boxes '-10' and '-35', transcription start site (as determined in (12)), and ErfA-binding site. (**E**) Conservation rates of P*exlBA*, *exlB* and *exlA* nucleotide sequences in 446 *Pseudomonas* strains. $n = 198\,470$ pairwise alignments for each sequence. Statistical analysis was performed using one-way ANOVA for *exlB* and *exlA* sequences against P*exlBA*. ****: $P < 0.0001$.

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

sively the *ergAB* promoter in *P. chlororaphis* PA23, confirming that differences in their protein sequences did not impact binding specificities. Furthermore, those data showed that the ErfA regulon is restricted to *ergAB* in *P. chlororaphis* PA23. ErfA^{IHMA87} was then used to determine EBSs on the genomes of *P. chlororaphis* PA23, *P. protegens* CHA0 and *P. putida* KT2440 (Figure 5G–I and Supplementary Table S3). This analysis further revealed the promoter of *ergAB* as the only target in the three species. Interestingly, while the *ergAB* operon is adjacent to *erfA* in *P. aeruginosa*, *P. chlororaphis* and *P. protegens*, it is 1.2 Mb apart in the chromosome of *P. putida*. This conserved regulation of ErfA stresses the indelible evolutionary conserved link between *erfA* and *ergAB*, and not *exlBA* (Figure 5J and K). In agreement with the DAP-seq data, we confirmed *in vivo* that, in *P. chlororaphis*, ErfA controls the expression of *ergAB* but not of *exlBA* (Figure 5L). Accordingly, the deletion of *erfA* affect neither cytotoxicity of *P. chlororaphis* on epithelial cells (Figure 5M) nor its virulence during *G. mellonella* larvae infection (Figure 5N). Altogether, ErfA represses ExlBA-dependent virulence specifically in *P. aeruginosa*.

exlBA promoter shows a wide diversity of regulatory sequences across *Pseudomonas* species

Interrogation of public databases showed that the *exlBA* operon is present in nearly 10% of all sequenced *Pseudomonas* genomes (Supplementary Table S4). In light of the differences observed in *exlBA* regulation between the experimentally tested species (*P. aeruginosa* versus *P. chlororaphis*, *P. putida* and *P. protegens*), we performed a phylogenetic analysis and compared *exlBA* promoter sequences (P*exlBA*) in all 446 *exlBA*$^+$ strains (Figure 6A). First, the analysis of their genomes revealed a strong co-occurrence of the three genes *erfA*, *ergA* and *ergB*, corroborating the functional relationship between them. In addition, all *erfA*$^+$ bacteria carry a conserved EBS at the *ergAB* promoter. On the contrary, EBS on *exlBA* promoter was found in only 6.9% of the scanned genomes, all corresponding to *P. aeruginosa* strains. Consequently, *ergAB* is the unique member of ErfA 'core' regulon, and *exlBA* is part of its 'accessory' regulon that would be *P. aeruginosa* specific. A similar imbalance was found with Vfr that is conserved in nearly all strains, as only *P. aeruginosa* strains, with a few exceptions, harbor a conserved VBS in the P*exlBA* region, further supporting the idea of an evolutionary acquired adapted regulatory network. Indeed, the Vfr protein is conserved in different lineages of *P. aeruginosa*; it was shown to be important in an acute murine lung infection model (38) and regulates a number of virulence-related factors including the T3SS (39,40), elastase, and exotoxin A (41). Additionally, ErfA might respond to other signals through its C-terminal cupin domain, probably linked to ErgA and ErgB functions, which are beneficial to *exlBA* expression specifically in *P. aeruginosa*. It is likely that environmental species respond to signals that are different from those encountered by a human pathogen such as *P. aeruginosa*, and that they control *exlBA* expression using unique and specific molecular mechanisms that are best suited for their particular niches. Also, the synteny of the *exlBA* locus is not conserved in different species, indicating that the genetic environment is variable and might influence the operon expression (Figure 6B). For instance, in *P. chlororaphis*, a plant-dwelling bacterium proposed as biocontrol agent, *exlBA* is found between genes encoding a two-component regulatory system and an operon involved in secondary messenger regulation. In *P. aeruginosa*, these two operons (*PA3040* to *PA3045*) are contiguous (Figure 6C) while *exlBA* is found in another unrelated location, along with a *P. aeruginosa*-specific sequence upstream of *exlBA* containing the promoter encompassing both ErfA- and Vfr-binding sites (Figure 6D). Furthermore, we found that P*exlBA* is much less conserved than *exlB* and *exlA* coding sequences, demonstrating the strain-specific evolution of promoter sequence after acquisition of *exlBA* (Figure 6E). It is established that DNA-binding regulatory proteins can readily acquire new targets, and we propose that the emergence of binding sites for Vfr and ErfA in the *exlBA* promoter region of *P. aeruginosa* wired the operon to the activator and the repressor regulons, balancing the operon transcription in response to signals specific to *P. aeruginosa* lifestyle.

DISCUSSION

We described a novel twist to HGT, where the expression of newly acquired genes is re-programed by adapting to the control of existing signaling pathways and TFs through concomitant evolution or capture of the appropriate regulatory sequences. This phenomenon arises from the fact that different organisms might need to acquire the same function, but do not need it in the same conditions. This is well exemplified by the difference in conservation rates between *exlBA* coding sequence and *exlBA* promoter. Indeed, while ExlBA proteins are well conserved, suggesting no major change in function between the different species, a much wider diversity of promoter sequences is found, including completely unrelated ones. In that regard, once a new trait is acquired, the evolutionary pressure might become stronger on its regulatory sequences so that it quickly falls under adequate expression control. The regulatory sequences found immediately upstream of the newly acquired structural genes may be modified by a series of mutations, leading to appearance of new binding sites for the corresponding TFs present in the recipient organism. This hypothesis is supported by the observed vast heterogeneity of *exlBA* promoter sequences suggesting a diversity of molecular mechanisms exerting the regulation of similar operons across *Pseudomonas* found in different environments. A broader examination of the conservation, or lack thereof, in regulatory sequences linked to genes acquired by HGT could reveal how common this adaptive mechanism shapes the outcome of molecular evolution giving an organism the ability to survive and thrive in a particular environment.

DATA AVAILABILITY

The genome of IHMA879472 is available at NCBI under the accession numbers CP041354 and CP041355 for the chromosome and plasmid sequences, respectively. The RNA-seq, ChIP-seq and DAP-seq data are available under the GEO accession numbers GSE137485, GSE137484 and GSE137648, respectively.

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Peter Panchev for his help with the *G. mellonella* experiments. We thank Eric Faudry and Stéphanie Bouillot for their help with cell culture and cytotoxicity assays. J.T. thanks François Parcy, Xuelei Lai and Arnaud Stigliani for their technical advices on DAP-seq. *Pseudomonas aeruginosa* strain IHMA879472 was kindly provided by International Health Management Association (IHMA; USA).

FUNDING

Agence Nationale de la Recherche [ANR-15-CE11-0018-01]; Laboratory of Excellence GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS [ANR-17-EURE-0003]; Fondation pour la Recherche Médicale [Team FRM 2017, DEQ20170336705]; French Ministry of Education and Research (to J.T.); CNRS, INSERM, CEA, and Grenoble Alpes University. Funding for open access charge: Fondation pour la Recherche Médicale [DEQ20170336705].
Conflict of interest statement. None declared.

REFERENCES

1. Price,M.N., Dehal,P.S. and Arkin,A.P. (2008) Horizontal gene transfer and the evolution of transcriptional regulation in Escherichia coli. *Genome Biol.*, **9**, R4.
2. Freschi,L., Vincent,A.T., Jeukens,J., Emond-Rheault,J.G., Kukavica-Ibrulj,I., Dupont,M.J., Charette,S.J., Boyle,B. and Levesque,R.C. (2019) The Pseudomonas aeruginosa Pan-Genome provides new insights on its population structure, horizontal gene transfer, and pathogenicity. *Genome Biol. Evol.*, **11**, 109–120.
3. Huber,P., Basso,P., Reboud,E. and Attree,I. (2016) Pseudomonas aeruginosa renews its virulence factors. *Environ. Microbiol. Rep.*, **8**, 564–571.
4. Elsen,S., Huber,P., Bouillot,S., Coute,Y., Fournier,P., Dubois,Y., Timsit,J.F., Maurin,M. and Attree,I. (2014) A type III secretion negative clinical strain of Pseudomonas aeruginosa employs a two-partner secreted exolysin to induce hemorrhagic pneumonia. *Cell Host Microbe*, **15**, 164–176.
5. Reboud,E., Elsen,S., Bouillot,S., Golovkine,G., Basso,P., Jeannot,K., Attree,I. and Huber,P. (2016) Phenotype and toxicity of the recently discovered exlA-positive Pseudomonas aeruginosa strains collected worldwide. *Environ. Microbiol.*, **18**, 3425–3439.
6. Roy,P.H., Tetu,S.G., Larouche,A., Elbourne,L., Tremblay,S., Ren,Q., Dodson,R., Harkins,D., Shay,R., Watkins,K. *et al.* (2010) Complete genome sequence of the multiresistant taxonomic outlier Pseudomonas aeruginosa PA7. *PLoS One*, **5**, e8842.
7. Freschi,L., Bertelli,C., Jeukens,J., Moore,M.P., Kukavica-Ibrulj,I., Emond-Rheault,J.G., Hamel,J., Fothergill,J.L., Tucker,N.P., McClean,S. *et al.* (2018) Genomic characterisation of an international Pseudomonas aeruginosa reference panel indicates that the two major groups draw upon distinct mobile gene pools. *FEMS Microbiol. Lett.*, **365**, doi:10.1093/femsle/fny120.
8. Ozer,E.A., Nnah,E., Didelot,X., Whitaker,R.J. and Hauser,A.R. (2019) The population structure of pseudomonas aeruginosa is characterized by genetic isolation of exoU+ and exoS+ Lineages. *Genome Biol. Evol.*, **11**, 1780–1796.
9. Boukerb,A.M., Marti,R. and Cournoyer,B. (2015) Genome Sequences of Three Strains of the Pseudomonas aeruginosa PA7 Clade. *Genome Announc.*, **3**, e01366-15.
10. Dingemans,J., Ye,L., Hildebrand,F., Tontodonati,F., Craggs,M., Bilocq,F., De Vos,D., Crabbe,A., Van Houdt,R., Malfroot,A. *et al.* (2014) The deletion of TonB-dependent receptor genes is part of the genome reduction process that occurs during adaptation of Pseudomonas aeruginosa to the cystic fibrosis lung. *Pathog. Dis.*, **71**, 26–38.
11. Kos,V.N., Deraspe,M., McLaughlin,R.E., Whiteaker,J.D., Roy,P.H., Alm,R.A., Corbeil,J. and Gardner,H. (2015) The resistome of Pseudomonas aeruginosa in relationship to phenotypic susceptibility. *Antimicrob. Agents Chemother.*, **59**, 427–436.
12. Berry,A., Han,K., Trouillon,J., Robert-Genthon,M., Ragno,M., Lory,S., Attree,I. and Elsen,S. (2018) cAMP and Vfr control exolysin expression and cytotoxicity of pseudomonas aeruginosa taxonomic outliers. *J. Bacteriol.*, **200**, e00135-18 .
13. Kanack,K.J., Runyen-Janecky,L.J., Ferrell,E.P., Suh,S.J. and West,S.E. (2006) Characterization of DNA-binding specificity and analysis of binding sites of the Pseudomonas aeruginosa global regulator, Vfr, a homologue of the Escherichia coli cAMP receptor protein. *Microbiology*, **152**, 3485–3496.
14. Wick,R.R., Judd,L.M., Gorrie,C.L. and Holt,K.E. (2017) Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.*, **13**, e1005595.
15. Seemann,T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, **30**, 2068–2069.
16. Yoon,S.H., Ha,S.M., Lim,J., Kwon,S. and Chun,J. (2017) A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek*, **110**, 1281–1286.
17. Darling,A.C., Mau,B., Blattner,F.R. and Perna,N.T. (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.*, **14**, 1394–1403.
18. Ngo,T.D., Ple,S., Thomas,A., Barette,C., Fortune,A., Bouzidi,Y., Fauvarque,M.O., Pereira de Freitas,R., Francisco Hilario,F., Attree,I. *et al.* (2019) Chimeric Protein-Protein interface inhibitors allow efficient inhibition of Type III secretion machinery and pseudomonas aeruginosa virulence. *ACS Infect. Dis*, **5** , 1843-1854 .
19. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
20. Anders,S., Pyl,P.T. and Huber,W. (2015) HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.
21. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
22. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
23. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
24. Machanick,P. and Bailey,T.L. (2011) MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics*, **27**, 1696–1697.
25. Bartlett,A., O'Malley,R.C., Huang,S.C., Galli,M., Nery,J.R., Gallavotti,A. and Ecker,J.R. (2017) Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.*, **12**, 1659–1672.
26. Winsor,G.L., Griffiths,E.J., Lo,R., Dhillon,B.K., Shay,J.A. and Brinkman,F.S. (2016) Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. *Nucleic Acids Res.*, **44**, D646–D653.
27. Sievers,F., Wilm,A., Dineen,D., Gibson,T.J., Karplus,K., Li,W., Lopez,R., McWilliam,H., Remmert,M., Soding,J. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.*, **7**, 539.
28. Letunic,I. and Bork,P. (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*, **23**, 127–128.
29. Nguyen,N.T.T., Contreras-Moreira,B., Castro-Mondragon,J.A., Santana-Garcia,W., Ossio,R., Robles-Espinoza,C.D., Bahin,M., Collombet,S., Vincens,P., Thieffry,D. *et al.* (2018) RSAT 2018: regulatory sequence analysis tools 20th anniversary. *Nucleic Acids Res.*, **46**, W209–W214.
30. Solovyev,V. and Salamov,A. (2011) *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies*. In: Li,RW (ed). Nova Science Publishers, NY. pp. 61–78.
31. Medina-Rivera,A., Defrance,M., Sand,O., Herrmann,C., Castro-Mondragon,J.A., Delerce,J., Jaeger,S., Blanchet,C.,

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

Vincens,P., Caron,C. *et al.* (2015) RSAT 2015: Regulatory sequence analysis tools. *Nucleic Acids Res.*, **43**, W50–W56.
32. Sood,U., Hira,P., Kumar,R., Bajaj,A., Rao,D.L.N., Lal,R. and Shakarad,M. (2019) Comparative genomic analyses reveal Core-Genome-Wide genes under positive selection and major regulatory hubs in outlier strains of pseudomonas aeruginosa. *Front. Microbiol.*, **10**, 53.
33. Browning,D.F., Butala,M. and Busby,S.J.W. (2019) Bacterial Transcription Factors: Regulation by Pick "N" Mix. *J. Mol. Biol.*, **431**, 4067–4077.
34. Schnell,R., Oehlmann,W., Sandalova,T., Braun,Y., Huck,C., Maringer,M., Singh,M. and Schneider,G. (2012) Tetrahydrodipicolinate N-succinyltransferase and dihydrodipicolinate synthase from Pseudomonas aeruginosa: structure analysis and gene deletion. *PLoS One*, **7**, e31133.
35. Impey,R.E., Panjikar,S., Hall,C.J., Bock,L.J., Sutton,J.M., Perugini,M.A. and Soares da Costa,T.P. (2019) Identification of two dihydrodipicolinate synthase isoforms from Pseudomonas aeruginosa that differ in allosteric regulation. *FEBS J.*, doi:10.1111/febs.15014.
36. Basso,P., Wallet,P., Elsen,S., Soleilhac,E., Henry,T., Faudry,E. and Attree,I. (2017) Multiple Pseudomonas species secrete exolysin-like toxins and provoke Caspase-1-dependent macrophage death. *Environ. Microbiol.*, **19**, 4045–4064.
37. Job,V., Bouillot,S., Gueguen,E., Robert-Genthon,M., Panchev,P., Elsen,S. and Attree,I. (2019) Pseudomonas two-partner secretion toxin Exolysin contributes to insect killing. bioRxiv doi: https://doi.org/10.1101/807867, 17 October 2019, preprint: not peer reviewed.
38. Smith,R.S., Wolfgang,M.C. and Lory,S. (2004) An adenylate cyclase-controlled signaling network regulates Pseudomonas aeruginosa virulence in a mouse model of acute pneumonia. *Infect. Immun.*, **72**, 1677–1684.
39. Wolfgang,M.C., Lee,V.T., Gilmore,M.E. and Lory,S. (2003) Coordinate regulation of bacterial virulence genes by a novel adenylate cyclase-dependent signaling pathway. *Dev. Cell*, **4**, 253–263.
40. Marsden,A.E., Intile,P.J., Schulmeyer,K.H., Simmons-Patterson,E.R., Urbanowski,M.L., Wolfgang,M.C. and Yahr,T.L. (2016) Vfr directly activates exsA transcription to regulate expression of the pseudomonas aeruginosa Type III secretion system. *J. Bacteriol.*, **198**, 1442–1450.
41. Fuchs,E.L., Brutinel,E.D., Jones,A.K., Fulcher,N.B., Urbanowski,M.L., Yahr,T.L. and Wolfgang,M.C. (2010) The Pseudomonas aeruginosa Vfr regulator controls global virulence factor expression through cyclic AMP-dependent and -independent mechanisms. *J. Bacteriol.*, **192**, 3553–3564.

Downloaded from https://academic.oup.com/nar/article/48/5/2388/5700547 by CEA grenoble user on 24 August 2020

Supplementary Information

Species-specific recruitment of transcription factors dictates toxin expression

Julian Trouillon[1]*, Erwin Sentausa[1], Michel Ragno[1], Mylène Robert-Genthon[1], Stephen Lory[2], Ina Attree[1], Sylvie Elsen[1]*

[1]Université Grenoble Alpes, CNRS ERL5261, CEA-IRIG-BCI, INSERM UMR1036, Grenoble, France
[2]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, USA

*To whom correspondence should be addressed. Tel: +33 0438783074 ; Email: julian.trouillon@cea.fr ; Email: sylvie.elsen@cea.fr

Present Address: Erwin Sentausa, Evotec ID (Lyon) SAS, Marcy l'Étoile, France

Supplementary Material & Methods

Plasmids and genetic manipulations. Plasmids and primers are listed in the Supplementary Table 6. For ErfA-His10 overproduction, the *erfA* sequence was amplified by PCR using either *P. aeruginosa* IHMA87 or *P. chlororaphis* PA23 genomic DNA as a template and appropriate primer pairs, and then integrated by Sequence- and Ligation-Independent Cloning (SLIC) (1) in pET52b cut with *Nco*I-*Sac*I. The resulting pET52b-erfA and pET52b-erfA-chloro plasmids were then transformed into *E. coli* BL21 Star (DE3).

To generate *P. aeruginosa* deletion mutants, approximately 500 bp of upstream and downstream sequences flanking regions of IHMA87 *exlBA*, *erfA* and *ergAB*, and PAO1 *erfA* were amplified using appropriate primer pairs (sF1/sR1 and sF2/sR2 for each construct). The two resulting, overlapping fragments were then cloned into *Sma*I-cut pEXG2 by SLIC. A similar strategy was used to introduce the sequence encoding VSV-G Tag (YTDIEMNRLGK) at the 3' end (before the stop codon) of both IHMA87 and PAO1 *erfA* genes, leading to pEXG2-erfA-VSVG and pEXG2-PA0225-VSVG plasmids. Transcriptional *erfA::lacZ* and *ergAB::lacZ* fusions in IHMA87 were created with *lacZ* integrated after the 3rd codon of the gene. The upstream region of *erfA* (or *ergB*), the *lacZ* gene, and the downstream region of *erfA* (or *ergB*) were first amplified using the three appropriate primer pairs. The resulting, overlapping fragments were then cloned into *Sma*I-cut pEXG2 by SLIC. To construct the P*popN-lacZ* transcriptional fusion, the 254 bp-long promoter fragment was PCR amplified using PAO1 chromosomal DNA as a template, sequenced, and cloned into the *Eco*RI-*Bam*HI sites of mini-CTX-lacZ.

Mutations were introduced in the ErfA-binding site of both *exlBA* (TAACAACACTTTGTGTCAAA) and *ergAB* (TAATGACACTTTGTGTCATA) promoters, directly in the chromosome of the strain IHMA87 *exlA::lacZ* and IHMA87 *ergAB::lacZ*, respectively. To do so, the upstream and the downstream regions flanking the EBS were amplified using appropriate primer pairs to introduce the mutation EBS2 (*exlBA*: TAACAACACTTTTCTAGAAA; *ergAB*: TAATGACACTTTTCTAGATA). The mutation created the *Xba*I restriction site (underlined) allowing selection of mutated genotype. The resulting, overlapping fragments were then cloned by SLIC into *Sma*I-cut pEXG2. For *P. chlororaphis erfA* mutagenesis, a 1048 bp fragment containing a truncated version of the *RS06185* gene with TGA incorporated after the 7th codon was synthesized by Genewiz and then subloned in EcoRI-XmaI of pEXG2, leading to pEXG2-Mut-Pchloro_erfA. For complementation, the miniCTX1-T*rrnB* plasmid was first modified to prevent transcriptional interference by integrating into the *Xho*I-*Kpn*I sites of mini-CTX1 plasmid the sequence of the strong *rrnB* terminator, amplified from plasmid pMMB67EH (2). The region containing the native IHMA87 *erfA* gene along with 750 bp upstream and 113 bp downstream of the gene was amplified by PCR and cloned by SLIC into the integrative miniCTX1-T*rrnB*.

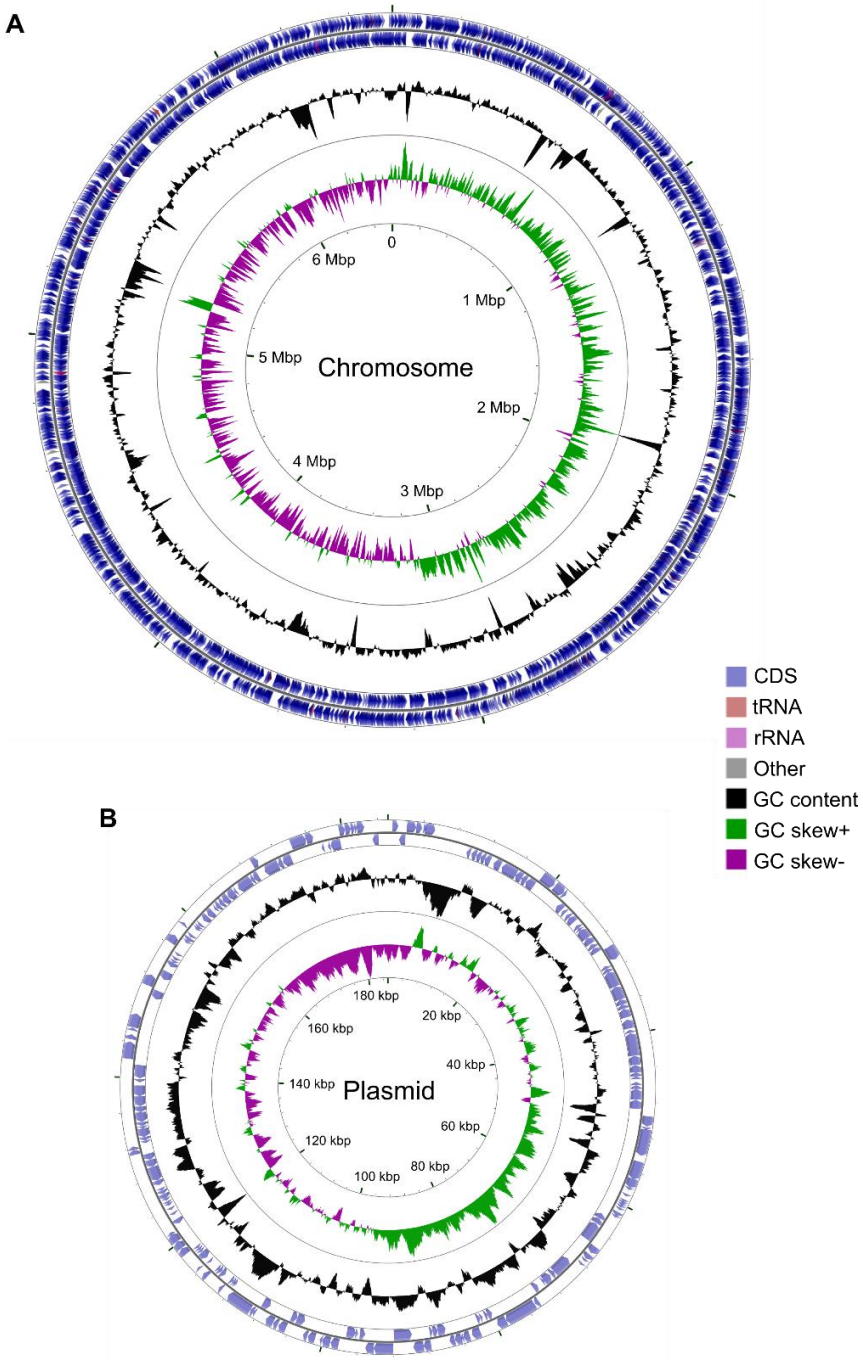
After SLIC, all the plasmids were transformed into competent TOP10 *E. coli* and verified by sequencing (Eurofins). The miniCTX- and pEXG2-derived vectors were transferred into *Pseudomonas* strains by triparental mating using pRK600 as a helper plasmid. For allelic exchange using the pEXG2 plasmids, merodiploids resulting from cointegration events were selected on LB plates containing irgasan and

gentamicin. Single colonies were then plated on NaCl-free LB agar plates containing 10% (wt/vol) sucrose to select for the loss of plasmid, and the resulting sucrose-resistant strains were checked for mutant genotype and gentamicin sensitivity. For complementation or transcriptional *lacZ* fusion integration using mini-CTX derived-plasmids, bacteria with plasmids inserted at the *att* site were selected on irgasan and tetracyclin. The PAO1 strains were further cured from the mini-CTX-lacZ backbone by excising FRT cassette following introduction of the pFLP2 plasmid as previously described (3).

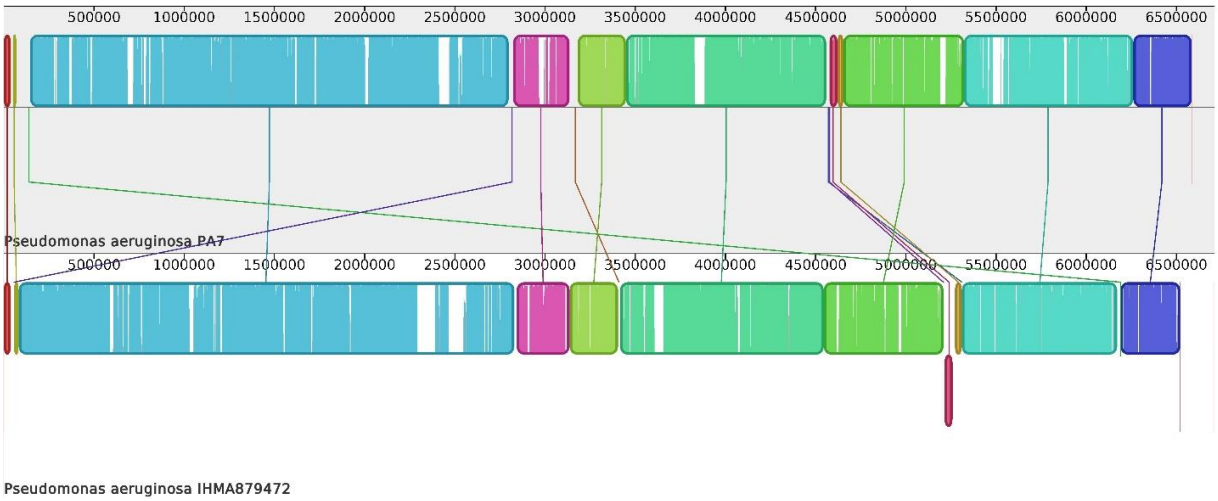
β-galactosidase activity assay. β-galactosidase activity was assessed when bacterial cultures reached indicated OD_{600} as previously described (4), with technical details reported in reference (5).

Protein purification. For ErfA-His10 overproduction, overnight cultures were diluted to OD_{600}=0.05 in LB medium containing 100 µg/ml ampicillin and expression induced at an OD_{600}=0.6 with 1 mM IPTG. After 3h of growth at 37°C, bacteria were harvested by centrifugation at 6,000 g for 10 min at 4°C and resuspended in L Buffer (25 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole, 1 mM PMSF, 5 % Glycerol, pH 8, containing Roche protease inhibitor cocktail). Bacteria were then lysed by sonication. After centrifugation at 66,500 g for 30 min at 4°C, the soluble fraction was directly loaded onto a 1-ml nickel column (Protino Ni-nitrotriacetic acid [NTA]; Macherey-Nagel). The column was washed with Wash Buffer (50 mM Tris-HCl, 500 mM NaCl, 5 % Glycerol, pH 8) containing increasing imidazole concentrations (20, 40, and 60 mM), and proteins were eluted with 200 mM imidazole. Aliquots from the peak protein fractions were analyzed by SDS-PAGE, and the fractions containing ErfA-His10 were pooled and dialyzed against ErfA Buffer (50 mM Tris-HCl, 250 mM NaCl, 50 mM KCl, 10 % Glycerol, 0.5 % Tween20, pH 7).

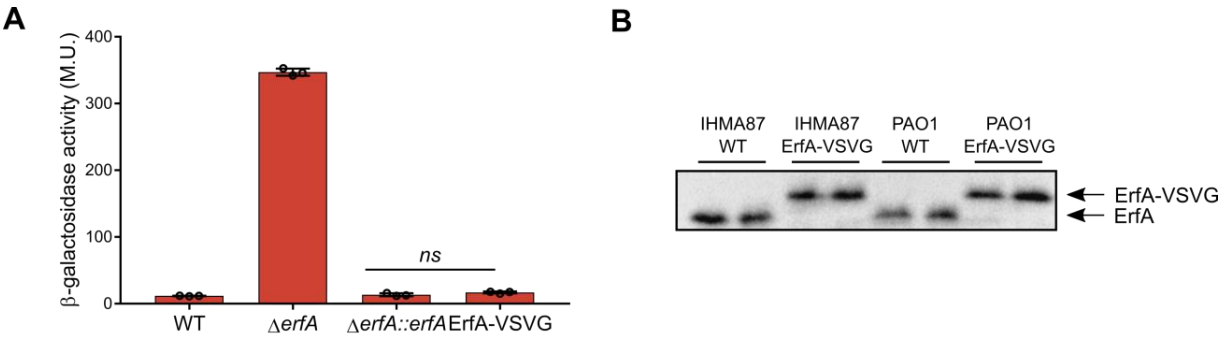
Western Blot. Recombinant ErfA-His10 protein was used to raise polyclonal antibodies in rabbit following manufacturer's recommendations (Biotem). To analyze ExlA secretion, 1.5 ml of a bacterial culture at OD_{600}=1.0 were centrifuged at 6,000 g for 5 min at 4°C. 15 µl of 2 % sodium deoxycholate (DOC) were added to supernatants and then incubated for 20 min at 4°C. 150 µl of trichloroacetic acid (TCA) were further added. After 40 min of incubation at 4°C, tubes were then centrifuged at 15,000 g for 15 min at 4°C. Pellets were resuspended in 30 µl of loading buffer and incubated at 100°C for 10 min. Proteins amount in total bacteria was evaluated from 1 ml of liquid cultures grown to OD_{600}=1 which was pelleted, resuspended in 100 µl of loading buffer and incubated at 100°C for 10 min. Samples were separated on denaturing 8 or 12 % PAGE and transferred onto polyvinyl difluoride (PVDF) membranes. Primary antibodies used for immunodetection consisted of a mixture of anti-ExlAΔCter (1:1,000 dilution) and anti-CterExlA (1:1,000) for ExlA detection, anti-ErfA (1:2,000), anti-EF-Tu (1:20,000), anti-FliC (1:10,000) and anti-VSVG (1:5,000). Secondary antibodies were anti-rabbit-HRP (Sigma) and anti-mouse-HRP (Sigma) diluted in PBS-Tween. The membranes were developed with Luminata Classico Western HRP (Millipore) substrate.



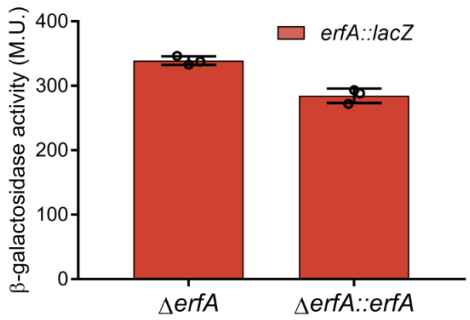
Supplementary Figure 1. Genome of the *P. aeruginosa* IHMA879472 (IHMA87) strain. Circular maps of the chromosome (**A**) and pIHMA87 plasmid (**B**) created using the CGView Server (6). Total sizes are 6,527,298 bp and 185,168 bp for the chromosome and the plasmid, respectively. From outside to center rings: CDS position and orientation, DNA GC content and GC skew.



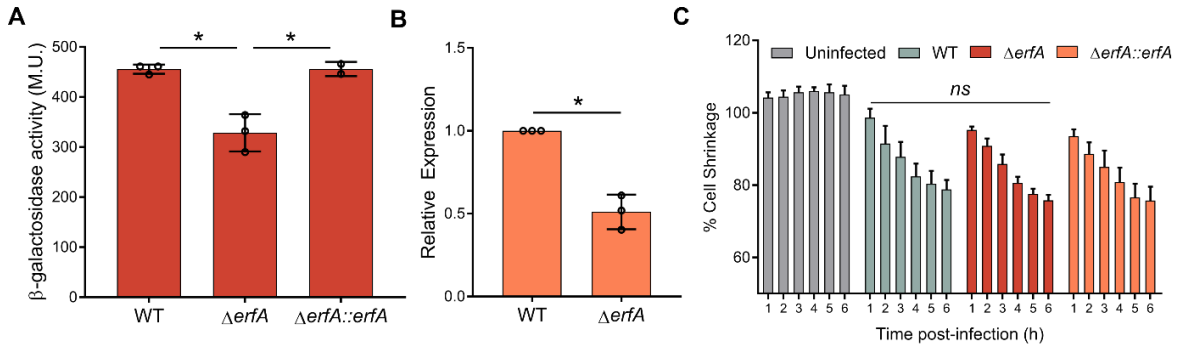
Supplementary Figure 2. Mauve visualization of locally collinear blocks identified between *P. aeruginosa* PA7 and IHMA87 genomes. Locally collinear blocks, regions without rearrangement of homologous backbone sequence are shown as contiguously colored regions. The height of the similarity profile inside the blocks corresponds to the average level of conservation in that region of the genome sequence. The two genomes are generally syntenic, although there is an inversion of about 4 kb in IHMA87 compared to PA7 (shown as a small red block below the horizontal line in IHMA87 genome) and there are sequence elements specific to each genome (shown as white areas inside the blocks). Furthermore, the IHMA87 plasmid sequence (white area flanked by red vertical lines at the end of IHMA87 genome) is completely absent from PA7.



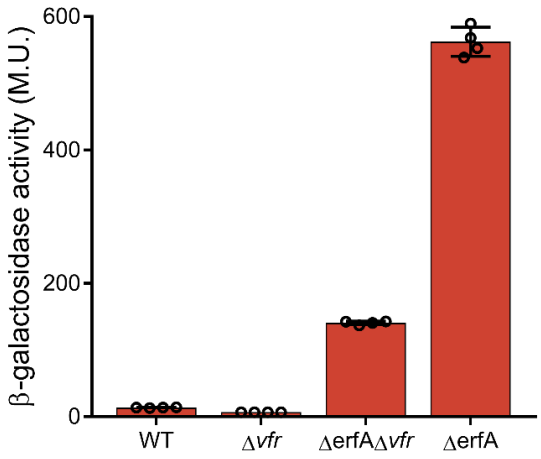
Supplementary Figure 3. Addition of the VSV-G tag does not affect ErfA function and expression. (**A**) β-galactosidase activities of IHMA87 strains harboring *exlBA::lacZ* transcriptional fusion. The strains were grown in LB and activities were measured at OD_{600}=1. Activities are represented as mean (in Miller Units) ± SD. from three independent experiments. Statistical significance was determined using two-tailed *T*-test. *ns*: not significant. Addition of the VSV-G tag does not impair ErfA inhibitory role on *exlBA*. (**B**) Western Blot with anti-ErfA antibodies on lysates of duplicate cultures used for ChIP-seq. Addition of the VSV-G tag does not change quantity of ErfA protein in bacteria.



Supplementary Figure 4. ErfA binding on P*ergAB* only has a mild effect on *erfA* transcription. β-galactosidase activities of IHMA87 strains harboring *erfA::lacZ* transcriptional fusion. Activities are expressed in Miller Units and are represented as mean ± SD from three independent experiments. ErfA binding on P*ergAB* in the intergenic region between *ergAB* and *erfA* strongly impacts *ergAB* transcription but only slightly diminishes *erfA* transcription. This mild effect does not reflect autoregulation but is probably due to steric hindrance between the RNA polymerase transcribing *erfA* and ErfA binding on P*ergAB* as the two promoters are facing each other.



Supplementary Figure 5. ErfA slightly impacts transcription of some T3SS genes, but not T3SS-dependent cytotoxicity. a, β-galactosidase activities of IHMA87 strains carrying a *popN::lacZ* transcriptional fusion measured after growth in LB up to OD_{600}=1. Activities are represented as mean (in Miller Units) ± SD from three independent experiments. Statistical significance was determined using two-tailed *T*-test. **b**, RT-qPCR analysis of *popB* relative expression (part of the *pcrGVHpopBD* transcriptional unit). Strains were grown in LB up to OD_{600}=1. Relative expressions are represented as mean ± SD from three independent experiments. Statistical significance was determined using two-tailed *T*-test (*p<0.05). **c**, Kinetics of bacterial cytotoxicity on A549 epithelial cells. Cell shrinkage was followed by monitoring Dii staining surface every hour after infection at a MOI of 1.



Supplementary Figure 6. ErfA hinders Vfr activation of *exlBA*. β-galactosidase activities of indicated IHMA87 strains carrying the *exlBA::lacZ* transcriptional fusion. The strains were grown in LB and activities were measured at OD_{600}=1.3. Activities are represented as mean (in Miller Units) ± SD from four independent experiments.

Supplementary Table 5 | List of bacterial strains and plasmids used in this study

Strain or plasmid	Genotype or relevant properties	Reference/source
Strains		
P. aeruginosa		
IHMA879472 (IHMA87)	wild-type strain (urinary infection)	(7)
IHMA87 Δ*exlBA*	IHMA87 with *exlBA* deletion	This work
IHMA87 *exlBA::lacZ*	IHMA87 with promoterless *lacZ* in *exlA*	(8)
IHMA87 *exlBA::lacZ* P*exlBA*-EBS2	IHMA87 with EBS2 mutation in P*exlBA::lacZ*	This work
IHMA87 Δ*erfA*	IHMA87 with *erfA* deletion	This work
IHMA87 Δ*erfA::erfA*	IHMA87 Δ*erfA* deletion with miniCTX-TrrnB-erfA in *attB* site (Tc^R)	This work
IHMA87 *erfA-vsvg*	IHMA87 with VSV-G tag-encoding sequence inserted at 3' end of *erfA*	This work
IHMA87 *erfA-vsvg exlBA::lacZ*	IHMA87 *erfA-vsvg* with *exlBA::lacZ* transcriptional fusion	This work
IHMA87 *erfA::lacZ*	IHMA87 with promoterless *lacZ* in *erfA*	This work
IHMA87 *::erfA erfA::lacZ*	IHMA87 with miniCTX-TrrnB-erfA in *attB* site and promoter less *lacZ* in *erfA* (Tc^R)	This work
IHMA87 Δ*erfA exlBA::lacZ*	IHMA87 *exlBA::lacZ* with *erfA* deletion	This work
IHMA87 Δ*vfr exlBA::lacZ*	IHMA87 *exlBA::lacZ* with *vfr* deletion	(8)
IHMA87 Δ*erfA* Δ*vfr exlBA::lacZ*	IHMA87 *exlBA::lacZ* with *vfr* and *erfA* deletion	This work
IHMA87 Δ*erfA::erfA exlBA::lacZ*	IHMA87 Δ*erfA exlBA::lacZ* with miniCTX-TrrnB-erfA in *attB* site (Tc^R)	This work
IHMA87 Δ*erfA ergAB::lacZ*	IHMA87 Δ*erfA* with *ergAB::lacZ in situ* transcriptional fusion	This work
IHMA87 Δ*ergAB*	IHMA87 with *ergAB* deletion	This work
IHMA87 Δ*erfA* Δ*ergAB*	IHMA87 with *erfA* and *ergAB* deletion	This work
IHMA87 Δ*ergAB exlBA::lacZ*	IHMA87 Δ*ergAB* with *exlBA::lacZ in situ* transcription fusion	This work
IHMA87 *ergAB::lacZ*	IHMA87 with promoterless *lacZ* in *ergB*	This work
IHMA87 *ergAB::lacZ* P*ergAB*-EBS2	IHMA87 with EBS2 mutation in P*ergAB::lacZ*	This work
PAO1F (PAO1)	Wild-type reference strain (wound isolate)	(9)
PAO1 *erfA-vsvg*	PAO1 with VSV-G tag-encoding sequence inserted at 3' end of *erfA*	This work
PAO1 Δ*erfA*	PAO1 with *erfA* deletion	This work
PAO1 Δ*erfA::erfA*	PAO1 Δ*erfA* with miniCTX-TrrnB-erfA in *attB* site (Tc^R)	This work
PAO1::*PpopN-lacZ*	PAO1 with P*popN-lacZ* fusion in attB site	This work
PAO1 Δ*erfA*::P*popN-lacZ*	PAO1Δ*erfA* with P*popN-lacZ* fusion in attB site	This work
PAO1 Δ*erfA::erfA* ::*PpopN-lacZ*	PAO1Δ*erfA::erfA* with P*popN-lacZ* fusion in attB site	This work
PA70	Non-CF Bronchiectasis isolate (expectoration)	(10)
PA70 Δ*erfA*	PA70 with *erfA* deletion	This work
AL-198	CF iolate (throat swab)	B. Tümmler
AL-198 Δ*erfA*	AL-198 with *erfA* deletion	This work
P. chlororaphis		
PA23	Soybean root isolate	(11)
PA23 Δ*exlBA*	PA23 with *exlBA* deletion (*RS20945*/*RS20950*)	(12)
PA23 Δ*erfA*	PA23 with *erfA* deletion (*RS06185*)	This work

P. putida		
KT2440	Soil isolate	(13)
P. protegens		
CHA0	Soil isolate	(14)
E. coli		
TOP10	Chemically competent cells	Invitrogen
SM10	Bi-parental mating (conjugal RP4 transfer functions)	Stephen Lory
BL21 Star (DE3)	F^- *ompT hsdSB* (rB^- mB^-) *gal dcm rne131*(DE3)	Invitrogen

Plasmids		
pRK2013	Helper plasmid with conjugative properties (Km^R)	(15)
pRK600	Helper plasmid with conjugative properties (Cm^R)	(16)
pBTK24	Plasmid carrying Himar-1 mariner transposon (Ap^R, Gm^R)	(17)
pEXG2	Allelic exchange vector (Gm^R)	(18)
pEXG2-mut-exlBA	pEXG2 carrying SLIC fragment for *exlBA* deletion (Gm^R)	This work
pEXG2-exlBA-lacZ	pEXG2 carrying *lacZ* integrated within *exlBA* fragment for integration into chromosome (Gm^R)	(8)
pEXG2-ergAB-lacZ	pEXG2 carrying SLIC fragment for *ergAB* deletion (Gm^R)	This work
pEXG2-erfA-lacZ		
pEXG2-mut-RS06750	pEXG2 carrying SLIC fragment for IHMA87 *erfA* deletion (Gm^R)	This work
pEXG2-mut-PA0225	pEXG2 carrying SLIC fragment for PAO1 *erfA* deletion (Gm^R)	This work
pEXG2-mut-IHMA87_00213-4	pEXG2 carrying SLIC fragment for IHMA87 *ergAB* deletion (Gm^R)	This work
pEXG2-erfA-VSVG	pEXG2 carrying SLIC fragment for integration of VSV-G Tag encoding sequence at 3'end of IHMA87 *erfA* gene (Gm^R)	This work
pEXG2-Mut-Pchloro_erfA	pEXG2 carrying *Eco*RI-*Xma*I fragment for integration for *P. chlororaphis erfA* deletion (Gm^R)	This work
pEXG2-PA0225-VSVG	pEXG2 carrying SLIC fragment for integration of VSV-G Tag encoding sequence at 3'end of PAO1 *erfA* gene (Gm^R)	This work
pEXG2-Mut-EBS2	pEXG2 carrying SLIC fragment for integrating EBS2 mutation within *pexlBA* promoter (Gm^R)	This work
pEXG2-Mut-ergA-EBS2	pEXG2 carrying SLIC fragment for integrating EBS2 mutation within *pergAB* promoter (Gm^R)	This work
mini-CTX1	Site-specific integrative plasmid (*attP* site, Tc^R)	(19)
miniCTX1-TrrnB	mini-CTX1 with strong *rrnB* terminator (*attP*, Tc^R)	This work
miniCTX-TrrnB-erfA	miniCTX1-TrrnB carrying SLIC fragment for *erfA* complementation (*attP*, Tc^R)	This work
mini-CTX-lacZ	Site-specific integrative plasmid with promoter less-*lacZ* (*attP* site, Tc^R)	(20)
miniCTX-*PpopN* lacZ	miniCTX-lacZ harboring the *popN* promoter fused to *lacZ* fusion (*attP*,*Tc*^{R})	This work
pFLP2	Source of Flp recombinase (Ap^R)	(3)
pET52b	Expression vector (Ap^R)	Novagen
pET52b-erfA	Expression vector of IHMA87 *erfA-his10* (Ap^R)	This work
pET52b-erfA-chloro	Expression vector of PA23 *erfA-his10* (Ap^R)	This work

Supplementary Table 6 | Primers used in this work

Name	Sequence (5' > 3')	Use
pEXG2-mut-exlBA-sF1	GGTCGACTCTAGAGGATCCCC ACCGTGGAGTATCAGGTCTG	*exlBA* deletion
pEXG2-mut-exlBA-sR1	TCA CAGCAGCAGGGCTCGGCA	*exlBA* deletion
pEXG2-mut-exlBA-sF2	TGCCGAGCCCTGCTGCTGTGAAGCCAGGAGCACGACCAGG	*exlBA* deletion
pEXG2-mut-exlBA-sR2	ACCGAATTCGAGCTCGAGCCC CCTCGAGCGCTACCACTGG	*exlBA* deletion
Mut-RS06750-sF1	GGTCGACTCTAGAGGATCCCC ACCAGCAGCAGGTTGCCGGC	IHMA87 & PAO1 *erfA* deletion
Mut-RS06750-sR1	GATCCGCTGGTTCATGAAGTCTCACAATTTCAAGCGCATGGACATT	IHMA87 & PAO1 *erfA* deletion
Mut-RS06750-sF2	GACTTCATGAACCAGCGGATCAT	IHMA87 & PAO1 *erfA* deletion
Mut-RS06750-sR2	ACCGAATTCGAGCTCGAGCCC AGCCTTCGAGGGCGACGCTG	IHMA87 *erfA* deletion
Mut-PA0225-sR2	ACCGAATTCGAGCTCGAGCCC GTCATGCGGATCAGCGTCAGG	PAO1 *erfA* deletion
pEXG2-mut-0223-4-sF1	GGTCGACTCTAGAGGATCCCC GCAATTTCAAGCGCATGGACAT	*ergAB* deletion
pEXG2-mut-0223-4-sR1	GAGCGTCGGGATCAGGTTCGGTCATGATGCAGTTGACGTCCGGG	*ergAB* deletion
pEXG2-mut-0223-4-sF2	CCGAACCTGATCCCGACGC	*ergAB* deletion
pEXG2-mut-0223-4-sR2	ACCGAATTCGAGCTCGAGCCC GAGAAATTCATCGCCTTCGCC	*ergAB* deletion
ErfA-vsvg-sF1	GGTCGACTCTAGAGGATCCCC TCCATGCGCTTGAAATTGCTG	IHMA87 *erfA* vsv-g tag insertion
ErfA-vsvg-sR1	ATTCATTTCGATATCGGTATAGCCGCCGCCTTCCTCGCTGTGG ATCACCA	IHMA87 *erfA* vsv-g tag insertion
ErfA-vsvg-sF2	TATACCGATATCGAAATGAATCGCCTGGGCAAATGAGGCGTG GGCTCGTCG	IHMA87 *erfA* vsv-g tag insertion
ErfA-vsvg-sR2	ACCGAATTCGAGCTCGAGCCC CTTCTCCACCGCGTCGCG	IHMA87 *erfA* vsv-g tag insertion
ErfA-vsvg-sF1	GGTCGACTCTAGAGGATCCCC TCCATGCGCTTGAAATTGCTG	PAO1 *erfA* vsv-g tag insertion
PA0225-vsvg-sR1	ATTCATTTCGATATCGGTATAGCCGCCGCCTTCGTCGCTGTGG ATCACCA	PAO1 *erfA* vsv-g tag insertion
PA0225-vsvg-sF2	TATACCGATATCGAAATGAATCGCCTGGGCAAATGAGGCGG CGGCTTCGGT	PAO1 *erfA* vsv-g tag insertion
ErfA-vsvg-sR2	ACCGAATTCGAGCTCGAGCCC CTTCTCCACCGCGTCGCG	PAO1 *erfA* vsv-g tag insertion
pCTXter-erfA-sF	GATATCGAATTCCTGCAGCCC CGATCACGCAGGCTTCCTC	IHMA87 *erfA* complementation
pCTX1ter-erfA-sR	TCTAGAACTAGTGGATCCCCC ACGTTAGCAAGCAACCGACC	IHMA87 *erfA* complementation
Mut-pexlBA-erfA-sF1	GGTCGACTCTAGAGGAT CCC GTCATATCTTCCACCGATACTTC	EBS2 P*exlBA* mutagenesis
Mut-pexlBA-erfA2-sR1	**CTAGA**AAAGTGTTGTTATGTAAACTTTGTGT	EBS2 P*exlBA* mutagenesis
Mut-pexlBA-erfA2-sF2	TACATAACAACACTTT**TCTAG**AAAGCTGTCGCTGAAACGTCCA	EBS2 P*exlBA* mutagenesis

Mut-pexlBA-erfA-sR2	ACCGAATTCGAGCTCGAGCCC GTGATCGCCGCCAGCAGACG	EBS2 P*exlBA* mutagenesis
Mut-pergAB-erfA-sF1	GGTCGACTCTAGAGGATCCCC GAGGATGATCCGCTGGTTCATG	EBS2 P*ergAB* mutagenesis
Mut-pergAB-erfA**2**-sR2	TAAATGACACGGATCAAGTTATAT**CTAGA**AAAGTGTCATTAGCA AGAAAAAGTTTT	EBS2 P*ergAB* mutagenesis
Mut-pergAB-erfA**2**-sF2	**TCTAG**ATATAACTTGATCCGTGTCATTTACA	EBS2 P*ergAB* mutagenesis
Mut-pergAB-erfA-sR2	ACCGAATTCGAGCTCGAGCCC TCGACAGCGCGGCGATGTGCA	EBS2 P*ergAB* mutagenesis
pEXG2-erfA-lacZ-sF1	TGCAGGTCGACTCTAGAGGATCCCC CGCCTTCTTCGTTGCCTACC	*erfA::lacZ* fusion
pEXG2-erfA-lacZ-sR1	AAGTTAAAATGCCGCGCCCCTACCT CATGGACATTGCGGACAGCC	*erfA::lacZ* fusion
P3 (SLIC_lacZ_F)	AGGTAGGGGCGCGGCATTTT	*erfA::lacZ* fusion
P4 (SLIC_lacZ_R)	TTATTTTTGACACCAGACCAACTG	*erfA::lacZ* fusion
pEXG2-erfA-lacZ-sF2	CCAGTTGGTCTGGTGTCAAAAATAA CGCTTGAAATTGCTGAGAAAA	*erfA::lacZ* fusion
pEXG2-erfA-lacZ-sR2	AGGTACCGAATTCGAGCTCGAGCCC CGATTAACAGTTCGATAATCGA	*erfA::lacZ* fusion
pEXG2-223-lacZ-sF1	TGCAGGTCGACTCTAGAGGATCCCCGACCTACTACACCCAGCAAC	*ergAB-lacZ* fusion
pEXG2-223-lacZ-sR1	AAGTTAAAATGCCGCGCCCCTACCTTGCGGACATTGGATGTACCC	*ergAB-lacZ* fusion
P3 (SLIC_lacZ_F)	AGGTAGGGGCGCGGCATTTT	*ergAB-lacZ* fusion
P4 (SLIC_lacZ_R)	TTATTTTTGACACCAGACCAACTG	*ergAB-lacZ* fusion
pEXG2-223-lacZ-sF2	CCAGTTGGTCTGGTGTCAAAAATAATCCATCCACGGCATCATCGG	*ergAB-lacZ* fusion
pEXG2-223-lacZ-sR2	AGGTACCGAATTCGAGCTCGAGCCCCGTCGGGATCAGGTTCGGC	*ergAB-lacZ* fusion
PSE-popN1	GAATTCGATGGCGTGGCGCATC	P*popN-lacZ* fusion
PSE-popN3	GGATCCCGCGGCGCGAGGG	P*popN-lacZ* fusion
pET52b-erfA-sF	TTAAGAAGGAGATATACC ATGTCCATGCGCTTGAAATTGA	IHMA87 *erfA* overexpression
pET52b-erfA-sR	CTACCGCGTGGCACCAGAGCGAG TTCCTCGCTGTGGATCACCA	IHMA87 *erfA* overexpression
pET52b-erfA-chloro-sF	TTAAGAAGGAGATATACC ATGTCTATCCGTTTGAAATTATTGA	PA23 *erfA* overpression
pET52b-erfA-chloro-sR	CTACCGCGTGGCACCAGAGCGAG TTCCTCGCCGCTGTGCACC	PA23 *erfA* overpression
rrnB-T12-F	GTCCGGTACCCCTAGGTTTGGCGGATGAGAGAAGATT	miniCTX-TrrnB construction
rrnB-T12-R	GTCCCTCGAGTGTAGAAACGCAAAAAGGCCA	miniCTX-TrrnB construction
Round-1 RndomPA-1	GGCCACGCGTCGACTAGTACNNNNNNNNNNCGATG	semi-random PCR
Round-1a	GGCCACGCGTCGACTAGTAGNNNNNNNNNCAGCAG	semi-random PCR
Round-1 pBTK	GAAGCTGTGGTATGGCTGTGCAGG	semi-random PCR
Round-2PA	GGCCACGCGTCGACTAGTAC	semi-random PCR
Round-2 pBTK	CGCACTCCCGTTCTGGATAATGTT	semi-random PCR
(Cy5) pexlB-40-EMSA-F	(Cy5-) ATGTAACAAACGGGTGAGGTCTG	P*exlBA* EMSA
pexlB+40-EMSA-4R	GACAGCTTTGACACAAAGTGTTGT	P*exlBA* EMSA
(Cy5-) pergAB-EMSA-F	(Cy5-) CAGCCTTCTCCCGATGGCAGT	P*ergAB* EMSA
pergAB-EMBSA-R	CCGGCTGTAAATGACACGGATC	P*ergAB* EMSA
rpoD F1	GCGCAACAGCAATCTCGTCT	RT-qPCR
rpoD R1	ATCCGGGGCTGTCTCGAATA	RT-qPCR

exlAqPCR F1	AGAGCGACAGCGTCATCCAG	RT-qPCR
exlAqPCR R1	GCGTCGACGTTGACCTTCAC	RT-qPCR
exlB qPCR up	CCTATGGCTACTGGACCTACA	RT-qPCR
exlB qPCR down	AGGTAGCTGTCGACATCCTTG	RT-qPCR
qPCR_ergA_up	ACTGGATCTCCACCCCGAAG	RT-qPCR
qPCR_ergA_down	CAGCTCAGGCAGTCGCTGTT	RT-qPCR
qPCR_ergB_up	CTACCTGAGCGACCCGGAAT	RT-qPCR
qPCR_ergB_down	AGGTCGGAGACGCTGACGAT	RT-qPCR
qPCR_popB_up	GATCGAGCTGGCGAAAATCAC	RT-qPCR
qPCR_popB_down	TCTTCTGCTGGTTGTCCTCCA	RT-qPCR
Adaptor A	CACGACGCTCTTCCGATCT	DAP-seq libraries
Adaptor B	P-GATCGGAAGAGCACACGTCTG	DAP-seq libraries

Supplementary references

1. Li, M.Z. and Elledge, S.J. (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Methods*, **4**, 251-256.
2. Furste, J.P., Pansegrau, W., Frank, R., Blocker, H., Scholz, P., Bagdasarian, M. and Lanka, E. (1986) Molecular cloning of the plasmid RP4 primase region in a multi-host-range tacP expression vector. *Gene*, **48**, 119-131.
3. Hoang, T.T., Karkhoff-Schweizer, R.R., Kutchma, A.J. and Schweizer, H.P. (1998) A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked Pseudomonas aeruginosa mutants. *Gene*, **212**, 77-86.
4. Miller, J.H. (1972) *Experiments in molecular genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
5. Thibault, J., Faudry, E., Ebel, C., Attree, I. and Elsen, S. (2009) Anti-activator ExsD forms a 1:1 complex with ExsA to inhibit transcription of type III secretion operons. *J Biol Chem*, **284**, 15762-15770.
6. Grant, J.R. and Stothard, P. (2008) The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Res*, **36**, W181-184.
7. Kos, V.N., Deraspe, M., McLaughlin, R.E., Whiteaker, J.D., Roy, P.H., Alm, R.A., Corbeil, J. and Gardner, H. (2015) The resistome of Pseudomonas aeruginosa in relationship to phenotypic susceptibility. *Antimicrob Agents Chemother*, **59**, 427-436.
8. Berry, A., Han, K., Trouillon, J., Robert-Genthon, M., Ragno, M., Lory, S., Attree, I. and Elsen, S. (2018) cAMP and Vfr Control Exolysin Expression and Cytotoxicity of Pseudomonas aeruginosa Taxonomic Outliers. *J Bacteriol*, **200**.
9. Bleves, S., Soscia, C., Nogueira-Orlandi, P., Lazdunski, A. and Filloux, A. (2005) Quorum sensing negatively controls type III secretion regulon expression in Pseudomonas aeruginosa PAO1. *J Bacteriol*, **187**, 3898-3902.
10. Reboud, E., Elsen, S., Bouillot, S., Golovkine, G., Basso, P., Jeannot, K., Attree, I. and Huber, P. (2016) Phenotype and toxicity of the recently discovered exlA-positive Pseudomonas aeruginosa strains collected worldwide. *Environ Microbiol*, **18**, 3425-3439.
11. Loewen, P.C., Villenueva, J., Fernando, W.G. and de Kievit, T. (2014) Genome Sequence of Pseudomonas chlororaphis Strain PA23. *Genome Announc*, **2**.

12. Job, V., Bouillot, S., Gueguen, E., Robert-Genthon, M., Panchev, P., Elsen, S. and Attree, I. (2019) Pseudomonas two-partner secretion toxin Exolysin contributes to insect killing. *bioRxiv* **807867**.
13. Nelson, K.E., Weinel, C., Paulsen, I.T., Dodson, R.J., Hilbert, H., Martins dos Santos, V.A., Fouts, D.E., Gill, S.R., Pop, M., Holmes, M. *et al.* (2002) Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440. *Environ Microbiol*, **4**, 799-808.
14. Voisard, C., Bull, C.T., Keel, C., Laville, J., Maurhofer, M., U., S., G., D. and D., H. (1994) In O'Gara, F., Dowling, D. N. and Boesten, B. (eds.), *Molecular Ecology of Rhizosphere Microorganisms: Biotechnology and the Release of GMO's*. VCH Weinheim, pp. 67-89.
15. 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. *Proc Natl Acad Sci U S A*, **76**, 1648-1652.
16. Kessler, B., de Lorenzo, V. and Timmis, K.N. (1992) A general system to integrate lacZ fusions into the chromosomes of gram-negative eubacteria: regulation of the Pm promoter of the TOL plasmid studied with all controlling elements in monocopy. *Mol Gen Genet*, **233**, 293-301.
17. Kulasekara, H.D., Ventre, I., Kulasekara, B.R., Lazdunski, A., Filloux, A. and Lory, S. (2005) A novel two-component system controls the expression of Pseudomonas aeruginosa fimbrial cup genes. *Mol Microbiol*, **55**, 368-380.
18. Rietsch, A., Vallet-Gely, I., Dove, S.L. and Mekalanos, J.J. (2005) ExsE, a secreted regulator of type III secretion genes in Pseudomonas aeruginosa. *Proc Natl Acad Sci U S A*, **102**, 8006-8011.
19. Hoang, T.T., Kutchma, A.J., Becher, A. and Schweizer, H.P. (2000) Integration-proficient plasmids for Pseudomonas aeruginosa: site-specific integration and use for engineering of reporter and expression strains. *Plasmid*, **43**, 59-72.
20. Becher, A. and Schweizer, H.P. (2000) Integration-proficient Pseudomonas aeruginosa vectors for isolation of single-copy chromosomal lacZ and lux gene fusions. *Biotechniques*, **29**, 948-950, 952.

4.1.3 Additional results

Search for *exlBA* upregulating conditions

Finding conditions that activate *exlBA* expression has been a long-lasting goal in the lab. While we did show that *exlBA* is slightly upregulated in response to calcium depletion in a Vfr-dependent manner (Berry et al., 2018), we still tried to find conditions with bigger impact on the operon expression. Additionally, as we learned during these experiments that the inhibition by ErfA was responsible for the small effect of Vfr in a wild-type strain, we were looking for conditions in which this inhibition would be alleviated. To that aim, I screened numerous growth conditions using the *exlBA::lacZ* fusion and found some potentially interesting leads that still require further investigation.

Among these is the temperature-dependent regulation of *exlBA*. Indeed, I found that *exlBA* seems to be slightly upregulated during growth at 25°C compared to 37°C (Figure 4.1). In light of the documented infections from *exlBA*$^+$ strains which happen in humans and animals, this result was not expected and reflects how little we still understand this secretion system. However, this could correspond to the recently discovered role of ExlBA in environmental species, which would thus need it at lower temperatures, such as during infection of insects or fungi (Basso et al., 2017a; Job et al., 2019).

Another potentially interesting regulatory feature is the upregulation of *exlBA* during anoxic growth. Both *exlBA* and *ergAB* were found to be upregulated during

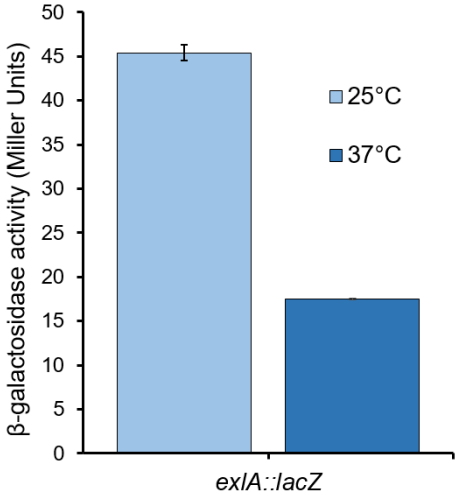


Figure 4.1: The effect of temperature on *exlBA* expression. β-galactosidase activities of IHMA87 strains harboring *exlBA::lacZ* transcriptional fusion. Strains were grown in LB from OD_{600}=0.1, and activities of the fusion were measured after 5 hours of growth. Enzyme activities are represented as mean ± standard deviation (SD) from three independent experiments.

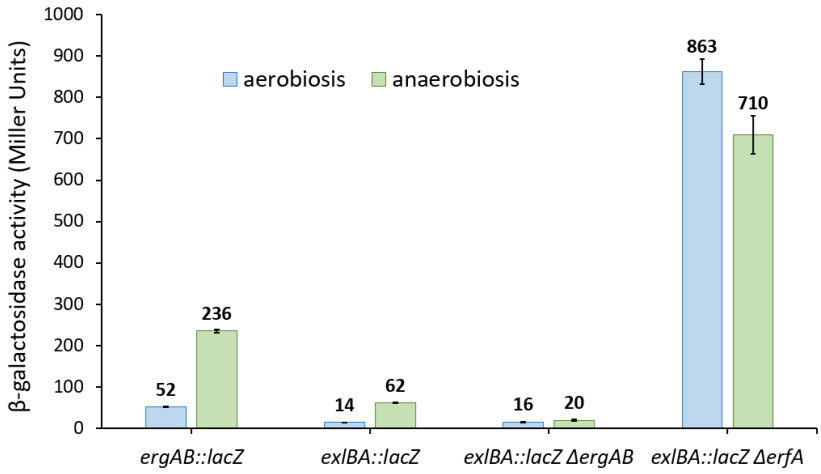


Figure 4.2: The effect of oxygen availability on ErfA activity. β-galactosidase activities of IHMA87 strains harboring *exlBA::lacZ* or *ergAB::lacZ* transcriptional fusion. Strains were grown at 37°C in LB containing 1% KNO_3 from OD_{600}=0.1, and activities of the fusion were measured after 16 hours of growth. Enzyme activities are represented as mean ± standard deviation (SD) from three independent experiments.

anoxic growth in a ErfA-dependent manner (Figure 4.2). Interestingly, this upregulation seemed to be abrogated in a *ergAB* mutant. This could suggest that ErgA and ErgB play a role in ErfA signal sensing or that ErfA senses a product metabolite of ErgA and ErgB, as it is the case for numerous metabolic regulators (Chubukov et al., 2014). Additionally, this upregulation was lost during growth in minimal medium (Figure 4.3), supporting the idea of a metabolite input processed by ErgA and ErgB. In this condition - anoxic growth in a rich medium -, ErfA could sense the corresponding product metabolite and consequently alleviate its inhibition of *ergAB*.

However, I still question these results as the observed upregulation could also be due to global transcriptional changes due to different growth phases or rates between aerobic and anaerobic growth. While the above-mentioned experiments were done on cultures that were all in stationary phases after overnight growth, a global upregulation, even in an *erfA* mutant, could be seen during anaerobiosis when measuring at the same optical density instead (Figure 4.4). Growth-rate dependent global regulation represents an additional layer of regulation that needs to be taken into account when studying individual regulatory circuits (Gerosa et al., 2013; Shahrezaei and Marguerat, 2015). In this case, it is still unclear whether the observed change in gene expression is a consequence of that phenomenon. More work is still required to fully apprehend this potential new regulatory feature, including the concomitant measure of other genes expression to assess and correct

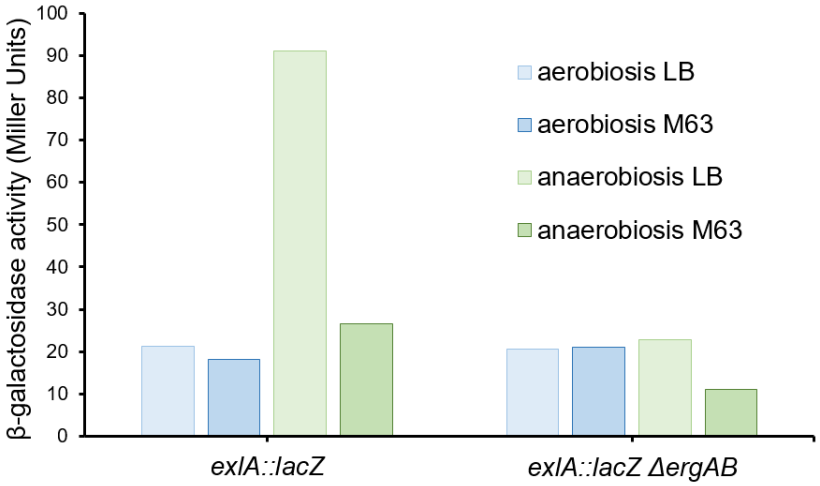


Figure 4.3: The effect of growth medium on *exlBA* anoxic upregulation. β-galactosidase activities of IHMA87 strains harboring *exlBA::lacZ* transcriptional fusion. Strains were grown at 37°C in LB or M63 medium containing 1% KNO_3 from OD_{600}=0.1, and activities of the fusion were measured after 16 hours of growth.

for global transcription regulation.

To study *P. aeruginosa* virulence, we often use simple infection models such as cytotoxicity assays on human epithelial cells (Trouillon et al., 2020a). However, depending on infection models and cell types, cell culture and infection can be done in many different ways and especially in many different media (Arora, 2013). I found that *exlBA* expression varies greatly when *P. aeruginosa* is grown in different

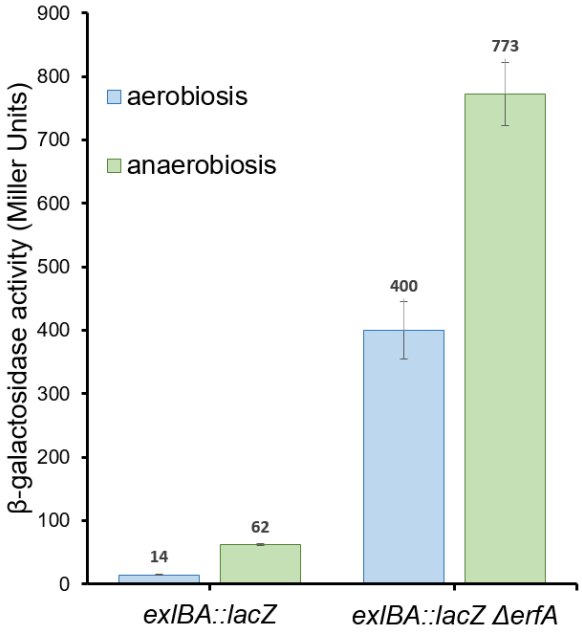


Figure 4.4: The effect of cell density on *exlBA* anoxic upregulation. β-galactosidase activities of IHMA87 strains harboring *exlBA::lacZ* transcriptional fusion. Strains were grown in LB containing 1% KNO_3 from OD_{600}=0.1, and activities of the fusion were measured at OD_{600}=1, which corresponded to 4 and 24 hours of aerobic and anaerobic growth, respectively. Enzyme activities are represented as mean ± standard deviation (SD) from three independent experiments.

cell culture media (Figure 4.5). This variation would probably result in very different cytotoxicity phenotypes and thus needs to be taken into account when looking at ExlBA-dependent virulence. Importantly, several media induced a much higher *exlBA* expression than what is measured during growth in LB medium, which we usually use for monitoring bacterial phenotypes. This could suggest that some of these media contain signal molecules involved in *exlBA* regulation, but this result could also come from growth-rate dependent global regulation, as explained above. Comparative gene expression measurements with control genes is now required to conclude on this question.

Even before my arrival in the lab, different attempts were conducted to try to monitor *exlBA* expression during infection, as infection would be the most obvious condition in which *exlBA* could be upregulated. This was attempted by RTqPCR or the use of a *gfp* transcriptional fusion during cell infection, but none of these methods yielded sufficient signal for a proper quantification of *exlBA* transcription. With the same goal in mind, I tried to use the *exlBA::lacZ* fusion during infection of *G. mellonella* larvae. Briefly, infected larvae were homogenized at different time points and β-galactosidase activity measured after rough washing of bacteria. However, we usually infect larvae with a very small number of bacteria (5 - 20 bacteria), and even when increasing this amount up to 10,000-fold I could not measure any detectable β-galactosidase activity from homogenized larvae. To overcome this problem, I tried using a mouse acute lung infection model. This allowed infection with 15 million bacteria, which was enough to obtain detectable β-galactosidase activity in homogenized mouse lungs. The inhalation resulted in 7 million bacteria detected

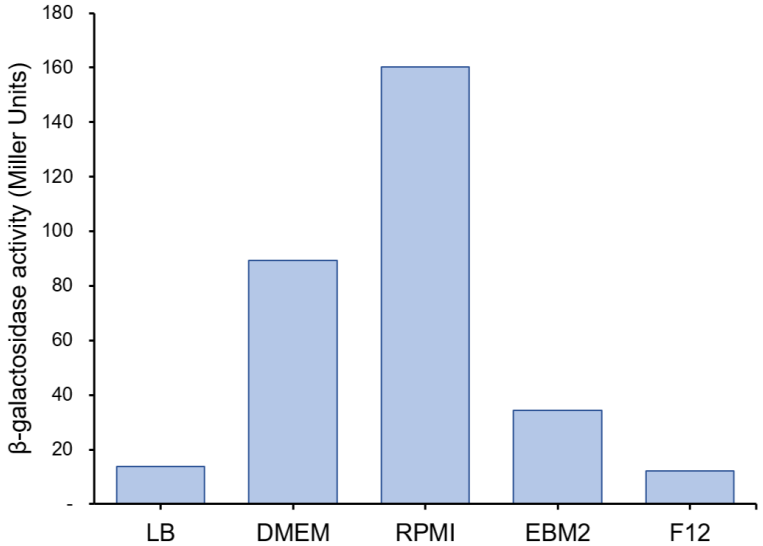


Figure 4.5: The effect of cell culture media on *exlBA* expression. β-galactosidase activities of IHMA87 harboring *exlBA::lacZ* transcriptional fusion. Activities were measured after overnight growth in the different media. Growth in LB serves as reference.

in the lungs right after infection (t0), and 13 million after two hours of infection, reflecting either bacterial growth or a delay in reaching the lungs. The fusion activity measurement showed a decrease in *exlBA* expression after two hours when compared to t0 (Figure 4.6). Here again, this result was quite surprising, but it could suggest that ExlBA is needed at other time points of infection than the was we used. More work would still be needed in order to get a better understanding of *exlBA* dynamic patterns of expression during the course of infection, which could help to better apprehend the role and regulation of this virulence factor.

ErfA expression or activity regulation

Hypothetically, if ExlBA is to be needed in quantity in a specific condition and the involved regulation happens at the transcription level, one obvious way to induce upregulation would be by preventing the binding of ErfA to *exlBA* promoter. There could be two different ways of doing so: reducing ErfA quantity or inhibiting its DNA-binding activity. In order to cover these two possibilities, I conducted two approaches.

To investigate a potential regulation of ErfA protein quantity, I first got interested in transcriptional regulation. To identify potential regulators of *erfA*, I created an *erfA::lacZ* transcriptional fusion in the chromosome and aimed at building and screening a transposon mutant library in that strain as I did for *exlBA*. However I was not able to perform transposon mutagenesis in this strain. The ability of IHMA87 to use conjugation has always been scarce and unstable, and even by

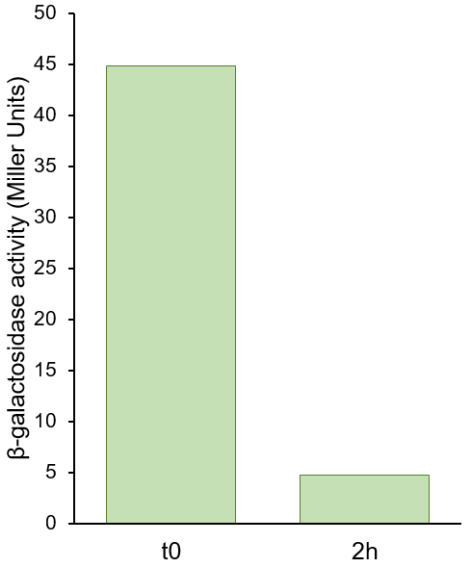


Figure 4.6: *exlBA* expression during acute lung infection in mice. β-galactosidase activities of IHMA87 harboring *exlBA::lacZ* transcriptional fusion. After infection by inhalation of C57BL/6 mice, activities were measured from 100 µl of homogenized lungs and normalized to activities measured in lungs infected with WT strains and to bacterial count as measured by concomitant plating and CFU counting.

testing three different transposon-containing conjugative plasmids and many different conjugation protocols, I was never able to obtain enough mutants for a proper screen. Secondly, there could be several possible means of inhibition of ErfA activity, including protein-protein interaction or binding of a small molecule to the Cupin domain of ErfA which could be a signal-sensor domain. To delve into the protein hypothesis, we obtained a library of bacterial two-hybrid plasmids from a collaborator to screen protein-protein interactions (Christophe Bordi, Marseille - France). This library was built with the fragmented gDNA of PAO1 inserted in the three possible open reading frames, fused to one of two adenylate cyclase subunit, and has already been successfully used to find interactions of proteins of interest with nearly all PAO1 proteins (Houot et al., 2012). I transformed the plasmids in a strain expressing an ErfA fused to the second adenylate cyclase subunit and containing the *lac* operon which contains a cAMP-dependent promoter and started to screen the obtained library on X-Gal plates. However, I screened over 40,000 mutants and did not find any with upregulation of the fusion, suggesting that no protein interacts with ErfA. Using this two-hybrid system, we were still able to show that ErfA interacts with itself (Figure 4.7), corroborating its predicted quaternary structure as a dimer.

The second hypothesis concerning the binding of a signal molecule to ErfA is supported by the prediction that the Cupin domain of ErfA has a putative signal-sensor function, as further discussed below (Trouillon et al., 2020b). We thus wanted to get a better understanding of the structure-function relation of the different parts of the ErfA protein, such as dimeric interaction and role in DNA-binding as well as the role of the Cupin domain. To that aim I tried to overproduce each of the two domains individually to look at the effect on different genetic background. But I was never able to get any detectable amount of both domains that way, and thus

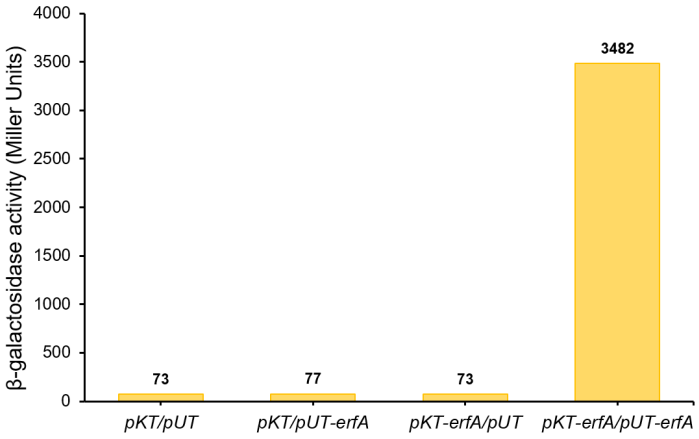


Figure 4.7: Formation of ErfA dimers in bacterial two-hybrid assay. β-galactosidase activities of *E. coli* DHM1 with different combination of pKT25 and pUT18c expression vectors containing the *erfA* gene or not. Measures were done after overnight growth at 37°C.

any phenotype, probably due to instability of the single domains.

Search for ErgAB function

While ErgA and ErgB are the most conserved targets of ErfA, we still have no clear idea of their function. Their sequence-based predicted functions are class II aldolase and dihydrodipicolinate synthase, respectively, but that was never confirmed. Even more, two independent studies tested ErgB activity and concluded that it does not exhibit any dihydrodipicolinate synthase activity (Schnell et al., 2012; Impey et al., 2020). Additionally, we showed that even though they are co-regulated with *exlBA*, no functional link was found between ErgAB and ExlBA (Trouillon et al., 2020a). In

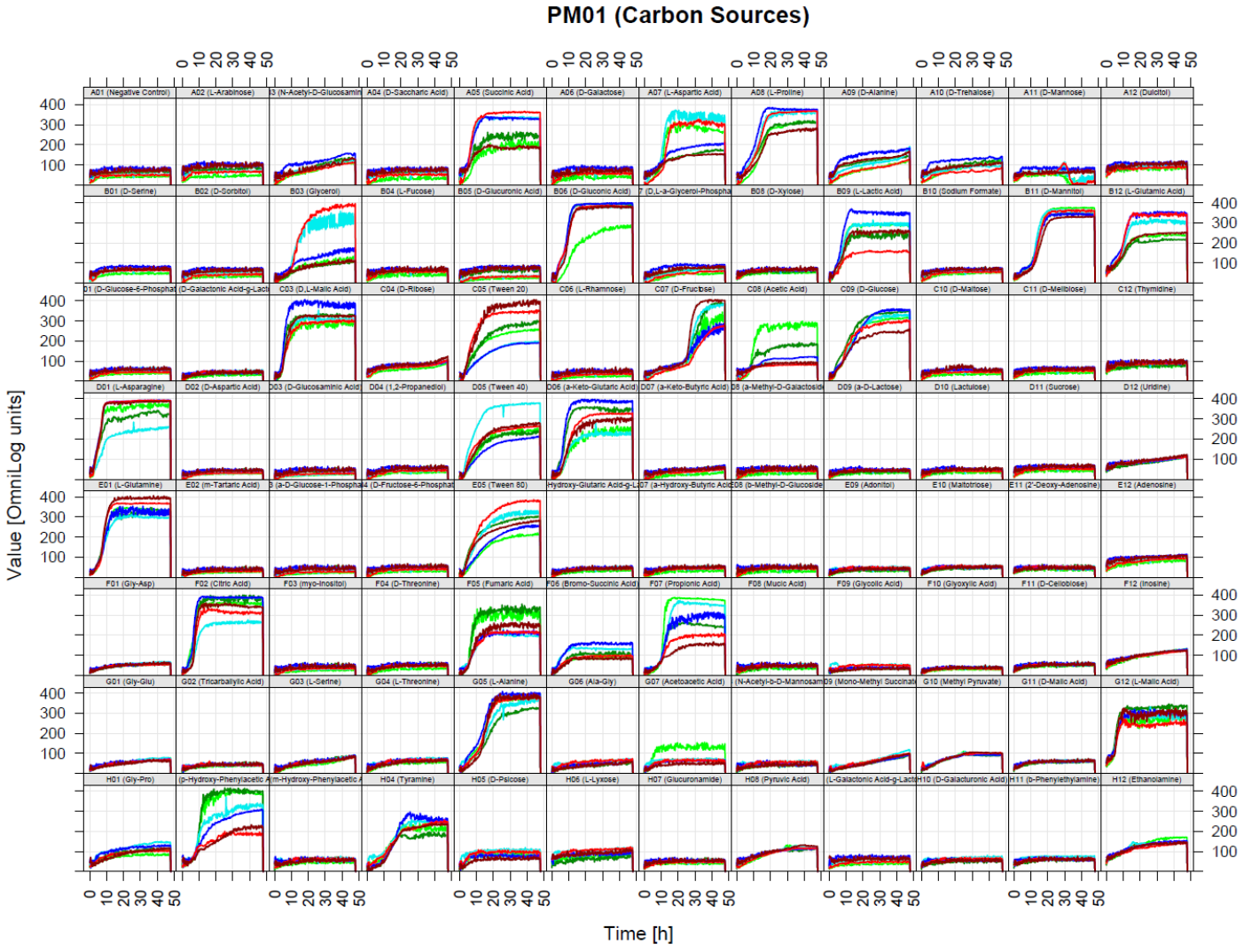


Figure 4.8: Growth fitness screen to assess ErgAB function. Biolog screen results on the PM01 plate using wild-type (reds), *erfA* (blues) and *erfA-ergAB* (greens) mutant strains in duplicates.

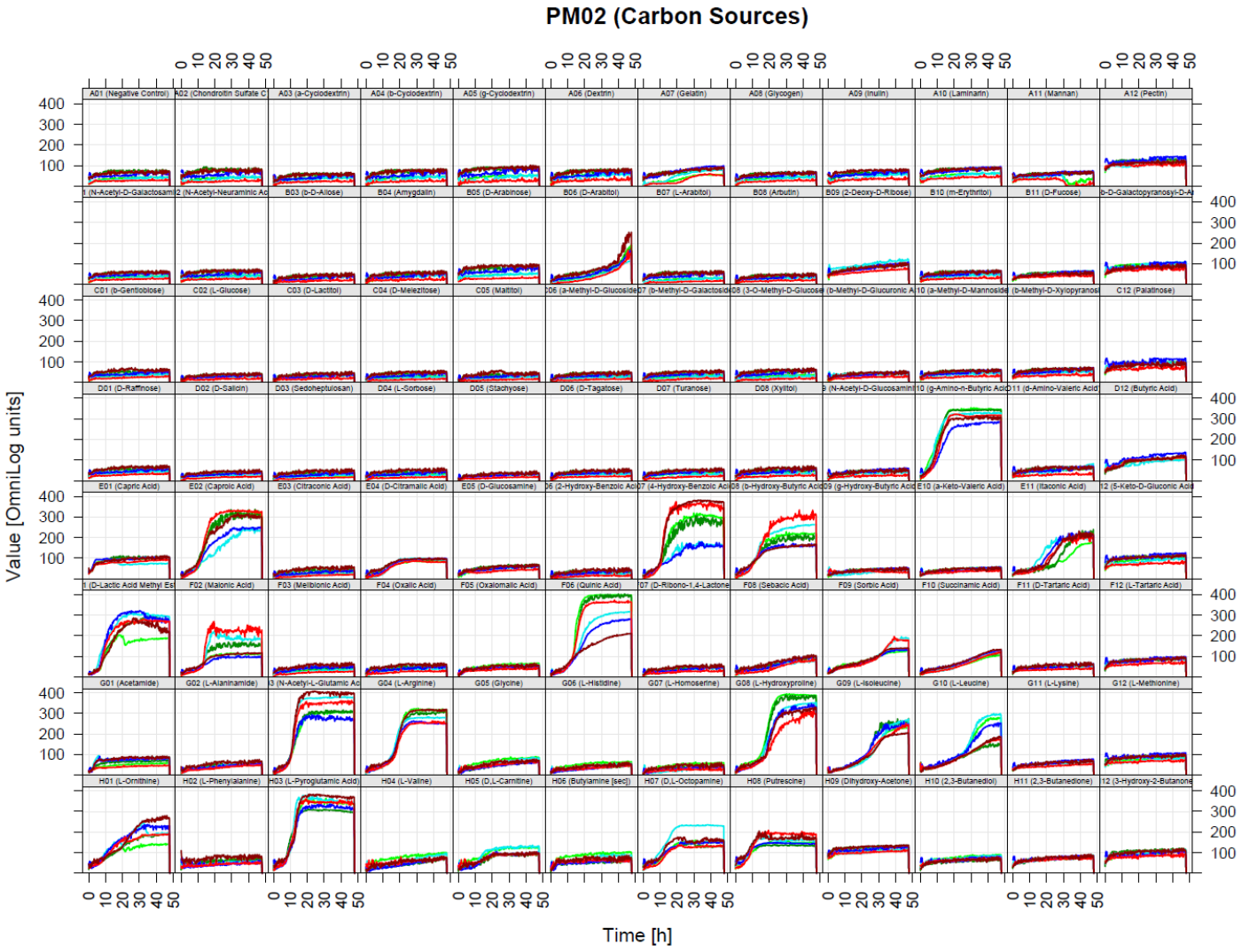


Figure 4.9: Growth fitness screen to assess ErgAB function. Biolog screen results on the PM02 plate using wild-type (reds), *erfA* (blues) and *erfA-ergAB* (greens) mutant strains in duplicates.

an attempt to narrow down ErgAB function, we screened for fitness changes during grown on 192 different carbon sources using wild-type, *erfA* and *erfA-ergAB* mutant strains (Figure 4.8 & 4.9). This screen was done using an Omnilog instrument (Biolog) which measures cell respiration (NADH production) using a tetrazolium dye reduced during respiration, allowing the monitoring of metabolic activity and growth over time. As a first screen of *ergAB* metabolic function we decided to screen fitness on different carbon sources (Biolog plates PM01 and PM02); this experiment was done in collaboration with the team of Benoît Cournoyer in Lyon. Several carbon sources induced small differences in growth between the strains but there was no major fitness differences. Importantly, no growth condition induced opposite

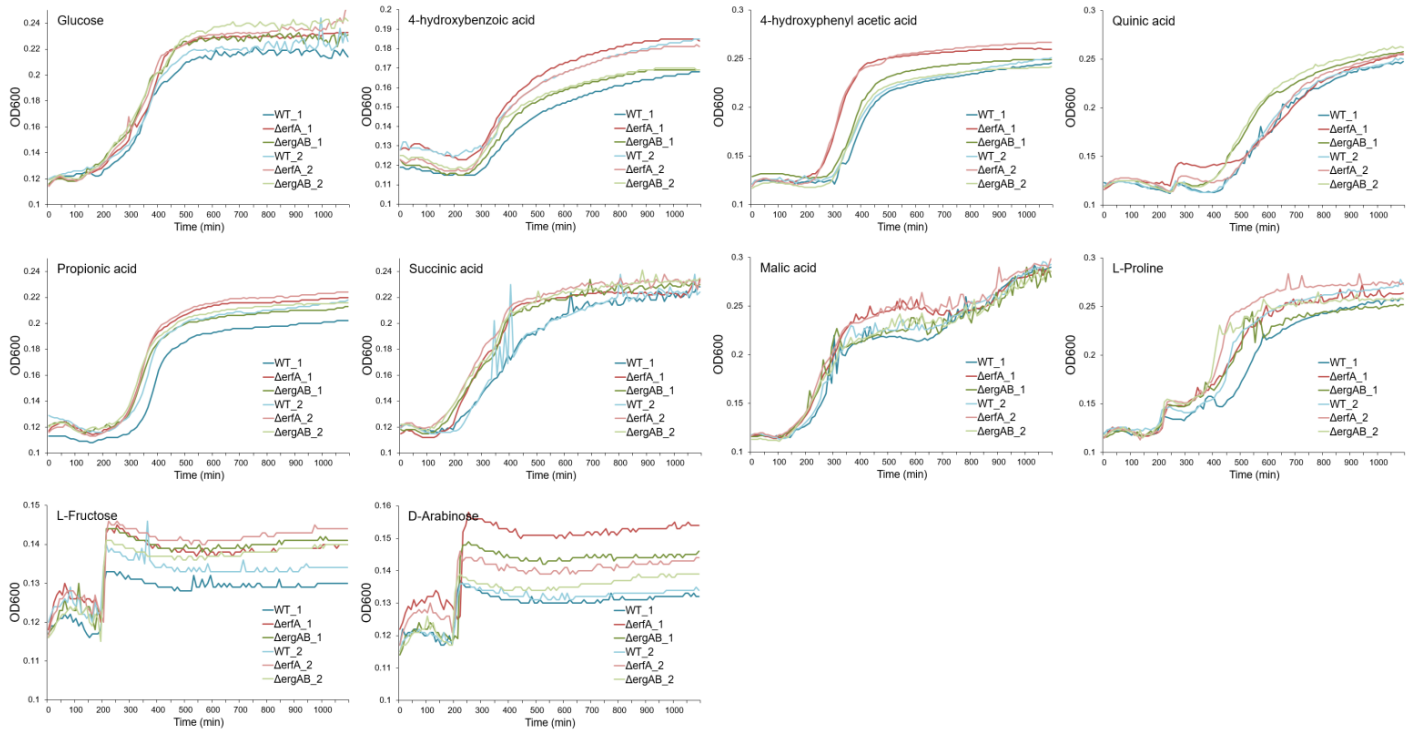


Figure 4.10: Growth fitness screen on 10 carbon sources. Growth curves of IHMA87 wild-type, *erfA* and *ergAB* mutant strains in M9 medium containing 0.4% of the indicated carbon source in duplicates. Growth on glucose serves as reference.

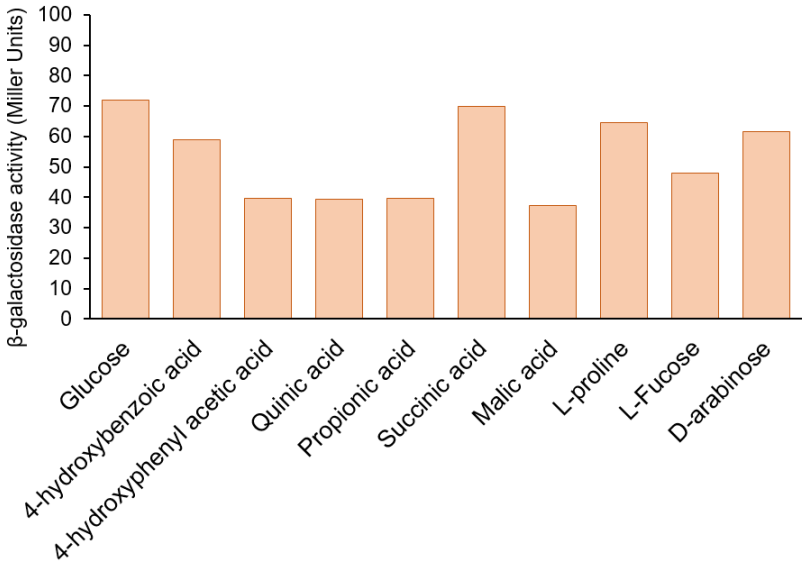


Figure 4.11: *ergAB* expression during growth on 10 carbon sources. β-galactosidase activities of IHMA87 strains harboring *ergAB::lacZ* transcriptional fusion. Bacteria were grown in M9 medium containing 0.4% of the indicated carbon source from OD_{600}=0.1, and activities of the fusion were measured after 4 hours of growth. Growth on glucose serves as reference.

growth change between the *erfA* and the *erfA-ergAB* mutants, as what would be expected for a ErgAB-dependent observation. To make sure of that result, I still selected 9 carbon sources that seemed to have small effects and tried to confirm these results during growth in minimal medium. However, I could not confirm any strong growth difference or any difference that seemed to be dependent of *ergAB* presence (Figure 4.10). Additionally, I checked *ergAB* expression in these conditions using a *ergAB::lacZ* transcriptional fusion and could not detect any major change in expression of the operon (Figure 4.11). Consequently, the function of ErgA and ErgB is still completely unknown to us.

Additional Materials & Methods

Anoxic growth. For anaerobic cultures, 1% KNO_3 and 0.05% of antifoam (Sigma) were added to LB or M9 media. Anaerobic cultures were performed in Hungate tubes containing 12 ml of medium which was deoxygenated by argon bubbling for 25 min before autoclaving as previously described (Pelosi et al., 2019). Hungate tubes were inoculated through the septum with 100 µl of precultures taken with disposable syringes and needles from closed Eppendorf tubes filled to the top grown with no agitation at 37°C overnight.

β-galactosidase activity measurement. β-galactosidase activities were measured as previously described (Trouillon et al., 2020a).

Mouse infection. Pathogen-free C57Bl/6 (WT) mice were infected by inhalation of 30 µl of a bacterial suspension containing 15 million bacteria as previously described (Bouillot et al., 2020). Mice were euthanized at the different time points and lungs were isolated and homogenized in PBS using a Polytron instrument (Philippe Huber, Grenoble - France).

Bacterial two-hybrid assay. The system used is based on *E. coli* cAMP signaling cascade (Karimova et al., 2000). The *erfA* gene of IHMA87 was cloned in pKT25 and pUT18c at the 3′ end of the genes encoding the two fragments of adenylate cyclase. The *E. coli* DHM1 cells, deficient in adenylate cyclase, were co-transformed with pKT25 and pUT18c derivatives. β-galactosidase activities were measured on cells grown overnight at 37°C.

4.1.4 TPS regulation

Julian Trouillon✉, Ina Attrée & Sylvie Elsen✉

Manuscript in preparation

Our lab has been working on all aspects of ExlBA over the last few years, which yielded an overall strong expertise on TPS systems and TPS regulation. While the functions and molecular mechanisms of TPS systems are reviewed every few years, it seems to us that the study of their regulation is often done individually in each species and not much effort has been yet made to try to bring this knowledge together to find patterns and differences that could help the field. We thus leveraged the 2020 covid-19 lockdown to write a review on TPS regulation. We found interesting common regulatory features between TPS systems and species, notably for TPS secreting cytolysins, including iron- or temperature-dependent regulations. Overall, this comparative analysis allowed the identification of common and specific mechanisms which we hope will help and guide future studies in the field. This manuscript is still in early stages of preparation.

How do bacteria cope with all their Two-Partner Secretion systems?

Julian Trouillon[1#], Ina Attrée[1], Sylvie Elsen[1#]

[1]Université Grenoble Alpes, CNRS ERL5261, CEA BIG-BCI, INSERM UMR1036, Grenoble, France

Address correspondence to julian.trouillon@cea.fr; sylvie.elsen@cea.fr

Abstract

Two-partner secretion (TPS) systems, also known as Type Vb secretion systems, allow the translocation of effector protein across the outer membrane of Gram-negative bacteria. Several subtypes of TPS systems exist, each secreting different families of proteins notably including cytolysins or adhesins. Often important for host infection or inter-bacterial competition, TPS systems have been found to be regulated by different mechanisms allowing their timely expression. Here, we review the ground of knowledge on TPS systems regulation and highlight both specific and common regulatory mechanisms across all TPS subfamilies. Several regulatory cues were found as common determinants of expression of TPS systems in different subfamilies, even across relatively distant species, revealing interesting conserved function-regulation relationships.

Introduction

The coordinate expression of membrane components is crucial for bacterial niche-specific survival and adaptation. Bacterial secretion systems have evolved to constitute a wide family of molecular machineries allowing the export of proteins or DNA (Green & Mecsas, 2016). In Gram-negative bacteria, several types of secretion systems allow export across the two membranes, each exhibiting a unique architecture, mechanism of secretion and nature of secreted cargo (for a review, see (Costa *et al.*, 2015)). Some bacterial species, such as *Pseudomonas aeruginosa*, potentially synthetize more than a dozen different secretion systems (Bleves *et al.*, 2010, Filloux, 2011), most of them being tightly regulated and assembled only in response to specific external signals.

Two-Partner Secretion (TPS) systems are simple export apparatus also referred to as subtype b of Type V Secretion Systems (T5SSs) (for a general review, see (Guerin *et al.*, 2017)). Members of this family include a secreted effector, TpsA, and its cognate outer membrane transporter TpsB, which can be fused into one dual-function protein in the family of autotransporters (ATs). Some TpsB proteins might also export more than one cargo proteins. T5SSs effectors are of major importance for bacterial physiology as they drive important phenotypes such as cytotoxicity, proteolysis, host-cell adhesion, biofilm formation and microbial competition. The two TPS systems ShlBA from *Serratia marcescens* and FhaBC from *Bordetella pertussis* were the first studied and were found to be critical actors of pathogenesis through toxin secretion (Poole *et al.*, 1988, Domenighini *et al.*, 1990). Since then, numerous other TPS systems have been characterized, often owing to their important role in pathogenesis. More recently, the Contact-Dependent Growth Inhibition (CDI) systems were

discovered as mediating processes such as cell-cell interactions, bacterial competition and biofilm formation (Ruhe *et al.*, 2013), and even virulence towards mammalian hosts (Allen *et al.*, 2020).

TPS systems, much like most secretion systems, are usually prone to be exchanged by horizontal gene transfer (HGT) (Rojas *et al.*, 2002), which in turn brings diversification of both function and regulation. As of today, more than 1,000 TPS systems have been identified in Gram-negative bacteria, especially in Proteobacteria, Cyanobacteria and Negativicutes where they are abundant, representing the third most common secretion system family behind ATs and type I secretion systems (Abby *et al.*, 2016). While the work on mechanisms of secretion and effector functions of TPS systems has been thoroughly reviewed (Leo *et al.*, 2012, Jacob-Dubuisson *et al.*, 2013, Fan *et al.*, 2016, Guerin *et al.*, 2017, Meuskens *et al.*, 2019), similar up-to date analyses are lacking concerning their regulation.

In this review, we aimed to assess the existence of common and specific regulatory mechanisms driving transcription of the most characterized TPS systems. Host or environmental signals (temperature, iron, inorganic phosphate) sensed through different mechanisms (signal-sensing transcription factors, Two-Component regulatory Systems) and bacterial physiological state (Lrp, secondary messengers cAMP and c-di-GMP, (p)ppGpp) are as many sensed information that can govern TPS system expression. In pathogenic bacteria, these TPS systems are frequently controlled by regulatory pathways that also drive the expression of other virulence factors. In addition, some TPS systems also have evolved specific regulatory mechanisms to adapt their synthesis to the physiology of different species.

Iron, temperature and FlhD are common regulators of cytolysins

A large family of TPS cargo proteins includes cytolysins/hemolysins which have the ability to disrupt membranes of host cells (Skals & Praetorius, 2013). Accordingly, conditions encountered within the host, as iron limitation and temperature changes, are common features that stimulate their expression (Table 1). Iron is indispensable for bacterial growth and hosts limit iron availability as part of their innate defense against invading bacteria (Weinberg, 2009). Therefore most bacterial pathogens (but not all, as seen below) sense iron-limiting conditions as a probable signal for arrival into host and respond by upregulating iron-acquisition systems and virulence genes, such as those encoding cytolysins. Temperature is also a common cue that modulates virulence factor expression (Konkel & Tilly, 2000). Different temperature changes can be sensed and indicate on the nature of the host, i.e. cold- or warm-blooded host, which in turn drives different regulatory responses. Finally, a complex regulatory interplay between flagellar and TPS gene expression exists in many bacterial pathogens, which relies on the master regulator FlhDC standing at the top of the flagellar regulatory cascade: this leads to coordinate synthesis of cytotoxic factors and flagellum, which also contributes to virulence through motility, chemotaxis, adhesion to and invasion of host surfaces .

Iron limitation has been shown to control the expression of two cytolysins in entomopathogenic bacteria, PhlA in *Photorhabdus luminescens* and XhlA in *Xenorhabdus nematophila. P. luminescens* lives symbiotically in the gut of entomopathogenic nematodes and kills insects when released in their hemolymph. Although the role of PhlA in *P. luminescens* lifestyle is still unknown, the operon *phlBA* is upregulated upon iron restriction, in agreement with a higher hemolytic activity measured in this condition (Brillard *et al.*, 2002). Several

Table 1 : common signals controlling TPS synthesis and their underlying mechanisms

Cue	Fonction	TPS	Organism	Mechanism	Reference
Iron	Cytolysin	PhlBA	*Photorhabdus luminescens*	Fur ? (putative binding-sites (BSs))	(Brillard *et al.*, 2002)
		XhlBA	*Xenorhabdus nematophila*	? FlhDC cascade	(Cowles & Goodrich-Blair, 2005) (Jubelin *et al.*, 2011)
		ShlBA	*Serratia marcescens*	Fur ? (putative BSs) Rss TCS	(Poole & Braun, 1988b) (Lin *et al.*, 2016)
		HpmBA	*Proteus mirabilis*	? (putative Fur BS)	(Uphoff & Welch, 1990)
		EthBA	*Edwardsiella tarda*	Fur	(Weinberg, 2009)
	Adhesin	EtpBA	*ETEC*	?	(Haines *et al.*, 2015)
	Iron-acquisition	HxuBC	*Haemophilus influenzae*	Fur	(Harrison *et al.*, 2013)
T°C	Cytolysin	ShlBA	*S. marcescens*	Rss TCS (via FlhDC)	(Lin *et al.*, 2016)
		EthBA	*E. tarda*	nucleoid protein Hha_{Et}	(Wang *et al.*, 2010)
	Adhesin	FhaBC	*Bordetella pertussis*	BvgSA TCS	(Cotter & Miller, 1994)
	Iron-acquisition	HxuBC	*H. influenzae*	?	(Alvarez-Estrada *et al.*, 2018)
FlhDC	Cytolysin	XhlBA	*X. nematophila*	FliZ (bimodale)	(Jubelin *et al.*, 2013)
		ShlBA	*S. marcescens*	Sigma FliA	(Di Venanzio *et al.*, 2014)
		HpmBA	*P. mirabilis*	?	(Fraser *et al.*, 2002)

putative binding sites for the Ferric Uptake Regulatory (Fur) protein, which controls iron homeostasis in many bacteria (Fillat, 2014), have been found in the *phlBA* promoter, suggesting its involvement in this regulation. The second entomopathogen, *X. nematophila*, kills insects either directly from the hemolymph or in cooperation with its symbiotic host nematode (for review (Forst *et al.*, 1997)). Among the virulence factors required for overcoming the insect immune system are at least three hemolysins, including the TpsA-like protein XhlA. The XhlBA (*X. nematophila* hemolysin) system has been reported to be required for full virulence towards *Manduca sexta* larvae, but not in mutualistic colonization of the nematodes, providing evidence that the hemolysin is active during insect infection (Cowles &

Goodrich-Blair, 2005). Probably secreted by XhlB, XhlA exhibits a surface-associated hemolytic activity that can also target granulocytes and plasmatocytes. Expression of *xhlBA* operon is higher in stationary phase and modulated by iron availability in a Fur-independent manner (Jubelin *et al.*, 2011). Strikingly, iron seems to exert two opposite effects on *xhlBA* operon depending on the growth phase; its limitation was first reported as upregulating the *xhlA* expression in exponentially growing cells but a repressive effect exerted through the flagellar regulatory cascade was observed in stationary phase (Cowles & Goodrich-Blair, 2005, Jubelin *et al.*, 2011). At the top of this cascade lies the master flagellar regulator $FlhD_4C_2$ which is required for full virulence by controlling both directly and indirectly flagellar genes and non-flagellar virulence genes (Givaudan & Lanois, 2000, Givaudan & Lanois, 2017), as observed in many other bacteria (see below). The master regulator controls positively XhlBA synthesis as mutation in *flhD* led to a reduction by half of the level of *xhlA* transcripts (Cowles & Goodrich-Blair, 2005). Its large regulon relies in part on its control over the *fliAZ* expression: this operon encodes the σ^{28} factor FliA and the regulator FliZ which in turn directly regulates expression of *flhDC* (positive feedback loop) and *xhlAB* among other genes (Lanois *et al.*, 2008, Givaudan & Lanois, 2017). A link between iron availability and flagellar regulatory cascade was found when both *flhDC* and *fliAZ* operons were shown downregulated during iron limitation, although not targeted by the Fur repressor (Jubelin *et al.*, 2011). Expression of the FliAZ regulon is thus iron-dependent and restricted to insect cadaver during infection. Production of hemolysins and other secreted products triggered in iron-rich sites of infection, such as in larvae cadavers, would facilitate dissemination of the bacteria and degradation of the tissues (Jubelin *et al.*, 2011). Besides integrating iron signals, FliZ was also reported to be a global regulator controlling positively or negatively over 278 genes and playing a role in stochastic phenotypic variation in *X*. *nematophila* (Jubelin *et al.*, 2013). Indeed, the expression of FliZ

target genes, including *xhlA*, was shown to be bimodal, leading to phenotypic heterogeneity in genetically identical bacteria: FliZ acts as a rheostat, its cellular amount controlling virulence factor expression and motility that might be decisive for bacterial survival in host (Jubelin *et al.*, 2013). Other regulators were found to control *xhlBA* expression, such as the global transcriptional regulator Leucine-responsive regulatory protein Lrp that directly binds to the *xhlBA* promoter *in vitro*, activating the operon expression (Cowles & Goodrich-Blair, 2005, Cowles *et al.*, 2007). Lrp is required for full virulence and coordinately regulates diverse metabolic pathways, antibiotic production, protease activity, motility, as well as the *l*ysR-like homolog *A* (*lrhA*) gene, which encodes a regulator controlling the expression of virulence (Cowles *et al.*, 2007, Richards *et al.*, 2008, Richards & Goodrich-Blair, 2009). It is to note that both Lrp and LrhA have a positive effect on transcription of the flagellar regulator FlhDC (Richards *et al.*, 2008), thus indirectly affecting *xhlBA* expression.

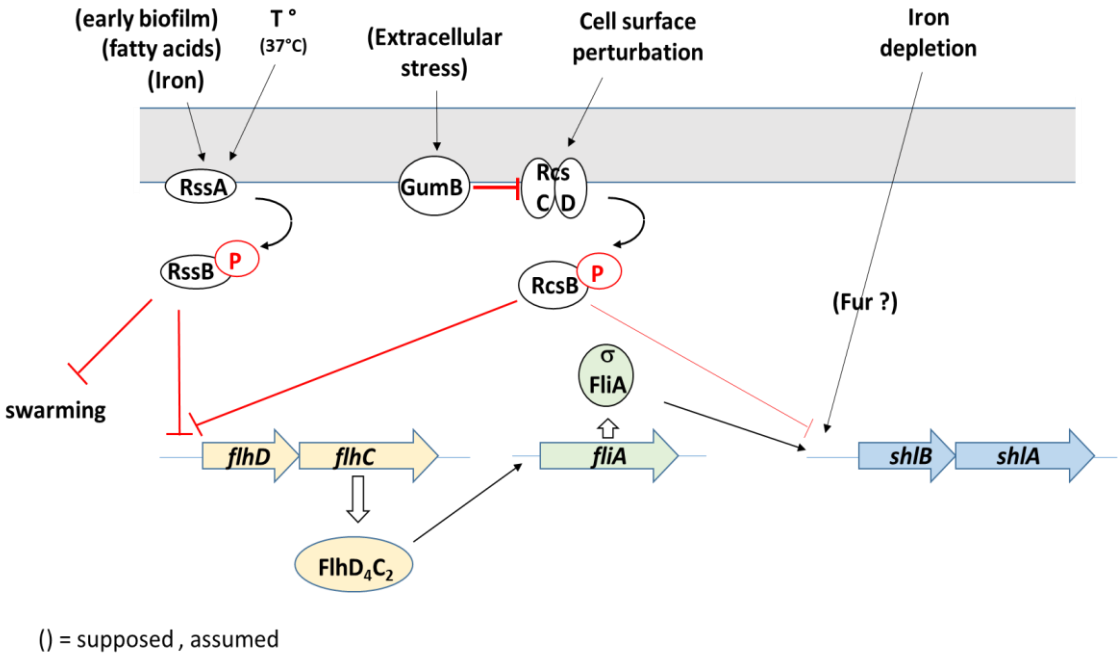


Figure 1: The regulation of *shlBA* in *Serratia marcescens.*

Serratia marcescens and *Proteus mirabilis*, two important opportunist Gram-negative bacteria belonging to the Enterobacteriaceae family, respectively encode the ShlBA and HpmBA TPS systems that are both regulated by iron and FlhDC. The ShlA hemolysin is a major virulence factor of *S. marcescens* and iron limitation was the first element shown to stimulate *shlBA* expression (Poole & Braun, 1988b). ShlA promotes lysis of red blood cells and thus the release of iron-containing hemoglobin, which probably facilitates iron acquisition by the pathogen. As for other cytolysins, Fur-binding sites are present in the *shlBA* promoter, inferring a direct negative control (Poole & Braun, 1988b). Another iron sensing mechanism relies on the RssAB two-component system (TCS) that negatively controls *shlBA* expression

(Lin *et al.*, 2010). Environmental Fe^{3+} ferric iron directly binds and stimulates autophosphorylation of the RssA sensor kinase and consequently induces the phosphorylation of the cognate response regulator RssB (Lin *et al.*, 2016). The RssAB system was shown to repress *shlAB* expression and swarming motility at 37°C, favoring transition toward a biofilm lifestyle (Lin *et al.*, 2010). The negative effect of temperature is in agreement with a swarming occurring at 30°C and not at 37°C (Lai *et al.*, 2005), and a 10-fold reduced hemolytic activity observed at 37°C *vs*. 30°C during exponential growth (Poole & Braun, 1988a). *S. marcescens* can infect both vertebrate and invertebrate hosts (Grimont & Grimont, 1978) and is only an opportunistic pathogen for humans, which might explain why optimal temperature for the expression of some virulence factors is not that of the human body. Through their control of RssAB signaling, other factors such as saturated fatty acids (Lai *et al.*, 2005), biofilm lifestyle (Tsai *et al.*, 2011) and iron availability (see above, (Lin *et al.*, 2010)) might also influence ShlA synthesis. RssAB effect on *shlBA* expression is not direct but carried out through $FlhD_4C_2$: indeed, the master flagellar regulator activates the synthesis of the sigma factor FliA that in turn binds to *shlBA* promoter and initiates the transcription (Lin *et al.*, 2010, Di Venanzio *et*

al., 2014). Therefore, activation of RssB leads to reduced levels of *shlBA* by inhibition of *flhDC* expression and consequent *fliA* downregulation. In agreement with this mechanism, *flhDC* expression is inhibited when temperature increases from 30°C to 37°C (Liu *et al.*, 2000). FlhDC synthesis is also under control of an envelope stress response through the Rcs phosphorelay (Di Venanzio *et al.*, 2014), a complex signal transduction system comprising inner membrane sensor proteins RcsC and RscD and the response regulator RcsB. Phosphorylated RcsB binds to *flhDC* and *shlBA* promoters, thus inhibiting *shlBA* expression both indirectly and directly (Di Venanzio *et al.*, 2014). Rcs signaling is triggered by alteration of cell surface, such as impaired synthesis of different components of the outer membrane (LPS or the glycolipid Enterobacterial Common Antigen "ECA") and high osmolarity (Castelli & Vescovi, 2011). Accordingly, alteration of ECA synthesis inhibits *shlBA* expression as well as hemolytic activity (Di Venanzio *et al.*, 2014). Recently, the Rcs phosphorelay was shown to control the ShlA-dependent induction of blebs in corneal epithelial cells, property shared with the cytolysin HpmA of *P. mirabilis* (Brothers *et al.*, 2019). The "blebbing" phenotype required the gene *gumB* encoding a member of the IgaA family of inner membrane proteins involved in bacterial stress response and controlling diverse phenotypes (Stella *et al.*, 2018). Inactivation of *gumB* leads to a strong down regulation of *shlBA* transcript through a mechanism that is Rcs-dependent (Brothers *et al.*, 2019). In agreement, GumB was reported to be required for survival and proliferation of *S. marcescens* in murine phagocytic cells and in *Galleria mellonella* larvae, probably due for a large part to its positive effect on ShlBA synthesis.

P. mirabilis is a frequent pathogen of the urinary tract and provokes infections in individuals with urinary tract abnormality or related with vesicular catheters. Among different virulence determinants, HpmA is a potent cytolysin able to mediate lysis of a wide range of cell types (for a review, (Armbruster *et al.*, 2018)). Both secretion and activation of HpmA

require the outer-membrane transporter HpmB (Uphoff & Welch, 1990). *hpmBA* expression is driven by a σ^{70} promoter that is upregulated during swarming (Fraser *et al.*, 2002) and several endonucleolytic cleavages occur along the unstable *hpmBA* mRNA transcript leading to an excess of stable *hpmA* mRNAs. This unbalanced mRNA quantities of the two-gene transcripts probably control the ratio of the two protein partners, producing more secreted cargos than transporters (Fraser *et al.*, 2002). Control by iron limitation of HpmBA synthesis is suggested by the presence of a putative Fur-binding site in the *hpmBA* promoter region (Uphoff & Welch, 1990). Both synthesis and secretion of HpmA are increased in hyper-flagellated swarming cells compared to vegetative cells, accompanied by a 20-fold higher hemolytic activity during swarming, hemolysin and flagellum synthesis gene expression being coordinated in swarming cells (Allison *et al.*, 1992, Gygi *et al.*, 1995). $FlhD_2C_2$, Lrp and UmoB are three regulators that control positively *hmpBA* expression and flagellar biosynthesis during swarming, even if only Lrp was shown to directly control *hpmBA* expression by binding to the operon promoter, probably at multiple sites (Fraser *et al.*, 2002). Like GumB of *S. marcescens*, UmoB is a membrane-bound positive regulator of the FlhDC master operon that belongs to the IgaA family of protein: it inhibits the Rcs phosphorelay that controls negatively fhlDC expression upon surface contact. However UmoB cannot complement functionally GumB, unlike other IgaA proteins (Morgenstein & Rather, 2012, Brothers *et al.*, 2019). Initially, it was shown that the inner membrane component of the flagellum export machine FlhA positively regulates flagellin-encoding *fliC* and *hpmA* in swarming cells (Gygi *et al.*, 1995). The reduced *fliC* and *hmpA* transcripts in the *flhA* mutant might result from the negative feedback loop from a defect in the membrane assembly of the flagellum on the *flhDC* expression (Dufour *et al.*, 1998).

The gram-negative opportunistic pathogen *Edwardsiella tarda* has a broad host range, causing hemorrhagic septicemia in fish and both gastro and extraintestinal infections in humans. Its virulence mainly relies on the EthA hemolysin as well as Type III (T3SS) and Type VI (T6SS) secretion systems. The expression of *ethB* was found upregulated more than 8-fold during infection of Japanese flounder (in liver and kidney) compared to *in vitro* growth conditions (Wang *et al.*, 2009). As many hemolysin-encoding genes, *ethA* is directly regulated by iron availability in a Fur-dependent manner (Hirono *et al.*, 1997, Wang *et al.*, 2009). A high-throughput genetic screen for *ethBA* regulators identified a direct repressor belonging to the GntR family called EthR (*eth*B Regulator). This regulator is also a direct activator of the *luxS* gene, modulating the LuxS/AI-2 quorum sensing (QS) system also involved in bacterial pathogenicity (Wang *et al.*, 2009). The EsrAB TCS negatively regulates EthA synthesis but activates both T3SS and T6SS in a process implying the AraC-family regulatory protein EsrC (Wang *et al.*, 2010, Zheng *et al.*, 2005). Deletion of the *esrB* gene permits a temperature-dependent regulation of *ethA* with a higher EthA-based hemolytic activity observed at 37°C than at 25°C (Wang *et al.*, 2010). The thermo-induction relies on the activity of the nucleoid protein Hha_{Et} that can bind upstream of *ethA* and thus directly modulates its expression probably by alteration of DNA topology (Wang *et al.*, 2010). Finally, *esrB* transcription is affected by an interplay between the temperature-responsive PhoQP system and EsrBA where PhoP upregulates *esrB* at 30°C (see refs in (Guijarro *et al.*, 2015)). To note, the EsrB-dependent upregulation of T3SS and T6SS gene clusters is higher at 25°C (Rao *et al.*, 2004) while hemolytic activity is reduced at this temperature (Wang *et al.*, 2010).

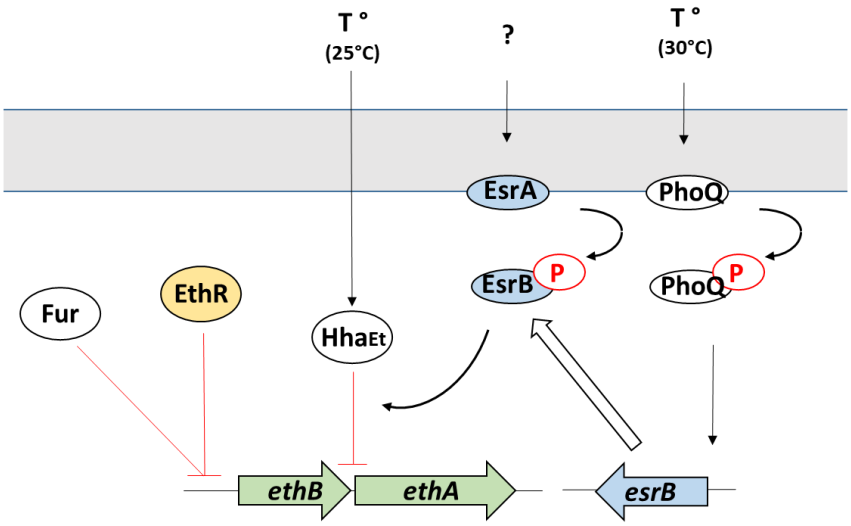


Figure 2: The regulation of *ethBA* in *Edwardsiella tarda.*

The gram-negative bacterium *P. aeruginosa* is able to infect various hosts ranging from plants, insects to humans and, as an efficient opportunistic pathogen, is a leading cause of nosocomial infections (Diggle & Whiteley, 2020). While its major virulence factor is the well-known T3SS (Klockgether & Tummler, 2017), a new taxonomic group - the PA7-like strains - contains *P. aeruginosa* strains that do not possess the T3SS-encoding genes nor those encoding associated exotoxins (Freschi *et al.*, 2019). Instead, they express ExlBA, a TPS secreting the cytotoxin ExlA (Elsen *et al.*, 2014, Huber *et al.*, 2016). Interestingly, these two secretion systems seem to be functionally incompatible, as no strain possessing both T3SS and ExlBA has ever been identified. ExlA is translocated through the outer membrane by ExlB and induces formation of pores in eukaryotic cell membranes, leading to membrane disruption and eventually cell death (Elsen *et al.*, 2014, Basso *et al.*, 2017a). This potent cytolytic effect makes ExlBA the major virulence factor of the PA7-like strains. Following the standard TPS genetic organization, the *exlB* and *exlA* genes are encoded as a single transcriptional unit and this operon has been found to be under the direct regulation of the global regulator Vfr (Berry

et al., 2018). Vfr is a homolog of the CRP regulator in *E. coli* and is known to regulate hundreds of genes, many of them encoding virulence factors (West *et al.*, 1994, Wolfgang *et al.*, 2003). Vfr is activated upon sensing of its cognate allosteric activator, the secondary messenger cAMP, and is thus involved in the cAMP- and c-di-GMP-dependent planktonic-biofilm lifestyle switch of *P. aeruginosa* (Moradali *et al.*, 2017). cAMP levels are tightly regulated by adenylate cyclases, such as CyaB, and are known to increase in response to calcium depletion or host cell contact (Lory *et al.*, 2004). The activation of *exlBA* by Vfr illustrates the key role of the regulator in virulence across the entire species; it has been thoroughly characterized in *T3SS*$^{+}$ strains where it activates the T3SS, and it is now known that even in the PA7-like lineage, which possesses ExlBA as major weapon instead of the T3SS, Vfr is also regulating this new virulence factor. A second regulator of *exlBA* has recently been identified, the Exolysin regulatory factor A (ErfA) transcription factor, which strongly inhibits the operon expression, thus counteracting the positive effect of cAMP-activated Vfr (Trouillon *et al.*, 2020). ErfA and Vfr bind to the *exlBA* promoter at two different locations, allowing independent bindings of both regulators. This configuration reminds the well-known *lac* operon regulation, where the non-overlapping bindings of CRP and LacI allow the independent signal integration of both the activator and the repressor (Dickson *et al.*, 1975). However, while Vfr activating signal is known, the existence and nature of a potential signal detected by ErfA is still to be elucidated in order to comprehensively apprehend virulence regulation in the PA7-like lineage. Interestingly, the ExlBA secretion system is also found in several other *Pseudomonas* species, including *P. chlororaphis*, *P. putida* and *P. protegens* (Basso *et al.*, 2017b) (Job *et al*., 2020). All these species thrive in very different environments and range from plant- and soil-dwelling biocontrol agents to human pathogens. To understand the underlying explanation behind the conservation of such a potent virulence factor in this wide array of different species, *exlBA*

promoter sequence was recently found to exhibit a high diversity across the *Pseudomonas* genus with the presence of both ErfA and Vfr binding sites being restricted to *P. aeruginosa* (Trouillon *et al.*, 2020). Indeed, after acquisition of *exlBA* by HGT, the operon was wired to the ErfA and Vfr regulons by evolution of their respective binding sites inside of the *P. aeruginosa* *exlBA* promoter. In other species, different putative binding sites are found, reflecting an evolutionary mechanism allowing the regulatory adaptation of newly acquired genes to better fit each individual niche (Trouillon *et al.*, 2020).

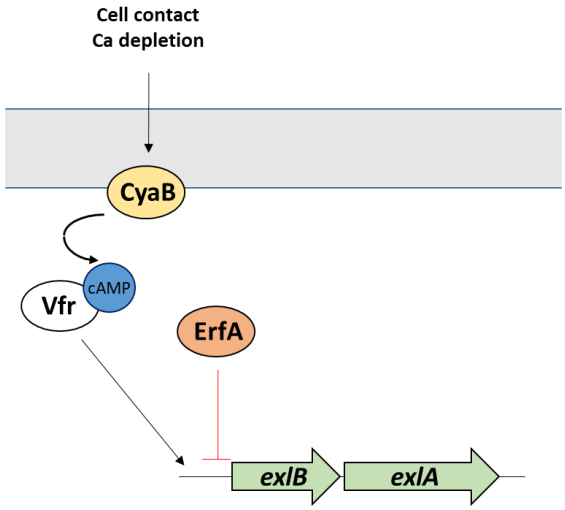


Figure 3: The regulation of *exlBA* in *Pseudomonas aeruginosa*.

TCS, phase variation and c-di-GMP regulate TPS adhesins

Many TpsA proteins exhibit adhesive properties. Adhesion is an essential step in pathogenesis, as it is required during host colonization; therefore, as observed with cytolysins, conditions encountered within the human host can stimulate their expression as temperature and iron. An example concerns EtpA, an adhesin produced by Enterotoxigenic *Escherichia coli* (ETEC) that is concentrated at the tip of the flagellum and required for adhesion to epithelial

intestinal cells and early small intestine colonization (Fleckenstein *et al.*, 2006, Roy *et al.*, 2009). This large glycosylated EtpA adhesin (~177 kDa) was shown to be upregulated under iron restriction in concert with the operon encoding the CFA/I fimbriae, even if the underlying mechanism for this regulation has not been established (Haines *et al.*, 2015). However, the information available for these TPS rather indicate species-specific and adhesin-specific regulatory mechanisms.

The best studied example is the filamentous hemagglutinin FHA of *Bordetella pertussis*, the causative agent of the whooping cough. FHA is a 220 kDa protein whose precursor, FhaB (367 kDa), is translocated by its cognate Transporter FhaC and processed during secretion by cleavage of its C-terminal prodomain. FHA folds into a rigid β-helix on bacterial surface and is ultimately released. In addition to adhesion to epithelial cells, the protein promotes biofilm formation, modulates immune responses (reviewed in (Scheller & Cotter, 2015)) as well as bacterial persistence in the lower respiratory tract of infected mice (Melvin *et al.*, 2015). Expression of FHA, as of other multiple virulence factors including the pertussis toxins PTX and Cya, is modulated by DNA gyrase-inhibiting drugs, indicating a role of DNA supercoiling (Graeff-Wohlleben *et al.*, 1995). The TPS-encoding genes are organized in a complex locus, *bvgSA-fhaB-fimA-fimBCDfhaC*, with *fhaC* being transcribed from the *fimB* promoter (Willems *et al.*, 1992, Scheller & Cotter, 2015). The *bvgAS* operon encodes the BvgSA TCS, a master regulatory system in numerous *Bordetella* species (Chen & Stibitz, 2019). BvgSA regulates hundreds of genes, notably encoding virulence factors such as fimbriae FIM and FHA, therefore controlling the transition between avirulent, intermediate and virulent phenotypic phases (Moon *et al.*, 2017). Virulence occurs when BvgS is active as a kinase: it autophosphorylates on a histidine residue, then the phosphate is transferred to the response regulator BvgA. The activated transcription factor in turn binds to multiple high- and low-

affinity binding sites within *fhaB* and *fimB* promoters as well as in the *bvgAS* promoter, activating their transcription (Boucher *et al.*, 1997, Boucher *et al.*, 2003, Cotter & Jones, 2003). Even if DNA-binding mechanism of BvgA has been thoroughly studied, all the signals controlling the TCS activity are still unknown. However it is well established that the Bvg system is active *in vivo* at 37°C during respiratory tract infection (Cotter & Miller, 1994). At 25°C, as well as in presence of the so-called *in vitro* "modulating signals" $MgSO_4$ and nicotinic acid, BvgS kinase activity is turned off and BvgA is unphosphorylated, presumably due to the reversal of the BvgSA phosphorelay (Boulanger *et al.*, 2013, Scarlato *et al.*, 1991). BvgS is composed of several linkers and sensory domains which each reflects a different potential regulatory role, illustrating the complex array of signals probably sensed by this TCS (Dupre *et al.*, 2015, Lesne *et al.*, 2018, Sobran & Cotter, 2019). Notably, oxidized ubiquinone is able to inhibit BvgS activity through its cytoplasmic PAS domain that harbors a probable quinone binding site (Bock & Gross, 2002), indicating that the domain might sense the redox state of ubiquinone. The BvgSA TCS can also integrate signals indirectly as its activity is controlled by another TCS, PrlSR, both *in vitro* and *in vivo* in the murine Lower Respiratory Tract (LRT) (Bone *et al.*, 2017). The histidine kinase PlrS (Persistence in the lower respiratory tract Sensor) is indeed able to increase the BvgSA-dependent virulence-associated phenotypes (such as hemolysis and adhesion) in response to elevated CO_2 *in vitro* and is required for full Bvg system activity in murine LRT. The PlrSR system might respond to cues found in LRT, such as low O_2 and high CO_2 levels, and controls important BvgSA-independent functions required for persistence in LRT (Bone *et al.*, 2017). Overall, this illustrates the complexity of the regulatory networks driving FhaBC synthesis by integrating many different signals, ensuring an adequate regulation of virulence phenotypes.

P. aeruginosa encodes one filamentous β-helical protein belonging to the FHA-like adhesin subfamily. Its name, CdrA, derives from cyclic diguanylate-regulated TPS partner A, as both its synthesis and outer membrane addressing are c-di-GMP-dependent (Borlee *et al.*, 2010, Rybtke *et al.*, 2015). CdrA attaches to the bacterial cell-surface through a putative cysteine-hook located at its C-terminus which keeps the adhesin anchored to the transporter, CdrB (Cooley *et al.*, 2016). This location ensures its function in bacterial auto-aggregration and contributes to biofilm structural integrity by binding and maintaining Psl (Borlee *et al.*, 2010). CdrA also prevents biofilm dispersion by tethering cell-cell interactions (Cherny & Sauer, 2020). c-di-GMP is a key second messenger controlling several important cellular traits, notably virulence of numerous pathogenic bacteria, and it is associated with biofilm lifestyle, with a high intracellular concentration stimulating the expression of extracellular matrix components and adhesion factors favoring the biofilm mode of growth (for reviews, see (Romling *et al.*, 2013, Valentini & Filloux, 2019). The *cdrAB* operon is under the direct control of FleQ, the transcriptional master regulator of flagellar gene expression that belongs to the family of Enhancer Binding Proteins (EBPs). Bacterial EBPs typically exhibit three domains with a central AAA+ ATPase domain that interact with the sigma factor σ^{54} (RpoN) and a C-terminal DNA-binding domain. *P. aeruginosa* FleQ activity is controlled by c-di-GMP binding to its N-terminal domain and direct interaction with another ATPase, the anti-activator FleN. Regulatory mechanisms used by FleQ are multiple as it acts both as a repressor and activator even on the same promoters, can bind to upstream sequences as EBPs but also around Transcription Start Sites (TSSs), and does not necessarily require RpoN binding (see references in (Matsuyama *et al.*, 2016)). The *cdrAB* operon is probably under the transcriptional dependence of σ^{70}-RNA polymerase, harboring RpoD-dependent -10/-35 boxes. The promoter driving *cdrAB* expression contains three FleQ boxes (GTCAATaaATTGAC) centered at positions

-109 (required for FleQ-mediated *cdrAB* repression when c-di-GMP is absent), +45 and +67 (required for FleQ-mediated *cdrAB* activation when c-di-GMP is present) relative to its TSS. Both c-di-GMP and FleN do not impair FleQ binding, but binding of c-di-GMP to the N-terminal FleQ domain triggers conformational modifications, leading to FleQ switches from a repressor to an activator on the *cdrAB* promoter (Baraquet *et al.*, 2012, Baraquet & Harwood, 2016). In addition to the impact on transcription, intracellular c-di-GMP levels regulate the CdrA secretion/release by membrane bound LapD/G system. LapG is a periplasmic protease sequestered by the membrane-anchored LapD when its EAL domain is loaded with available c-di-GMP. Upon the drop of c-di-GMP levels, LapD releases LapG, which in turn cleaves the periplasmic cysteine-hook in CdrA, thus releasing it into surrounding media. This mechanism triggers the release of the bacterium from the adhered surface and consequently biofilm disintegration (Cooley *et al.*, 2016). The LapD/G system is widely conserved in many bacterial species, including important pathogens, and post-transcriptionally control the location of other adhesins such as LapA and MapA of *Pseudomonas fluorescens* and BrtA of *Bordetella bronchiseptica* (Navarro *et al.*, 2011, Newell *et al.*, 2011, Ambrosis *et al.*, 2016, Collins *et al.*, 2020). Overall, c-di-GMP ensures both the synthesis and the activity of the CdrA adhesin by stimulating the expression of its operon and preventing its proteolysis, allowing its retention within the outer membrane and different stimuli affecting c-di-GMP intracellular levels can in turn affect CdrA status.

The synthesis of many bacterial adhesins is under the control of the so-called “phase variation” mechanism, which is a reversible ON/OFF switch of gene expression leading to different levels of protein abundance (Ahmad *et al.*, 2017). Phase variation is a potent adaptive mechanism increasing pathogenicity of bacteria by creating a phenotypic heterogeneity in a clonal population. In the case of adhesins, this generates an adhesion gradient and a balance

between efficient adhesion and efficient immune evasion (Dawid *et al.*, 1999). Other extracellular protein structures can exhibit bimodal expression, as we already mentioned for the XhlA hemolysin and the flagellum of *X*. *nematophila* (Jubelin *et al.*, 2013). Two surface-exposed High-Molecular-Weight (HMW) proteins, and potential vaccine candidates (Winter & Barenkamp, 2016), are regulated by phase variation in *H. influenza* NTHi, a worldwide leading cause of respiratory tract infections (Davis *et al.*, 2014). These adhesins confer an advantage for the attachment step but are also a target of the immune system, as they tend to be antigenic (Davis *et al.*, 2014). Even if they share significant sequence similarity, HMWA1 and HMWA2 exhibit different cellular binding specificities; indeed their internal binding domains, that are surface-exposed in the mature proteins, share only 40 % sequence identity (Barenkamp & Leininger, 1992, St Geme *et al.*, 1993, Buscher *et al.*, 2004). The two proteins are encoded within two loci, *hmw1A1B1C* and *hmw2A2B2C*, which probably originated from gene duplication. Around 75% of *H. influenzae* strains possess the two loci but the HMW adhesion genes display a large genetic diversity in their binding domain-encoding portions (Buscher *et al.*, 2004, Giufre *et al.*, 2006). The TpsB-like HmwB proteins (St Geme & Grass, 1998) and the glycosyltransferases HmwC (McCann & St Geme, 2014) are required for adhesin processing and secretion, and the two homologs are conserved and interchangeable. Both transcription and post-translational modifications of the two HMWA are under phase variation control. One mechanism controlling the expression relies on 7-bp simple sequence repeats (SSRs) located in the promoter regions (Barenkamp & Leininger, 1992, Dawid *et al.*, 1999). The repeats are located between the two promoters identified upstream from *hmwA* and their numbers (ranging from 9 to 28 copies) were shown to vary spontaneously both *in vitro* and *in vivo* (animal model and human infections) in a Rec-independent manner (Dawid *et al.*, 1999, Giufre *et al.*, 2008). The repeat location suggests a slipped-strand mispairing and

DNA polymerase slippage mechanism during DNA replication. The number of repeats correlates with the mRNA abundance and thus adhesin synthesis; an increase of SSRs leads to a reduction of mRNAs quantities, either due to an effect on mRNA stability or on the transcription activity of the promotor (Dawid *et al.*, 1999). The *hmw* expression relies also on the N6-adenine DNA-methyl transferase ModA (Atack *et al.*, 2015). The *modA* expression itself is regulated by phase variation and 20 distinct *modA* alleles encode proteins, which target and methylate different sequences, controlling different targets, among them are the *hmw* genes and genes involved in iron acquisition. This phase variation mechanism was shown to be important in bacterial virulence and niche adaptation (Atack *et al.*, 2015). Finally, the SSRs are also present in promoters controlling the synthesis of HmwC1 and HmwC2, the two enzymes affecting HmwA glycosylation, implying distribution of heterogeneous proteoforms on the bacterial surface and probably efficient immune evasion (Elango & Schulz, 2020).

PfhB1 and PfhB2 (also named, FhaB1 and FhaB2 respectively) are putative filamentous hemagglutinins of *Pasteurella multocida*, the causative agent of fowl cholera and several animal diseases (Wilson & Ho, 2013). They are both required for *in vivo* fitness and virulence (Fuller *et al.*, 2000, Tatum *et al.*, 2005 , Guo *et al.*, 2014). PfhB1/B2 are encoded in the locus *lspB-pfhB1-lspB-pfhB2* also comprising genes encoding two TpsB proteins (May *et al.*, 2001). Hemagglutinin biosynthesis is coordinated with that of the capsule by two important regulatory proteins in the bacterial physiology, namely Fis and Hfq. The global regulator Fis is a growth–phase dependent nucleoid-associated transcriptional factor that controls positively *lspB-fhaB2* and capsular genes expression (Steen *et al.*, 2010). On the other hand, Hfq is a RNA-binding protein that mainly facilitates translation regulation *via* interaction with small noncoding RNAs (ncRNA) and their mRNA targets (Vogel & Luisi, 2011). This post-transcriptional regulator plays a crucial role in regulating virulence factors and its inactivation

reduces *P. multocida* virulence and fitness in mice, as observed with other bacteria (Megroz *et al.*, 2016). Among the numerous virulence factors it controls, it was shown that abundance of PfhB2 and LspB2 proteins was reduced by around 6 fold in a *hfq* mutant. As expression of *pfhB2* and *lspB2* genes was also slightly reduced, Hfq was also proposed to probably act positively at the transcriptional level (Megroz *et al.*, 2016), although an impact on mRNA stability could also explain the transcriptomic data. However, the exact mechanism has to be deciphered and the potential ncRNA(s) involved in PfhB2 regulation discovered. Of note, this regulatory network is even more complex, as Fis seems to also positively regulate Hfq expression (Megroz *et al.*, 2016).

In *P. aeruginosa*, the TpsA-like CupB5 protein is a putative surface-exposed adhesin encoded within the *cupB* operon (Ruer *et al.*, 2008), that was shown to play a regulatory role in alginate production and mucoid conversion (de Regt *et al.*, 2014). In addition to *cupB5*, the six-gene *cupB* operon codes for a chaperone-usher pathway (cup) assembling fimbriae at the bacterial surface (Ruer *et al.*, 2008). CupB5 belongs to a multi-substrate TPS system, being the second cargo of LepB, the transporter of the LepBA system (Garnett *et al.*, 2015). The "large extracellular protease" LepA protein contains a trypsin-like serine protease motif and is endowed with proteolytic activity (Kida *et al.*, 2008): to the best of our knowledge, it is the only TpsA protease reported in the literature. By activating the NF-κB pathway through protease-activated receptors (PARS), LepA induces inflammatory responses in human cells, modulates host response against infection (Kida *et al.*, 2008), and degrades hemoglobin, a source of heme iron and peptides for a pathogen (Kida *et al.*, 2011). Interestingly, the expression of both *lepA* and *cupB* operon is controlled by RocS1, a membrane-bound unorthodox sensor histidine kinase, pointing to the coordinate synthesis of the two TPS cargoes (Garnett *et al.*, 2015). RocS1 was also reported to promote biofilm maturation, reduce

antibiotic resistance and T3SS-encoding genes by modulating the phosphorylation status of several response regulators, among them RocA1 and RocA2 (Mikkelsen *et al.*, 2011, Sivaneson *et al.*, 2011, Francis *et al.*, 2017). The molecular mechanism used by RocS1 to control *cupB* and *lep* genes has not been clarified yet. We need to mention that the RocS1A1R phosphorelay system of *P. aeruginosa* is homologous to the BvgSAR system that regulates many virulence traits in *Bordetella* species (Mikkelsen *et al.*, 2011), notably the adhesin ShlA of *B. pertussis*.

Iron controls hemopexin-binding protein required for iron acquisition

Some bacteria have evolved TPS systems to capture molecules required for their survival. An example is *Haemophilus influenzae* that is unable to synthesize iron-containing heme (Evans *et al.*, 1974), an essential cofactor required for the function of many cellular proteins. This obligate human commensal/pathogen, responsible for a large range of infections, possesses several heme acquisition systems; one of them comprises the HxuBA TPS and the HxuC protein that permit the use of host hemopexin as an important source of both heme and iron (Cope *et al.*, 1995, Cope *et al.*, 1998, Fournier *et al.*, 2011). The cell surface-exposed HxuA binds hemopexin and allows heme uptake by HxuC, a TonB-dependent receptor. Expression of the *hxuCBA* operon is upregulated in iron-heme restricted medium in a Fur-dependent manner (Whitby *et al.*, 2006, Harrison *et al.*, 2013). In *Haemophilus parasuis* serovar 5, a transcriptomic analysis revealed that *hxuCBA* operon was upregulated *in vitro* in conditions mimicking infection, which are iron limitation and high temperature (Alvarez-Estrada *et al.*, 2018). Recently, as already mentioned for the Hmw adhesins and other outer membrane-proteins/complexes (see above), the expression of the *hxuCBA* operon was found to be biphasic due to ModA, the N6-adenine DNA-methyl transferase whose expression is

submitted to phase variation that plays an important rolein niche adaptation of *Haemophilus* NTHi (Atack *et al.*, 2015).

Regulation of TpsA proteins with unknown function

Although TPS substrates are usually easy to identify from their genetic organization and specific domains, the prediction of their exact function is less obvious. Nevertheless, for several TpsAs without attributed function, we are starting to apprehend their regulation.

This is notably the case for the RscBA system of *Yersinia enterocolitica,* a pathogen to a wide range of hosts, provoking acute gastroenteritis in humans (Drummond *et al.*, 2012). The chromosome-located *rscABC* locus encodes RscB that would transport the RscA protein, which is highly similar to the HmvA adhesin. The locus, as well as the *ompF* gene, has been shown to be under the positive control of RscR (regulator of systemic colonization), a LysR type regulator expressed early during infection. Accordingly, inactivation of either *rscR* or *rscA* leads to increased bacterial splenic dissemination in a mouse model of infection (Nelson *et al.*, 2001). To date, the coinducer to which RscR responds has not been identified.

More information is known concerning the regulation of the *P. aeruginosa* protein PdtA, a probable β-helical protein related to the HMW-like adhesin subfamily (Faure *et al.*, 2014). This 430-kDa protein is processed during secretion and is anchored and exposed to the cell-surface. PdtA was shown to be involved in the pathogenesis of *P. aeruginosa* during infection of *Caenorhabditis elegans* (Faure *et al.*, 2014) and it is synthesized during acute and chronic human infections, as some patient's sera contained antibodies directed against PdtA (Llamas *et al.*, 2009). Interestingly, the two partners are coded by two genes that are not contiguous but separated by the gene *phdA*. In addition, they belong to two different

transcriptional units, one consisting of the single gene *pdtA* and the other comprising 11 genes and starting with *phdA* (Faure *et al.*, 2013, Quesada *et al.*, 2016). However, the two units are co-regulated and show higher expression in condition of inorganic phosphate (Pi) limitation, giving rise to the name "Phosphate depletion regulated TPS" or Pdt (Faure *et al.*, 2013, Faure *et al.*, 2014). Pi limitation is sensed by the Pst system and transmitted to the PhoRB TCS (Chekabab *et al.*, 2014). Among the Pho regulon is the operon encoding the PUMA3 system, a Cell-Surface Signaling (CSS) system comprising the extracytoplasmic function (ECF) sigma factor VreI, the anti-sigma VreR and the receptor VreA (Faure *et al.*, 2013, Llamas *et al.*, 2009, Llamas *et al.*, 2014). In condition of Pi limitation, the phosphorylated response regulator PhoB activates the expression of the *vreAIR* operon and, consequently, σ^{VreI} activates, among others, the expression of the two transcriptional units containing *pdtA* and *pdtB* ((Llamas *et al.*, 2009, Faure *et al.*, 2013). In addition PhoB also controls directly expression of these genes, as it binds to *phoB box* identified just upstream the -35 element of the *pdtA* and *phdA* promoters; binding of PhoB was shown to be required for the recruitment of σ^{VreI} on the *pdtA* promoter (Quesada *et al.*, 2016). Of note, the importance of σ^{VreI} in virulence has recently been highlighted in a zebrafish model of infection, where Pi starvation increases *P. aeruginosa* virulence in embryo in a VreI- and VreR-dependent manner (Otero-Asman *et al.*, 2020).

The Lsp system of *Haemophilus ducreyi*, a pathogen responsible for the sexually transmitted chancroid infection, is involved in phagocytosis inhibition (Vakevainen *et al.*, 2003, Dodd *et al.*, 2014). LspB is the outer membrane transporter allowing the secretion of two substrates, LspA1 and LspA2 (Ward *et al.*, 2004), which share 86% of identity. LspB and LspA2 are encoded in a bicistronic operon whereas *lspA1* is distantly located. Only the *lspB-A2* operon is upregulated during infection in pustules compared to *in vitro* growth (Gangaiah *et al.*, 2016). While the *lspA1* gene is constitutively expressed *in vitro*, *lspA2* expression was

shown to be growth phase-dependent, with higher expression in stationary phase. Hfq contributes to the regulation of around 40% of the genes differentially expressed in exponential phase relative to stationary phase and can be considered as a major contributor of stationary phase gene regulation in *H. ducreyi* that is devoid of RpoS homolog. Hfq positively affects the transcript levels of several virulence factors including *lspBA2*, likely by mRNA stabilization (Gangaiah *et al.*, 2014). The presence of fetal calf serum, that might be encountered during infection, is required for the expression of *lspA2* and the LspB-dependent release of LspA1 (Labandeira-Rey *et al.*, 2009). The *lspBA2* locus is also negatively regulated by CpxAR, the cell envelope stress TCS, as many other key virulence factors involved in chancroid formation (Gangaiah *et al.*, 2013). CpxR was shown to bind directly to the *lspBA2* promoter *in vitro* (Labandeira-Rey *et al.*, 2009). Deletion of the membrane fusion protein MtrC of the MTR efflux transporter leads to the activation of CpxAR system and thus to a decrease of *lspB* mRNA levels (Rinker *et al.*, 2011). Another regulatory element is the nucleoid-associated protein Fis whose deletion decreases the expression of the *lspB/lspA2* operon but does not abolish phagocytosis inhibition, probably due to the unaffected *lspA1* (Labandeira-Rey *et al.*, 2013). In *H. ducreyi*, the expression pattern of *fis* is different from that observed in *E. coli* and it has a positive effect on *gyrB* transcription, potentially increasing DNA supercoiling throughout the growth phase. The last important pathway, intertwined with the ones already described, concerns the alarmone (p)ppGpp and DksA, known to regulate *H. ducreyi* virulence in humans (Holley *et al.*, 2014, Holley *et al.*, 2015). During stringent response, DksA helps stabilize the interaction between (p)ppGpp and RNA polymerase, enhancing (p)ppGpp-controlled transcription. DksA can also affect transcription by acting directly on RNA polymerase. An increased expression of *lspB* was observed in mutants unable to produce (p)ppGpp at mid-log phase, which may explain the growth phase-dependent reduction of

phagocytosis observed in the mutants (Holley *et al.*, 2014). DksA also plays a role in resistance to phagocytic killing and resistance to oxidative stress as DksA- and (p)ppGpp-deficient transcriptomes significantly overlap at stationary phase in *H. ducreyi* (Holley *et al.*, 2015) .

Regulation of Contact-Dependent Inhibition systems

Contact-dependent growth inhibition (CDI) systems are contact-mediated competition systems of the TPS family. Akin to T6SSs, CDI systems allow the direct injection of antimicrobial effectors into bacterial competitor cells (Aoki *et al.*, 2010). They are composed of an outer-membrane specific transporter (CdiB) that allows the recruitment of a large exoprotein (CdiA) to the cell surface. In turn, CdiA integrates its C-terminal toxin domain into a neighboring competitor cell leading to growth inhibition or cell death (Aoki *et al.*, 2005, Morse *et al.*, 2012). CdiA toxins activity include nucleases, adenosine deaminases, ADP-ribosyl cyclases, and metallopeptidases. In order to protect themselves and their sibling cells, bacteria that possess CDI systems also encode an immunity protein (CdiI) which neutralizes CdiA toxin domain activity. In addition to their role in bacterial competition, CDI systems have been shown to mediate community-associated behaviors, called contact-dependent signaling (CDS), through gene expression regulation upon delivery of the toxin into immune bacteria (Garcia *et al.*, 2016, Danka *et al.*, 2017). By doing so, CDI systems are involved in cell-cell communication notably driving biofilm formation and high cell density phenotypes (Guilhabert & Kirkpatrick, 2005, Garcia *et al.*, 2013). CDI systems are generally encoded as a single transcriptional unit of three genes following a *cdiBAI* organization (Guerin *et al.*, 2017). They have been discovered in *E. coli*, where they showed a constitutive expression (Aoki *et al.*, 2005). Since then, various CDI regulatory mechanisms have been unraveled.

There are two different CDI systems *in P. aeruginosa*, CdiBAI1 and CdiBAI2, which exhibit C-terminal domain variability between strains and lineages (Allen & Hauser, 2019). Indeed, numerous different domains were found, associated to several putative functions. Major classes encompass domain with unknown functions, deaminases, nucleases and tRNases. While this diversity in *P. aeruginosa* CDI systems has been described, the knowledge on their regulation is still sparse as only one study addressed it yet. This work showed that both CDIs are post-transcriptionally inhibited by the RNA-binding regulator RmsA (Mercy *et al.*, 2016). This regulation is thought to be direct through prevention of the binding of the 30S ribosomal subunit to *cdi* mRNAs by RmsA, inhibiting the translation-dependent stabilization of mRNAs and promoting mRNA degradation. The regulation was abrogated during static growth, suggesting a condition-dependent regulation by RmsA. RmsA is known to be a major regulator of the switch between planktonic and biofilm lifestyles. It is controlled by the complex GacSA-LadS-RetS regulatory pathway that governs the switch between planktonic lifestyle and biofilm formation (Moradali *et al.*, 2017, Brencic & Lory, 2009) while it has been shown in *P. aeruginosa* and other species that CDI systems play key roles in biofilm formation and community structure (Anderson *et al.*, 2014, Mercy *et al.*, 2016). This link suggests a function-related regulation of *P. aeruginosa* CDI systems and raises the need for further characterization of these competition systems in regards to the well-described lifestyle switch of this major pathogen.

As *P. aeruginosa*, *Acinetobacter* species possess two types of CDI systems that have different C-terminal domains (De Gregorio *et al.*, 2019). *Acinetobacter* CDI systems of type I (CDI_1) are abundant in *A. pittii* and type II (CDI_2) in *A. nosocomialis*, while both are found in the opportunistic pathogen *A. baumannii.* Both types of CDI were found to inhibit growth in non-immune cells in *A. baylyi,* but to not affect biofilm formation or contact-dependent signaling

(De Gregorio *et al.*, 2018). Surprisingly, the CDI_1 system was shown to inhibit biofilm formation and have a cytoplasmic DNase activity in *A. baumannii*. In this species, the CDI_2 system is not expressed during growth in rich media (Roussin *et al.*, 2019). However, the $cdiB_2$ promoter exhibits a putative PhoB binding site, suggesting a phosphate-dependent regulation of this system, similarly to what was shown for *pdtA* in *P. aeruginosa* (Quesada *et al.*, 2016). Additionally, in the *A. baumannii* V15 strain, one cdi_1 locus is conserved and has been shown to be under the regulation of the BfmRS TCS (Krasauskas *et al.*, 2019). BfmRS is known to be involved in biofilm formation and was shown to inhibit CDI_1 in this strain. It is thought that this downregulation diminishes bacterial competition during biofilm formation, allowing the incorporation of other *Acinetobacter* species, which might favor survival.

The BcpABI CDI system is present in several *Burkholderia* species, including the opportunistic human pathogen *B. pseudomallei* that possesses up to 10 CDI systems (Nikolakakis *et al.*, 2012). BcpABI has been first identified in *B. thailandensis*, where it mediates biofilm formation independently of its CDI activity (Garcia *et al.*, 2013). Within biofilms, BcpABI was shown to be stochastically regulated, as only a small subset of cells expresses the system. The action of BcpA onto sibling cells was found to induce transcriptional and phenotypic changes, including a decrease expression of the *bcpAIOB* operon (Garcia *et al.*, 2016); this mechanism is probably at the origin of the stochastic nature of the operon expression. Furthermore, BcpABI has been found to be upregulated through QS by acyl-homoserine lactones (AHLs) and the associated transcriptional regulator BtaR1, highlighting the importance of this CDI system in high cell density (Majerczyk *et al.*, 2014a). Interestingly, QS does not regulate BcpABI in *B. pseudomallei* and *B. mallei,* revealing regulatory differences between different *Burkholderia* species ((Majerczyk *et al.*, 2014b) and illustrates the regulatory diversity that can be found in a genus: all species were phylogenetically related,

but lived in very different environments with one saprophyte, one opportunistic pathogen and one host-restricted pathogen, explaining the need for different regulations.

Neisseria meningitidis is the causative agent of meningococcal meningitis and possesses several adhesins required for virulence as type IV pili and several non-fimbrial adhesins (Opa, Opc, NhhA, and NadA). Although HrpA was first identified as a TpsA adhesin contributing to epithelial cell adhesion (Schmitt *et al.*, 2007) , it was then demonstrated to be essential for the survival of the bacteria in the cell and their escape from the intracellular space after invasion (Tala *et al.*, 2008). Its role in biofilm formation was further demonstrated (Neil & Apicella, 2009). However, HrpA was found to exhibit growth-inhibitory properties important for antibacterial competition (Arenas *et al.*, 2013). Its gene is located in the *hrpBA*(*IS*) island, with *hrpB* coding for the outer membrane transporter that secretes HrpA. Downstream from *hrpA* are found *hrpIS* repeats with several *hrpS* interspaced with cognate immunity-encoding genes. The *hrpS* genes code for the TpsC cassettes which are 5'-truncated HrpA proteins exhibiting variable toxic domains: they provide to HrpA a possible repertoire of toxic domains gained through genetic recombination (Tommassen & Arenas, 2017). The *hrp* operon was found upregulated in anaerobic conditions (2.4- to 8.7-fold), after 5 h of cell association in culture plates (*hrpB* 11.1-fold, *hrpA* 5.1-fold) and when bacteria were detaching from biofilms on HBE cells (*hrpB* 14.4-fold, *hrpA* 11.34-fold), revealing a complex regulatory mechanism (Neil & Apicella, 2009). The expression of the *hrp* operon was also found downregulated in a mutant of *rpoB* that functionally mimics stringent response, exhibiting a reduced survival in human monocytes cells and global transcriptional changes (Colicchio *et al.*, 2015).

CDI systems are also found in plant pathogens. In *Dickeya dadantii*, the HecBA CDI system was shown to be important for pathogenesis during infection of *Nicotiana clevelandii*

seedlings (Rojas *et al.*, 2002). More recently, the nucleoid-associated regulator Fis was found to activate *hecBA* expression under high DNA supercoiling conditions (Jiang *et al.*, 2015). As DNA supercoiling has already been associated with virulence and phytopathogenicity in *D. dadantii* (Lautier & Nasser, 2007, Ouafa *et al.*, 2012), this result highlights the role of HecBA during infection. Indeed, the Fis-dependent activation of virulence genes under high DNA supercoiling probably reflects what is happening during infection and might be triggered by signals found inside of the plant host.

In *Xylella fatidiosa*, the HfxAB CDI system is important for colonization and cell-cell aggregation during plant infection (Guilhabert & Kirkpatrick, 2005). Interestingly, the *hfxA* mutant exhibits a higher virulence but lower colonization and biofilm formation, attributing this CDI system to milder infection phenotypes but again to phenotypes linked to high cell density environments. *X. fatidiosa* is a major plant pathogen causing diseases with important economic outcomes, with several past oubreaks decimating crops of olives or grapevines ((Kyrkou *et al.*, 2018, Saponari *et al.*, 2019). RpfF, an enzyme involved in the synthesis of a fatty acid diffusible signaling factor (DSF), was shown to upregulate *hfxA* and *hfxB* expression (Wang *et al.*, 2012). DSF mediates cell-cell communication in biofilm and a decrease in DSF production was linked to a lower expression of *hfxA* and *hfxB*, which is consequently both regulated by and regulating biofilm-associated cell-cell communication. This regulation was also linked to c-di-GMP levels, highlighting again the role of HfxAB in biofilm formation and signaling (Chatterjee *et al.*, 2010).

The *Xanthomonas* genus comprises several plant pathogen species, notably responsible for the common bacterial blight of bean, a seed borne disease (Audy *et al.*, 1994). In this bacteria, the TPS system FhaBC has been redefined as a CDI system that mediates cell-

cell adhesions and is important for tissue colonization during plant infection (Darsonval *et al.*, 2009, Gottig *et al.*, 2009). Accordingly, this system was shown to be upregulated both *in planta* during infection (Gottig *et al.*, 2009), and during growth in XVM2 medium which mimics plant intracellular environment (Astua-Monge *et al.*, 2005).

References

Abby, S.S., J. Cury, J. Guglielmini, B. Neron, M. Touchon & E.P. Rocha, (2016) Identification of protein secretion systems in bacterial genomes. *Sci Rep* **6**: 23080.

Ahmad, S., M. Ahmad, S. Khan, F. Ahmad, S. Nawaz & F.U. Khan, (2017) An overview on phase variation, mechanisms and roles in bacterial adaptation. *J Pak Med Assoc* **67**: 285-291.

Allen, J.P. & A.R. Hauser, (2019) Diversity of Contact-Dependent Growth Inhibition Systems of Pseudomonas aeruginosa. *J Bacteriol* **201**.

Allen, J.P., E.A. Ozer, G. Minasov, L. Shuvalova, O. Kiryukhina, K.J.F. Satchell & A.R. Hauser, (2020) A comparative genomics approach identifies contact-dependent growth inhibition as a virulence determinant. *Proc Natl Acad Sci U S A* **117**: 6811-6821.

Allison, C., N. Coleman, P.L. Jones & C. Hughes, (1992) Ability of Proteus mirabilis to invade human urothelial cells is coupled to motility and swarming differentiation. *Infect Immun* **60**: 4740-4746.

Alvarez-Estrada, A., C.B. Gutierrez-Martin, E.F. Rodriguez-Ferri & S. Martinez-Martinez, (2018) Transcriptomics of Haemophilus (Glasserella) parasuis serovar 5 subjected to culture conditions partially mimetic to natural infection for the search of new vaccine antigens. *BMC Vet Res* **14**: 326.

Ambrosis, N., C.D. Boyd, O.T. GA, J. Fernandez & F. Sisti, (2016) Homologs of the LapD-LapG c-di-GMP Effector System Control Biofilm Formation by Bordetella bronchiseptica. *PLoS One* **11**: e0158752.

Anderson, M.S., E.C. Garcia & P.A. Cotter, (2014) Kind discrimination and competitive exclusion mediated by contact-dependent growth inhibition systems shape biofilm community structure. *PLoS Pathog* **10**: e1004076.

Aoki, S.K., E.J. Diner, C.T. de Roodenbeke, B.R. Burgess, S.J. Poole, B.A. Braaten, A.M. Jones, J.S. Webb, C.S. Hayes, P.A. Cotter & D.A. Low, (2010) A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature* **468**: 439-442.

Aoki, S.K., R. Pamma, A.D. Hernday, J.E. Bickham, B.A. Braaten & D.A. Low, (2005) Contact-dependent inhibition of growth in Escherichia coli. *Science* **309**: 1245-1248.

Arenas, J., K. Schipper, P. van Ulsen, A. van der Ende & J. Tommassen, (2013) Domain exchange at the 3' end of the gene encoding the fratricide meningococcal two-partner secretion protein A. *BMC Genomics* **14**: 622.

Armbruster, C.E., H.L.T. Mobley & M.M. Pearson, (2018) Pathogenesis of Proteus mirabilis Infection. *EcoSal Plus* **8**.

Astua-Monge, G., J. Freitas-Astua, G. Bacocina, J. Roncoletta, S.A. Carvalho & M.A. Machado, (2005) Expression profiling of virulence and pathogenicity genes of Xanthomonas axonopodis pv. citri. *J Bacteriol* **187**: 1201-1205.

Atack, J.M., Y.N. Srikhanta, K.L. Fox, J.A. Jurcisek, K.L. Brockman, T.A. Clark, M. Boitano, P.M. Power, F.E. Jen, A.G. McEwan, S.M. Grimmond, A.L. Smith, S.J. Barenkamp, J. Korlach, L.O. Bakaletz & M.P. Jennings, (2015) A biphasic epigenetic switch controls immunoevasion, virulence and niche adaptation in non-typeable Haemophilus influenzae. *Nat Commun* **6**: 7828.

Audy, P., A. Laroche, G. Saindon, H.C. Huang & R.L. Gilbertson, (1994) Detection of the Bean Common Blight Bacteria, Xanthomonas-Campestris Pv Phaseoli and X-C Phaseoli Var Fuscans, Using the Polymerase Chain-Reaction. *Phytopathology* **84**: 1185-1192.

Baraquet, C. & C.S. Harwood, (2016) FleQ DNA Binding Consensus Sequence Revealed by Studies of FleQ-Dependent Regulation of Biofilm Gene Expression in Pseudomonas aeruginosa. *J Bacteriol* **198**: 178-186.

Baraquet, C., K. Murakami, M.R. Parsek & C.S. Harwood, (2012) The FleQ protein from Pseudomonas aeruginosa functions as both a repressor and an activator to control gene expression from the pel operon promoter in response to c-di-GMP. *Nucleic Acids Res* **40**: 7207-7218.

Barenkamp, S.J. & E. Leininger, (1992) Cloning, expression, and DNA sequence analysis of genes encoding nontypeable Haemophilus influenzae high-molecular-weight surface-exposed proteins related to filamentous hemagglutinin of Bordetella pertussis. *Infect Immun* **60**: 1302-1313.

Basso, P., M. Ragno, S. Elsen, E. Reboud, G. Golovkine, S. Bouillot, P. Huber, S. Lory, E. Faudry & I. Attree, (2017a) Pseudomonas aeruginosa Pore-Forming Exolysin and Type IV Pili Cooperate To Induce Host Cell Lysis. *mBio* **8**.

Basso, P., P. Wallet, S. Elsen, E. Soleilhac, T. Henry, E. Faudry & I. Attree, (2017b) Multiple Pseudomonas species secrete exolysin-like toxins and provoke Caspase-1-dependent macrophage death. *Environ Microbiol* **19**: 4045-4064.

Berry, A., K. Han, J. Trouillon, M. Robert-Genthon, M. Ragno, S. Lory, I. Attree & S. Elsen, (2018) cAMP and Vfr Control Exolysin Expression and Cytotoxicity of Pseudomonas aeruginosa Taxonomic Outliers. *J Bacteriol* **200**.

Bleves, S., V. Viarre, R. Salacha, G.P. Michel, A. Filloux & R. Voulhoux, (2010) Protein secretion systems in Pseudomonas aeruginosa: A wealth of pathogenic weapons. *Int J Med Microbiol* **300**: 534-543.

Bock, A. & R. Gross, (2002) The unorthodox histidine kinases BvgS and EvgS are responsive to the oxidation status of a quinone electron carrier. *Eur J Biochem* **269**: 3479-3484.

Bone, M.A., A.J. Wilk, A.I. Perault, S.A. Marlatt, E.V. Scheller, R. Anthouard, Q. Chen, S. Stibitz, P.A. Cotter & S.M. Julio, (2017) Bordetella PlrSR regulatory system controls BvgAS activity and virulence in the lower respiratory tract. *Proc Natl Acad Sci U S A* **114**: E1519-E1527.

Borlee, B.R., A.D. Goldman, K. Murakami, R. Samudrala, D.J. Wozniak & M.R. Parsek, (2010) Pseudomonas aeruginosa uses a cyclic-di-GMP-regulated adhesin to reinforce the biofilm extracellular matrix. *Mol Microbiol* **75**: 827-842.

Boucher, P.E., A.E. Maris, M.S. Yang & S. Stibitz, (2003) The response regulator BvgA and RNA polymerase alpha subunit C-terminal domain bind simultaneously to different faces of the same segment of promoter DNA. *Mol Cell* **11**: 163-173.

Boucher, P.E., K. Murakami, A. Ishihama & S. Stibitz, (1997) Nature of DNA binding and RNA polymerase interaction of the Bordetella pertussis BvgA transcriptional activator at the fha promoter. *J Bacteriol* **179**: 1755-1763.

Boulanger, A., Q. Chen, D.M. Hinton & S. Stibitz, (2013) In vivo phosphorylation dynamics of the Bordetella pertussis virulence-controlling response regulator BvgA. *Mol Microbiol* **88**: 156-172.

Brencic, A. & S. Lory, (2009) Determination of the regulon and identification of novel mRNA targets of Pseudomonas aeruginosa RsmA. *Mol Microbiol* **72**: 612-632.

Brillard, J., E. Duchaud, N. Boemare, F. Kunst & A. Givaudan, (2002) The PhlA hemolysin from the entomopathogenic bacterium Photorhabdus luminescens belongs to the two-partner secretion family of hemolysins. *J Bacteriol* **184**: 3871-3878.

Brothers, K.M., J.D. Callaghan, N.A. Stella, J.M. Bachinsky, M. AlHigaylan, K.L. Lehner, J.M. Franks, K.L. Lathrop, E. Collins, D.M. Schmitt, J. Horzempa & R.M.Q. Shanks, (2019) Blowing epithelial cell bubbles with GumB: ShlA-family pore-forming toxins induce blebbing and rapid cellular death in corneal epithelial cells. *PLoS Pathog* **15**: e1007825.

Buscher, A.Z., K. Burmeister, S.J. Barenkamp & J.W. St Geme, 3rd, (2004) Evolutionary and functional relationships among the nontypeable Haemophilus influenzae HMW family of adhesins. *J Bacteriol* **186**: 4209-4217.

Castelli, M.E. & E.G. Vescovi, (2011) The Rcs signal transduction pathway is triggered by enterobacterial common antigen structure alterations in Serratia marcescens. *J Bacteriol* **193**: 63-74.

Chatterjee, S., N. Killiny, R.P. Almeida & S.E. Lindow, (2010) Role of cyclic di-GMP in Xylella fastidiosa biofilm formation, plant virulence, and insect transmission. *Mol Plant Microbe Interact* **23**: 1356-1363.

Chekabab, S.M., J. Harel & C.M. Dozois, (2014) Interplay between genetic regulation of phosphate homeostasis and bacterial virulence. *Virulence* **5**: 786-793.
Chen, Q. & S. Stibitz, (2019) The BvgASR virulence regulon of Bordetella pertussis. *Curr Opin Microbiol* **47**: 74-81.
Cherny, K.E. & K. Sauer, (2020) Untethering and Degradation of the Polysaccharide Matrix Are Essential Steps in the Dispersion Response of Pseudomonas aeruginosa Biofilms. *J Bacteriol* **202**.
Colicchio, R., C. Pagliuca, G. Pastore, A.G. Cicatiello, C. Pagliarulo, A. Tala, E. Scaglione, J.C. Sammartino, C. Bucci, P. Alifano & P. Salvatore, (2015) Fitness Cost of Rifampin Resistance in Neisseria meningitidis: In Vitro Study of Mechanisms Associated with rpoB H553Y Mutation. *Antimicrob Agents Chemother* **59**: 7637-7649.
Collins, A.J., A.B. Pastora, T.J. Smith & G.A. O'Toole, (2020) MapA, a Second Large RTX Adhesin Conserved across the Pseudomonads, Contributes to Biofilm Formation by Pseudomonas fluorescens. *J Bacteriol* **202**.
Cooley, R.B., T.J. Smith, W. Leung, V. Tierney, B.R. Borlee, G.A. O'Toole & H. Sondermann, (2016) Cyclic Di-GMP-Regulated Periplasmic Proteolysis of a Pseudomonas aeruginosa Type Vb Secretion System Substrate. *J Bacteriol* **198**: 66-76.
Cope, L.D., S.E. Thomas, Z. Hrkal & E.J. Hansen, (1998) Binding of heme-hemopexin complexes by soluble HxuA protein allows utilization of this complexed heme by Haemophilus influenzae. *Infect Immun* **66**: 4511-4516.
Cope, L.D., R. Yogev, U. Muller-Eberhard & E.J. Hansen, (1995) A gene cluster involved in the utilization of both free heme and heme:hemopexin by Haemophilus influenzae type b. *J Bacteriol* **177**: 2644-2653.
Costa, T.R., C. Felisberto-Rodrigues, A. Meir, M.S. Prevost, A. Redzej, M. Trokter & G. Waksman, (2015) Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol* **13**: 343-359.
Cotter, P.A. & A.M. Jones, (2003) Phosphorelay control of virulence gene expression in Bordetella. *Trends Microbiol* **11**: 367-373.
Cotter, P.A. & J.F. Miller, (1994) BvgAS-mediated signal transduction: analysis of phase-locked regulatory mutants of Bordetella bronchiseptica in a rabbit model. *Infect Immun* **62**: 3381-3390.
Cowles, K.N., C.E. Cowles, G.R. Richards, E.C. Martens & H. Goodrich-Blair, (2007) The global regulator Lrp contributes to mutualism, pathogenesis and phenotypic variation in the bacterium Xenorhabdus nematophila. *Cell Microbiol* **9**: 1311-1323.
Cowles, K.N. & H. Goodrich-Blair, (2005) Expression and activity of a Xenorhabdus nematophila haemolysin required for full virulence towards Manduca sexta insects. *Cell Microbiol* **7**: 209-219.
Danka, E.S., E.C. Garcia & P.A. Cotter, (2017) Are CDI Systems Multicolored, Facultative, Helping Greenbeards? *Trends Microbiol* **25**: 391-401.
Darsonval, A., A. Darrasse, K. Durand, C. Bureau, S. Cesbron & M.A. Jacques, (2009) Adhesion and fitness in the bean phyllosphere and transmission to seed of Xanthomonas fuscans subsp. fuscans. *Mol Plant Microbe Interact* **22**: 747-757.
Davis, G.S., S. Marino, C.F. Marrs, J.R. Gilsdorf, S. Dawid & D.E. Kirschner, (2014) Phase variation and host immunity against high molecular weight (HMW) adhesins shape population dynamics of nontypeable Haemophilus influenzae within human hosts. *J Theor Biol* **355**: 208-218.
Dawid, S., S.J. Barenkamp & J.W. St Geme, 3rd, (1999) Variation in expression of the Haemophilus influenzae HMW adhesins: a prokaryotic system reminiscent of eukaryotes. *Proc Natl Acad Sci U S A* **96**: 1077-1082.
De Gregorio, E., E.P. Esposito, R. Zarrilli & P.P. Di Nocera, (2018) Contact-Dependent Growth Inhibition Proteins in Acinetobacter baylyi ADP1. *Curr Microbiol* **75**: 1434-1440.
De Gregorio, E., R. Zarrilli & P.P. Di Nocera, (2019) Contact-dependent growth inhibition systems in Acinetobacter. *Sci Rep* **9**: 154.

de Regt, A.K., Y. Yin, T.R. Withers, X. Wang, T.A. Baker, R.T. Sauer & H.D. Yu, (2014) Overexpression of CupB5 activates alginate overproduction in Pseudomonas aeruginosa by a novel AlgW-dependent mechanism. *Mol Microbiol* **93**: 415-425.

Di Venanzio, G., T.M. Stepanenko & E. Garcia Vescovi, (2014) Serratia marcescens ShlA pore-forming toxin is responsible for early induction of autophagy in host cells and is transcriptionally regulated by RcsB. *Infect Immun* **82**: 3542-3554.

Dickson, R.C., J. Abelson, W.M. Barnes & W.S. Reznikoff, (1975) Genetic regulation: the Lac control region. *Science* **187**: 27-35.

Diggle, S.P. & M. Whiteley, (2020) Microbe Profile: Pseudomonas aeruginosa: opportunistic pathogen and lab rat. *Microbiology (Reading)* **166**: 30-33.

Dodd, D.A., R.G. Worth, M.K. Rosen, S. Grinstein, N.S. van Oers & E.J. Hansen, (2014) The Haemophilus ducreyi LspA1 protein inhibits phagocytosis by using a new mechanism involving activation of C-terminal Src kinase. *mBio* **5**: e01178-01114.

Domenighini, M., D. Relman, C. Capiau, S. Falkow, A. Prugnola, V. Scarlato & R. Rappuoli, (1990) Genetic characterization of Bordetella pertussis filamentous haemagglutinin: a protein processed from an unusually large precursor. *Mol Microbiol* **4**: 787-800.

Drummond, N., B.P. Murphy, T. Ringwood, M.B. Prentice, J.F. Buckley & S. Fanning, (2012) Yersinia enterocolitica: a brief review of the issues relating to the zoonotic pathogen, public health challenges, and the pork production chain. *Foodborne Pathog Dis* **9**: 179-189.

Dufour, A., R.B. Furness & C. Hughes, (1998) Novel genes that upregulate the Proteus mirabilis flhDC master operon controlling flagellar biogenesis and swarming. *Mol Microbiol* **29**: 741-751.

Dupre, E., J. Herrou, M.F. Lensink, R. Wintjens, A. Vagin, A. Lebedev, S. Crosson, V. Villeret, C. Locht, R. Antoine & F. Jacob-Dubuisson, (2015) Virulence regulation with Venus flytrap domains: structure and function of the periplasmic moiety of the sensor-kinase BvgS. *PLoS Pathog* **11**: e1004700.

Elango, D. & B.L. Schulz, (2020) Phase-Variable Glycosylation in Nontypeable Haemophilus influenzae. *J Proteome Res* **19**: 464-476.

Elsen, S., P. Huber, S. Bouillot, Y. Coute, P. Fournier, Y. Dubois, J.F. Timsit, M. Maurin & I. Attree, (2014) A type III secretion negative clinical strain of Pseudomonas aeruginosa employs a two-partner secreted exolysin to induce hemorrhagic pneumonia. *Cell Host Microbe* **15**: 164-176.

Evans, N.M., D.D. Smith & A.J. Wicken, (1974) Haemin and nicotinamide adenine dinucleotide requirements of Haemophilus influenzae and Haemophilus parainfluenzae. *J Med Microbiol* **7**: 359-365.

Fan, E., N. Chauhan, D. Udatha, J.C. Leo & D. Linke, (2016) Type V Secretion Systems in Bacteria. *Microbiol Spectr* **4**.

Faure, L.M., S. Garvis, S. de Bentzmann & S. Bigot, (2014) Characterization of a novel two-partner secretion system implicated in the virulence of Pseudomonas aeruginosa. *Microbiology (Reading)* **160**: 1940-1952.

Faure, L.M., M.A. Llamas, K.C. Bastiaansen, S. de Bentzmann & S. Bigot, (2013) Phosphate starvation relayed by PhoB activates the expression of the Pseudomonas aeruginosa sigmavreI ECF factor and its target genes. *Microbiology (Reading)* **159**: 1315-1327.

Fillat, M.F., (2014) The FUR (ferric uptake regulator) superfamily: diversity and versatility of key transcriptional regulators. *Arch Biochem Biophys* **546**: 41-52.

Filloux, A., (2011) Protein Secretion Systems in Pseudomonas aeruginosa: An Essay on Diversity, Evolution, and Function. *Front Microbiol* **2**: 155.

Fleckenstein, J.M., K. Roy, J.F. Fischer & M. Burkitt, (2006) Identification of a two-partner secretion locus of enterotoxigenic Escherichia coli. *Infect Immun* **74**: 2245-2258.

Forst, S., B. Dowds, N. Boemare & E. Stackebrandt, (1997) Xenorhabdus and Photorhabdus spp.: bugs that kill bugs. *Annu Rev Microbiol* **51**: 47-72.

Fournier, C., A. Smith & P. Delepelaire, (2011) Haem release from haemopexin by HxuA allows Haemophilus influenzae to escape host nutritional immunity. *Mol Microbiol* **80**: 133-148.

Francis, V.I., E.C. Stevenson & S.L. Porter, (2017) Two-component systems required for virulence in Pseudomonas aeruginosa. *FEMS Microbiol Lett* **364**.

Fraser, G.M., L. Claret, R. Furness, S. Gupta & C. Hughes, (2002) Swarming-coupled expression of the Proteus mirabilis hpmBA haemolysin operon. *Microbiology (Reading)* **148**: 2191-2201.

Freschi, L., A.T. Vincent, J. Jeukens, J.G. Emond-Rheault, I. Kukavica-Ibrulj, M.J. Dupont, S.J. Charette, B. Boyle & R.C. Levesque, (2019) The Pseudomonas aeruginosa Pan-Genome Provides New Insights on Its Population Structure, Horizontal Gene Transfer, and Pathogenicity. *Genome Biol Evol* **11**: 109-120.

Fuller, T.E., M.J. Kennedy & D.E. Lowery, (2000) Identification of Pasteurella multocida virulence genes in a septicemic mouse model using signature-tagged mutagenesis. *Microb Pathog* **29**: 25-38.

Gangaiah, D., M. Labandeira-Rey, X. Zhang, K.R. Fortney, S. Ellinger, B. Zwickl, B. Baker, Y. Liu, D.M. Janowicz, B.P. Katz, C.A. Brautigam, R.S. Munson, Jr., E.J. Hansen & S.M. Spinola, (2014) Haemophilus ducreyi Hfq contributes to virulence gene regulation as cells enter stationary phase. *mBio* **5**: e01081-01013.

Gangaiah, D., X. Zhang, B. Baker, K.R. Fortney, H. Gao, C.L. Holley, R.S. Munson, Jr., Y. Liu & S.M. Spinola, (2016) Haemophilus ducreyi Seeks Alternative Carbon Sources and Adapts to Nutrient Stress and Anaerobiosis during Experimental Infection of Human Volunteers. *Infect Immun* **84**: 1514-1525.

Gangaiah, D., X. Zhang, K.R. Fortney, B. Baker, Y. Liu, R.S. Munson, Jr. & S.M. Spinola, (2013) Activation of CpxRA in Haemophilus ducreyi primarily inhibits the expression of its targets, including major virulence determinants. *J Bacteriol* **195**: 3486-3502.

Garcia, E.C., M.S. Anderson, J.A. Hagar & P.A. Cotter, (2013) Burkholderia BcpA mediates biofilm formation independently of interbacterial contact-dependent growth inhibition. *Mol Microbiol* **89**: 1213-1225.

Garcia, E.C., A.I. Perault, S.A. Marlatt & P.A. Cotter, (2016) Interbacterial signaling via Burkholderia contact-dependent growth inhibition system proteins. *Proc Natl Acad Sci U S A* **113**: 8296-8301.

Garnett, J.A., D. Muhl, C.H. Douse, K. Hui, A. Busch, A. Omisore, Y. Yang, P. Simpson, J. Marchant, G. Waksman, S. Matthews & A. Filloux, (2015) Structure-function analysis reveals that the Pseudomonas aeruginosa Tps4 two-partner secretion system is involved in CupB5 translocation. *Protein Sci* **24**: 670-687.

Giufre, M., A. Carattoli, R. Cardines, P. Mastrantonio & M. Cerquetti, (2008) Variation in expression of HMW1 and HMW2 adhesins in invasive nontypeable Haemophilus influenzae isolates. *BMC Microbiol* **8**: 83.

Giufre, M., M. Muscillo, P. Spigaglia, R. Cardines, P. Mastrantonio & M. Cerquetti, (2006) Conservation and diversity of HMW1 and HMW2 adhesin binding domains among invasive nontypeable Haemophilus influenzae isolates. *Infect Immun* **74**: 1161-1170.

Givaudan, A. & A. Lanois, (2000) flhDC, the flagellar master operon of Xenorhabdus nematophilus: requirement for motility, lipolysis, extracellular hemolysis, and full virulence in insects. *J Bacteriol* **182**: 107-115.

Givaudan, A. & A. Lanois, (2017) Flagellar Regulation and Virulence in the Entomopathogenic Bacteria-Xenorhabdus nematophila and Photorhabdus luminescens. *Curr Top Microbiol Immunol* **402**: 39-51.

Gottig, N., B.S. Garavaglia, C.G. Garofalo, E.G. Orellano & J. Ottado, (2009) A filamentous hemagglutinin-like protein of Xanthomonas axonopodis pv. citri, the phytopathogen responsible for citrus canker, is involved in bacterial virulence. *PLoS One* **4**: e4358.

Graeff-Wohlleben, H., H. Deppisch & R. Gross, (1995) Global regulatory mechanisms affect virulence gene expression in Bordetella pertussis. *Mol Gen Genet* **247**: 86-94.

Green, E.R. & J. Mecsas, (2016) Bacterial Secretion Systems: An Overview. *Microbiol Spectr* **4**.

Grimont, P.A. & F. Grimont, (1978) The genus Serratia. *Annu Rev Microbiol* **32**: 221-248.

Guerin, J., S. Bigot, R. Schneider, S.K. Buchanan & F. Jacob-Dubuisson, (2017) Two-Partner Secretion: Combining Efficiency and Simplicity in the Secretion of Large Proteins for Bacteria-Host and Bacteria-Bacteria Interactions. *Front Cell Infect Microbiol* **7**: 148.

Guijarro, J.A., D. Cascales, A.I. Garcia-Torrico, M. Garcia-Dominguez & J. Mendez, (2015) Temperature-dependent expression of virulence genes in fish-pathogenic bacteria. *Front Microbiol* **6**: 700.

Guilhabert, M.R. & B.C. Kirkpatrick, (2005) Identification of Xylella fastidiosa antivirulence genes: hemagglutinin adhesins contribute a biofilm maturation to X. fastidios and colonization and attenuate virulence. *Mol Plant Microbe Interact* **18**: 856-868.

Guo, D.C., Y. Sun, A.Q. Zhang, J.S. Liu, Y. Lu, P.X. Liu, D.W. Yuan, Q. Jiang, C.D. Si & L.D. Qu, (2014) Construction and Virulence of Filamentous Hemagglutinin Protein B1 Mutant of Pasteurella multocida in Chickens. *J Integr Agr* **13**: 2268-2275.

Gygi, D., M.J. Bailey, C. Allison & C. Hughes, (1995) Requirement for FlhA in flagella assembly and swarm-cell differentiation by Proteus mirabilis. *Mol Microbiol* **15**: 761-769.

Haines, S., N. Arnaud-Barbe, D. Poncet, S. Reverchon, J. Wawrzyniak, W. Nasser & G. Renauld-Mongenie, (2015) IscR Regulates Synthesis of Colonization Factor Antigen I Fimbriae in Response to Iron Starvation in Enterotoxigenic Escherichia coli. *J Bacteriol* **197**: 2896-2907.

Harrison, A., E.A. Santana, B.R. Szelestey, D.E. Newsom, P. White & K.M. Mason, (2013) Ferric uptake regulator and its role in the pathogenesis of nontypeable Haemophilus influenzae. *Infect Immun* **81**: 1221-1233.

Hirono, I., N. Tange & T. Aoki, (1997) Iron-regulated haemolysin gene from Edwardsiella tarda. *Mol Microbiol* **24**: 851-856.

Holley, C., D. Gangaiah, W. Li, K.R. Fortney, D.M. Janowicz, S. Ellinger, B. Zwickl, B.P. Katz & S.M. Spinola, (2014) A (p)ppGpp-null mutant of Haemophilus ducreyi is partially attenuated in humans due to multiple conflicting phenotypes. *Infect Immun* **82**: 3492-3502.

Holley, C.L., X. Zhang, K.R. Fortney, S. Ellinger, P. Johnson, B. Baker, Y. Liu, D.M. Janowicz, B.P. Katz, R.S. Munson, Jr. & S.M. Spinola, (2015) DksA and (p)ppGpp have unique and overlapping contributions to Haemophilus ducreyi pathogenesis in humans. *Infect Immun* **83**: 3281-3292.

Huber, P., P. Basso, E. Reboud & I. Attree, (2016) Pseudomonas aeruginosa renews its virulence factors. *Environ Microbiol Rep* **8**: 564-571.

Jacob-Dubuisson, F., J. Guerin, S. Baelen & B. Clantin, (2013) Two-partner secretion: as simple as it sounds? *Res Microbiol* **164**: 583-595.

Jiang, X., P. Sobetzko, W. Nasser, S. Reverchon & G. Muskhelishvili, (2015) Chromosomal "stress-response" domains govern the spatiotemporal expression of the bacterial virulence program. *mBio* **6**: e00353-00315.

Jubelin, G., A. Lanois, D. Severac, S. Rialle, C. Longin, S. Gaudriault & A. Givaudan, (2013) FliZ is a global regulatory protein affecting the expression of flagellar and virulence genes in individual Xenorhabdus nematophila bacterial cells. *PLoS Genet* **9**: e1003915.

Jubelin, G., S. Pages, A. Lanois, M.H. Boyer, S. Gaudriault, J.B. Ferdy & A. Givaudan, (2011) Studies of the dynamic expression of the Xenorhabdus FliAZ regulon reveal atypical iron-dependent regulation of the flagellin and haemolysin genes during insect infection. *Environ Microbiol* **13**: 1271-1284.

Kida, Y., Y. Higashimoto, H. Inoue, T. Shimizu & K. Kuwano, (2008) A novel secreted protease from Pseudomonas aeruginosa activates NF-kappaB through protease-activated receptors. *Cell Microbiol* **10**: 1491-1504.

Kida, Y., T. Shimizu & K. Kuwano, (2011) Cooperation between LepA and PlcH contributes to the in vivo virulence and growth of Pseudomonas aeruginosa in mice. *Infect Immun* **79**: 211-219.

Klockgether, J. & B. Tummler, (2017) Recent advances in understanding Pseudomonas aeruginosa as a pathogen. *F1000Res* **6**: 1261.

Konkel, M.E. & K. Tilly, (2000) Temperature-regulated expression of bacterial virulence genes. *Microbes Infect* **2**: 157-166.

Krasauskas, R., J. Skerniskyte, J. Armalyte & E. Suziedeliene, (2019) The role of Acinetobacter baumannii response regulator BfmR in pellicle formation and competitiveness via contact-dependent inhibition system. *BMC Microbiol* **19**: 241.

Kyrkou, I., T. Pusa, L. Ellegaard-Jensen, M.F. Sagot & L.H. Hansen, (2018) Pierce's Disease of Grapevines: A Review of Control Strategies and an Outline of an Epidemiological Model. *Front Microbiol* **9**: 2141.

Labandeira-Rey, M., D.A. Dodd, C.A. Brautigam, K.R. Fortney, S.M. Spinola & E.J. Hansen, (2013) The Haemophilus ducreyi Fis protein is involved in controlling expression of the lspB-lspA2 operon and other virulence factors. *Infect Immun* **81**: 4160-4170.

Labandeira-Rey, M., J.R. Mock & E.J. Hansen, (2009) Regulation of expression of the Haemophilus ducreyi LspB and LspA2 proteins by CpxR. *Infect Immun* **77**: 3402-3411.

Lai, H.C., P.C. Soo, J.R. Wei, W.C. Yi, S.J. Liaw, Y.T. Horng, S.M. Lin, S.W. Ho, S. Swift & P. Williams, (2005) The RssAB two-component signal transduction system in Serratia marcescens regulates swarming motility and cell envelope architecture in response to exogenous saturated fatty acids. *J Bacteriol* **187**: 3407-3414.

Lanois, A., G. Jubelin & A. Givaudan, (2008) FliZ, a flagellar regulator, is at the crossroads between motility, haemolysin expression and virulence in the insect pathogenic bacterium Xenorhabdus. *Mol Microbiol* **68**: 516-533.

Lautier, T. & W. Nasser, (2007) The DNA nucleoid-associated protein Fis co-ordinates the expression of the main virulence genes in the phytopathogenic bacterium Erwinia chrysanthemi. *Mol Microbiol* **66**: 1474-1490.

Leo, J.C., I. Grin & D. Linke, (2012) Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philos Trans R Soc Lond B Biol Sci* **367**: 1088-1101.

Lesne, E., E. Dupre, M.F. Lensink, C. Locht, R. Antoine & F. Jacob-Dubuisson, (2018) Coiled-Coil Antagonism Regulates Activity of Venus Flytrap-Domain-Containing Sensor Kinases of the BvgS Family. *mBio* **9**.

Lin, C.S., J.T. Horng, C.H. Yang, Y.H. Tsai, L.H. Su, C.F. Wei, C.C. Chen, S.C. Hsieh, C.C. Lu & H.C. Lai, (2010) RssAB-FlhDC-ShlBA as a major pathogenesis pathway in Serratia marcescens. *Infect Immun* **78**: 4870-4881.

Lin, C.S., Y.H. Tsai, C.J. Chang, S.F. Tseng, T.R. Wu, C.C. Lu, T.S. Wu, J.J. Lu, J.T. Horng, J. Martel, D.M. Ojcius, H.C. Lai & J.D. Young, (2016) An iron detection system determines bacterial swarming initiation and biofilm formation. *Sci Rep* **6**: 36747.

Liu, J.H., M.J. Lai, S. Ang, J.C. Shu, P.C. Soo, Y.T. Horng, W.C. Yi, H.C. Lai, K.T. Luh, S.W. Ho & S. Swift, (2000) Role of flhDC in the expression of the nuclease gene nucA, cell division and flagellar synthesis in Serratia marcescens. *J Biomed Sci* **7**: 475-483.

Llamas, M.A., F. Imperi, P. Visca & I.L. Lamont, (2014) Cell-surface signaling in Pseudomonas: stress responses, iron transport, and pathogenicity. *FEMS Microbiol Rev* **38**: 569-597.

Llamas, M.A., A. van der Sar, B.C. Chu, M. Sparrius, H.J. Vogel & W. Bitter, (2009) A Novel extracytoplasmic function (ECF) sigma factor regulates virulence in Pseudomonas aeruginosa. *PLoS Pathog* **5**: e1000572.

Lory, S., M. Wolfgang, V. Lee & R. Smith, (2004) The multi-talented bacterial adenylate cyclases. *Int J Med Microbiol* **293**: 479-482.

Majerczyk, C., M. Brittnacher, M. Jacobs, C.D. Armour, M. Radey, E. Schneider, S. Phattarasokul, R. Bunt & E.P. Greenberg, (2014a) Global analysis of the Burkholderia thailandensis quorum sensing-controlled regulon. *J Bacteriol* **196**: 1412-1424.

Majerczyk, C.D., M.J. Brittnacher, M.A. Jacobs, C.D. Armour, M.C. Radey, R. Bunt, H.S. Hayden, R. Bydalek & E.P. Greenberg, (2014b) Cross-species comparison of the Burkholderia pseudomallei, Burkholderia thailandensis, and Burkholderia mallei quorum-sensing regulons. *J Bacteriol* **196**: 3862-3871.

Matsuyama, B.Y., P.V. Krasteva, C. Baraquet, C.S. Harwood, H. Sondermann & M.V. Navarro, (2016) Mechanistic insights into c-di-GMP-dependent control of the biofilm regulator FleQ from Pseudomonas aeruginosa. *Proc Natl Acad Sci U S A* **113**: E209-218.

May, B.J., Q. Zhang, L.L. Li, M.L. Paustian, T.S. Whittam & V. Kapur, (2001) Complete genomic sequence of Pasteurella multocida, Pm70. *Proc Natl Acad Sci U S A* **98**: 3460-3465.
McCann, J.R. & J.W. St Geme, 3rd, (2014) The HMW1C-like glycosyltransferases--an enzyme family with a sweet tooth for simple sugars. *PLoS Pathog* **10**: e1003977.
Megroz, M., O. Kleifeld, A. Wright, D. Powell, P. Harrison, B. Adler, M. Harper & J.D. Boyce, (2016) The RNA-Binding Chaperone Hfq Is an Important Global Regulator of Gene Expression in Pasteurella multocida and Plays a Crucial Role in Production of a Number of Virulence Factors, Including Hyaluronic Acid Capsule. *Infect Immun* **84**: 1361-1370.
Melvin, J.A., E.V. Scheller, C.R. Noel & P.A. Cotter, (2015) New Insight into Filamentous Hemagglutinin Secretion Reveals a Role for Full-Length FhaB in Bordetella Virulence. *mBio* **6**.
Mercy, C., B. Ize, S.P. Salcedo, S. de Bentzmann & S. Bigot, (2016) Functional Characterization of Pseudomonas Contact Dependent Growth Inhibition (CDI) Systems. *PLoS One* **11**: e0147435.
Meuskens, I., A. Saragliadis, J.C. Leo & D. Linke, (2019) Type V Secretion Systems: An Overview of Passenger Domain Functions. *Front Microbiol* **10**: 1163.
Mikkelsen, H., M. Sivaneson & A. Filloux, (2011) Key two-component regulatory systems that control biofilm formation in Pseudomonas aeruginosa. *Environ Microbiol* **13**: 1666-1681.
Moon, K., R.P. Bonocora, D.D. Kim, Q. Chen, J.T. Wade, S. Stibitz & D.M. Hinton, (2017) The BvgAS Regulon of Bordetella pertussis. *mBio* **8**.
Moradali, M.F., S. Ghods & B.H. Rehm, (2017) Pseudomonas aeruginosa Lifestyle: A Paradigm for Adaptation, Survival, and Persistence. *Front Cell Infect Microbiol* **7**: 39.
Morgenstein, R.M. & P.N. Rather, (2012) Role of the Umo proteins and the Rcs phosphorelay in the swarming motility of the wild type and an O-antigen (waaL) mutant of Proteus mirabilis. *J Bacteriol* **194**: 669-676.
Morse, R.P., K.C. Nikolakakis, J.L. Willett, E. Gerrick, D.A. Low, C.S. Hayes & C.W. Goulding, (2012) Structural basis of toxicity and immunity in contact-dependent growth inhibition (CDI) systems. *Proc Natl Acad Sci U S A* **109**: 21480-21485.
Navarro, M.V., P.D. Newell, P.V. Krasteva, D. Chatterjee, D.R. Madden, G.A. O'Toole & H. Sondermann, (2011) Structural basis for c-di-GMP-mediated inside-out signaling controlling periplasmic proteolysis. *PLoS Biol* **9**: e1000588.
Neil, R.B. & M.A. Apicella, (2009) Role of HrpA in biofilm formation of Neisseria meningitidis and regulation of the hrpBAS transcripts. *Infect Immun* **77**: 2285-2293.
Nelson, K.M., G.M. Young & V.L. Miller, (2001) Identification of a locus involved in systemic dissemination of Yersinia enterocolitica. *Infect Immun* **69**: 6201-6208.
Newell, P.D., C.D. Boyd, H. Sondermann & G.A. O'Toole, (2011) A c-di-GMP effector system controls cell adhesion by inside-out signaling and surface protein cleavage. *PLoS Biol* **9**: e1000587.
Nikolakakis, K., S. Amber, J.S. Wilbur, E.J. Diner, S.K. Aoki, S.J. Poole, A. Tuanyok, P.S. Keim, S. Peacock, C.S. Hayes & D.A. Low, (2012) The toxin/immunity network of Burkholderia pseudomallei contact-dependent growth inhibition (CDI) systems. *Mol Microbiol* **84**: 516-529.
Otero-Asman, J.R., J.M. Quesada, K.K. Jim, A. Ocampo-Sosa, C. Civantos, W. Bitter & M.A. Llamas, (2020) The extracytoplasmic function sigma factor sigma(VreI) is active during infection and contributes to phosphate starvation-induced virulence of Pseudomonas aeruginosa. *Sci Rep* **10**: 3139.
Ouafa, Z.A., S. Reverchon, T. Lautier, G. Muskhelishvili & W. Nasser, (2012) The nucleoid-associated proteins H-NS and FIS modulate the DNA supercoiling response of the pel genes, the major virulence factors in the plant pathogen bacterium Dickeya dadantii. *Nucleic Acids Res* **40**: 4306-4319.
Poole, K. & V. Braun, (1988a) Influence of growth temperature and lipopolysaccharide on hemolytic activity of Serratia marcescens. *J Bacteriol* **170**: 5146-5152.
Poole, K. & V. Braun, (1988b) Iron regulation of Serratia marcescens hemolysin gene expression. *Infect Immun* **56**: 2967-2971.
Poole, K., E. Schiebel & V. Braun, (1988) Molecular characterization of the hemolysin determinant of Serratia marcescens. *J Bacteriol* **170**: 3177-3188.

Quesada, J.M., J.R. Otero-Asman, K.C. Bastiaansen, C. Civantos & M.A. Llamas, (2016) The Activity of the Pseudomonas aeruginosa Virulence Regulator sigma(VreI) Is Modulated by the Anti-sigma Factor VreR and the Transcription Factor PhoB. *Front Microbiol* **7**: 1159.

Rao, P.S., Y. Yamada, Y.P. Tan & K.Y. Leung, (2004) Use of proteomics to identify novel virulence determinants that are required for Edwardsiella tarda pathogenesis. *Mol Microbiol* **53**: 573-586.

Richards, G.R. & H. Goodrich-Blair, (2009) Masters of conquest and pillage: Xenorhabdus nematophila global regulators control transitions from virulence to nutrient acquisition. *Cell Microbiol* **11**: 1025-1033.

Richards, G.R., E.E. Herbert, Y. Park & H. Goodrich-Blair, (2008) Xenorhabdus nematophila lrhA is necessary for motility, lipase activity, toxin expression, and virulence in Manduca sexta insects. *J Bacteriol* **190**: 4870-4879.

Rinker, S.D., M.P. Trombley, X. Gu, K.R. Fortney & M.E. Bauer, (2011) Deletion of mtrC in Haemophilus ducreyi increases sensitivity to human antimicrobial peptides and activates the CpxRA regulon. *Infect Immun* **79**: 2324-2334.

Rojas, C.M., J.H. Ham, W.L. Deng, J.J. Doyle & A. Collmer, (2002) HecA, a member of a class of adhesins produced by diverse pathogenic bacteria, contributes to the attachment, aggregation, epidermal cell killing, and virulence phenotypes of Erwinia chrysanthemi EC16 on Nicotiana clevelandii seedlings. *Proc Natl Acad Sci U S A* **99**: 13142-13147.

Romling, U., M.Y. Galperin & M. Gomelsky, (2013) Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol Mol Biol Rev* **77**: 1-52.

Roussin, M., S. Rabarioelina, L. Cluzeau, J. Cayron, C. Lesterlin, S.P. Salcedo & S. Bigot, (2019) Identification of a Contact-Dependent Growth Inhibition (CDI) System That Reduces Biofilm Formation and Host Cell Adhesion of Acinetobacter baumannii DSM30011 Strain. *Front Microbiol* **10**: 2450.

Roy, K., G.M. Hilliard, D.J. Hamilton, J. Luo, M.M. Ostmann & J.M. Fleckenstein, (2009) Enterotoxigenic Escherichia coli EtpA mediates adhesion between flagella and host cells. *Nature* **457**: 594-598.

Ruer, S., G. Ball, A. Filloux & S. de Bentzmann, (2008) The 'P-usher', a novel protein transporter involved in fimbrial assembly and TpsA secretion. *EMBO J* **27**: 2669-2680.

Ruhe, Z.C., D.A. Low & C.S. Hayes, (2013) Bacterial contact-dependent growth inhibition. *Trends Microbiol* **21**: 230-237.

Rybtke, M., J. Berthelsen, L. Yang, N. Hoiby, M. Givskov & T. Tolker-Nielsen, (2015) The LapG protein plays a role in Pseudomonas aeruginosa biofilm formation by controlling the presence of the CdrA adhesin on the cell surface. *Microbiologyopen* **4**: 917-930.

Saponari, M., A. Giampetruzzi, G. Loconsole, D. Boscia & P. Saldarelli, (2019) Xylella fastidiosa in Olive in Apulia: Where We Stand. *Phytopathology* **109**: 175-186.

Scarlato, V., A. Prugnola, B. Arico & R. Rappuoli, (1991) The bvg-dependent promoters show similar behaviour in different Bordetella species and share sequence homologies. *Mol Microbiol* **5**: 2493-2498.

Scheller, E.V. & P.A. Cotter, (2015) Bordetella filamentous hemagglutinin and fimbriae: critical adhesins with unrealized vaccine potential. *Pathog Dis* **73**: ftv079.

Schmitt, C., D. Turner, M. Boesl, M. Abele, M. Frosch & O. Kurzai, (2007) A functional two-partner secretion system contributes to adhesion of Neisseria meningitidis to epithelial cells. *J Bacteriol* **189**: 7968-7976.

Sivaneson, M., H. Mikkelsen, I. Ventre, C. Bordi & A. Filloux, (2011) Two-component regulatory systems in Pseudomonas aeruginosa: an intricate network mediating fimbrial and efflux pump gene expression. *Mol Microbiol* **79**: 1353-1366.

Skals, M. & H.A. Praetorius, (2013) Mechanisms of cytolysin-induced cell damage -- a role for auto- and paracrine signalling. *Acta Physiol (Oxf)* **209**: 95-113.

Sobran, M.A. & P.A. Cotter, (2019) The BvgS PAS Domain, an Independent Sensory Perception Module in the Bordetella bronchiseptica BvgAS Phosphorelay. *J Bacteriol* **201**.

St Geme, J.W., 3rd, S. Falkow & S.J. Barenkamp, (1993) High-molecular-weight proteins of nontypable Haemophilus influenzae mediate attachment to human epithelial cells. *Proc Natl Acad Sci U S A* **90**: 2875-2879.

St Geme, J.W., 3rd & S. Grass, (1998) Secretion of the Haemophilus influenzae HMW1 and HMW2 adhesins involves a periplasmic intermediate and requires the HMWB and HMWC proteins. *Mol Microbiol* **27**: 617-630.

Steen, J.A., P. Harrison, T. Seemann, I. Wilkie, M. Harper, B. Adler & J.D. Boyce, (2010) Fis is essential for capsule production in Pasteurella multocida and regulates expression of other important virulence factors. *PLoS Pathog* **6**: e1000750.

Stella, N.A., K.M. Brothers, J.D. Callaghan, A.M. Passerini, C. Sigindere, P.J. Hill, X. Liu, D.J. Wozniak & R.M.Q. Shanks, (2018) An IgaA/UmoB Family Protein from Serratia marcescens Regulates Motility, Capsular Polysaccharide Biosynthesis, and Secondary Metabolite Production. *Appl Environ Microbiol* **84**.

Tala, A., C. Progida, M. De Stefano, L. Cogli, M.R. Spinosa, C. Bucci & P. Alifano, (2008) The HrpB-HrpA two-partner secretion system is essential for intracellular survival of Neisseria meningitidis. *Cell Microbiol* **10**: 2461-2482.

Tatum, F.M., A.G. Yersin & R.E. Briggs, (2005) Construction and virulence of a Pasteurella multocida fhaB2 mutant in turkeys. *Microb Pathog* **39**: 9-17.

Tommassen, J. & J. Arenas, (2017) Biological Functions of the Secretome of Neisseria meningitidis. *Front Cell Infect Microbiol* **7**: 256.

Trouillon, J., E. Sentausa, M. Ragno, M. Robert-Genthon, S. Lory, I. Attree & S. Elsen, (2020) Species-specific recruitment of transcription factors dictates toxin expression. *Nucleic Acids Res* **48**: 2388-2400.

Tsai, Y.H., J.R. Wei, C.S. Lin, P.H. Chen, S. Huang, Y.C. Lin, C.F. Wei, C.C. Lu & H.C. Lai, (2011) RssAB signaling coordinates early development of surface multicellularity in Serratia marcescens. *PLoS One* **6**: e24154.

Uphoff, T.S. & R.A. Welch, (1990) Nucleotide sequencing of the Proteus mirabilis calcium-independent hemolysin genes (hpmA and hpmB) reveals sequence similarity with the Serratia marcescens hemolysin genes (shlA and shlB). *J Bacteriol* **172**: 1206-1216.

Vakevainen, M., S. Greenberg & E.J. Hansen, (2003) Inhibition of phagocytosis by Haemophilus ducreyi requires expression of the LspA1 and LspA2 proteins. *Infect Immun* **71**: 5994-6003.

Valentini, M. & A. Filloux, (2019) Multiple Roles of c-di-GMP Signaling in Bacterial Pathogenesis. *Annu Rev Microbiol* **73**: 387-406.

Vogel, J. & B.F. Luisi, (2011) Hfq and its constellation of RNA. *Nat Rev Microbiol* **9**: 578-589.

Wang, F., M. Zhang, Y.H. Hu, W.W. Zhang & L. Sun, (2009) Regulation of the Edwardsiella tarda hemolysin gene and luxS by EthR. *J Microbiol Biotechnol* **19**: 765-773.

Wang, N., J.L. Li & S.E. Lindow, (2012) RpfF-dependent regulon of Xylella fastidiosa. *Phytopathology* **102**: 1045-1053.

Wang, X., Q. Wang, J. Xiao, Q. Liu, H. Wu & Y. Zhang, (2010) Hemolysin EthA in Edwardsiella tarda is essential for fish invasion in vivo and in vitro and regulated by two-component system EsrA-EsrB and nucleoid protein HhaEt. *Fish Shellfish Immunol* **29**: 1082-1091.

Ward, C.K., J.R. Mock & E.J. Hansen, (2004) The LspB protein is involved in the secretion of the LspA1 and LspA2 proteins by Haemophilus ducreyi. *Infect Immun* **72**: 1874-1884.

Weinberg, E.D., (2009) Iron availability and infection. *Biochim Biophys Acta* **1790**: 600-605.

West, S.E., A.K. Sample & L.J. Runyen-Janecky, (1994) The vfr gene product, required for Pseudomonas aeruginosa exotoxin A and protease production, belongs to the cyclic AMP receptor protein family. *J Bacteriol* **176**: 7532-7542.

Whitby, P.W., T.M. Vanwagoner, T.W. Seale, D.J. Morton & T.L. Stull, (2006) Transcriptional profile of Haemophilus influenzae: effects of iron and heme. *J Bacteriol* **188**: 5640-5645.

Willems, R.J., H.G. van der Heide & F.R. Mooi, (1992) Characterization of a Bordetella pertussis fimbrial gene cluster which is located directly downstream of the filamentous haemagglutinin gene. *Mol Microbiol* **6**: 2661-2671.

Wilson, B.A. & M. Ho, (2013) Pasteurella multocida: from zoonosis to cellular microbiology. *Clin Microbiol Rev* **26**: 631-655.

Winter, L.E. & S.J. Barenkamp, (2016) Naturally Acquired HMW1- and HMW2-Specific Serum Antibodies in Adults and Children Mediate Opsonophagocytic Killing of Nontypeable Haemophilus influenzae. *Clin Vaccine Immunol* **23**: 37-46.

Wolfgang, M.C., V.T. Lee, M.E. Gilmore & S. Lory, (2003) Coordinate regulation of bacterial virulence genes by a novel adenylate cyclase-dependent signaling pathway. *Dev Cell* **4**: 253-263.

Zheng, J., S.L. Tung & K.Y. Leung, (2005) Regulation of a type III and a putative secretion system in Edwardsiella tarda by EsrC is under the control of a two-component system, EsrA-EsrB. *Infect Immun* **73**: 4127-4137.

4.2 The family of XRE-cupin regulators

4.2.1 The XRE-cupin regulators

Julian Trouillon✉, Michel Ragno, Victor Simon, Ina Attrée & Sylvie Elsen✉
Posted on *bioRxiv* in 2020. Currently in revision in *mSystems*.

While studying ErfA and ErgAB, I also got interested into similar regulators in order to help me better understand them. Early on, I identified several other "ErfA-like" regulators in *P. aeruginosa*, including PsdR and PauR, two TFs studied for their roles in dipeptides and polyamines metabolism, respectively. At that time, my project on ErfA was close to its end and I was starting to grasp the potential of DAP-seq for the large scale study of entire families of TFs. As ErfA and ErgAB still left us with unanswered questions, we decided to study the entire family of 8 TFs sharing ErfA domain architecture - XRE and Cupin domains - in an attempt to better understand this type of regulators. To that aim, I combined DAP-seq and RNA-seq to determine the regulons of these TFs and found that they are local, specialized TFs that act as inhibitors of metabolism-related genes. The analysis of their conservation also revealed that this family of TFs is widespread and diverse across the *Pseudomonas* genus, comprising TFs that locally cluster with metabolic genes. Finally, we were planning on studying in more details the different phenotypes and mechanisms of signal sensing of these TFs but most planned experiments could unfortunately not be done due to the covid-19 lockdown.
In this work, I performed all experiments and analyses except for RT-qPCRs (SE), some of the clonings (SE), the antibiotic assay (VS) and about half of the protein purifications (MR).

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

Transcription inhibitors with XRE DNA-binding and cupin signal-sensing domains drive metabolic diversification in *Pseudomonas*

Julian Trouillon[1#], Michel Ragno[1], Victor Simon[1], Ina Attrée[1], Sylvie Elsen[1#]

[1]Université Grenoble Alpes, CNRS ERL5261, CEA BIG-BCI, INSERM UMR1036, Grenoble, France

Address correspondence to julian.trouillon@cea.fr; sylvie.elsen@cea.fr

Running Title: The XRE-cupin inhibitors family of *Pseudomonas*

KEYWORDS : Pseudomonas aeruginosa, transcription factors, XRE-cupin, RNA-seq, DAP-seq

ABSTRACT

Transcription factors (TFs) are instrumental in the bacterial response to new environmental conditions. They can act as direct signal sensors and subsequently induce changes in gene expression leading to physiological adaptation. Here, by combining RNA-seq and DAP-seq, we studied a family of eight TFs in *Pseudomonas aeruginosa*. This family, encompassing TFs with XRE-like DNA-binding and cupin signal-sensing domains, includes the metabolic regulators ErfA, PsdR and PauR and five so far unstudied TFs. The genome-wide delineation of their regulons identified 39 regulatory interactions with genes mostly involved in metabolism. We found that the XRE-cupin TFs are inhibitors of their neighboring genes, forming local, functional units encoding proteins with functions in condition-specific metabolic pathways. The phylogenetic analysis of this family of regulators across the *Pseudomonas* genus revealed a wide diversity of such metabolic regulatory modules and identified species with potentially higher metabolic versatility. Numerous uncharacterized XRE-cupin TFs were found near metabolism-related genes, illustrating the need of further systematic characterization of transcriptional regulatory networks in order to better understand the mechanisms of bacterial adaptation to new environments.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

IMPORTANCE

Bacteria of the *Pseudomonas* genus, including the major human pathogen *P. aeruginosa,* are known for their complex regulatory networks and high number of transcription factors, which contribute to their impressive adaptive ability. However, even in the most studied species, most of the regulators are still uncharacterized. With the recent advances in high-throughput sequencing methods, it is now possible to fill this knowledge gap and help understanding how bacteria adapt and thrive in new environments. By leveraging these methods, we provide an example of a comprehensive analysis of an entire family of transcription factors and bring new insights into metabolic and regulatory adaptation in the *Pseudomonas* genus.

INTRODUCTION

The regulation of gene transcription is a key mechanism in the evolution and adaptation of bacteria to a wide array of environments. Transcriptional regulation involves the interaction between *trans*-acting regulatory proteins called transcription factors (TFs) and *cis*-regulatory elements found on the promoters of regulated genes. Many external and internal signals, such as metabolite concentrations, oxygen, temperature, pH levels or surface contact, are integrated into regulatory networks through different sensing mechanisms in order to provide an appropriate transcriptional response allowing the bacterium to adapt effectively to the new environment. TFs are classified in several families depending on the structural similarities in their DNA-binding domains (DBDs) which can either stand alone or be present in multi-domain proteins containing variable arrangements with sensing or oligomerization domains (Ulrich *et al.*, 2005). This allows the integration of countless different input and output signals into regulatory networks. Although TFs are classified in defined families (Sanchez *et al.*, 2020), they each control specific sets of genes to coordinate bacterial responses, and the targets and

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

functions of each newly identified TF cannot be inferred but need to be experimentally established. One-component systems (OCS) are the most diverse TFs that can directly sense intracellular or imported extracellular signals through their signal-sensing domain and transmit the signal, probably through a conformational change, to their DBD (Ulrich *et al.*, 2005). Here we explored one family of OCS in *Pseudomonas aeruginosa,* a gram-negative opportunistic human pathogen.

P. aeruginosa possesses a considerable metabolic versatility and one of the most complex regulatory networks found in the bacterial kingdom, with roughly 10% of all genes dedicated to transcription regulation (Stover *et al.*, 2000, Rodrigue *et al.*, 2000, Galan-Vasquez *et al.*, 2011). This regulatory network contains about 500 predicted TFs. While some regulatory pathways have been thoroughly studied, the vast majority of *P. aeruginosa* TFs are still uncharacterized. With the development of NGS-based methods dedicated to the genome-wide characterization of TFs, the view of complex interplays between and inside different families of TFs is starting to emerge (Huang *et al.*, 2019, Rajeev *et al.*, 2020). We recently characterized a transcriptional inhibitor, ErfA, recruited to down-regulate the expression of an horizontally transferred operon encoding a major virulence factor, ExlA, specifically in one *P. aeruginosa* lineage (Trouillon *et al.*, 2020). ErfA belongs to the large family of regulators with a XRE-like DBD that usually binds their DNA targets as homodimers. Besides its N-terminal DBD, ErfA possesses a C-terminal putative sensor domain with a “cupin” fold (Trouillon *et al.*, 2020). While ErfA was rewired to also regulate the horizontally acquired *exlBA* operon in *P. aeruginosa*, its main conserved target across *Pseudomonas* species is a metabolic operon, located adjacent to the *erfA* gene, with no role in bacterial virulence. ErfA was proposed to act as a sensory switch on this transcription unit in response to yet unknown conditions. While signal-sensing TFs are widely distributed in bacteria (Ulrich *et al.*, 2005), TFs sharing ErfA specific domain architecture

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

have not been studied, raising the question of whether ErfA-like regulators could be common and important in metabolic and/or virulence regulation.

In this work, we focused on the subfamily of eight *P. aeruginosa* TFs that share ErfA architecture in order to examine their genome-wide regulatory targets. In addition to ErfA (Trouillon *et al.*, 2020), the family comprises two members with incompletely-defined regulons and five uncharacterized TFs. The combination of RNA-seq and DAP-seq approaches allowed us to define the regulons of six of these TFs. Each family member has specific targets ranging from one to twelve binding sites in promoters of genes related to small molecule uptake or processing. Regulators with XRE-cupin domains were found as local, specialized inhibitors that are widespread across the *Pseudomonas* genus. While many uncharacterized XRE-cupin TFs were identified, some species (such as *P. putida*) harbored up to ten of these regulators whereas others (i.e. *P. stutzeri*) possess only on, potentially reflecting different metabolic versatilities.

RESULTS

Eight TFs of the XRE-like family share similar architectures. The *P. aeruginosa* PAO1 genome encodes 19 proteins with a helix-turn-helix DNA-binding motif similar to that of the CI and Cro repressors of the phage λ that features the members of the XRE-like family (Wood *et al.*, 1990, Ohlendorf *et al.*, 1983) (Figure 1A). The XRE-like domains of about 60 residues adopt a well-characterized helical conformation with helices 2 and 3 involved in DNA binding (Figure 1B). The DBD was found alone or associated to either a peptidase S24 domain (AlpR and PrtR) or to a predicted sensor cupin domain at the C-terminus of proteins (Figure 1C). The versatile cupin domain folds into a 6-stranded β-barrel and is associated to a wide range of function (Dunwell

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

et al., 2004). In addition to ErfA (PA0225), seven TFs share the XRE-cupin architecture with only two having partially attributed functions: PA0535, PA1359, PA1884, PA2312, PA4499 (PsdR), PA4987 and PA5301 (PauR). Using targeted approaches, PsdR and PauR were previously shown to regulate dipeptide metabolism (Kiely *et al.*, 2008, Asfahl *et al.*, 2015) and polyamines metabolism (Chou *et al.*, 2013), respectively. By regulating independent metabolic pathways, these two TFs play an essential role for bacterial growth in their corresponding environmental conditions, which predicts potentially important functions for the five so far unstudied TFs of the family.

TFs of the XRE-cupin family are local, highly specialized repressors. To get a global view on target specificities and regulatory networks of *P. aeruginosa* XRE-cupin TFs, we undertook to determine the regulons of all members of this family, except for ErfA that we already comprehensively studied (Trouillon *et al.*, 2020). To that aim, the corresponding genes were deleted in the genome of PAO1 and the transcriptomes of the engineered mutants were compared to that of the wild-type strain by RNA-seq. In parallel, we expressed and purified the seven recombinant regulators, and their targets were identified *in vitro* on fragmented *P. aeruginosa* PAO1 genome by DAP-seq (Bartlett *et al.*, 2017, Trouillon *et al.*, 2020). Altogether, we determined transcriptomes and direct DNA targets for six of the seven regulators (Figure 2A, Tables S3 and S4). The PA2312 protein was less soluble and stable than any of the other proteins and was thus probably inactive and/or aggregated resulting in no significant peak detected in DAP-seq. In addition, the *PA2312* gene seems not expressed in the tested condition, as revealed by the few reads covering the CDS observed by RNA-seq in the wild-type strain. Overall, the family of XRE-cupin TFs regulates 39 genes in the PAO1 strain through direct binding to 19 genomic regions in total (Figure 2B).

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

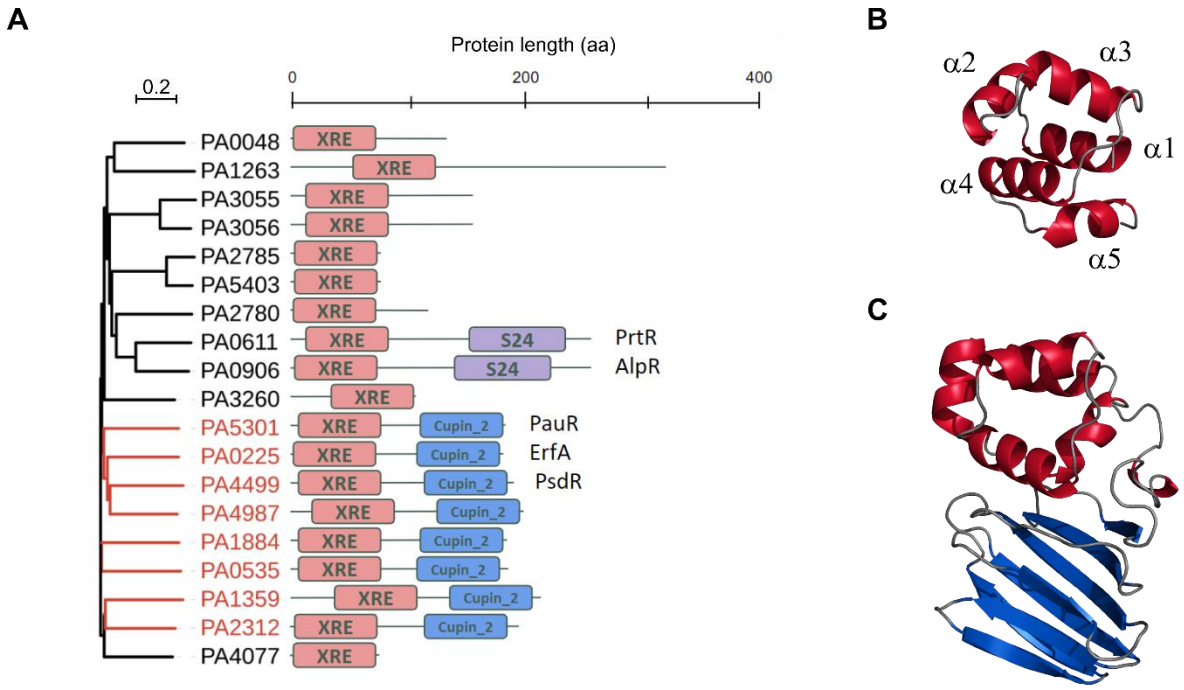


FIGURE 1. *P. aeruginosa* TFs belonging to the XRE family. (**A**) Maximum likelihood phylogenetic tree of the 19 proteins possessing a XRE DNA-binding domain in *P. aeruginosa* PAO1 (Winsor *et al.*, 2016). The leaves of the eight XRE-cupin regulators are colored in red in the tree. (**B**) Stereo ribbon representation of the model of the XRE DNA-binding domain of ErfA. The five alpha helices are annotated. (**C**) Stereo ribbon representation of the model of ErfA structure. The XRE domain is in red and the cupin domain in blue. Model prediction was done using the SWISS-Model tool (Biasini *et al.*, 2014) using the structure with PDB ID 1Y9Q as template.

In all cases, there was a strong correlation between *in vivo* and *in vitro* targets; the majority of DAP-seq peaks were centered on the promoters of genes that were also significantly dysregulated in RNA-seq. Most binding events occurred in the core promoter regions of regulated genes (Figure 3A) and each target gene was found upregulated in the mutant of the corresponding regulator (Figure 3B), showing that the XRE-cupin regulators are inhibitors of transcription. Our analyses confirmed the known regulon of PsdR and extended our knowledge on PauR, as discussed below. For most regulators, we found one or two direct targets, with the exception of PauR (PA5301) that directly regulates 29 genes (Figure 2B). Some genes were found slightly downregulated in all mutants with no corresponding binding sites found, probably due to experimental conditions or genetic manipulations (Figure 2A).

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

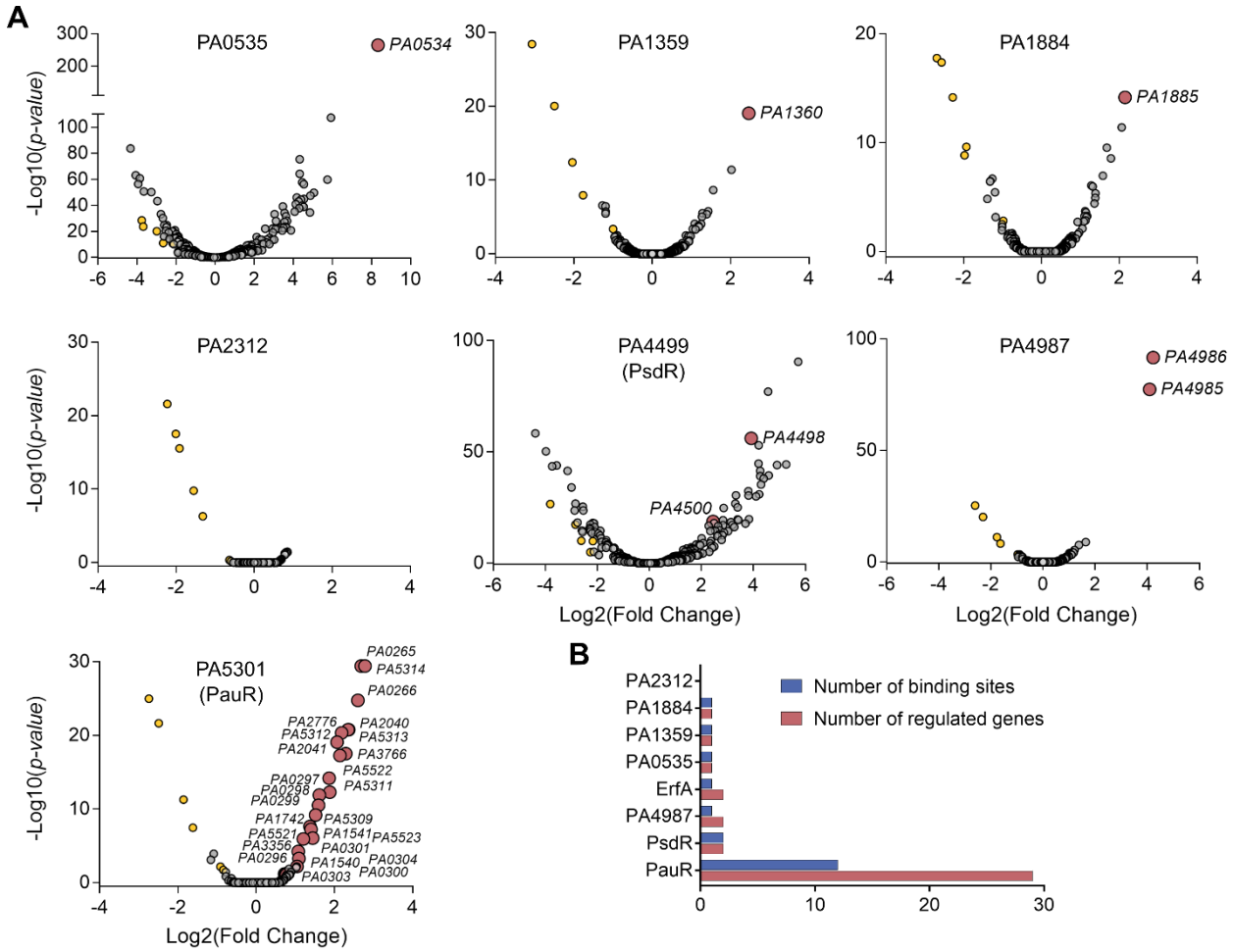


FIGURE 2. Determination of the regulons of the XRE-cupin regulators. (**A**) Volcano plots displaying the RNA-seq results of the genes differentially expressed in the respective XRE-cupin mutants versus parental strain PAO1. Genes for which a DAP-seq binding peak was identified in the promoter are represented by red circles and annotated with their gene ID. The five genes (yellow circles) found commonly downregulated in all mutants represent artefacts probably due to genetic manipulation. (**B**) Summary of number of regulatory targets per TF in *P. aeruginosa* PAO1.

At least one regulatory target per TF was further verified by RT-qPCR and electrophoretic mobility shift assays (EMSA) and in all cases the RNA-seq and DAP-seq results were confirmed by these targeted approaches (Figure 4).

Strikingly, all the XRE-cupin TFs bind to at least one intergenic region directly adjacent to their own gene (Figure 5). PauR strikes as an outlier with its 29 regulated genes scattered across twelve different genomic locations, while all other XRE-cupin repressors have one or two

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

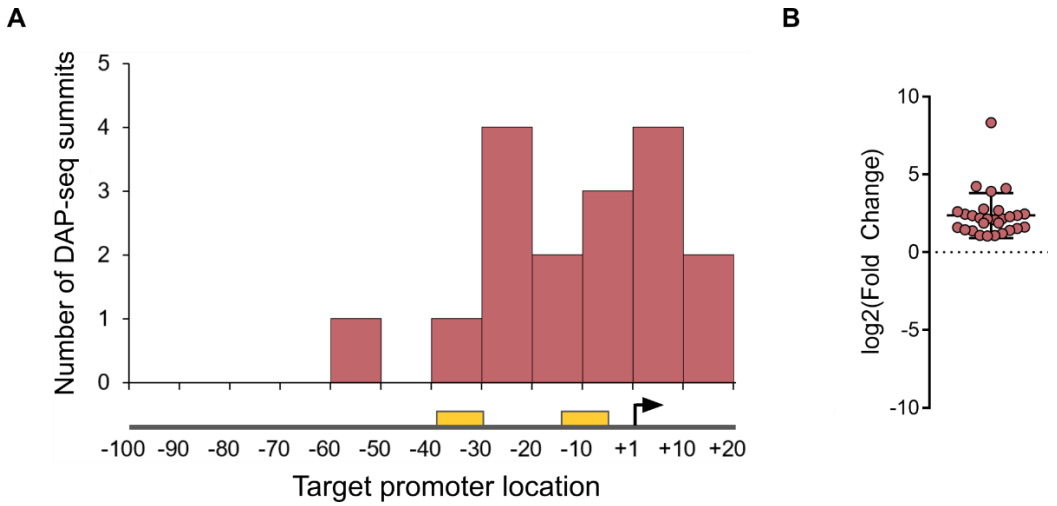


FIGURE 3. The XRE-cupin regulators are inhibitors of transcription. (**A**) Repartition of binding sites identified by DAP-seq within target promoters. RNA polymerase binding sites were either inferred from experimentally determined transcription start sites (Wurtzel *et al.*, 2012) or predicted using BPROM (http:// softberry.com) if no data was available. The transcription start site is shown as a black arrow, and the -10 and -35 boxes as yellow rectangles. (**B**) RNA-seq expression fold changes of target genes in regulators mutants compared to the wild-type strain.

targets, always found in the direct vicinity of their own genes, showing that they are local, specialized regulators that form local functional units with their regulated genes. Horizontal gene transfer (HGT) is thought to shape the evolution of these gene groups and their regulatory relationships, and might be at the origin of their differentiation from one another (Hershberg *et al.*, 2005, Price *et al.*, 2008).

Based on DAP-seq peaks, DNA-binding motifs could be generated for PsdR and PauR, which have more than one binding site (Figure S1). These motifs exhibit a palindromic architecture, as the previously determined ErfA consensus (Trouillon *et al.*, 2020), supporting DNA binding of XRE-cupin TFs as dimers. Mutations introduced into the binding site prevented PauR binding, further validating the inferred consensus (Figure S1). Interestingly, mutation of only half the palindromic sequence recognized by PsdR led to a faster mobility fragment than that obtained with the wild-type sequence. This might correspond to the binding of one monomer to the mutated probe and supports the model in which the usual two-band shift

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

observed with most XRE-cupin TFs corresponds to the binding of a monomer (lower, fainter band) and of a dimer (higher, stronger band) at higher concentration (Figures 4 and S1). The three determined binding consensus are quite different, probably due to differences in the amino acid composition of their DNA binding interface observed among the TF family (Figure S2). This supports their different DNA binding preferences and regulon specificities, as no common targets were revealed for any of the regulators.

Altogether, we defined the regulons of seven TFs, completing our knowledge of the XRE-cupin TF family in *P. aeruginosa* (Figure 5A). Our results show that all the members of the family are inhibitors of transcription, most likely acting by occluding the promoter and preventing RNA polymerase recruitment. Despite of their common features (domain architecture, local action, inhibitory function), each XRE-cupin TF targets specific regulons without overlapping roles.

XRE-cupin TFs regulate specific metabolic pathways. To determine the functions of each TF, we scrutinized the roles of their regulons in details (Figures 5A and 5B). That of ErfA was already mentioned and encompasses the *ergAB* operon in PAO1, playing a putative role in amino acid or carbohydrate metabolism (Trouillon *et al.*, 2020). PA0535 has only one target: it inhibits the *pauB1* gene and potentially its own expression, the two genes being predicted as an operon (Winsor *et al.*, 2016) (Figures 4A and 5A). PauB1 is one of the four redundant PauB proteins, which are FAD-dependent oxidoreductases working in the complex γ-glutamylation pathway required for polyamine utilization in *P. aeruginosa* (Chou *et al.*, 2013, Luengo & Olivera, 2020). Although expression of its gene was shown to be increased in presence of putrescine (Chou *et al.*, 2008), PauB1 is required for cadaverine catabolism (Chou *et al.*, 2013). We thus identified PA0535 as the missing regulator of the γ-glutamylation pathway, most of the genes being under the control of PauR, or BauR (Luengo & Olivera, 2020).

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

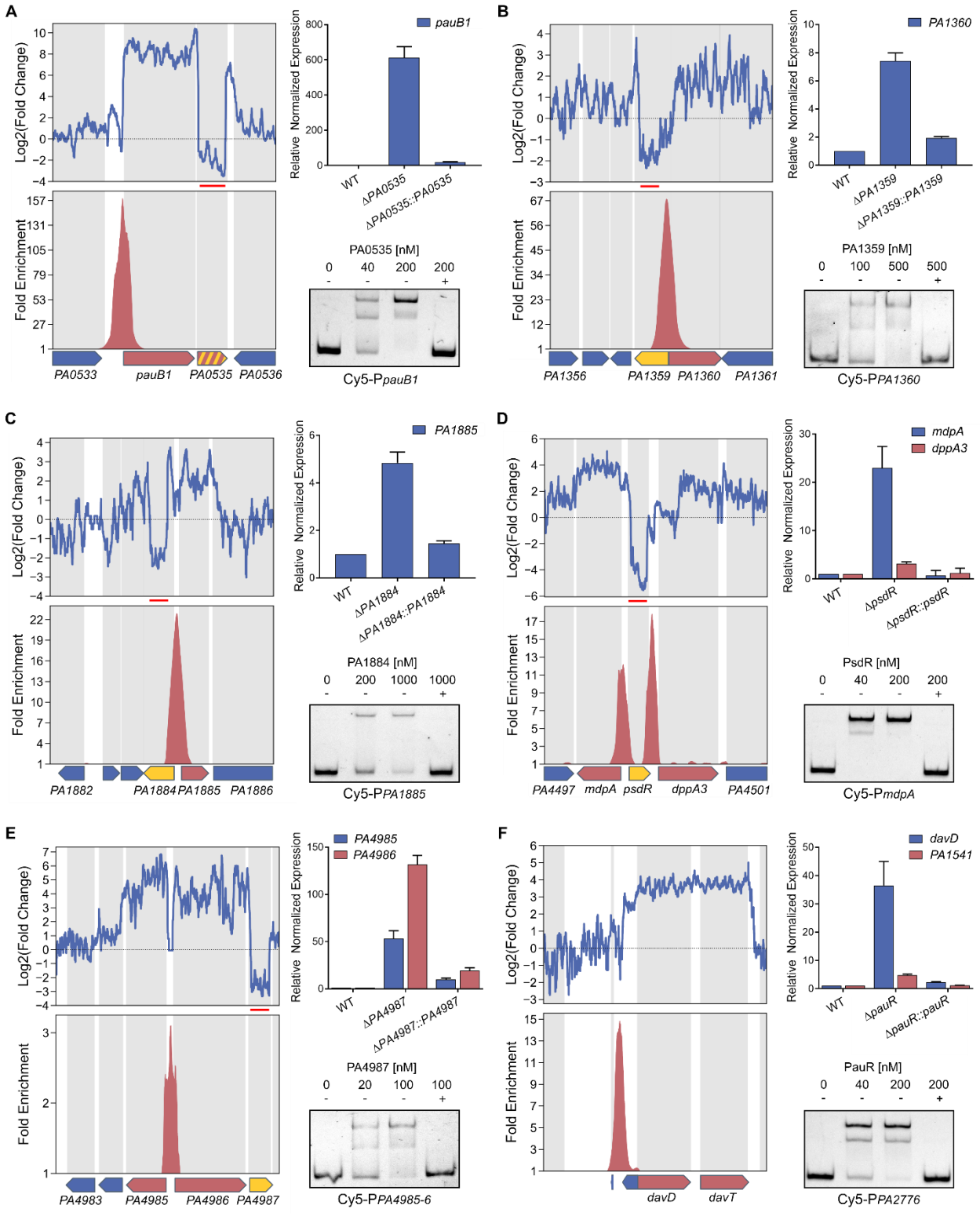


FIGURE 4. RT-qPCR and EMSA confirm the genome-wide results. Selected TFs and targets were PA0535 and *pauB1* (*PA0534*) (**A**), PA1359 and *PA1360* (**B**), PA1884 and *PA1885* (**C**), PsdR and *mdpA* (*PA4498*) and *dppA3* (*PA4450*) (**D**), PA4987 and *PA4985-PA4986* (**E**), PauR and *davD* (*PA0265*), *PA1541* (by RT-qPCR), and *PA2776* (by EMSA) (**F**). Upper left panels: local RNA-seq read abundance fold changes in the corresponding mutants compared to parental strain. Red lines show the deleted region in the mutant strain. Lower left panels: local fold enrichments in the corresponding regulator obtained by DAP-seq compared to negative controls. Target genes are shown as red arrows, genes encoding the studied TF as yellow arrows and others as blue arrows. Suspected autoregulation are denoted with dashed yellow-red

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

arrows. Upper right panels: RT-qPCR showing the regulation of the target genes in the corresponding regulatory mutants and complemented strains. Experiments were performed in triplicates and normalized to the *rpoD* transcripts. Error bars indicate the SD. Lower right panels: EMSA on target binding sites. Recombinant XRE-cupin-His_{10} proteins were incubated with 0.5 nM Cy5-labeled probes for 15 min before electrophoresis. For competition assays, excess of unlabeled probes (100 nM) is denoted '+'.

Indeed the key regulator PauR was shown to bind *in vitro* to eight sites and thus to potentially affect expression of 18 genes (Chou *et al.*, 2013). As mentioned above, our results confirmed all known DNA targets of PauR and RNA-seq provided further information concerning the extent of gene expression perturbation resulting from the TF binding. Two previously reported PauR binding sites were predicted to control the expression of six genes (Chou *et al.*, 2013), while our genome-wide expression data show that they actually impact the expression of 13 genes, *pauA1* and the *spuABCDEFGH operon* for one, and the two divergently transcribed two-gene operons, *pauC-PA5311* and *gabT2-PA5314* for the other. Furthermore, we also extended the PauR regulon by identifying four additional binding sites, one located upstream of *pauR* gene, strongly supporting an autoregulation mechanism. Three new identified targets were the *PA3766* gene encoding a probable amino acid/polyamine transporter and two operons, *davD-davT* and *PA1541-40,* which we confirmed by RT-qPCR (Figure 4F). While *PA1541* encodes a probable transporter, the DavD and DavT proteins are enzymes involved in the conversion from 5-aminovalerate (AMV) to glutarate (Luengo & Olivera, 2020). Cadaverine is converted into AMV through the γ-glutamylation pathway, showing that PauR regulation extends beyond this metabolic pathway to further downstream steps. The PA1359 regulator negatively controls the expression of the *PA1360* gene (Figure 4B), coding for a putative drug/metabolite transporter similar to the threonine exporter RhtA in *E. coli* (Livshits *et al.*, 2003). PA1884 has also only one target, the *PA1885* gene (Figure 4C), which product is a putative acyltransferase with a GNAT

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

(Gcn5-related N-acetyltransferases) domain, which could either confer antibiotic resistance or have potential metabolic functions.

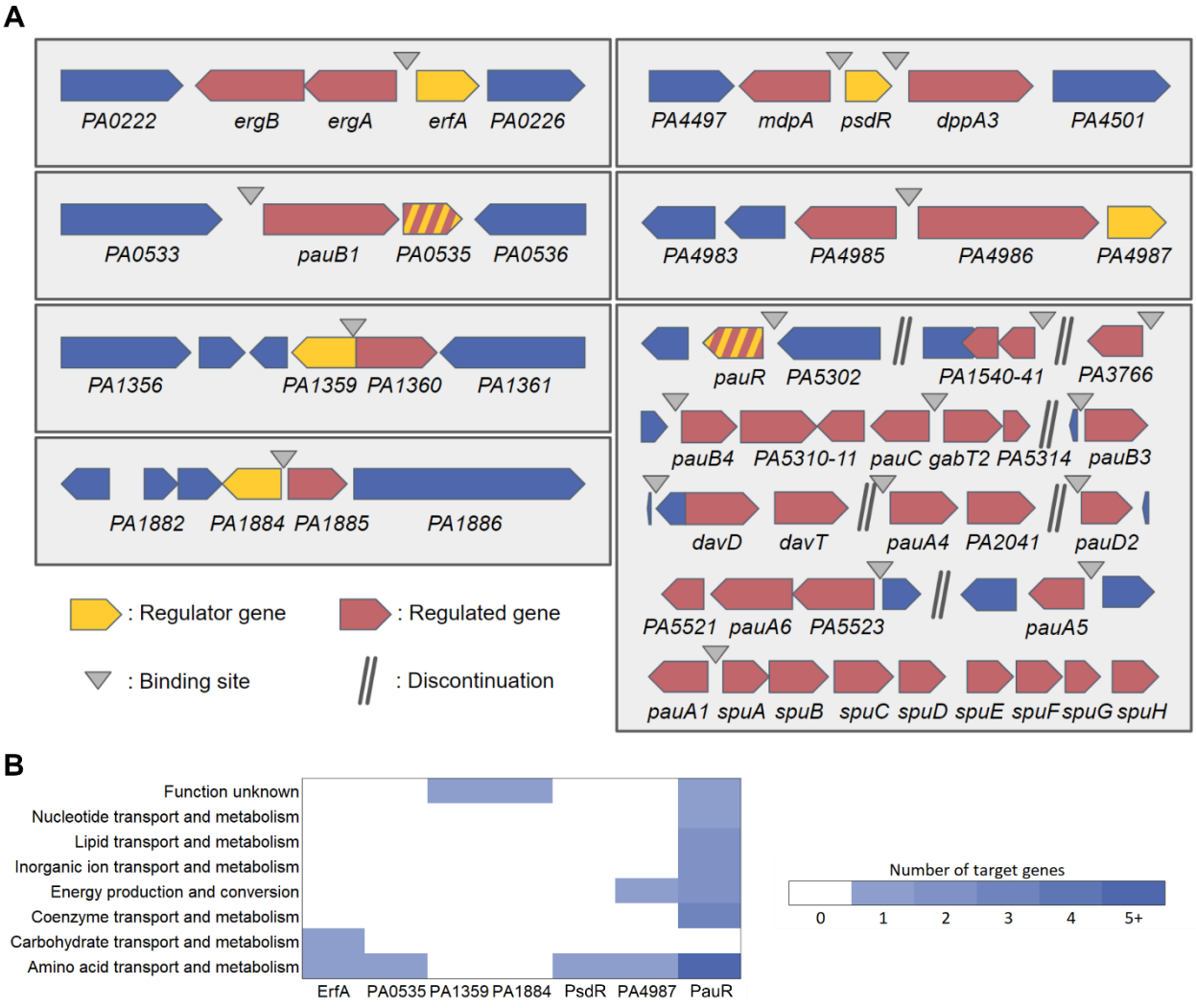


FIGURE 5. The complete regulons of the XRE-cupin regulators. (**A**) Schematic views of the local targeted regions of the seven XRE-cupin with determined regulons in *P. aeruginosa* PAO1. (**B**) Functional annotation of XRE-cupin regulatory targets. COG functional annotations were retrieved from PAO1 genome on the *Pseudomonas* database (Winsor *et al.*, 2016) for all target genes.

Even if *PA1885* is predicted in operon with the downstream *polB* gene (Winsor *et al.*, 2016), RNA-seq data indicated no difference on this gene, excluding a control by the inhibitor and suggesting the existence of a *polB*-specific promoter (Figure 4C). The PsdR regulator was already identified as the repressor of *dppA3* and *mdpA* (Kiely *et al.*, 2008), involved in the uptake and utilization of small peptides, our results completes the scheme by demonstrating that the

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

control is direct and involves two binding sites surrounding the *psdR* gene (Figure 4D). DppA3 is a substrate-binding protein delivering tripeptides/dipeptides to the ABC transporter DppBCDF (Pletzer *et al.*, 2014), and MdpA is a metallo-dipeptidase involved in their processing (Kiely *et al.*, 2008). Finally, PA4987 has one binding site in the intergenic sequence of *PA4985* and *PA4986*, downregulating expression of both genes as confirmed by RT-qPCR (Figure 4E). Although predicted as an operon with *PA4985*, we did not observe any effect on *PA4984* expression (Figure 4E). PA4986 is a putative oxidoreductase and PA4985 a possible periplasmic spermidine/putrescine-binding protein. Overall, the XRE-cupin regulators seem to regulate functions related to specific amino acids or small molecule uptake or processing (Figure 5B).

Based on our knowledge of their different regulons, we attempted to further investigate the XRE-cupin TFs roles through two global phenotypic assays. We first took advantage of the *Galleria mellonella* infection model frequently used to assess overall bacterial fitness and virulence (Tsai *et al.*, 2016, Andrejko *et al.*, 2014, Hernandez *et al.*, 2019). No significant difference could be observed in survival curves between larvae infected with wild-type and mutant strains (Figure S3). Therefore, the overexpression of genes regulated by the XRE-cupin family of repressors did not provide *P. aeruginosa* any advantage or disadvantage in this simple animal model. On the other hand, we assessed the strains antibiotics resistance to 24 different clinically relevant antibiotics and again found no differences (Figure S3). These results reflect the specificity of the XRE-cupin regulators and their functional units towards precise conditions. Indeed, *psdR* inactivation was shown to provide a growth fitness during proteolytic growth in caseinate medium, as this provides the dipeptide substrates for the derepressed dipeptides uptake and degradation (Asfahl *et al.*, 2015). Also, *mdpA* expression is induced in presence of X-pro dipeptides that might explain the faster growth in their presence due to their higher metabolism (Kiely *et al.*, 2008). The fact that they do not affect fitness in global phenotypic

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

assays illustrates that TFs of this family, along with their target genes, are niche-specific functional units allowing the bacteria to detect and respond to defined conditions.

It is hypothesized that XRE-cupin TFs are able to detect small signal molecules and detach from their DNA binding sites upon sensing in order to induce expression of the target genes. PauR was indeed shown to be involved in the derepression of genes of the γ-glutamylation pathway in presence of putrescine and cadaverine (Chou *et al.*, 2013). In addition, the *E. coli* PauR homolog PuuR is known to dissociate from its binding site upon direct sensing of putrescine (Nemoto *et al.*, 2012). In the same line, *mdpA* expression was shown to be induced in presence of different dipeptides, probably sensed by PsdR (Kiely *et al.*, 2008). In all these cases, the signals were metabolites that are part or the precursors of the regulated metabolic pathways. The cupin domain in these regulators is predicted as a signal-sensing domain and thus could be responsible for the specific sensing of each regulon-related metabolites, as it is often the case (Fernandez-Lopez *et al.*, 2015). The crystal structure of a XRE-cupin regulator from *Vibrio cholera* has been solved and shows a binding pocket in the cupin domain with bound D-methionine (PDB ID: 1Y9Q). The structure modelization of all eight *P. aeruginosa* XRE-cupin regulators also led to the identification of specific binding pockets in each cupin domain (Figure S4). Similarly, to what was found for their DNA-binding interfaces (Figure S2), each regulator exhibits a different amino-acid composition inside of this putative signal-sensing pocket. These variabilities at the two major functional regions of the proteins explain the difference in DNA- and signal molecule-binding preferences and may result from evolutionary processes involving HGT that led to the multiplication and differentiation of these regulators.

XRE-cupin regulators are diverse and differently conserved between *Pseudomonas* species. As XRE-cupin TFs seem to regulate small individual metabolic pathways that could reflect the

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

different environments encountered by the bacteria and their ability to adapt, we assessed the conservation of this family across the *Pseudomonas* genus. We first examined the conservation of the eight regulators studied here and found different conservation between *Pseudomonas* species (Figure 6A). Interestingly, PA1359 and PauR were conserved in nearly all strains, suggesting a more central role or ancestral origin. The six other regulators are less conserved between strains and species. We can notice that although PA2312 seems not to be active in the PAO1 strain, its gene is conserved within *P. aeruginosa* strains and is present in other species, underlining its physiological importance. Here again, the conservation of the TFs and their regulatory targets probably depends on the environments encountered by each species which might require or not the associated metabolic functional units.

To apprehend the importance of these regulators, we investigated the presence of other XRE-cupin TFs in all strains. Using all XRE-cupin sequences available on NCBI (*n* = 23,072) as templates, we created and validated a Hidden Markov Model for the automated detection of such regulators. The model was able to specifically detect all XRE-cupin TFs in PAO1 (Figure S5). Using this model, we screened nearly three million proteins retrieved from all *Pseudomonas* complete genomes for the presence of XRE-cupin regulators and identified 3,147 of them, with between zero and ten present per strain, depending on the strain or species (Figures 6A and 6B). Some species were found with only few XRE-cupin TFs, less than two or three on average for some of them as for example in *P. mendocina* or *P. stutzeri*. This could reflect the fact that these species are more niche-specific and may not encounter a wide variety of environments and thus need less of these optional metabolic response functional units. On the other hand, species like *P. aeruginosa*, *P. protegens* or *P. putida*, possess around eight and up to ten

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

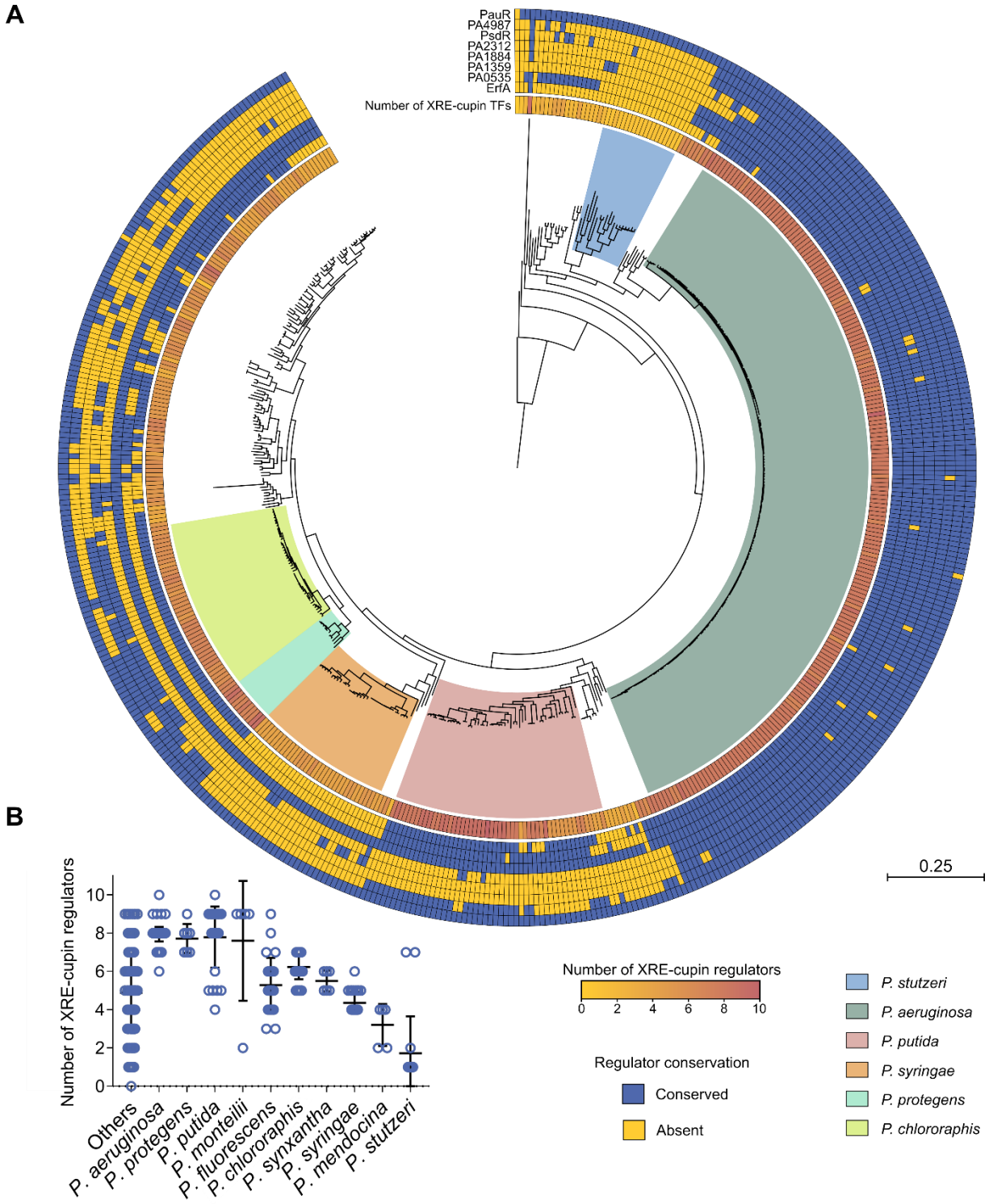


FIGURE 6. Phylogenetic analysis of the conservation of the XRE-cupin regulatory family across the *Pseudomonas* genus. (**A**) Maximum-Likelihood phylogenetic tree of 503 *Pseudomonas* complete genomes. The tree was generated from the multiple alignment of the concatenated sequences of 66 core genes for each strain with 100 bootstraps. The most represented species are delineated with a colored background. The number of XRE-cupin proteins detected through the Hidden Markov Model search is shown as the inner circle yellow-to-red heatmap. The outer circles show the results of homolog search by reciprocal best blast hit search for the eight regulators studied here. (**B**) Dot plot showing the distribution of number of XRE-cupin regulators per strain for each species. Species represented by less than five strains are grouped in the "Others" column.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

XRE-cupin regulators. All three of these species are known for their versatility and capacity to adapt to a wide array of environmental conditions. The total number of XRE-cupin regulators per strain often encompasses other TFs than the eight studied here, showing that much more regulators of this family are present across the *Pseudomonas* genus, probably associated to as many more different metabolic functional units. For instance, while only five of the eight TFs studied here are conserved in *P. putida*, most strains of this species possess eight or nine XRE-cupin TFs. This exemplifies the diversity of this family of TFs and shows that many species or strains have evolved new functional units including new XRE-cupin regulators to respond to the specific environmental niches they might encounter.

XRE-cupin regulators control neighboring enzyme-coding and metabolism-related genes. Six out of the seven XRE-cupin TFs experimentally characterized here regulate at least one gene adjacent to their own gene, forming local functional units. To investigate whether the local target regulation is universal for XRE-cupin TFs across the *Pseudomonas* genus, we first investigated the conservation of their neighboring genes (Figure 7A). We found a high conservation of regulator/regulated genes pairs, highlighting the local nature of these functional units and the fact that they are exchanged by HGT as a group of genes between bacteria. To go further on the characterization of all XRE-cupin regulators, we assessed the functions of the neighbor genes of all 3,147 regulators identified here. Based on GO Terms annotations of all 6,294 direct genetic neighbors, the vast majority of them encodes either enzymes or small molecule transporters (Figure 7B). This result strongly corroborates the fact that local functional units comprising XRE-cupin regulators act as metabolic modules for the response to the presence of specific metabolites.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

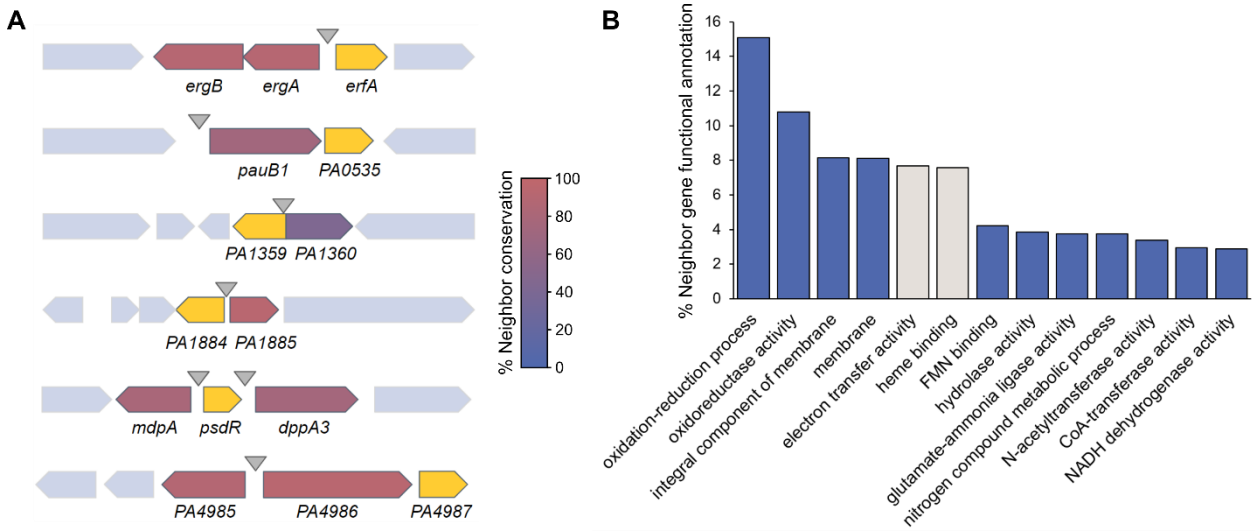


FIGURE 7. Gene synteny and genetic environment of the XRE-cupin regulators. (**A**) Conservation of XRE-cupin neighboring target genes. Regulator-coding genes are shown in yellow. Target genes are colored depending on how often they were found as conserved neighbors of their associated XRE-cupin TF. (**B**) Histogram showing the proportion of the most represented (>2%) GO functional annotations. Functional annotations were obtained from Pfam results of Interproscan search on 6,294 genes neighboring XRE-cupin regulators from 503 *Pseudomonas* genomes. Two categories are shown in grey as they correspond to the *cycB* gene, a conserved neighbor of *pauR* not regulated by PauR, and thus are not representing XRE-cupin regulatory targets.

DISCUSSION

In the present study, we characterized a family of eight TFs sharing the same domain architecture (XRE-cupin TFs) in *P. aeruginosa* and identified their regulatory targets. The XRE-cupin family members are exclusive inhibitors of the transcription of metabolism-related genes probably needed in defined, specific conditions. The current model is that these TFs act as regulatory switches that keep the transcription of their neighboring target genes down until their products are needed for a given metabolic pathway. Upon sensing of the precursor metabolite of the regulated pathway through their cupin domain, they disengage from the DNA, in order to allow transcription of the genes coding for enzymes or transporters that will process the sensed molecule. This could include catabolism for nutrient source, as seen for PsdR, or detoxification via export of sensed molecules, as predicted for PA1359.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

In most cases, the target genes were found located in the direct vicinity of the regulator gene, revealing the local nature of these regulatory interactions. Neighbor regulation is a common feature in bacteria, which has been explained by different models (Hershberg *et al.*, 2005, Lawrence, 2003). Genes that are functionally related, including regulators and regulated targets, tend to regroup together on the chromosome in order to be a complete functional unit when transferred by HGT. Different types of neighbor regulation are known (Hershberg *et al.*, 2005), three of which were found in this study; simple neighbor regulation and co-regulation of neighbors in both *cis* and *trans*, with some instances of autoregulation. In all cases, the total number of targets was small, allowing the formation of these local functional units. The very specific functions of the XRE-cupin regulons explain the small number of regulated genes and the fact that these units are local and probably easily transferrable gene clusters.

We found a high conservation of some of the studied TFs and their neighboring target genes across the *Pseudomonas* genus. These functional units may have been exchanged by HGT between bacteria that encounter similar environmental conditions. Additionally, a large number of different XRE-cupin regulators, each associated with their specific genetic neighborhood, were identified across the *Pseudomonas* genus. Out of the eight TFs studied here, none shared any common regulatory targets, and the two known sensed signals (dipeptides and polyamines) are different. This is explained by differences found at both DNA-binding and signal-sensing interfaces. Such differences are thought to happen through evolution after and during repurposing or duplication events and explain the presence of different XRE-cupin regulators in each species. The large number of this type of regulators illustrates their ancestral nature and their importance as condition-specific, transferrable means of metabolic diversification.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

Pseudomonas species are famous for their wide metabolic versatility, encompassing many condition-specific metabolic pathways (Silby *et al.*, 2011). While the prediction of metabolism-related functions works well, experimental studies are still needed to pinpoint the exact pathways represented by uncharacterized genes. As illustrated by the target genes found here which mostly have only predicted functions, there is still a great need of phenotypic characterization of putative new metabolic pathways. Such studies would complete our knowledge on the mode of action of the XRE-cupin family of regulators and help understand both the function and sensed signals of these class of TFs. We also believe that family-wide and inter-species studies of TFs should become more common in order to be able to decipher global regulatory networks, especially in bacteria where they are still scarce.

MATERIALS AND METHODS

Bacterial strains. The bacterial strains used in this study are listed in Supplementary Table 1. *P. aeruginosa* and *E. coli* strains were grown in Lysogeny Broth (LB) at 37°C under agitation (300 rpm). *P. aeruginosa* strains were selected on Pseudomonas Isolation Agar (PIA). Antibiotics for *P. aeruginosa* were added when needed at the following concentrations: 200 µg/ml carbenicillin, 200 µg/ml gentamicin and 200 µg/ml tetracycline, and for *E. coli*: 100 µg/ml ampicillin, 50 µg/ml gentamycin and 10 µg/ml tetracycline.

Plasmids and genetic manipulations. Plasmids and primers are listed in the Supplementary Tables S1 and S2, respectively. For overproduction of recombinant His10-tagged proteins, each gene sequence was amplified by PCR using PAO1 genomic DNA as a matrix and appropriate primer pairs, then integrated by SLIC in pET-52b cut with *Nco*I-*Sac*I and sequenced.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

To generate *P. aeruginosa* deletion mutants, upstream and downstream flanking regions of each gene were amplified using appropriate primer pairs (sF1/sR1 and sF2/sR2 for each construct). The two resulting, overlapping fragments were then cloned into *Sma*I-cut pEXG2 by Sequence- and Ligation-Independent Cloning (SLIC) (Li & Elledge, 2007) and sequenced.

For complementation of *PA1359*, *PA1884*, *PA4499* and *PA5301* mutants, a fragment encompassing the promoter (around 500 bp upstream from the start codon) and the CDS was amplified by PCR and cloned by SLIC into the mini-CTX1 plasmid cut by *Sma*I. When the gene was predicted as the second gene of an operon (like *PA0535*, *PA2312* and *PA4987*) (Winsor *et al.*, 2016), the operon promoter was directly fused to the regulators gene. To do so, two fragments were generated, the upstream fragment carrying around 500 bp upstream and the 3 first codons of the first gene of the operon, and the second one the few last codons and the entire CDS of the gene of interest. Then the two overlapping fragments were cloned by SLIC in pEXG2 and sequenced.

All mini-CTX1- and pEXG2-derived plasmids were transferred by triparental mating into *P. aeruginosa*, using the helper plasmid pRK600. For allelic exchange, merodiploids resulting from cointegration events were selected on PIA plates containing gentamicin. Colonies were then plated on NaCl-free LB agar plates containing 10% sucrose to select for the loss of plasmid. The resulting sucrose-resistant strains were checked for gentamicin sensitivity and the mutant genotypes were determined by PCR. For complementation using mini-CTX1-derived plasmids, bacteria with *att* site-inserted plasmid were selected on PIA plates containing tetracycline, then they were cured from the mini-CTX1 backbone by excising FRT cassette with pFLP2 plasmid as previously described (Hoang *et al.*, 1998).

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

Structure modeling. Protein sequences from the eight *P. aeruginosa* XRE-cupin TFs were aligned against the RCSB Protein Data Bank (Berman et al., 2000) in order to find an experimentally obtained 3D structure to use as a template for modeling. The 1.90Å 3D structure of a XRE-cupin transcription factor from *Vibrio cholerae* (PDB id: 1Y9Q) was used as template as it permitted the best modelling confidence using the SWISS-MODEL tool (Biasini *et al.*, 2014).

Protein purification. All the pET-52b plasmids were transformed into *E. coli* BL21 Star (DE3). For protein overproduction, overnight cultures were diluted to an OD_{600} of 0.05 in LB medium containing 100 µg/ml ampicillin and expression induced at an OD_{600} of 0.6 with 1 mM IPTG. After 3h of growth at 37°C, bacteria were harvested by centrifugation at 6,000 g for 10 min at 4°C and resuspended in L Buffer (25 mM Tris-HCl, 500 mM NaCl, 10 mM Imidazole, 1 mM PMSF, 5 % Glycerol, pH 8, containing Roche protease inhibitor cocktail). Bacteria were then lysed by sonication. After centrifugation at 66,500 g for 30 min at 4°C, the soluble fraction was directly loaded onto a 1-ml nickel column (Protino Ni-nitrotriacetic acid [NTA]; Macherey-Nagel). The column was washed with Wash Buffer (50 mM Tris-HCl, 500 mM NaCl, 5 % Glycerol, pH 8) containing increasing Imidazole concentrations (20, 40, and 60 mM), and proteins were eluted with 200 mM Imidazole. Aliquots from the peak protein fractions were analyzed by SDS-PAGE, and the fractions containing the proteins of interest were pooled and dialyzed against ErfA Buffer (50 mM Tris-HCl, 250 mM NaCl, 50 mM KCl, 10 % Glycerol, 0.5 % Tween20, pH 7).

DAP-seq experimental procedure, sequencing and data analysis. DAP-seq was carried out in triplicates on PAO1 genomic DNA exactly as previously described (Trouillon *et al.*, 2020). Sequencing was performed by the high-throughput sequencing core facility of I2BC (http://www.i2bc.paris-saclay.fr) using an Illumina NextSeq500 instrument. Approximately 8 million single-end reads per sample were generated on average with >90% of reads uniquely

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

aligning to the PAO1 genome. Data analysis was performed as previously described (Trouillon *et al.*, 2020).

RNA isolation. *P. aeruginosa* strains were grown from overnight cultures diluted to an OD_{600} of 0.1 in 3 ml of fresh LB medium at 37°C under agitation in duplicate. Total RNA was isolated at an OD_{600} of 1.0 using hot phenol-chloroform extraction as previously described (Trouillon *et al.*, 2020).

RNA-seq libraries construction, sequencing and data analysis. After RNA isolation and DNAse treatment, RNA sample quality was assessed on an Agilent Bioanalyzer, yielding RINs of 9 or higher. Then ribosomal RNAs were depleted using the RiboMinus Transcriptome Isolation Kit (Thermofisher) following manufacturer instructions. The cDNA libraries were then constructed from 50 ng of depleted RNA using the NEBNext Ultra II Directional RNA library prep kit following manufacturer instructions (NEB). Libraries were size-selected to 200-700 bp using SPRIselect beads and quality was assessed on the Agilent Bioanalyzer using High Sensitivity DNA chips. Sequencing were done on an Illumina NextSeq500 and approximately 12 million single-end reads per sample were generated. Data analysis was performed as previously described (Trouillon *et al.*, 2020).

RT-qPCR. After total RNA isolation and DNase treatment, cDNA synthesis was carried out using 2 µg of RNA as previously described (Trouillon *et al.*, 2020). The experiments were performed with 3 biological replicates for each strain, and the relative expression of mRNAs was analyzed with the CFX Manager software (Bio-Rad) using the Pfaffl method relative to *uvrD* reference Cq values. Statistical analyses were performed by T-test. The sequences of primers are listed in Supplementary Table 2.

Electrophoretic Mobility Shift Assay. Genomic DNA was used as a matrix to generate wild-type sequence probes. To amplify probes bearing mutagenized binding sites, three different matrices

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

were created: the one with mutated binding site in P_{PA0534} resulted from the fusion of two overlapping fragments generated by primer pairs pEXG2-mut-PA0534-BS-sF1/sR1 and pEXG2-mut-PA0534-BS-sF2/sR2. PCR fragments with modified binding sites in P_{PA2776} and P_{mdpA} were produced by using primer pairs CTX-PA2776-lacZ-sF/PA2776-mut-lacZ-sR and CTX-mdpA- lacZ-sF/mdAmut-lacZ sR, respectively. Then all fluorescent DNA probes were generated by PCR in two steps : a first PCR using specific primer pairs amplified the target DNA regions (83-96 bp) flanked by a 21 bp-region which is targeted in a second PCR by a single Cy5-labelled primer. The resulting labelled probes were purified on DNA Clean up columns (NEB) then incubated at 0.5 nM for 5 min at 37°C in EMSA Buffer (10 mM Tris-HCl, 50 mM KCl, 10 mM $MgCl_2$, 10 % Glycerol, 0.1 mg/ml BSA, pH 8) containing 25 ng/µl poly(dI-dC). For competition assays, 100 nM unlabelled DNA probes (200 fold excess) were incubated with the labelled probes. Recombinant proteins were added at the indicated concentration in a final reaction volume of 20 µl and incubated for an additional 15 min at 25°C. Samples were then loaded on a native 8 % Tris-Acetate-EDTA (TAE) polyacrylamide gel and run at 100 V and 4°C in cold 1X TAE Buffer. Fluorescence imaging was performed using a Chemidoc MP.

Phylogenetic analysis and XRE-cupin identification. All *Pseudomonas* complete genomes (*n*=503) were retrieved from the *Pseudomonas* Genome Database (Winsor *et al.*, 2016). The sequences from 66 core genes were concatenated for each genome and a multiple alignment was performed with MAFFT Galaxy version 7.221.3 (Katoh & Standley, 2013) using default settings. The resulting alignment was used to build a Maximum-Likelihood phylogenetic tree using MEGA X with 100 bootstraps which was visualized and annotated using iTOL v5 (Kumar *et al.*, 2018, Letunic & Bork, 2019). XRE-cupin homolog identification was performed by Reciprocal Best Blast Hit (RBBH) analysis (Cock *et al.*, 2015) on the European Galaxy server (Jalili *et al.*, 2020)

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

using protein sequences from *P. aeruginosa* PAO1 against all protein sequences from the other 502 genomes.

For XRE-cupin identification, ErfA protein sequence was used to retrieve all protein sequences (*n*=23,072) with a XRE-cupin domain architecture (HTH_3 and Cupin_2) from the NCBI Conserved Domain Architecture Retrieval Tool (Marchler-Bauer *et al.*, 2015). A multiple alignment was performed on these sequences using MAFFT Galaxy version 7.221.3 (Katoh & Standley, 2013) with default settings. The resulting alignment was used to build a hidden Markov model using hmmbuild from HMMER (Potter *et al.*, 2018). All 2,862,589 protein sequences from the 503 genomes were then scanned through the model using hmmscan from HMMER (Potter *et al.*, 2018). A bimodal distribution of HMMER scores could be observed which led to the determination of a true positive XRE-cupin threshold (HMMER score >100), that was validated by functional prediction from both sides (Figure S5). All proteins above this threshold (*n*=3,147) were considered XRE-cupin regulators.

Genetic neighbor functional prediction. Gene IDs from all proteins matching the HMM model were used to identify the direct upstream and downstream neighbor genes from all 503 genomes annotation files using Python 3.7. The corresponding 6,294 protein sequences were then used for functional prediction analysis using the Interproscan functional prediction tool (Jones *et al.*, 2014). Go annotations associated with Pfam predicted functions were then used for functional prediction.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

ACKNOWLEDGMENTS

We are grateful to Peter Panchev, Yvan Caspar, Eric Faudry and Viviana Job for their help with *G. mellonella* experiments, antibiotics resistance assays, biochemistry and structural modelling advices, respectively. This work was supported by grants from Agence Nationale de la Recherche [ANR-15-CE11-0018-01], the Laboratory of Excellence GRAL financed within the Grenoble Alpes University graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS [ANR-17-EURE-0003], and the Fondation pour la Recherche Medicale [Team FRM 2017, DEQ20170336705]. Julian Trouillon received a Ph.D. fellowship from French Ministry of Education and Research. We further acknowledge support from CNRS, INSERM, CEA, and Grenoble Alpes University.

CONFLICT OF INTEREST

The authors declare no competing financial interests.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: J.T., I.A., S.E. Performed the experiments: J.T., M.R., V.S., S.E. Analyzed the data: J.T., I.A., S.E. Wrote the initial draft: J.T., S.E. Manuscript finalization: J.T., I.A., S.E.

REFERENCES

Andrejko, M., A. Zdybicka-Barabas & M. Cytrynska, (2014) Diverse effects of Galleria mellonella infection with entomopathogenic and clinical strains of Pseudomonas aeruginosa. *J Invertebr Pathol* **115**: 14-25.

Asfahl, K.L., J. Walsh, K. Gilbert & M. Schuster, (2015) Non-social adaptation defers a tragedy of the commons in Pseudomonas aeruginosa quorum sensing. *ISME J* **9**: 1734-1746.

Bartlett, A., R.C. O'Malley, S.C. Huang, M. Galli, J.R. Nery, A. Gallavotti & J.R. Ecker, (2017) Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat Protoc* **12**: 1659-1672.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

555 Biasini, M., S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. Gallo
556 Cassarino, M. Bertoni, L. Bordoli & T. Schwede, (2014) SWISS-MODEL: modelling
557 protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids*
558 *Res* **42**: W252-258.
559 Chou, H.T., D.H. Kwon, M. Hegazy & C.D. Lu, (2008) Transcriptome analysis of agmatine and
560 putrescine catabolism in Pseudomonas aeruginosa PAO1. *J Bacteriol* **190**: 1966-1975.
561 Chou, H.T., J.Y. Li, Y.C. Peng & C.D. Lu, (2013) Molecular characterization of PauR and its role
562 in control of putrescine and cadaverine catabolism through the gamma-glutamylation
563 pathway in Pseudomonas aeruginosa PAO1. *J Bacteriol* **195**: 3906-3913.
564 Cock, P.J., J.M. Chilton, B. Gruning, J.E. Johnson & N. Soranzo, (2015) NCBI BLAST+
565 integrated into Galaxy. *Gigascience* **4**: 39.
566 Dunwell, J.M., A. Purvis & S. Khuri, (2004) Cupins: the most functionally diverse protein
567 superfamily? *Phytochemistry* **65**: 7-17.
568 Favrot, L., J.S. Blanchard & O. Vergnolle, (2016) Bacterial GCN5-Related N-Acetyltransferases:
569 From Resistance to Regulation. *Biochemistry* **55**: 989-1002.
570 Fernandez-Lopez, R., R. Ruiz, F. de la Cruz & G. Moncalian, (2015) Transcription factor-based
571 biosensors enlightened by the analyte. *Front Microbiol* **6**: 648.
572 Galan-Vasquez, E., B. Luna & A. Martinez-Antonio, (2011) The Regulatory Network of
573 Pseudomonas aeruginosa. *Microb Inform Exp* **1**: 3.
574 Hernandez, R.J., E. Hesse, A.J. Dowling, N.M. Coyle, E.J. Feil, W.H. Gaze & M. Vos, (2019)
575 Using the wax moth larva Galleria mellonella infection model to detect emerging bacterial
576 pathogens. *PeerJ* **6**: e6150.
577 Hershberg, R., E. Yeger-Lotem & H. Margalit, (2005) Chromosomal organization is shaped by
578 the transcription regulatory network. *Trends Genet* **21**: 138-142.
579 Hoang, T.T., R.R. Karkhoff-Schweizer, A.J. Kutchma & H.P. Schweizer, (1998) A broad-host-
580 range Flp-FRT recombination system for site-specific excision of chromosomally-located
581 DNA sequences: application for isolation of unmarked Pseudomonas aeruginosa
582 mutants. *Gene* **212**: 77-86.
583 Huang, H., X. Shao, Y. Xie, T. Wang, Y. Zhang, X. Wang & X. Deng, (2019) An integrated
584 genomic regulatory network of virulence-related transcriptional factors in Pseudomonas
585 aeruginosa. *Nat Commun* **10**: 2931.
586 Jalili, V., E. Afgan, Q. Gu, D. Clements, D. Blankenberg, J. Goecks, J. Taylor & A. Nekrutenko,
587 (2020) The Galaxy platform for accessible, reproducible and collaborative biomedical
588 analyses: 2020 update. *Nucleic Acids Res*.
589 Jones, P., D. Binns, H.Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A.
590 Mitchell, G. Nuka, S. Pesseat, A.F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S.Y.
591 Yong, R. Lopez & S. Hunter, (2014) InterProScan 5: genome-scale protein function
592 classification. *Bioinformatics* **30**: 1236-1240.
593 Katoh, K. & D.M. Standley, (2013) MAFFT multiple sequence alignment software version 7:
594 improvements in performance and usability. *Mol Biol Evol* **30**: 772-780.
595 Kiely, P.D., J. O'Callaghan, A. Abbas & F. O'Gara, (2008) Genetic analysis of genes involved in
596 dipeptide metabolism and cytotoxicity in Pseudomonas aeruginosa PAO1. *Microbiology*
597 **154**: 2209-2218.
598 Kumar, S., G. Stecher, M. Li, C. Knyaz & K. Tamura, (2018) MEGA X: Molecular Evolutionary
599 Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**: 1547-1549.
600 Lawrence, J.G., (2003) Gene organization: selection, selfishness, and serendipity. *Annu Rev*
601 *Microbiol* **57**: 419-440.
602 Letunic, I. & P. Bork, (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new
603 developments. *Nucleic Acids Res* **47**: W256-W259.
604 Li, M.Z. & S.J. Elledge, (2007) Harnessing homologous recombination in vitro to generate
605 recombinant DNA via SLIC. *Nat Methods* **4**: 251-256.
606 Livshits, V.A., N.P. Zakataeva, V.V. Aleshin & M.V. Vitushkina, (2003) Identification and
607 characterization of the new gene rhtA involved in threonine and homoserine efflux in
608 Escherichia coli. *Res Microbiol* **154**: 123-135.
609 Luengo, J.M. & E.R. Olivera, (2020) Catabolism of biogenic amines in Pseudomonas species.
610 *Environ Microbiol* **22**: 1174-1192.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.29.226225. this version posted July 29, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC 4.0 International license.

611 Marchler-Bauer, A., M.K. Derbyshire, N.R. Gonzales, S. Lu, F. Chitsaz, L.Y. Geer, R.C. Geer, J.
612 He, M. Gwadz, D.I. Hurwitz, C.J. Lanczycki, F. Lu, G.H. Marchler, J.S. Song, N. Thanki,
613 Z. Wang, R.A. Yamashita, D. Zhang, C. Zheng & S.H. Bryant, (2015) CDD: NCBI's
614 conserved domain database. *Nucleic Acids Res* **43**: D222-226.
615 Nemoto, N., S. Kurihara, Y. Kitahara, K. Asada, K. Kato & H. Suzuki, (2012) Mechanism for
616 regulation of the putrescine utilization pathway by the transcription factor PuuR in
617 Escherichia coli K-12. *J Bacteriol* **194**: 3437-3447.
618 Ohlendorf, D.H., W.F. Anderson, M. Lewis, C.O. Pabo & B.W. Matthews, (1983) Comparison of
619 the structures of cro and lambda repressor proteins from bacteriophage lambda. *J Mol*
620 *Biol* **169**: 757-769.
621 Pletzer, D., C. Lafon, Y. Braun, T. Kohler, M.G. Page, M. Mourez & H. Weingart, (2014) High-
622 throughput screening of dipeptide utilization mediated by the ABC transporter DppBCDF
623 and its substrate-binding proteins DppA1-A5 in Pseudomonas aeruginosa. *PLoS One* **9**:
624 e111311.
625 Potter, S.C., A. Luciani, S.R. Eddy, Y. Park, R. Lopez & R.D. Finn, (2018) HMMER web server:
626 2018 update. *Nucleic Acids Res* **46**: W200-W204.
627 Price, M.N., P.S. Dehal & A.P. Arkin, (2008) Horizontal gene transfer and the evolution of
628 transcriptional regulation in Escherichia coli. *Genome Biol* **9**: R4.
629 Rajeev, L., M.E. Garber & A. Mukhopadhyay, (2020) Tools to map target genes of bacterial two-
630 component system response regulators. *Environ Microbiol Rep* **12**: 267-276.
631 Rodrigue, A., Y. Quentin, A. Lazdunski, V. Mejean & M. Foglino, (2000) Two-component systems
632 in Pseudomonas aeruginosa: why so many? *Trends Microbiol* **8**: 498-504.
633 Sanchez, I., R. Hernandez-Guerrero, P.E. Mendez-Monroy, M.A. Martinez-Nunez, J.A. Ibarra &
634 E. Perez-Rueda, (2020) Evaluation of the Abundance of DNA-Binding Transcription
635 Factors in Prokaryotes. *Genes (Basel)* **11**.
636 Silby, M.W., C. Winstanley, S.A. Godfrey, S.B. Levy & R.W. Jackson, (2011) Pseudomonas
637 genomes: diverse and adaptable. *FEMS Microbiol Rev* **35**: 652-680.
638 Stover, C.K., X.Q. Pham, A.L. Erwin, S.D. Mizoguchi, P. Warrener, M.J. Hickey, F.S. Brinkman,
639 W.O. Hufnagle, D.J. Kowalik, M. Lagrou, R.L. Garber, L. Goltry, E. Tolentino, S.
640 Westbrock-Wadman, Y. Yuan, L.L. Brody, S.N. Coulter, K.R. Folger, A. Kas, K. Larbig,
641 R. Lim, K. Smith, D. Spencer, G.K. Wong, Z. Wu, I.T. Paulsen, J. Reizer, M.H. Saier,
642 R.E. Hancock, S. Lory & M.V. Olson, (2000) Complete genome sequence of
643 Pseudomonas aeruginosa PAO1, an opportunistic pathogen. *Nature* **406**: 959-964.
644 Trouillon, J., E. Sentausa, M. Ragno, M. Robert-Genthon, S. Lory, I. Attree & S. Elsen, (2020)
645 Species-specific recruitment of transcription factors dictates toxin expression. *Nucleic*
646 *Acids Res* **48**: 2388-2400.
647 Tsai, C.J., J.M. Loh & T. Proft, (2016) Galleria mellonella infection models for the study of
648 bacterial diseases and for antimicrobial drug testing. *Virulence* **7**: 214-229.
649 Ulrich, L.E., E.V. Koonin & I.B. Zhulin, (2005) One-component systems dominate signal
650 transduction in prokaryotes. *Trends Microbiol* **13**: 52-56.
651 Winsor, G.L., E.J. Griffiths, R. Lo, B.K. Dhillon, J.A. Shay & F.S. Brinkman, (2016) Enhanced
652 annotations and features for comparing thousands of Pseudomonas genomes in the
653 Pseudomonas genome database. *Nucleic Acids Res* **44**: D646-653.
654 Wood, H.E., K.M. Devine & D.J. McConnell, (1990) Characterisation of a repressor gene (xre)
655 and a temperature-sensitive allele from the Bacillus subtilis prophage, PBSX. *Gene* **96**:
656 83-88.
657

Supplementary Information

Transcription inhibitors with XRE DNA-binding and cupin signal-sensing domains drive metabolic diversification in *Pseudomonas*

Julian Trouillon[1#], Michel Ragno[1], Victor Simon[1], Ina Attrée[1], Sylvie Elsen[1#]

[1]Université Grenoble Alpes, CNRS ERL5261, CEA BIG-BCI, INSERM UMR1036, Grenoble, France

Address correspondance to julian.trouillon@cea.fr; sylvie.elsen@cea.fr

Supplementary methods

***Galleria mellonella* infection assay**. *G. mellonella* larvae infections were performed as previously described (Trouillon *et al.*, 2020). Animals injected with sterile PBS served as a control for physical trauma. CFUs for each dilution were systematically counted from the insulin pen and were of 7±1 CFUs per injection. Twenty larvae were injected per condition. Infection development was followed for 24 h at 37°C and the animals were considered dead when not reacting to touch. Strains were independently randomized and blinded ensuring no bias in CFU spotting and counting, and in animal death counting which was also done by a different person. Statistical significance was assessed using a Log-rank test.

Antibiotics resistance screening. Antimicrobial susceptibility was assessed by broth microdilution using the automated BD Phoenix system (Becton Dickinson, France) according to the manufacturers' recommendations. In brief, colonies were suspended in BD Phoenix™ ID broth to obtain approximately a 0.5 McFarland ($1.5.10^8$ UFC/ml) suspension. Broth was then placed on BD Phoenix™ AP system that automatically adjusts the optical density and further dilutes the bacterial suspension in BD Phoenix™ AST broth (a cation-adjusted formulation of Mueller-Hinton broth containing 0.01% Tween 80) to obtain a final bacterial concentration in the AST broth of approximately 5.10^5 CFU/ml. A redox indicator was then added to the AST broth. Inoculated AST broth was then poured in BD Phoenix™ NMIC-417 panel, loaded into the BD Phoenix™ M50 instrument and incubated at 35˚C ± 1˚C. Results were interpreted by the BD EpiCenter™ system according to CASFM-EUCAST (European Committee on Antimicrobial Susceptibility Testing) 2019 V1 breakpoints.

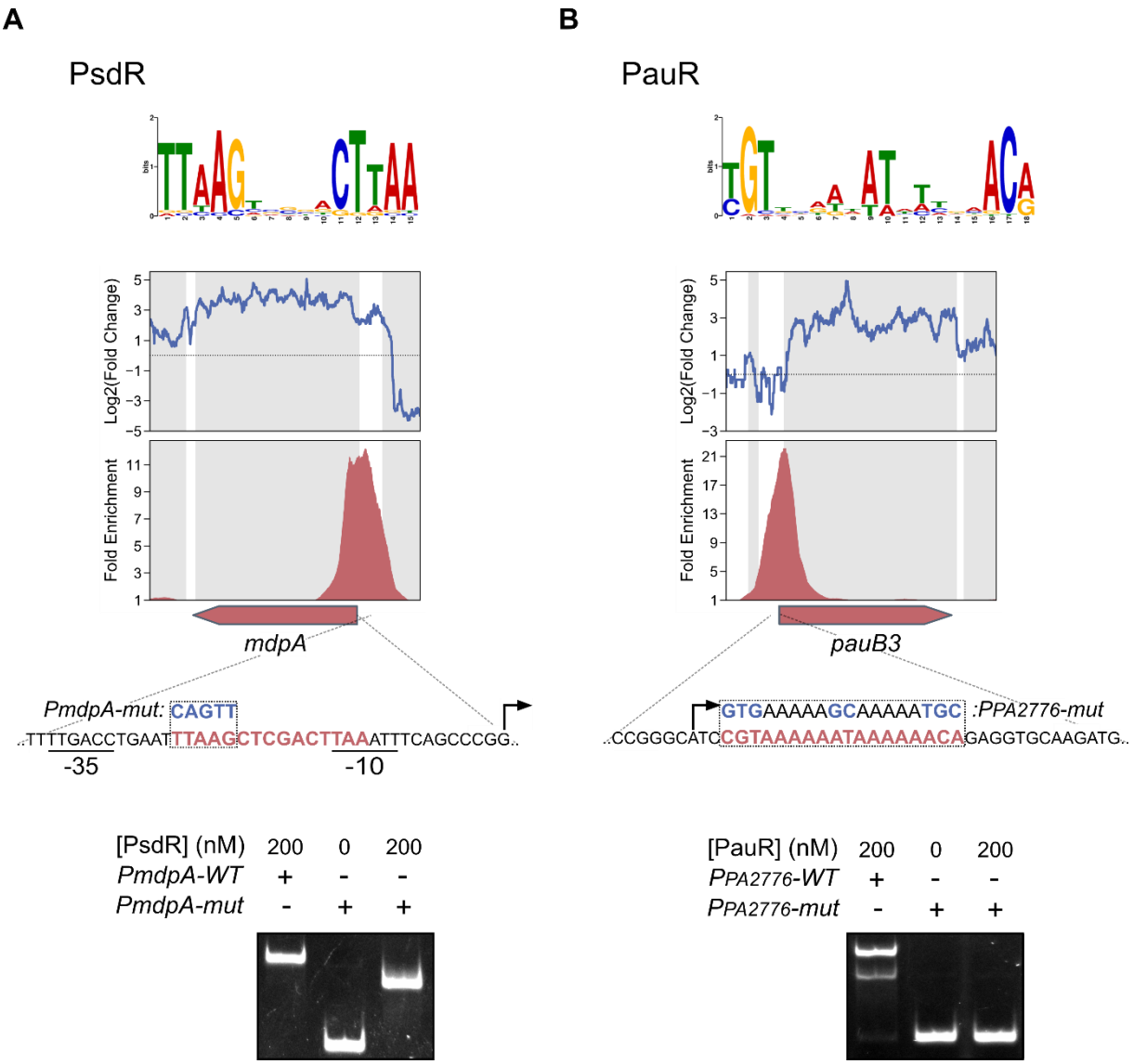


FIGURE S1 : Validation of the DNA binding sites of PsdR and PauR. For PsdR (**A**) and PauR (**B**), from upper to lower panels: (i) Enriched DNA motif obtained with MEME-ChIP in top DAP-seq peaks, (ii) local RNA-seq read abundance fold changes in the corresponding regulator mutants compared to the parental strain, (iii) local DAP-seq fold enrichments in the corresponding regulator experiments compared to negative controls, (iv) zoom view of the bound promoter region; mutated probes used in EMSA are shown with exchanged nucleotides in blue, and (v) EMSA of the corresponding region with either wild-type sequence or mutated probes. The mutation of half of the palindromic motif recognized by PsdR led to the shift of a lower band than with the wild-type probe, potentially representing the binding of only one monomer to the non-mutated half site. On the other hand, the mutation of the entire site for PauR led to a complete loss of binding.

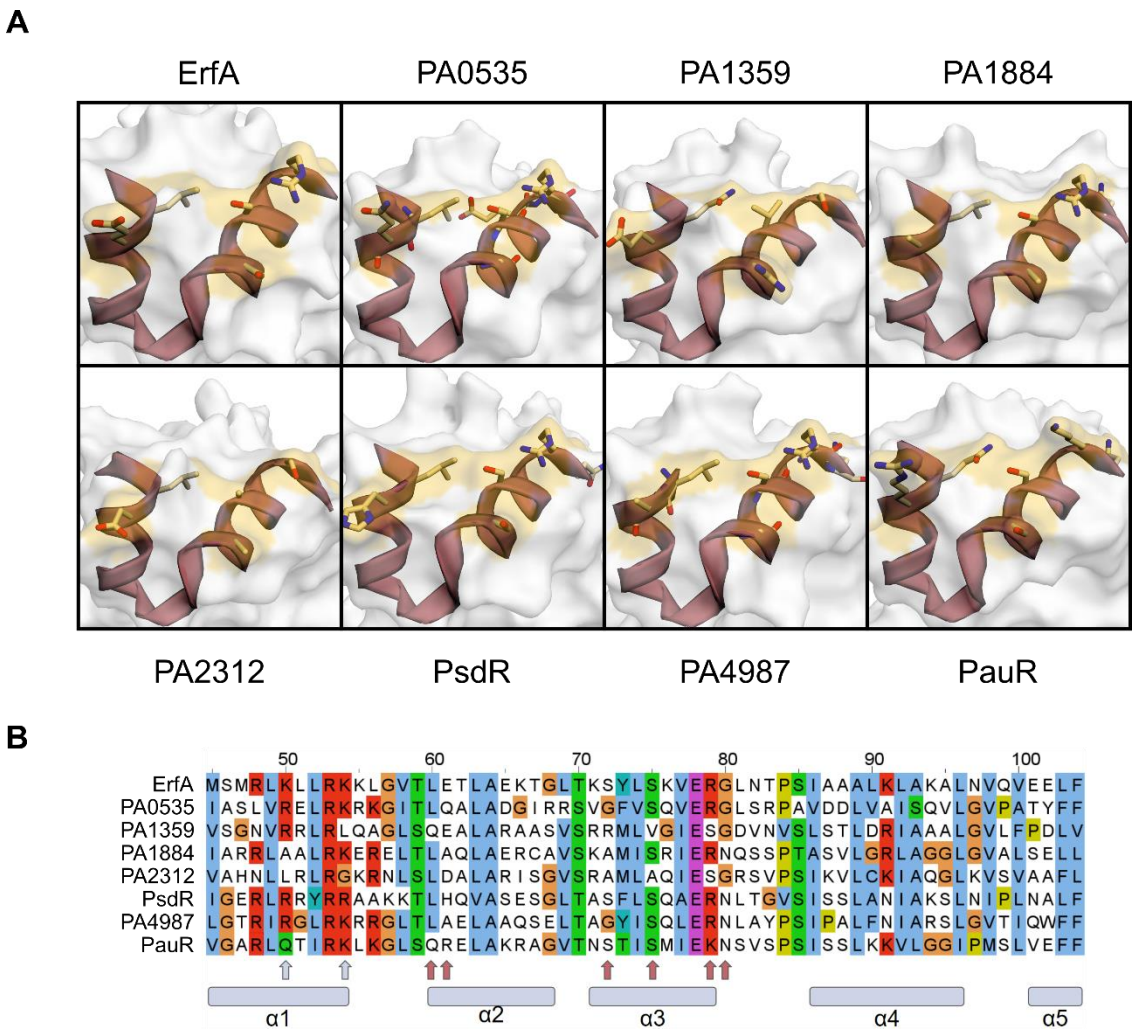


FIGURE S2 : Predicted structures of DNA-binding interfaces of the eight XRE-cupin regulators. (**A**) Stereo ribbon representation of the predicted DNA-binding interfaces of the eight XRE-cupin regulators. Model prediction was done using the SWISS-Model tool (Waterhouse *et al.*, 2018) and the structure with PDB ID 1Y9Q as template. Amino acids predicted to be involved in specific DNA-binding are depicted in yellow. (**B**) Sequence alignment of the eight XRE domains. The five alpha helices are denoted by grey rectangles. Amino acids predicted to be involved in specific and non-specific DNA-binding are shown by red and grey arrows, respectively.

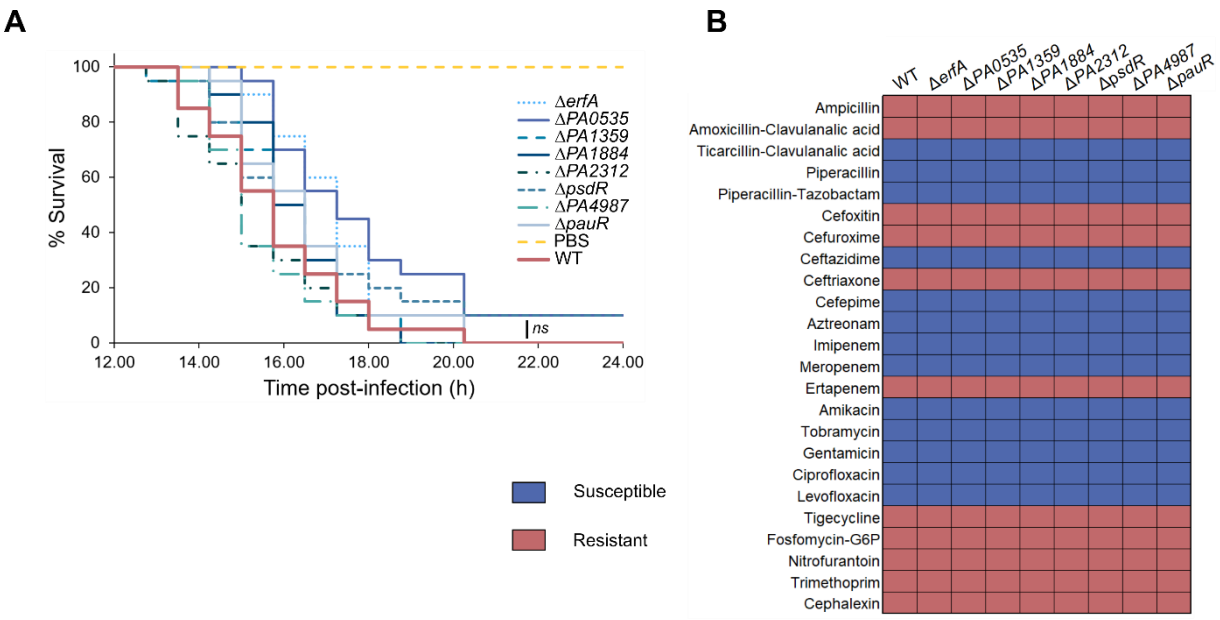


FIGURE S3. The XRE-cupin regulons do not affect bacterial fitness in *G. mellonella*, nor antibiotics resistance. (**A**) Survival curves of *Galleria mellonella* larvae infected with an average of 7-10 bacteria per larva. Twenty larvae were infected per strain. Significance testing was performed using log-rank. (**B**) Antibiotics resistance phenotypes of the XRE-cupin regulatory mutants. Resistance to 24 clinically relevant antibiotics was assessed in liquid using the automated BD Phoenix system.

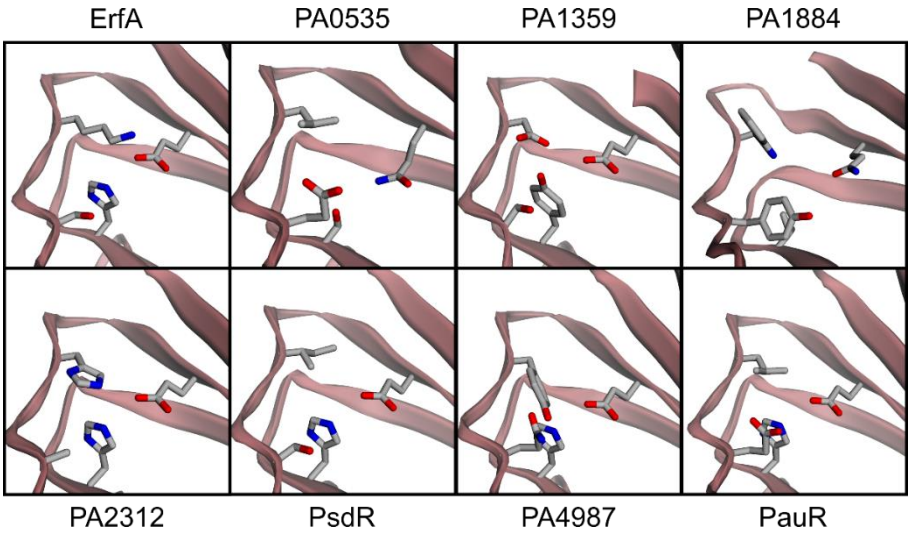


FIGURE S4. Structure modeling of the XRE-cupin putative signal-sensing binding pocket. Model prediction was done using the SWISS-Model tool (Waterhouse *et al.*, 2018) using the structure with PDB ID 1Y9Q as template. Inwards amino acids potentially involved in molecule binding are drawn.

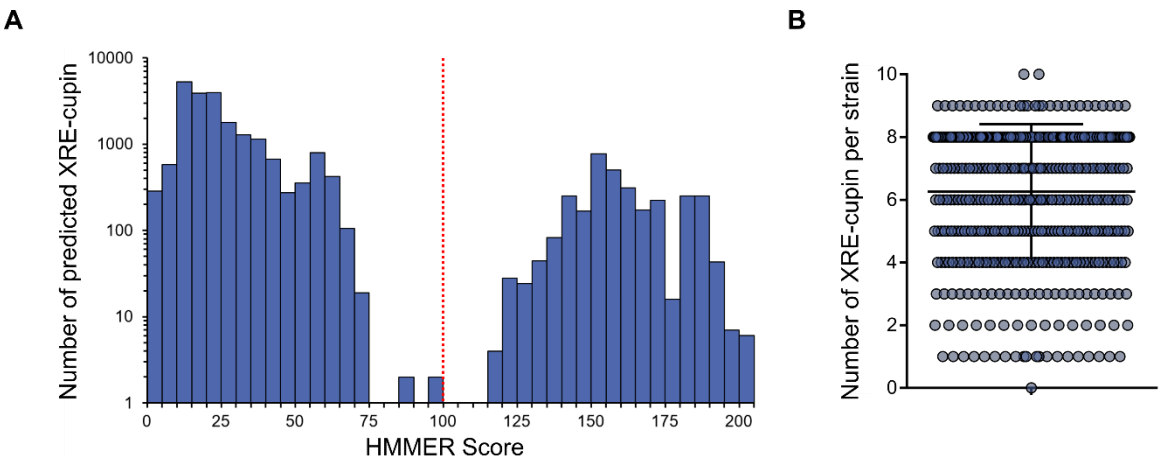


FIGURE S5 : HMMER prediction of XRE-cupin regulators. (**A**) Repartition of HMMER scores of all 23,974 proteins found to match the XRE-cupin HMM model. The bimodal distribution clearly delineates between false and true positives. All eight known *P. aeruginosa* XRE-cupin regulators had scores higher than 140 and all manually inspected predictions with scores < 100 were found not to be XRE-cupin proteins. Consequently, a threshold of > 100 (denoted by a red line) was used to consider true XRE-cupin regulators. (**B**) Repartition of the number of predicted XRE-cupin per strain.

Table S1 | List of bacterial strains and plasmids used in this study

Strain or plasmid	Genotype or relevant properties	Reference/source
Strains		
P. aeruginosa		
PAO1	Wound isolate, sequenced laboratory strain	J. Mougous
PAO1 Δ*erfA*	PAO1 with *erfA* (*PA0225*) deletion	(Trouillon *et al.*, 2020)
PAO1 Δ*erfA::erfA*	PAO1 Δ*erfA* with *erfA* in *attB* site	(Trouillon *et al.*, 2020)
PAO1 Δ*PA0535*	PAO1 with *PA0535* deletion	This work
PAO1 Δ*PA0535::PA0535*	PAO1 Δ*PA0535* with *PA0535* in *attB* site	This work
PAO1 Δ*PA1359*	PAO1 with *PA1359* deletion	This work
PAO1 Δ*PA1359::PA1359*	PAO1 Δ*PA1359* with *PA1359* in *attB* site	This work
PAO1 Δ*PA1884*	PAO1 with *PA1884* deletion	This work
PAO1 Δ*PA1884::PA1884*	PAO1 Δ*PA1884* with *PA1884* in *attB* site	This work
PAO1 Δ*PA2312*	PAO1 with *PA2312* deletion	This work
PAO1 Δ*PA2312::PA2312*	PAO1 Δ*PA2312* with *PA2312* in *attB* site	This work
PAO1 Δ*psdR*	PAO1 with *psdR* (*PA4499*) deletion	This work
PAO1 Δ*psdR::psdR*	PAO1 Δ*psdR* with *psrD* in *attB* site	This work
PAO1 Δ*PA4987*	PAO1 with *PA4987* deletion	This work
PAO1 Δ*PA4987::PA4987*	PAO1 Δ*PA4987* with *PA4987* in *attB* site	This work
PAO1 Δ*pauR*	PAO1 with *pauR* (*PA5301*) deletion	This work
PAO1 Δ*pauR::pauR*	PAO1 Δ*pauR* with *pauR* in *attB* site	This work
E. coli		
TOP10	Chemically competent cells	Invitrogen
BL21 Star (DE3)	F^- *ompT hsdSB* (rB^- mB^-) *gal dcm rne131*(DE3)	Invitrogen
Plasmids		
pRK600	Helper plasmid with conjugative properties (Cm^R)	(Kessler *et al.*, 1992)
pEXG2	Allelic exchange vector (Gm^R)	(Rietsch *et al.*, 2005)
pEXG2-mut-PA0535	pEXG2 carrying SLIC fragment for *PA0535* deletion (Gm^R)	This work
pEXG2-mut-PA1359	pEXG2 carrying SLIC fragment for *PA1359* deletion (Gm^R)	This work
pEXG2-mut-PA1884	pEXG2 carrying SLIC fragment for *PA1884* deletion (Gm^R)	This work
pEXG2-mut-PA2312	pEXG2 carrying SLIC fragment for *PA2312* deletion (Gm^R)	This work
pEXG2-mut-PA4499	pEXG2 carrying SLIC fragment for *PA4499* deletion (Gm^R)	This work
pEXG2-mut-PA4987	pEXG2 carrying SLIC fragment for *PA4987* deletion (Gm^R)	This work
pEXG2-mut-PA5301	pEXG2 carrying SLIC fragment for *PA5301* deletion (Gm^R)	This work
pFLP2	Source of Flp recombinase (Ap^R)	(Hoang *et al.*, 1998)
mini-CTX1	Site-specific integrative plasmid (*attP* site, Tc^R)	(Hoang *et al.*, 2000)
miniCTX-TrrnB-erfA	miniCTX1-TrrnB carrying *erfA* gene for complementation (*attP*, Tc^R)	(Trouillon *et al.*, 2020)
mini-CTX-PA0535	mini-CTX1 carrying *PA0535* gene for complementation (*attP*, Tc^R)	This work
mini-CTX-PA1359	mini-CTX1 carrying *PA1359* gene for complementation (*attP*, Tc^R)	This work
mini-CTX-PA1884	mini-CTX carrying *PA1884* gene for complementation (*attP*, Tc^R)	This work
mini-CTX-PA2312	mini-CTX1 carrying *PA2312* gene for complementation (*attP*, Tc^R)	This work
mini-CTX-PA4499	mini-CTX1 carrying *PA4499* gene for complementation (*attP*, Tc^R)	This work
mini-CTX-PA4987	mini-CTX1 carrying *PA4987* gene for complementation (*attP*, Tc^R)	This work
mini-CTX-PA5301	mini-CTX1 carrying *PA5301* gene for complementation (*attP*, Tc^R)	This work
pET52b	Expression vector (Ap^R)	Novagen
pET52b-PA0535	Expression vector of *PA0535-his10* (Ap^R)	This work
pET52b-PA1359	Expression vector of *PA1359-his10* (Ap^R)	This work
pET52b-PA1884	Expression vector of *PA1884-his10* (Ap^R)	This work
pET52b-PA2312	Expression vector of *PA2312-his10* (Ap^R)	This work
pET52b-PA4499	Expression vector of *PA4499-his10* (Ap^R)	This work
pET52b-PA4987	Expression vector of *PA4987-his10* (Ap^R)	This work
pET52b-PA5301	Expression vector of *PA5301-his10* (Ap^R)	This work

Table S2 | Primers used in this work

Name	Sequence (5' => 3')	Use
pEXG2-mut-PA0535-sF1	**GGTCGACTCTAGAGGATCCCC**TCGTCCTCGACTACTTCCGC	*PA0535* deletion
pEXG2-mut-PA0535-sR1	**GCGTATACCCAGAGCACCCTG<u>TCA</u>**GTCGTCGGCTACTTCGCTCA	*PA0535* deletion
pEXG2-mut-PA0535-sF2	CAGGGTGCTCTGGGTATACG	*PA0535* deletion
pEXG2-mut-PA0535-sR2	**ACCGAATTCGAGCTCGAGCCC**TCGGTGTCTATGTGGCGGTC	*PA0535* deletion
pEXG2-mut-PA1359-sF1	**GGTCGACTCTAGAGGATCCCC** GGTGCGCACCGAGAGATAGA	*PA1359* deletion
pEXG2-mut-PA1359-sR1	**CTGGCGAAGGCATGGAAGTCG<u>TCA</u>**GCCGGAGACATGTTCGAGGA	*PA1359* deletion
pEXG2-mut-PA1359-sF2	CGACTTCCATGCCTTCGCCA	*PA1359* deletion
pEXG2-mut-PA1359-sR2	**ACCGAATTCGAGCTCGAGCCC**ACGCTGGCTGAAGCCGTAGT	*PA1359* deletion
pEXG2-mut-PA1884-sF1	**GGTCGACTCTAGAGGATCCCC**GTTGGTCGAGGAGCAGTTGC	*PA1884* deletion
pEXG2-mut-PA1884-sR1	**CGTTGCAGGTGGAACGACTGC<u>TCA</u>**GATGCGCGAGATCATCGCCT	*PA1884* deletion
pEXG2-mut-PA1884-sF2	GCAGTCGTTCCACCTGCAAC	*PA1884* deletion
pEXG2-mut-PA1884-sR2	**ACCGAATTCGAGCTCGAGCCC**TGCCTGCTCATGCTGGGGTT	*PA1884* deletion
pEXG2-mut-PA2312-sF1	**GGTCGACTCTAGAGGATCCCC**GCAGGTACGCGAGAGCCTTT	*PA2312* deletion
pEXG2-mut-PA2312-sR1	**TAGAACAGGATCGAGTCGCCG<u>TCA</u>**GACGCGCTGGCCAATCAGAT	*PA2312* deletion
pEXG2-mut-PA2312-sF2	CGGCGACTCGATCCTGTTCT	*PA2312* deletion
pEXG2-mut-PA2312-sR2	**ACCGAATTCGAGCTCGAGCCC**CTGGCGATGATGGGCATGTC	*PA2312* deletion
pEXG2-mut-PA4499-sF1	**GGTCGACTCTAGAGGATCCCC**TCACTGGGGTGCCATTGCAC	*PA4499* deletion
pEXG2-mut-PA4499-sF1	**CAGGCCCATGGTGGTGATGGT<u>TCA</u>**TTCGCTGGCGACTTGGTGCA	*PA4499* deletion
pEXG2-mut-PA4499-sF1	ACCATCACCACCATGGGCCT	*PA4499* deletion
pEXG2-mut-PA4499-sF1	**ACCGAATTCGAGCTCGAGCCC**TAGTCGGTGCCGGTGGTGTA	*PA4499* deletion
pEXG2-mut-PA4987-sF1	**GGTCGACTCTAGAGGATCCCC**GCTGTATGTGCGCGACCAGT	*PA4987* deletion
pEXG2-mut-PA4987-sF1	**GGTGATCACCCAGAGCACCAG<u>TCA</u>**GGTGCCGAGGAAGTGCGTTT	*PA4987* deletion
pEXG2-mut-PA4987-sF1	CTGGTGCTCTGGGTGATCAC	*PA4987* deletion
pEXG2-mut-PA4987-sF1	**ACCGAATTCGAGCTCGAGCCC**TGCTGCTCGGCGATACCATG	*PA4987* deletion
pEXG2-mut-PA5301-sF1	**GGTCGACTCTAGAGGATCCCC** TACAGCACCGAGCGTAGCCA	*PA5301* deletion
pEXG2-mut-PA5301-sF1	**GTGGTGGCGCTGATCAAGCGT<u>TCA</u>**TTGCAGACGAGCACCGACGT	*PA5301* deletion
pEXG2-mut-PA5301-sF1	ACGCTTGATCAGCGCCACCA	*PA5301* deletion
pEXG2-mut-PA5301-sF1	**ACCGAATTCGAGCTCGAGCCC**CTTGCCGACGATGTCTTCGC	*PA5301* deletion
Comp-CTX-PA0535-sF1	**GATATCGAATTCCTGCAGCCC**CAAGGAGATCGCCAAGGAACTG	*PA0535* complementation
Comp-CTX-PA0535-sR1	CAGGAACATCGCGACCTCCGT	*PA0535* complementation

Comp-CTX-PA0535-sF2	**ACGGAGGTCGCGATGTTCCTG**CGTTTGCGCGACCTGTTCTGA	*PA0535* complementation
Comp-CTX-PA0535-sR2	**TCTAGAACTAGTGGATCCCCC**CGACAGGGTAACGCGCCT	*PA0535* complementation
Comp-CTX-PA1359-sF	**GATATCGAATTCCTGCAGCCC**CGAAGAGGATGTACAGCGCCC	*PA1359* complementation
Comp-CTX-PA1359-sR	**TCTAGAACTAGTGGATCCCCC**TCCTCTGGAAGGCTGGCTCGA	*PA1359* complementation
Comp-CTX-PA1884-sF	**GATATCGAATTCCTGCAGCCC**GCTGCCGCTGACCGTGTCCA	*PA1884* complementation
Comp-CTX-PA1884-sR	**TCTAGAACTAGTGGATCCCCC**CCTCGACTGGTCGCCGGAG	*PA1884* complementation
Comp-CTX-PA2312-sF1	**GATATCGAATTCCTGCAGCCC**TGAAGGCGATCTGCTGCGCC	*PA2312* complementation
Comp-CTX-PA2312-sR1	GCGGTCCATGAGGCATTCCTC	*PA2312* complementation
Comp-CTX-PA2312-sF2	**GAGGAATGCCTCATGGACCGC**CGGGCCTGACCGATGAACCT	*PA2312* complementation
Comp-CTX-PA2312-sR2	**TCTAGAACTAGTGGATCCCCC**TCTTCGTCCATGGCAGCAAGG	*PA2312* complementation
Comp-CTX-PA4499-sF	**GATATCGAATTCCTGCAGCCC**TTCCTCCCAGGTGTGGATGGC	*PA4499* complementation
Comp-CTX-PA4499-sR	**TCTAGAACTAGTGGATCCCCC**GATGGGAAGGTGCGCTACGGA	*PA4499* complementation
Comp-CTX-PA4987-sF1	**GATATCGAATTCCTGCAGCCC**CCTCGACCTGGACCAGGTCC	*PA4987* complementation
Comp-CTX-PA4987-sR1	GAAATTCACGGTTGGCGTCCTT	*PA4987* complementation
Comp-CTX-PA4987-sF2	**AGGACGCCAACCGTGAATTTC**AAGGCCGCCTGGAGCATCTGA	*PA4987* complementation
Comp-CTX-PA4987-sR2	**TCTAGAACTAGTGGATCCCCC**CTGGCCCTGCATCGCGACAG	*PA4987* complementation
Comp-CTX-PA5301-sF	**GATATCGAATTCCTGCAGCCC**CGCTACAGCACCGAGCGTAG	*PA5301* complementation
Comp-CTX-PA5301-sR	**TCTAGAACTAGTGGATCCCCC**CTTGCGGCATTAGGTCGCAGC	*PA5301* complementation
pET52b-PA0535-sF	**TTAAGAAGGAGATATACC**ATGAGCGAAGTAGCCGACGA	*PA0535* overexpression
pET52b-PA0535-sR	**CTACCGCGTGGCACCAGAGCGAG**GGCGTATACCCAGAGCACC	*PA0535* overexpression
pET52b-PA1359-sF	**TTAAGAAGGAGATATACC**ATGGAAGAGGTCCGGCAGTG	*PA1359* overexpression
pET52b-PA1359-sR	**CTACCGCGTGGCACCAGAGCGAG**GCTCACCACGTTGCGGGTG	*PA1359* overexpression

pET52b-PA1884-sF	**TTAAGAAGGAGATATACC**ATGGATATCGACGAACTGATTG	*PA1884* overexpression
pET52b-PA1884-sR	**CTACCGCGTGGCACCAGAGCGAG**GTCGTGGAGGATCACCAGC	*PA1884* overexpression
pET52b-PA2312-sF	**TTAAGAAGGAGATATACC**ATGCATACCGAACCCGATGATC	*PA2312* overexpression
pET52b-PA2312-sR	**CTACCGCGTGGCACCAGAGCGAG**GTCCAGGCGCTCCGGGTA	*PA2312* overexpression
pET52b-PA4499-sF	**TTAAGAAGGAGATATACC**ATGACCGTAGACCGCATCGG	*PA4499* overexpression
pET52b-PA4499-sR	**CTACCGCGTGGCACCAGAGCGAG**GGGCGTCGGATGGTCGTC	*PA4499* overexpression
pET52b-PA4987-sF	**TTAAGAAGGAGATATACC**ATGCCCCGCCCGTCACCG	*PA4987* overexpression
pET52b-PA4987-sR	**CTACCGCGTGGCACCAGAGCGAG**GAACGTCGGCGGGGTGATC	*PA4987* overexpression
pET52b-PA5301-sF	**TTAAGAAGGAGATATACC**ATGGACGTCGGTGCTCGTCT	*PA5301* overexpression
pET52b-PA5301-sR	**CTACCGCGTGGCACCAGAGCGAG**GAAATTTGCGGGCGTGGTGG	*PA5301* overexpression
pEXG2-mut-PA0534-BS-sF1	**GGTCGACTCTAGAGGATCCCC**GTTCGACCTCATCTGCCAGG	EMSA (P*PA0534* mut BS)
pEXG2-mut-PA0534-BS-sR1	**GGTTATTTGAGCTCAAAAATAA**CACTTAAAATTTCACGACACCAG	EMSA (P*PA0534* mut BS)
pEXG2-mut-PA0534-BS-sF2	**TTATTTTTGAGCTCAAATAACC**AATATAATTTCACGGAGGTCGCG	EMSA (P*PA0534* mut BS)
pEXG2-mut-PA0534-BS-sR2	**ACCGAATTCGAGCTCGAGCCC**ACGAAGCGTAGCCCCTCGTA	EMSA (P*PA0534* mut BS)
CTX-PA2776-lacZ-sF	**GATATCGAATTCCTGCAGCCC**GGGAAGCGTGAAGAAGTCCG	EMSA (P*PA2776* mut BS)
CTX-PA2776-mut-lacZ-sR	**GCTAGTTAGTTAGGATCCCCC**TGCGGCATCTTGCACCTCGCATTTTT GCTTTTTCAC GATGCCCGGATGCTACCCGA	EMSA (P*PA2776* mut BS)
CTX-mdpA-lacZ-sF	**GATATCGAATTCCTGCAGCCC**CCGCATTGATCAGGTTGCCG	EMSA (P*mdpA* mut BS)
CTX-mdpA-mut-lacZ-sR	**GCTAGTTAGTTAGGATCCCCC**TGTTCTCCTCGTCTGGACCGGGCTGA AATTTAAGTCGAGAACTGATTCAGGTCAAAACAAATTTTCAGT	EMSA (P*mdpA* mut BS)
(Cy5-) pergAB-EMSA-F	(Cy5-) CAGCCTTCTCCCGATGGCAGT	EMSA (Trouillon *et al.*, 2020)
pPA0534-EMSA-F	CAGCCTTCTCCCGATGGCAGT CTCGGGACTGGTGTCGTGAA	EMSA
pPA0534-EMSA-R	CAGCCTTCTCCCGATGGCAGT AGGAACATCGCGACCTCCGT	EMSA
pPA1360-EMSA-F	CAGCCTTCTCCCGATGGCAGT GACATGTTCGAGGACGCTGG	EMSA
pPA1360-EMSA-R	CAGCCTTCTCCCGATGGCAGT CGGGAAACGGGATGTGCACT	EMSA

pPA1885-EMSA-F	CAGCCTTCTCCCGATGGCAGT TATCCATGAGTGGCCCTGGC	EMSA
pPA1885-EMSA-R	CAGCCTTCTCCCGATGGCAGT GGGCAAGCTCCAATGCAAGG	EMSA
pPA2776-EMSA-F	CAGCCTTCTCCCGATGGCAGT CGATTTCGACTCGGGTAGCAT	EMSA
pPA2776-EMSA-2-R	CAGCCTTCTCCCGATGGCAGT TGCGGCATCTTGCACCTCTG	EMSA
pPA4498-EMSA-F	CAGCCTTCTCCCGATGGCAGT CCCGAACCCCTATTTTTAAGTG	EMSA
pPA4498-EMSA-R	CAGCCTTCTCCCGATGGCAGT TGTTCTCCTCGTCTGGACCG	EMSA
pPA4985-6-EMSA-F	CAGCCTTCTCCCGATGGCAGT ACGGGACAGAAACTATCCAGTC	EMSA
pPA4985-6-EMSA-R	CAGCCTTCTCCCGATGGCAGT GCGCTTCAGAAACAGGATATGA	EMSA
uvrD up	CATATCCTGGTGGACGAGTTCC	RT-qPCR
uvrD down	CGCTGAACTGCTGGATGTTCTC	RT-qPCR
qPCR_pauB1_up	ACCTATGGACGGCGATCCTG	RT-qPCR
qPCR_pauB1_down	ACGAAGCGTAGCCCCTCGTA	RT-qPCR
qPCR_fumC1_up2	TCGGGCAACTTCGAACTGAA	RT-qPCR
qPCR_fumC1_down2	GAGCTTGCCCTGGTTGACCT	RT-qPCR
qPCR_PA0265_up2	TGTTCCGCTTCAAGGACGAG	RT-qPCR
qPCR_PA0265_down2	CCATGCCGTACTCCAGTTGC	RT-qPCR
qPCR_PA1360_up	TCCCTGGCGAAAAGCATGT	RT-qPCR
qPCR_PA1360_down	AGGAGCAGCAGCAGGATCAG	RT-qPCR
qPCR_PA1541_up	TCGCCTACGCCCTTTGGGAA	RT-qPCR
qPCR_PA1541_down	TACAGGCCGATGCTTTCGCC	RT-qPCR
qPCR_PA1885_up2	CCGAACGGCTGTACCTGAAG	RT-qPCR
qPCR_PA1885_down2	CAGCCAGGCGCTTGTAGAAA	RT-qPCR
qPCR_PA4498_up2	GCATCAGCACCACCGAAGTC	RT-qPCR
qPCR_PA4498_down2	ACCGAACAGGACGATGCAGA	RT-qPCR
qPCR_PA4500_up	CAACGCCGACGATGTGCTGT	RT-qPCR
qPCR_PA4500_down	TTGTCCAGGCCCATGTCGGT	RT-qPCR
qPCR_PA4985_up	CGTGGTGATGACGTCCACCT	RT-qPCR
qPCR_PA4985_down	ACCACCGCCCAGTAGTCCAT	RT-qPCR
qPCR_PA4986_up	CTGTTCAGCCCCCTGGAAAT	RT-qPCR
qPCR_PA4986_down	GTGGCTCGTTGACCAGGTTG	RT-qPCR

REFERENCES

Hoang, T.T., R.R. Karkhoff-Schweizer, A.J. Kutchma & H.P. Schweizer, (1998) A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked Pseudomonas aeruginosa mutants. *Gene* 212: 77-86.

Hoang, T.T., A.J. Kutchma, A. Becher & H.P. Schweizer, (2000) Integration-proficient plasmids for Pseudomonas aeruginosa: site-specific integration and use for engineering of reporter and expression strains. *Plasmid* 43: 59-72.

Kessler, B., V. de Lorenzo & K.N. Timmis, (1992) A general system to integrate lacZ fusions into the chromosomes of gram-negative eubacteria: regulation of the Pm promoter of the TOL plasmid studied with all controlling elements in monocopy. *Mol Gen Genet* 233: 293-301.

Rietsch, A., I. Vallet-Gely, S.L. Dove & J.J. Mekalanos, (2005) ExsE, a secreted regulator of type III secretion genes in Pseudomonas aeruginosa. *Proc Natl Acad Sci U S A* 102: 8006-8011.

Trouillon, J., E. Sentausa, M. Ragno, M. Robert-Genthon, S. Lory, I. Attree & S. Elsen, (2020) Species-specific recruitment of transcription factors dictates toxin expression. *Nucleic Acids Res* 48: 2388-2400.

Waterhouse, A., M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F.T. Heer, T.A.P. de Beer, C. Rempfer, L. Bordoli, R. Lepore & T. Schwede, (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46: W296-W303.

4.2.2 Additional results

One of our major goals for this project, which still remains unachieved, was to better characterize the signal-sensing mechanism of these TFs. Unfortunately, the covid-19 lockdown dramatically shortened my lab time available for the needed experiments. Before the lockdown, I was starting these experiments and firstly tried to prove the involvement of the cupin domain in specific signal sensing. As a proof of concept, I wanted to use PauR, which has been proposed to change its DNA binding activity upon sensing of putrescine (Chou et al., 2013). To determine the role of the cupin domain in that sensing mechanism, I constructed three different chimeric genes, containing the DNA-binding domain of ErfA fused to the cupin domain of PauR, each one with a different fusion site determined based on predicted structure in order to improve chances to obtain a stable protein. The use of *exlBA* transcriptional fusion in presence or absence of putrescine could have confirmed this engineered putrescine-dependent *exlBA* regulation and proved the functional specificity between DNA-binding activity of the XRE domain and the potential signal-sensing activity of the cupin domain. However, I never achieved having any of the chimera to be expressed and stable, as assessed by western blotting.

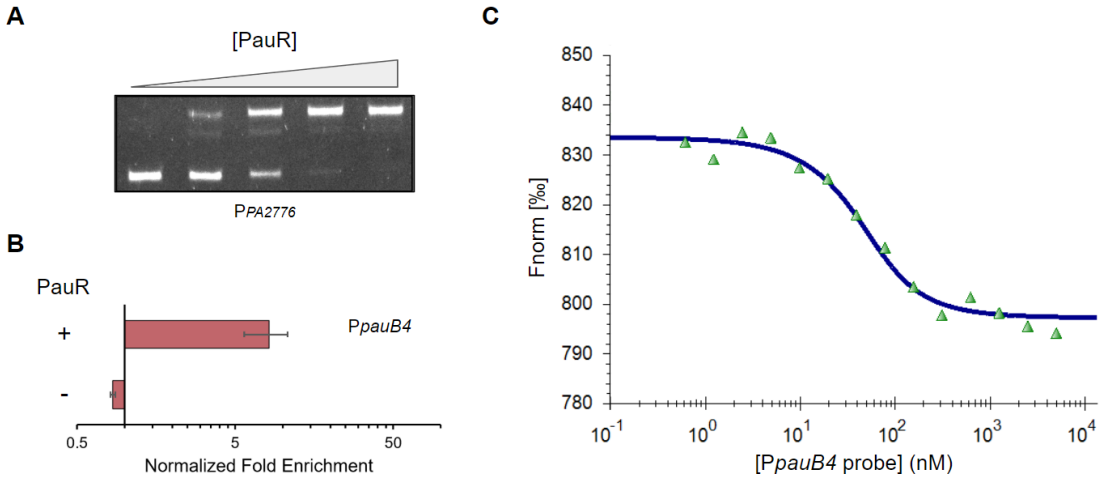


Figure 4.12: Measuring PauR DNA-binding. (A) EMSA binding assay with purified PauR protein (0, 10, 50, 100 and 200 nM) incubated with 0.5 nM Cy5-labeled *PA2776* promoter DNA probe for 15 min before electrophoresis. (B) DAP-qPCR on PAO1 sheared genomic DNA with purified PauR (0 or 500 ng). Results are shown as fold enrichment of *pauB4* promoter region compared to *uvrD* coding region and represented as mean ± standard deviation (SD) from three independent experiments. (C) Microscale thermophoresis assay using 100 nM PauR and increasing concentrations of DNA probe containing *pauB4* promoter with a MST power of 40% and LED on at 95%. Binding was done for 15 min at room temperature before measurement.

In related experiments, I was also trying to confirm the impact of putrescine sensing on PauR activity. To that aim, I tried several approaches, as each was failing one after another. Briefly, I was always able to confirm specific PauR-DNA interactions using EMSA, DAP-qPCR and microscale thermophoresis (Figure 4.12). Microscale thermophoresis is an *in vitro* assay based on measuring temperature-induced movements and fluorescence changes of potentially interacting molecules (Wienken et al., 2010). However, all three approaches failed to prove the PauR-putrescine interaction or any direct effect of putrescine on PauR DNA-binding activity (Figure 4.13). These results were surprising as a previous study showed that putrescine induces PauR-dependent expression changes of PauR targets *in vivo*, and suggested a PauR-putrescine interaction *in vitro* by chemical crosslinking (Chou et al., 2013). Additionally, the *E. coli* homolog of PauR, PuuR, was shown to directly interact with putrescine which leads to a decrease in DNA binding activity upon sensing (Nemoto et al., 2012b). In this report, the interaction had a relatively small effect on DNA binding and seems to happen in short ranges of concentrations which might explain why I missed it. Another explanation could be that PauR interacts with a product metabolite derived from putrescine instead of putrescine itself, which could expain the *in vivo* effect and why the PauR-putrescine interaction was hard to detect.

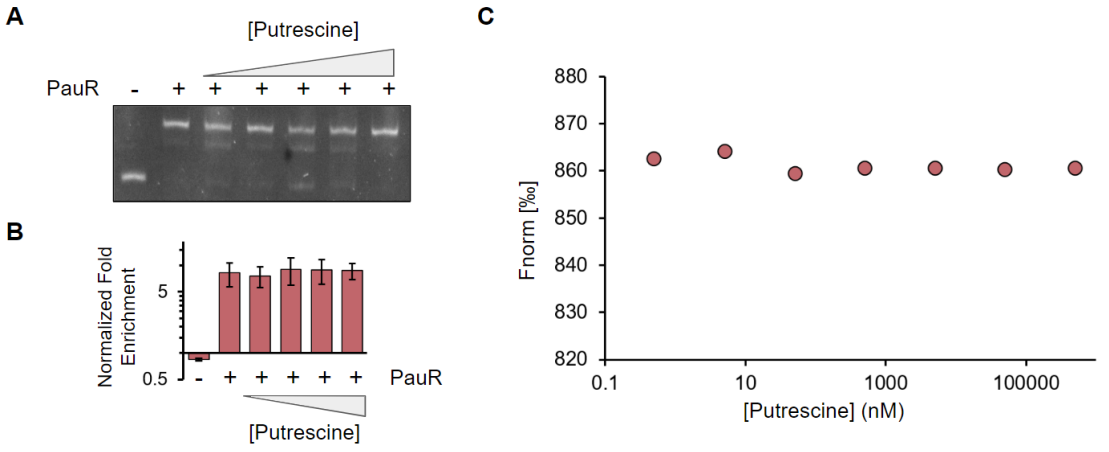


Figure 4.13: Assessing PauR-putrescine interaction. (A) EMSA binding assay with purified PauR protein (0 or 100 nM) incubated with 0.5 nM Cy5-labeled *PA2776* promoter DNA probe and increasing concentration of putrescine for 15 min before electrophoresis. (B) DAP-qPCR on PAO1 sheared genomic DNA with purified PauR (0 or 500 ng) and increasing concentration of putrescine. Results are shown as fold enrichment of *pauB4* promoter region compared to *uvrD* coding region and represented as mean ± standard deviation (SD) from three independent experiments. (C) Microscale thermophoresis assay using 100 nM PauR, 500nM DNA probe containing *pauB4* promoter and increasing concentration of putrescine with a MST power of 40% and LED on at 95%. Binding was done for 15 min at room temperature before measurement.

Additional Materials & Methods

Electromobility Shift Assays. EMSAs were performed as previously described (Trouillon et al., 2020b).

DAP-qPCR. DAP was performed as previously described (Trouillon et al., 2020b). The resulting eluted DNA was then used as input for qPCR using the Luna Universal qPCR kit (NEB) and primers specific to the *pauB4* promoter region recognized by PauR and to the *uvrD* gene for normalization. DAP were performed in triplicates, and no PauR protein was added in negative controls.

Microscale thermophoresis. Purified PsdR was dialyzed to a buffer containing 50 mM HEPES, 250 mM NaCl, 50 mM KCl and 5% Glycerol at pH 7.2 and then labeled with the Monolith Protein Labeling RED-NHS Kit following manufacturer's instructions. Putrescine solutions were prepared in ligand buffer containing 50 mM Tris, 50 mM KCl, 20 mM $MgCl_2$ and 5 % Glycerol at pH 7. Putative interactions were monitored by mixing labeled PsdR at a final concentration of 100 nM with putrescine solution at a 1:1 volume ratio and incubating at room temperature for 15 minutes. Microscale thermophoresis was performed as previously described using a Monolith NT.115 device with premium capillaries (Ngo et al., 2020).

4.3 The Response Regulators of *P. aeruginosa*

Julian Trouillon✉, Lionel Imbert, Anne-Marie Villard, Thierry Vernet, Ina Attrée & Sylvie Elsen✉

Work in progress

After setting up DAP-seq and realizing its potential for the study of several TFs at once, we decided to use it to improve our knowledge on *P. aeruginosa* regulatory networks at larger scales. Indeed, most of *P. aeruginosa* TFs are still uncharacterized or poorly characterized and DAP-seq can be used at the scale of dozens of TFs. In the case of TCSs, the fact that they are only active in response to specific signals, unknown in most cases, makes it challenging to assess they role *in vivo*, let alone determining their entire regulons (Rajeev et al., 2020). In order to provide a first global view on TCS regulatory networks in *P. aeruginosa*, we obtained an IDEX (UGA) funding and worked in collaboration with two platforms from IBS, Grenoble: the RobioMol platform for plasmid generation and the cell-free platform for protein expression. In this project, we expressed and purified all DNA-binding response regulators of PAO1 and used DAP-seq, which circumvents the limitation of having to know the signal for the study of TCSs, to determine their regulons. This is still an ongoing project which will comprise about 400 DAP-seq experiments, with 312 that have already been done, of which 104 (representing all 51 DNA-binding RRs in PAO1) have been sequenced and analyzed and are presented here. In PAO1, we identified a total of 6,290 binding sites in the promoter regions of 66 % of all genes, illustrating the major importance of TCSs in global regulatory networks. Additionally, the characterization of RRs is currently being done in two more strains (PA14 and IHMA87), which represent with PAO1 the three major *P. aeruginosa* phylogenetic lineages. This comparative approach will provide valuable information about the conservation and specificity of the TCS regulatory network in *P. aeruginosa*.

In this project, I performed all DAP-seq experiments and data analysis, I did part of the plasmid clonings (along with AMV, TV and SE) and LI performed cell-free protein expressions.

As stated in section 2.2.1, *P. aeruginosa* possesses one of the highest number of TCSs, with 72 RRs in PAO1 including 51 with DNA-binding activity (Table 4.1), as followed:

- 4 RRs have been thoroughly studied with combinations of RNA-seq, ChIP-seq and confirmation experiments (GacA, AlgR, PhoB and BfmR), as parts of the PAGnet regulatory network (Huang et al., 2019).
- the effect on global gene expression has been assessed for 19 RRs (mostly microarrays)
- 8 RRs have few targets identified (through targeted approaches such as EMSAs or footprints)
- 20 RRs have no confirmed targets identified

Consequently, even though the RR family is one of the most studied families, we still don't known the complete regulons of most of them. Additionally, the vast majority of those for which we have target information have been studied only in one strain. However, regulatory networks are not identical between bacteria and even between strains of the same species. This regulatory plasticity can occur from differences in promoter sequences or in TF or target genes content and could explain phenotypic differences between strains. In this project, in order to fully characterize the RRs role across *P. aeruginosa* lineages and to get a precise view of regulatory networks plasticity, we expressed and used in DAP-seq all 51 PAO1 DNA-binding RRs on the genomes of three different *P. aeruginosa* strains. The libraries are currently being sequenced for PA14 and IHMA87, and the PAO1 results are presented below.

In the ErfA and XRE-cupin regulators projects, we were expressing the TFs in *E. coli* and purifying them before use for DAP-seq (Trouillon et al., 2020a; Trouillon et al., 2020b). While *E. coli* expression ensures proper purity of the proteins, it is also time-consuming and thus cannot be easily used at the scale of dozens of TFs. Consequently, we used cell-free protein expression for this project, as described in the original DAP-seq protocol (Bartlett et al., 2017). To that aim, we cloned all 51 RR genes from PAO1 into the pIVEX2.4d cell-free expression vector (Roche) with the help of the RoBioMol molecular biology platform (Anne-Marie Villard and Thierry Vernet, IBS - Grenoble). This plasmid allows the addition of a N-terminal His-tag, which is ideal for RRs as their DNA-binding domain is at the C-terminal end. All

Table 4.1: List of PAO1 RRs.

ID	Name	PAO1 ID	Protein length (aa)	Family
1	PA0034	PA0034	208	NarL-like
2	CreB	PA0463	230	OmpR-like
3	AgtR	PA0601	211	NarL-like
4	TctD	PA0756	224	OmpR-like
5	PirR	PA0929	240	OmpR-like
6	FleR	PA1099	474	NtrC-like
7	PA1157	PA1157	237	OmpR-like
8	PhoP	PA1179	226	OmpR-like
9	AauR	PA1335	426	NtrC-like
10	PA1397	PA1397	211	NarL-like
11	PA1437	PA1437	230	OmpR-like
12	KdpE	PA1637	231	OmpR-like
13	ParR	PA1799	236	OmpR-like
14	ErbR	PA1978	222	NarL-like
15	EraR	PA1980	226	NarL-like
16	PA2376	PA2376	214	NarL-like
17	PA2479	PA2479	227	OmpR-like
18	CzcR	PA2523	225	OmpR-like
19	GacA	PA2586	215	NarL-like
20	BqsR	PA2657	224	OmpR-like
21	PfeR	PA2686	306	OmpR-like
22	CopR	PA2809	227	OmpR-like
23	AtvR	PA2899	213	NarL-like
24	RocA2	PA3045	208	NarL-like
25	CprR	PA3077	224	OmpR-like
26	GltR	PA3192	243	OmpR-like
27	CpxR	PA3204	226	OmpR-like
28	ErdR	PA3604	218	NarL-like
29	TtsR	PA3714	214	NarL-like
30	NarL	PA3879	220	NarL-like
31	RocA1	PA3948	210	NarL-like
32	PA4032	PA4032	239	OmpR-like
33	PA4080	PA4080	215	NarL-like
34	BfmR	PA4101	247	OmpR-like
35	BfiR	PA4196	215	NarL-like
36	PprB	PA4296	276	NarL-like
37	ColR	PA4381	228	OmpR-like
38	RoxR	PA4493	187	ActR-like
39	PilR	PA4547	446	NtrC-like
40	CbrB	PA4726	479	NtrC-like
41	PmrA	PA4776	222	OmpR-like
42	IrlR	PA4885	230	OmpR-like
43	AruR	PA4983	245	OmpR-like
44	NtrC	PA5125	477	NtrC-like
45	DctD	PA5166	463	NtrC-like
46	AmgR	PA5200	248	OmpR-like
47	AlgR	PA5261	249	LytTR-like
48	PhoB	PA5360	230	OmpR-like
49	PA5364	PA5364	301	OmpR-like
50	AlgB	PA5483	450	NtrC-like
51	MifR	PA5511	448	NtrC-like

cell-free protein expressions were then performed at the cell-free expression platform (Lionel Imbert, IBS - Grenoble).

4.3.1 DAP-seq optimization

In this project, I had to adapt my DAP-seq protocol to (i) the high number of samples to be processed, (ii) cell-free expressed proteins and (iii) the inherent need of RRs to be activated. To test how the protocol modifications described below affect DAP-seq output, we also cloned *erfA* into a pIVEX2.3d (equivalent of pIVEX2.4d but with a C-terminal tag), expressed it in cell-free along with the others and used it as positive control.

First, I scaled up the protocol by lowering incubation volumes previously used (Trouillon et al., 2020a) to perform the experiment in 96-well plates (see method section 4.3.4), which allowed us to perform within one day all protein expressions and DAP-seq for one genome (104 samples for PAO1: 51 proteins + 1 negative control, in duplicates).

As stated above, RRs usually need to be phosphorylated to be active. In *in vitro* experiments, this can be overcome by the use of small phosphate donors, such as acetyl phosphate (AcP), that allow autophosphorylation of the RR receiver domain and thus RR activation (Bourret, 2010). Another approach used has been to express and purify the cognate HK and use it for RR phosphorylation *in vitro* (Yamamoto et al., 2005). As AcP as been successfully used in recent RR DAP-seq experiments (Garber et al., 2018; Zhang et al., 2020) and its use is much more convenient than having to purify all HKs, we decided to use it for RR activation in this project. In addition to AcP, 10 mM $MgCl_2$ were also added to the binding buffer as divalent cations are necessary for RR receiver domain phosphorylation (Bourret, 2010). To determine the volume of cell-free expression extracts needed and assess the effect of AcP addition, I used DAP-qPCR with ErfA and two RRs from the two most abundant RR families (BfiR for NarL-like and ParR for OmpR-like RRs) and primers specific to one of their known binding sites. While target enrichment increased with increasing protein amounts, showing that DAP was working, there was no benefit from the addition of 50µM AcP (Figure 4.14A). This has already been reported in a RR DAP-seq experiment for which 2 RRs (CopR1 and CopR2) bound similarly to DNA with or without AcP (Garber et al., 2018). Additionally, a large increase in AcP concentration did not result in further increase of target enrichment (Figure 4.14B).

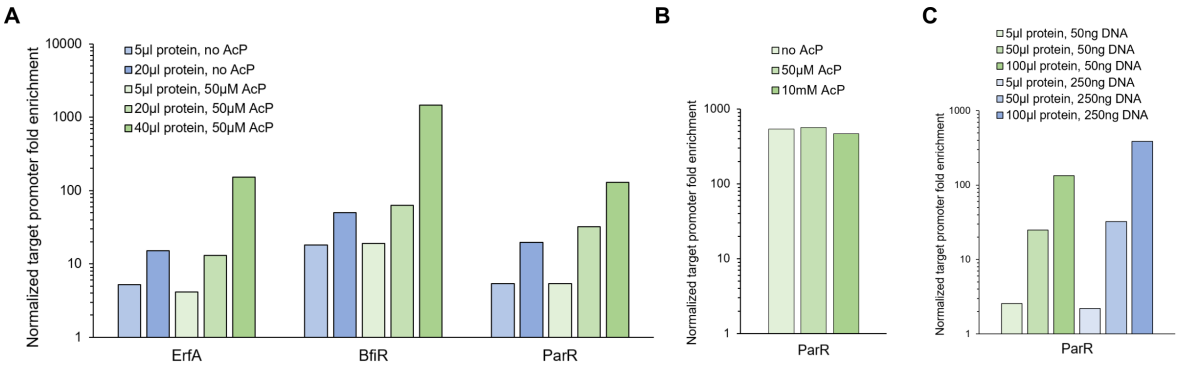


Figure 4.14: Setting up of RR DAP-seq experiments. DAP-qPCR on PAO1 sheared genomic DNA with cell-free soluble protein extracts. Results are shown as fold enrichment of target promoter region between TF protein samples and control protein samples (untagged Gfp protein) normalized to the control *uvrD* coding region with varying protein volumes and AcP concentration (**A**), varying AcP concentration (**B**) or varying protein volumes and DNA input amounts (**C**). Target regions were promoter regions of *ergAB*, *cafA* and *PA1797* for ErfA, BfiR and ParR, respectively.

There was also no effect of the increase of input genomic DNA, but increasing protein extract volumes up to 100 µl induced the best target enrichment (Figure 4.14C). The fact that AcP had no effect on DAP-qPCR target enrichment could be explained in two ways: (i) as the cell-free protein expression mixture already contains a phosphate donor (creatine phosphate), RRs might already be phosphorylated and thus wouldn't need AcP in the binding reaction; however we could not test this hypothesis as creatine phosphate is necessary for protein expression, and (ii) RRs might not need to be phosphorylated *in vitro* to induce enrichment of target DNA. Indeed, it has already been shown that unphosphorylated RRs are still able to bind their DNA targets *in vitro*, although they bind with lower affinity than when phosphorylated (Siam and Marczynski, 2000; Jones et al., 2005). We are still trying to decipher which of these hypotheses is true and will, to that aim, soon check cell-free expressed RR phosphorylation status with or without AcP using SDS-PAGE gels sensitive to phosphorylated residues (Phos-tag). Altogether, although we are still unsure about the phosphorylation status of our cell-free expressed RR proteins, there still was very good enrichment of DNA target regions in DAP-qPCR which allowed us to move to the actual DAP-seq experiments. Even though there was no effect with the tested RRs, we decided to keep AcP and $MgCl_2$ in the buffer in case it might still help with some of the so far untested RRs.

4.3.2 Expression of the RRs

After the cloning of all 51 DNA-binding RR genes in expression plasmids, we expressed all proteins in *E. coli*-derived cell-free extracts (Figure 4.15). Overall, the proteins were well expressed and soluble. Among few exceptions with low-to-no protein amount detected, the RRs number 30, 41, 42 and 47 all resulted in failed DAP-seq experiment as stated below. We recently found out that the expression plasmids of RRs number 30 and 42 carried PCR error in their RR sequence, that were somehow missed in our previous plasmid sequencing verification, which might explain the lack of protein expression. The two corresponding plasmids are currently being remade to perform DAP-seq again for these RRs. RRs number 41 and 47 (PmrA and AlgR, respectively) might be unstable in our conditions. After protein expression, I assessed the ability of all RRs to bind to the magnetic cobalt beads and their purity (Figure 4.16). Overall, proteins retrieved from beads were relatively pure. Some differences visible between western blot and silver staining results were later disproved by redoing the experiment which resulted in proteins detected with both approaches.
Altogether, expression and purity were overall very satisfactory and we thus went

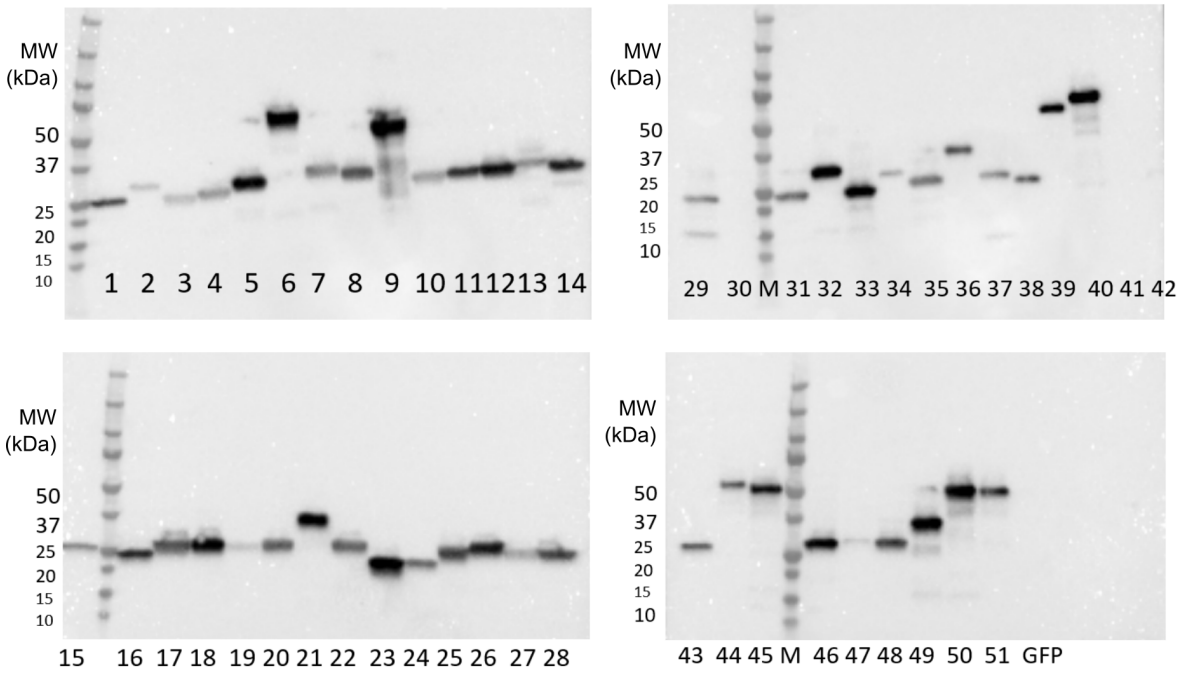


Figure 4.15: Western Blot analysis of soluble protein extracts after expression of the 51 PAO1 RRs. Immunodetection of RRs in supernatants of cell-free protein extracts using anti-His antibodies. RRs are numbered by ID as described in Table 4.1. Performed by Lionel Imbert.

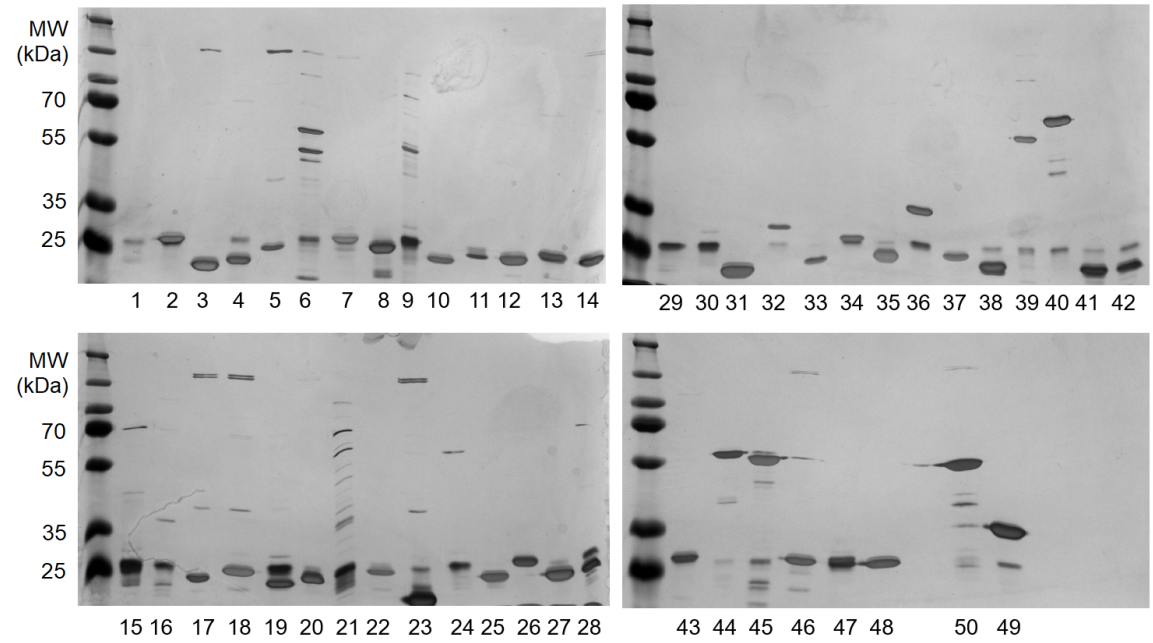


Figure 4.16: Silver stain analysis of 51 PAO1 purified RRs. The protein-bead binding step and first wash were done as in DAP-seq followed by protein elution, SDS-PAGE and silver staining analysis. RR number 51 is missing in this experiment, but its correct bead binding and purity wer later confirmed. RRs are numbered by ID as described in Table 4.1.

on with DAP-seq.

4.3.3 The TCSs transcriptional network of PAO1

Overall, the DAP-seq analysis of all 51 DNA-binding RRs in PAO1 allowed the detection of reproducible peaks for 43 RRs, 41 of which had peaks in promoter regions (as defined in method section 4.3.4), including 38 for which a conserved DNA-binding motif could be identified. Out of the failed experiments, two could be explained by poor protein production (PmrA (30) and AlgR (47)) and four are not explained and could be unstable or inactive in our conditions (PA0034 AtvR, PA4080 and AruR). Interestingly, out of these four, two (PA0034 and PA4080) are orphan predicted RRs (with no associated HK) and one (AtvR) is the only atypical RR lacking phosphorylatable aspartate in PAO1. However, two are explained by plasmid issue as stated above (NarL and IrlR) and two are explained by an insufficient number of sequenced reads (AauR and PA5364), the DAP-seq experiments will consequently be redone for these 4 RRs.

Reproducible DAP-seq peaks that pass our quality thresholds and are found

within promoter regions (see Method section 4.3.4) were assigned to the corresponding transcriptional unit (TUs) (single genes or entire operons) to identify direct TF-target genes interactions. Out of 41 RRs, there were between 1 and 428 binding sites identified per RR (Figure 4.17A). Among the targeted TUs, most were targeted by only one RR but numerous TUs were targeted by several different RRs, revealing major cross-regulatory interactions (Figure 4.17B). Overall, we identified 6,290 regulatory interactions involving 41 RRs and 2,174 target TUs comprising a total of 3,776 genes, yielding the first near-complete view of the complex *P. aeruginosa* TCS regulatory network (Figure 4.18). By comparing with a dataset of 40 known interactions reported by EMSAs found in the literature involving 17 of the 51 RRs, we found 31 in our DAP-seq results, showing an overall good recovery of known targets and thus confirming the ability of DAP-seq to identify RR regulatory targets. By searching for conserved motifs in DAP-seq peaks, we identified putative DNA-binding sites for 38 RRs across all three major RRs families (Figure 4.19). Notably, OmpR-like RRs are known to mostly bind to direct repeats and we identify direct repeats for several OmpR-like RRs, or in some cases at least the half site.

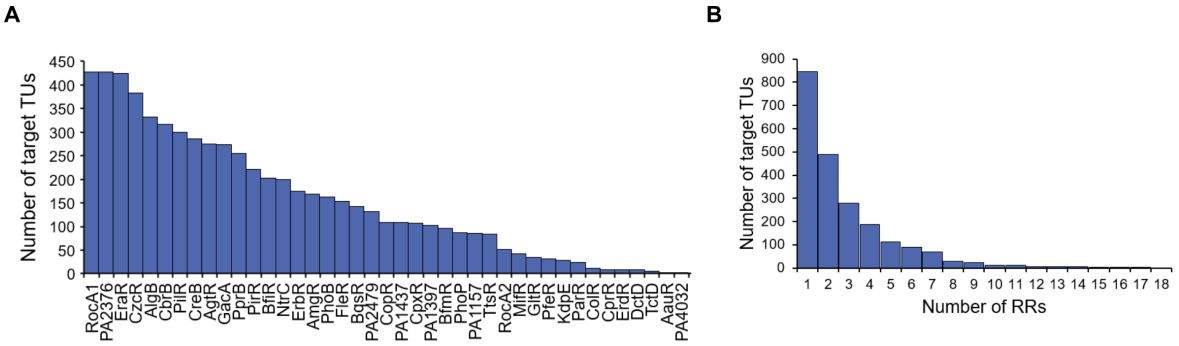


Figure 4.17: Numbers of regulatory interactions per RR or target TU. (**A**) Numbers of regulatory interactions as identified by DAP-seq are shown for each RR. (**B**) Distribution of target TUs based on the number of RRs binding to their promoters.

In this work, we retrieved numerous known direct regulatory interactions between RRs and target TUs. However, some were missed, which illustrates the limitation of our approach. Indeed, as an *in vitro* method which only involves one purified protein at a time, DAP-seq results do not reflect the impact of cellular content, DNA topology or other proteins, which could all affect TF binding and thus DAP-seq results. In the case of RRs, protein-protein interactions have been shown to impact RR activity (Pannen et al., 2016), even between two RRs as suggested for PhoB and TctD in *P. aeruginosa* (Bielecki et al., 2015), suggesting that DAP-seq

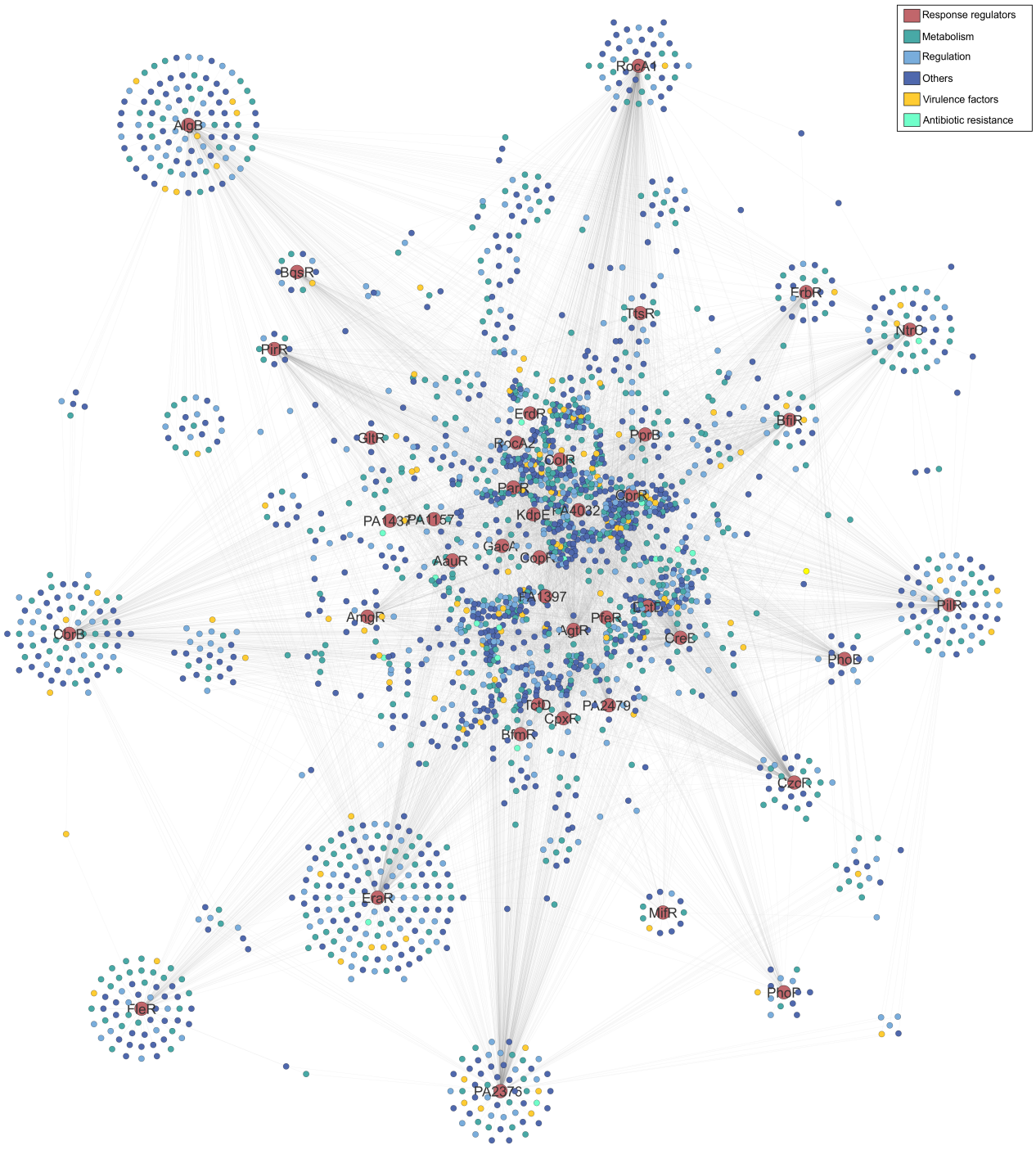


Figure 4.18: The near-complete *P. aeruginosa* TCS regulatory network. Regulatory interactions identified by DAP-seq for 41 RRs are shown as a network comprising 2,215 nodes and 6,290 edges. RRs are colored in red and target genes are colored depending on their COG predicted function. The "Metabolism" category comprises all COGs containing the words "metabolism" or "catabolism" and the "Energy production" COG. The "Regulation" category comprises the COGs Transcription, Translation, Post-translational modification and Signal transduction. The "Virulence Factor" and "Antibiotic resistance" categories comprise genes annotated as such in the Pseudomonas Database (Winsor et al., 2016). For operons, the annotation of the first gene is given. The network was visualized using Cytoscape 3.8 using the yFiles Organic layout.

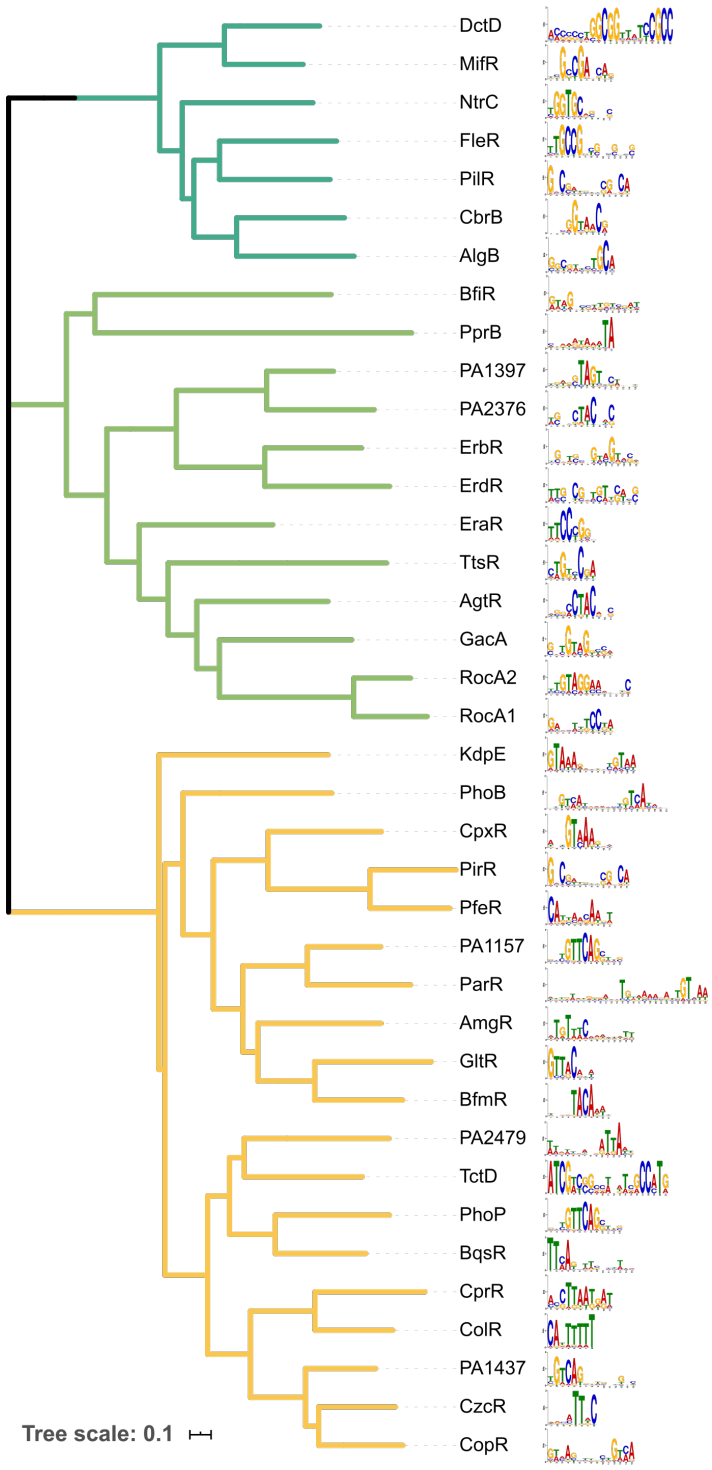


Figure 4.19: Comparison of RRs sequences and DNA-binding preferences. Maximum-Likelihood phylogenetic tree of 38 RRs PAO1 protein sequences done with 100 bootstraps. Putative DNA-binding motifs are shown as predicted using MEME-ChIP. Tree branches are colored in blue, green and yellow for NtrC-, NarL- and OmpR-like RRs, respectively.

alone cannot capture 100% of regulatory interactions for some RRs. Recently, a variation of the DAP-seq approach was developed to assess the genome-wide binding of heterodimeric TF complexes using sequential selection of heterodimers, which could be used to decipher such mechanisms (Lai et al., 2020). Moreover, the *in vitro* conditions used here, especially in regard to RR phosphorylation status, might impact DNA binding. Indeed, phosphorylation has been shown to affect RR binding differently between target binding sites (Gao et al., 2005), suggesting that for some RRs, the unphosphorylated form might not bind to all targets. The extend of this phenomenon in our results will be determined once we further characterized the phosphorylation status of all RRs in our conditions.

The results presented here represent the first delineation of the near-complete TCS regulatory network in PAO1. We studied all 51 DNA-binding RRs of PAO1; however, they are not all conserved across *P. aeruginosa* strains and thus do not fully reflect the regulatory networks found across the species. Indeed, differences in gene content can induce the rewiring of regulatory networks through new or modified regulatory interactions (Perez and Groisman, 2009a). Notably, the accessory genome, which differ between strains, represents a large number of potential strain-specific regulatory interactions as target genes might not be conserved across the species. While gene content can change, there also are differences in the sRNA content between strains, which could also bring regulatory differences. Additionally, when looking at three *P. aeruginosa* reference strains (PAO1, PA14 and IHMA87), representing the three major phylogenetic lineages, we see that the vast majority (48) of the 51 RRs studied here are conserved among strains but there are a few differences, with a total of 55 unique DNA-binding RRs (Figure 4.20). Indeed, some RRs are only present in one or two of the three strains. In most cases, the entire gene is simply present or missing, but in others, small mutations can cause inactivation, such as with PfeR which carries an 8-bp deletion in its gene in PA14 and is thus inactive only in that strain. Other such examples include the deletion

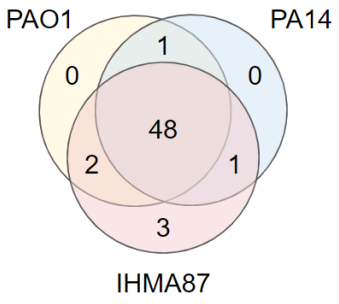


Figure 4.20: Conservation of DNA-binding RRs between three *P. aeruginosa* reference strains. The determination of RRs was done by crossing NCBI, Interpro, P2CS and P2RP databases prediction for each strain. Homology between the detected RRs was assessed using reciprocal best blast hit.

of the GacS HK in the *P. aeruginosa* CHA strain which led to the disturbance of the entire Gac/Rsm pathway (Sall et al., 2014), or the mutation of the LadS HK in PA14 (Mikkelsen et al., 2011a). Such evolutionary events then induce regulatory rewiring of the targets of the missing TF, creating regulatory diversity even between strains of the same species. For that reason, in addition to PAO1, we performed DAP-seq on all 51 RRs presented here in PA14 and IHMA87 as well. We are also currently cloning the genes of the 4 remaining RRs (3 IHMA87-specific and 1 not present in PAO1) to conduct DAP-seq with them too. This comparative analysis will give a first global view of TCS regulatory network plasticity across the *P. aeruginosa* species.

Overall, we identified numerous known and new target genes across the entire family of *P. aeruginosa* DNA-binding RRs. I present here the first results obtained with PAO1 and a first crude global analysis. For this project, more data will be obtained, as the PA14 and IHMA87 results are pending and I will perform few more DAP-seq experiments for the ones that did not work for technical reasons in all three genomes. Also, the results presented here come from a first brief analysis of the overall DAP-seq results since I obtained these results very recently, but more work will be done to refine the analysis and go into more detail in the investigation of this major regulatory network. Interestingly, there were a lot of genes that are targeted by more than one TF, revealing the large complexity of this regulatory network. Additionally, the comparative analysis between all three genomes will yield very interesting insights in terms of regulatory conservation between closely related strains.

4.3.4 Materials & Methods

Plasmid cloning. The 51 RR genes were PCR amplified from purified PAO1 genomic DNA and cloned into *Not*I-*Sal*I-cut pIVEX2.4d linearized vectors by SLIC as previously described (Trouillon et al., 2020a). All plasmids were transformed into competent TOP10 *E. coli* and verified by sequencing (Eurofins).

Cell-free protein expression. Cell-free protein expression was performed as previously described for 3 hours at 37°C at the cell-free platform of IBS - Grenoble (Bertrand et al., 2020).

DAP-seq. DAP-seq was performed as previously described (Trouillon et al., 2020a) with the following modifications: all steps were performed in 96-well PCR plates. 100 µl of cell-free protein extracts were diluted in Binding Buffer (sterile PBS supplemented with 10mM $MgCl_2$, 0.01% Tween 20 and 50µM acetyl phosphate) containing 10 µl of pre-washed magnetic cobalt beads for a final volume of 200 µl and incubated for 40 min at room temperature. Washing were done with 200 µl of Binding Buffer. Negative control experiments were done using a pIVEX vector expressing an untagged GFP protein. All DAP-seq experiments were performed in duplicates for a total of 104 samples.

Sequencing & Data analysis. Sequencing was performed by the high throughput sequencing core facility of I2BC (Centre de Recherche de Gif – http://www.i2bc.paris-saclay.fr) using an Illumina NextSeq500 instrument. An average of 2.6 million single-end 75-bp reads per sample were generated with >95% of reads uniquely aligning to the PAO1 genome using Bowtie2 (Langmead and Salzberg, 2012). Peak calling was done using MACS2 (Zhang et al., 2008) for each duplicate against the two negative control samples. Peaks found in the two replicates were selected using the Intersect tool from BEDTools (Quinlan and Hall, 2010) with a minimum overlap of 50% on each compared peaks. Peaks were then filtered by Irreproducible Discovery Rate (IDR) with a IDR threshold of 0.005 using IDR Galaxy version 2.0.3 on the European Galaxy server (Li et al., 2011; Afgan et al., 2018). Promoter regions were defined as the 400-bp upstream region of each transcription units, as defined in the "Operon" annotation of the Pseudomonas Database (Winsor et al., 2016). Peaks having a summit position within a promoter region were identified using the Intersect tool from BEDTools (Quinlan and Hall, 2010). Peaks in promoter regions were used for generation of the network and

inference of regulatory interactions. MEME-ChIP (Machanick and Bailey, 2011) was used with default settings to find conserved motifs on all peaks selected by IDR.

4.4 Promoter diversity

Julian Trouillon

Work in progress

Since I have worked on ErfA and found the differences in *exlBA* promoter between *Pseudomonas* species, I have always wanted to find out how common it is for genes to have evolved different promoters between closely related species. While the covid-19 lockdown made me loose time on the experimental side, it also gave me some computer time to try to answer that question. To that aim, I firstly defined the pan-genome of all complete *Pseudomonas* genomes (n = 503) and thus identified ortholog groups for all genes. By grouping genes in operons and retrieving both coding and promoter sequences, I then computed 540 million phylogenetic distances to cluster genes depending on their local promoter diversity. I found that hundreds of genes exhibit promoter diversity levels comparable to that of *exlBA*. To correlate that to actual promoter regulatory motifs diversity, I generated over 170,000 DNA motifs across all promoters, clustered them using a custom motif conservation metric and found many genes with high promoter DNA motif diversity. The comparison of this result with experimental data on known TF regulons and binding motifs revealed that most regulatory interactions are not conserved even if both TF and target are. Finally, this is an ongoing work for which I still need to verify and strengthen part of the methods used as phylogeny and such analyses are initially not in my domain of expertise.
In this work, I performed all analyses.

Promoter sequence diversification as a major mean for bacterial regulatory networks evolution

Abstract

Transcription factors (TFs) participate in bacterial response and adaptation to environmental cues. While most studies focus on analyzing gene content differences to explain bacterial phenotypic and niche specificities, relatively little is known on the prevalence of regulatory network rewiring through promoter evolution in this mechanism. By analyzing 2.6 million sequences across 503 complete genomes, we determined the *Pseudomonas* pan-genus promoter pool and quantified promoter diversification for each ortholog group. The search and conservation analysis of both *de novo*-generated and known regulatory motifs revealed that the vast majority of DNA motifs are not globally conserved across orthologous promoters. In most cases, direct regulatory interactions were not maintained in all strains possessing both TF and target genes. These findings give a first global insight at promoter diversification and illustrate its major role in transcriptional regulation evolution between closely related bacterial species, revealing major flaws in our way of apprehending transcriptional regulation and phenotypic evolution.

Main

To adapt to new environments, cells require sensing and response mechanisms, many of which rely on transcription factors (TFs) (Mejía-Almonte *et al.*, 2020). TFs input signals generally encompass environmental cues such as metabolites, ions or temperature which then induce output regulatory responses through TFs DNA-binding activity (Chubukov *et al.*, 2014). It is well established that the gain and loss of transcription factor DNA binding sites (TFBS), as opposed to TF genes themselves, is a major mechanism contributing to morphological and developmental evolution in higher eukaryotes (Wittkopp & Kalay, 2012). This is known because most

investigations have focused on *cis*-acting promoter sequences due to the highly similar gene content between related eukaryotic species. On the other hand, led by the much more diverse gene content found across bacterial species, most studies on evolution of prokaryotic regulatory networks have focused on genes and TFs diversity. Additionally, the investigation of transcription regulation is often restricted to a small number of bacterial species, and among each of them, very few reference strains. This often leads to the false assumption that regulatory interactions are conserved in closely related species or strains as long as both TF and target genes are. Although these biases have been pointed out as misleading in the past (Perez & Groisman, 2009) and several recent studies have shown examples of important phenotypic implications of TFBS diversity for single TFs (Trouillon *et al*., 2020a; Connolly *et al*., 2020; O'Boyle *et al*., 2020), there is still a major lack of knowledge on the importance of TFBS evolution in bacterial phenotypic and niche specificities.

In order to provide a first global view on the importance and frequency of promoter diversification, we studied promoter conservation across the *Pseudomonas* genus, one of the most diverse bacterial genus which contains more than 200 species (Gomila *et al*., 2015). To that aim, we analyzed all complete *Pseudomonas* genomes (*n* = 503) from the Pseudomonas Database (Supplementary Figure 1 and Supplementary Table 1) (Winsor *et al*., 2016) and determined all gene ortholog groups. After integrating operons structures, we then retrieved all promoter and coding sequences, performed 6894 multiple alignments and generated >540 million pairwise phylogenetic distance values to assess both promoter and gene diversity in each ortholog group (Pipeline described in Methods section and Supplementary Figure 2). Using *k*-means clustering analysis, we then identified clusters of genes that exhibit high or low promoter diversity (Figure 1A and Supplementary Table 2). Although the Pseudomonas Database is biased towards *P. aeruginosa* genomes (*n* = 197 complete genomes), there was no correlation between either gene or promoter diversity scores and the size of ortholog groups (Supplementary Figure 3). Overall, gene and promoter diversity correlated relatively well, with a Pearson coefficient of 0.76, and promoter/gene diversity ratios were similar over all clusters with an average of 1.98, showing a global higher conservation of coding over promoter sequences (Figure 1B). Clusters 1 to 3 show

relatively high sequence conservation while clusters 4 to 6 are more diverse. Although higher gene diversity, such as seen in cluster 6, could indicate functional changes between orthologs and thus explain the need for different regulations, there still is a large number of genes with relatively high conservation that exhibit very diverse promoters, as seen in cluster 5. This cluster (*n* = 614 genes and operons), and cluster 4 to a lesser extent, thus represent a large niche of potential regulatory diversity among orthologs with conserved functions.

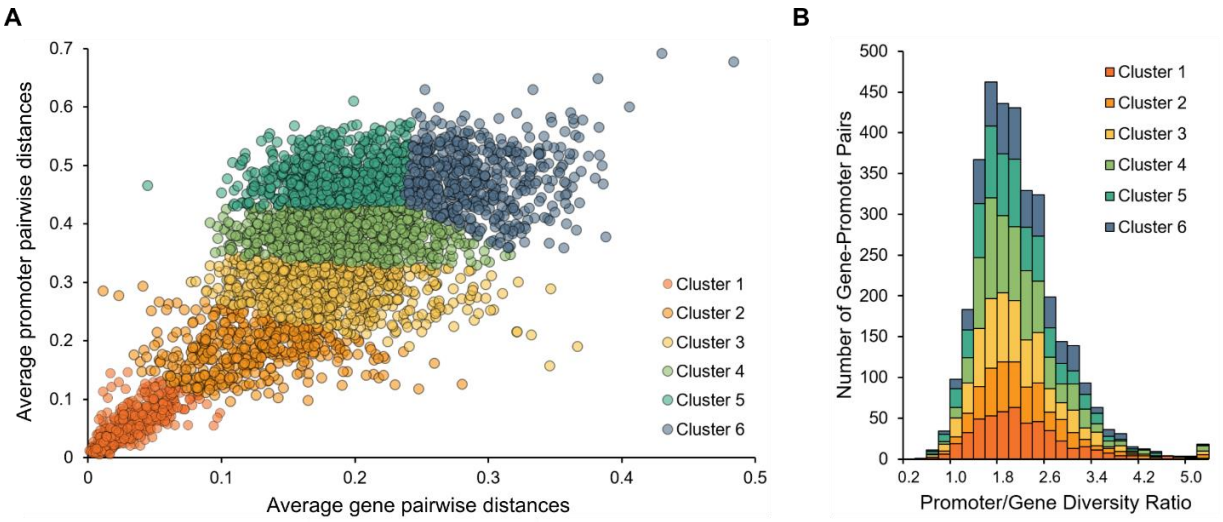


Figure 1: Clustering of ortholog groups by gene and promoter diversity. (**A**) Average pairwise distances of gene and promoter sequences for each ortholog group (*n* = 3,447). (**B**) Sorting of ortholog groups by ratio of promoter over gene average pairwise distances. Groups are colored by clusters, as determined by *k*-means clustering.

As a mean to correlate promoter sequence diversity to lower DNA motifs conservation and thus infer regulatory differences, we generated and measured the conservation of all *de novo* detected DNA motifs in each promoter ortholog groups (~ 170,000 motifs) (Pipeline described in Methods section and Supplementary Figure 4). To assess differences, we created a metric to measure motif conservation called Top Motifs average Conservation (TMC) that represents the average conservation of the most significant DNA motifs detected in each promoter ortholog groups (see Methods). Overall, TMC values were found to correlate negatively with promoter pairwise distances and to display the same trend of promoter diversity through the previously described clusters (Figure 2A and Supplementary Figure 5). Here again, in clusters 4 to 6, there were hundreds of genes with TMC values below 70 % (and going down as low as 24 %) which

means that even the most conserved DNA motifs in these promoters show relatively low conservation, strongly suggesting regulatory differences among strains or species for these genes.

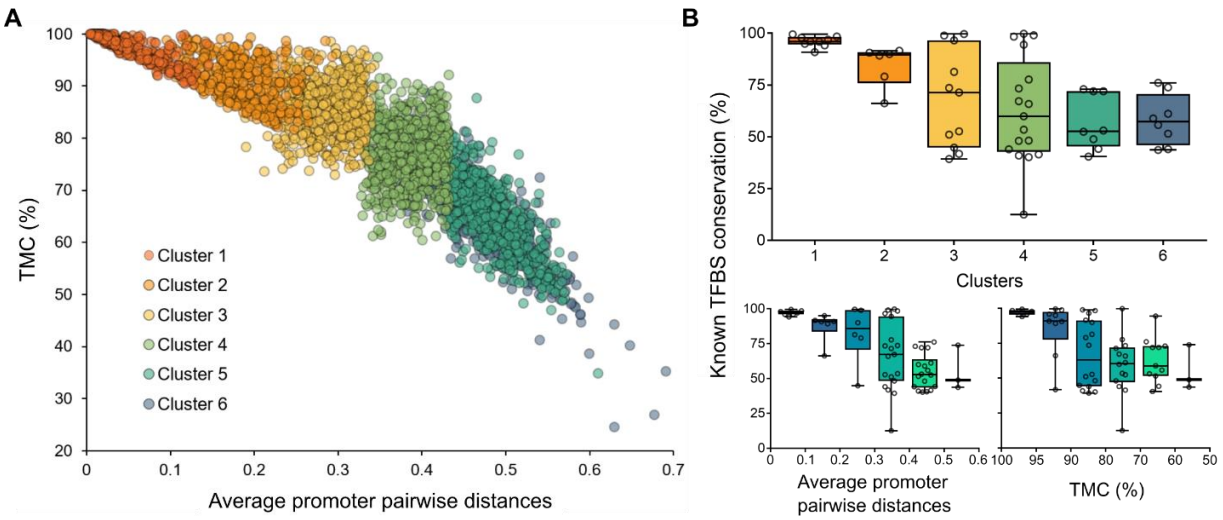


Figure 2: Conservation of *de novo* detected motifs and known TFBS across promoter ortholog groups. (**A**) Relation between average promoter sequence diversity and conservation of *de novo* detected DNA motifs (shown as TMC) across promoter ortholog groups. (**B**) Conservation analysis of known TFBS involved in regulatory interactions experimentally confirmed in *P. aeruginosa* PAO1 (as described in the Methods section) across orthologs, compared to the previously determined clusters (upper panel), promoter diversity (lower left) and TMC values (lower right). Groups are colored by clusters when relevant.

To confirm this hypothesis, we leveraged high quality experimental data combining genome-wide binding and expression measurements in *P. aeruginosa* from several studies to assess the conservation of regulatory interactions involving different TFs (see Methods). These known TFBSs (n = 50) were found to exhibit diverse levels of conservation across ortholog groups, corroborating with previous metrics and clusters (Figure 2B). Strikingly, regulatory interactions - through known TFBSs - with target genes from clusters 4 to 6 were found to be conserved in only ~50-60 % of strains possessing both TF and target genes on average, revealing major regulatory diversification for these genes. On the other hand, genes from cluster 1 show TFBSs conserved in nearly all strains and species. Interestingly, more specialized or local TFs seem to have more conserved regulatory interactions than global TFs with broader regulatory impacts (Supplementary Figure 6). Specialized TFs possess relatively small regulons that usually encompass genes involved in the same function and often respond to signals corresponding to

that function, creating strong evolutionary bonds between TFs and target genes. Global regulators tend to respond to broader signals and concomitantly regulate various functions that can be specific to each bacterium's physiology or environment, explaining the need for different regulatory interactions.

Regulatory rewiring through modification of gene content (TF and/or target genes) is a widespread mechanism in bacteria, which resulted in the overlooking of the implication of TFBS diversification in this process (Perez & Groisman, 2009). Here, in an attempt to strengthen the idea that conservation of target and TF genes does not necessarily means conservation of their regulatory interactions in closely related strains or species, we provided a first global view of TFBS conservation and showed that DNA motifs found in promoters are widely diverse. Indeed, very few TFBSs were found conserved in 100 % of strains possessing both TF and target, suggesting that promoter diversification is an extremely broad and common mechanism. DNA motifs were found to diverge at different rates depending on genes, allowing the delineation of different groups of genes with low, medium or high promoter diversity. Overall, these results bring a novel view on the frequency of this mechanism of regulation evolution and reveal the need for future studies to assess regulatory conservation when studying TFs functions in single strains.

Methods

Sequences and genomes annotations retrieving. All sequences and annotations were retrieved from the Pseudomonas Genome database v18.1 (Winsor *et al*., 2016). All complete *Pseudomonas* genomes (*n* = 527) were filtered for duplicates or deprecated genomes which resulted in a final dataset comprising 503 genomes (Supplementary Table 1). For each genome, the complete DNA sequence, all protein sequences and all gene annotations were directly downloaded from the website.

Ortholog groups determination and genomes phylogenetic analysis. Ortholog identification was performed by Reciprocal Best Blast Hit (RBBH) analysis (Cock et al., 2015) on the European Galaxy

server (Jalili et al., 2020) using all protein sequences from *P. aeruginosa* PA14 against all protein sequences from the other 502 genomes. To maximize chances of function conservation inside ortholog groups, orthologs were defined as RBBH hits that share >90 % alignment coverage and >50 % sequence identity, both on protein sequences. For the global phylogenetic analysis of *Pseudomonas* genomes (Supplementary Figure 1), sequences from all core genes conserved in all 503 genomes (n = 66) were concatenated for each genome and a multiple alignment was performed on the concatenated sequences with MAFFT Galaxy version 7.221.3 (Katoh & Standley, 2013) using default settings. The resulting alignment was used to build a Maximum-Likelihood phylogenetic tree using MEGA X with 100 bootstraps which was visualized and annotated using iTOL v5 (Kumar *et al*., 2018; Letunic & Bork, 2019).

Genes and promoters sequence diversity measurements and clustering. Gene coordinates were obtained from genome annotations and used to retrieve coding DNA sequences. For operons, the first gene was used for diversity comparison. Operons were set as experimentally determined in *P. aeruginosa* PA14 (Wurtzel *et al*., 2012). For promoter region sequences, the 200 bp upstream of genes start codon were used, as regions of this length were shown to encompass >95 % transcription start sites in *P. aeruginosa* (Wurtzel *et al*., 2012). In each ortholog group, two multiple alignments were performed (for gene and promoter sequences, respectively) with MAFFT Galaxy version 7.221.3 (Katoh & Standley, 2013) using default settings. Pairwise phylogenetic distances were then computed from the resulting multiple alignments using dist.seqs from Mothur tools (Schloss *et al*., 2009) on the European Galaxy server. Averages were then computed from Mothur output files using Python 3.7. *k*-means clustering was performed using BioVinci v1.1.5 (BioTuring Inc.) and *k* was determined using the elbow method.

***de novo* DNA motifs search and TMC calculation.** *de novo* DNA motif search was performed on each promoter ortholog group using DREME (Bailey, 2011) with default settings on the European Galaxy server. TMC determination was done by averaging the percentage of strains possessing each of the 10 most significantly detected motifs for each ortholog group.

Known TFBS conservation analysis. Experimental data on *P. aeruginosa* PAO1 were obtained from several studies (Jones *et al*., 2014; Huang *et al*., 2019; Trouillon *et al*., 2020b). TFs were selected based on several criteria: (i) both genome-wide binding (ChIP-seq or DAP-seq) and expression (RNA-seq) data were available and a high-confidence regulon was determined and at least partly confirmed with targeted approaches, (ii) a DNA-binding motif with high information content was inferred from this regulon, (iii) by reanalyzing the binding data ourselves, we could infer a similar binding motif as the one proposed in the original study, and (iv) the TF is conserved in >98 % of the studied 503 genomes. For each TF, the binding site was generated using reported peaks with MEME-ChIP (Machanick *et al*., 2011) and the presence of the binding site was then searched in the ortholog groups of each confirmed target gene using FIMO (Grant *et al*., 2011). Ortholog groups for which the TFBS was not detected in *P. aeruginosa* PAO1 strain (where the interaction was experimentally confirmed) were excluded from the analysis.

Acknowledgments

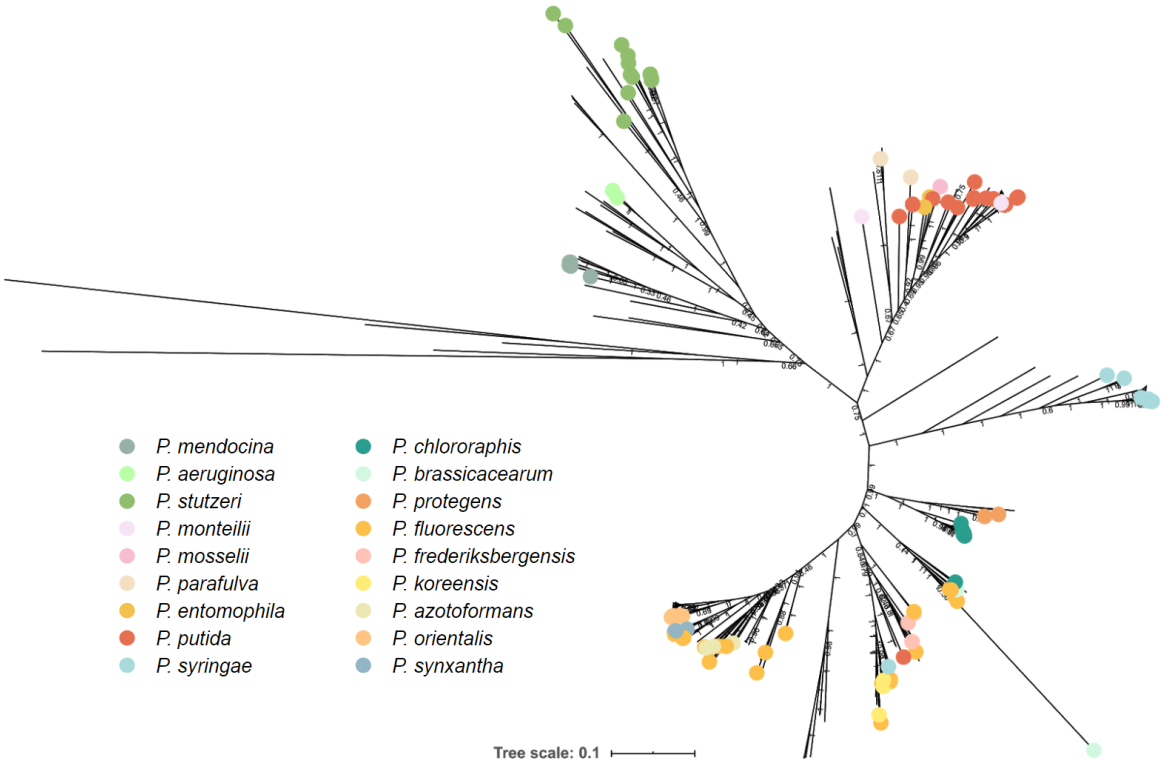
I am grateful to Björn Grüning, who helped solving issues with the submission of large numbers of jobs on the European Galaxy server for this work.

References

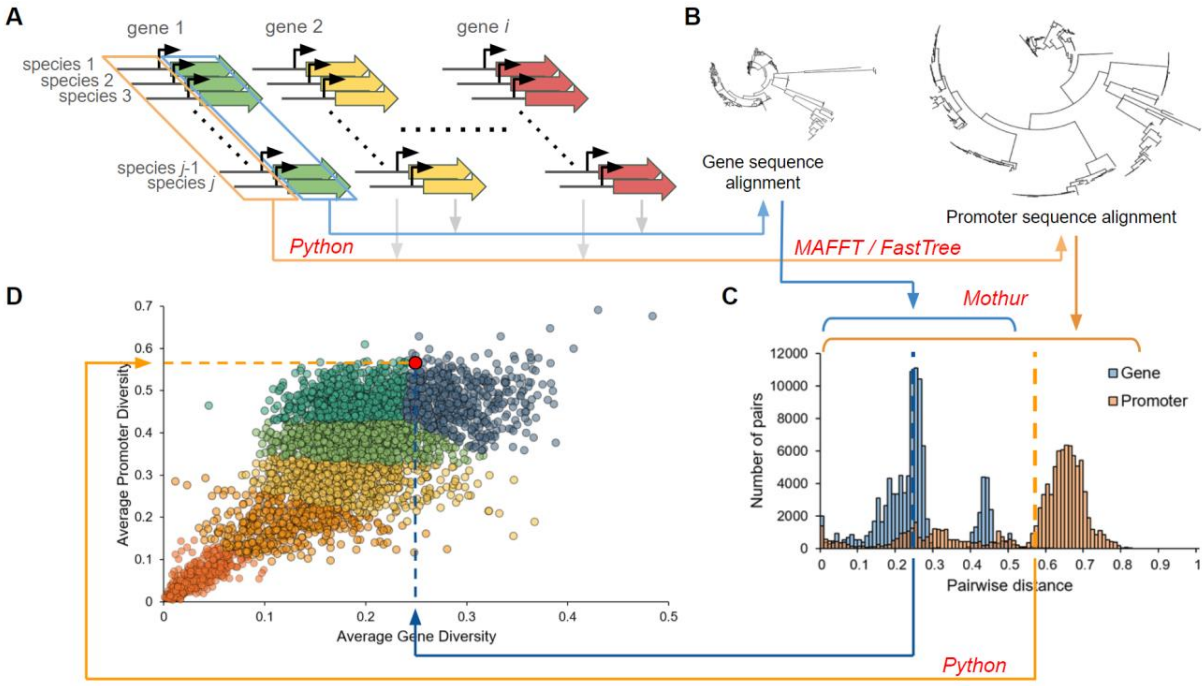
- Bailey, T. L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics*, 27(12), 1653-1659.
- Chubukov, V., Gerosa, L., Kochanowski, K., & Sauer, U. (2014). Coordination of microbial metabolism. *Nature Reviews Microbiology*, 12(5), 327-340.
- Cock, P. J., Chilton, J. M., Grüning, B., Johnson, J. E., & Soranzo, N. (2015). NCBI BLAST+ integrated into Galaxy. *Gigascience*, 4(1), s13742-015.
- Connolly, J. P., O'Boyle, N., & Roe, A. J. (2020). Widespread strain-specific distinctions in chromosomal binding dynamics of a highly conserved Escherichia coli transcription factor. *Mbio*, 11(3).
- Gomila, M., Peña, A., Mulet, M., Lalucat, J., & García-Valdés, E. (2015). Phylogenomics and systematics in Pseudomonas. *Frontiers in microbiology*, 6, 214.
- Grant, C. E., Bailey, T. L., & Noble, W. S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics*, 27(7), 1017-1018.

- Huang, H., Shao, X., Xie, Y., Wang, T., Zhang, Y., Wang, X., & Deng, X. (2019). An integrated genomic regulatory network of virulence-related transcriptional factors in Pseudomonas aeruginosa. *Nature communications*, 10(1), 1-13.
- Jalili, V., Afgan, E., Gu, Q., Clements, D., Blankenberg, D., Goecks, J., ... & Nekrutenko, A. (2020). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. *Nucleic Acids Research*.
- Jones, C. J., Newsom, D., Kelly, B., Irie, Y., Jennings, L. K., Xu, B., ... & Wozniak, D. J. (2014). ChIP-Seq and RNA-Seq reveal an AmrZ-mediated mechanism for cyclic di-GMP synthesis and biofilm development by Pseudomonas aeruginosa. *PLoS Pathogens*, 10(3), e1003984.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4), 772-780.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution*, 35(6), 1547-1549.
- Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic acids research*, 47(W1), W256-W259.
- Machanick, P., & Bailey, T. L. (2011). MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics*, 27(12), 1696-1697.
- Mejía-Almonte, C., Busby, S. J., Wade, J. T., van Helden, J., Arkin, A. P., Stormo, G. D., ... & Collado-Vides, J. (2020). Redefining fundamental concepts of transcription initiation in bacteria. *Nature Reviews Genetics*, 1-16.
- O'Boyle, N., Turner, N. C., Roe, A. J., & Connolly, J. P. (2020). Plastic Circuits: Regulatory Flexibility in Fine Tuning Pathogen Success. *Trends in Microbiology*, 28(5), 360-371.
- Perez, J. C., & Groisman, E. A. (2009). Evolution of transcriptional regulatory circuits in bacteria. *Cell*, 138(2), 233-244.
- Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., ... & Sahl, J. W. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology*, 75(23), 7537-7541.
- Trouillon, J., Sentausa, E., Ragno, M., Robert-Genthon, M., Lory, S., Attrée, I., & Elsen, S. (2020a). Species-specific recruitment of transcription factors dictates toxin expression. *Nucleic acids research*, 48(5), 2388-2400.
- Trouillon, J., Ragno, M., Simon, V., Attrée, I., & Elsen, S. (2020b). Transcription inhibitors with XRE DNA-binding and cupin signal-sensing domains drive metabolic diversification in Pseudomonas. *bioRxiv*.

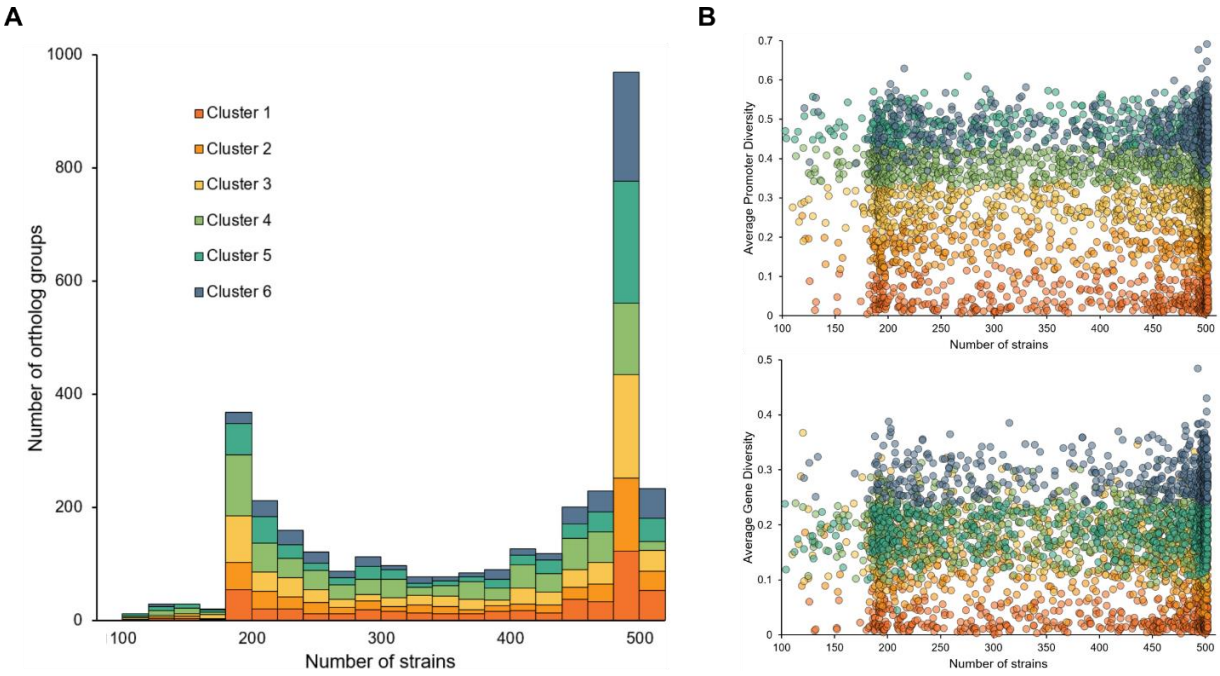
- Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A., & Brinkman, F. S. (2016). Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. *Nucleic acids research*, 44(D1), D646-D653.
- Wittkopp, P. J., & Kalay, G. (2012). Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics*, 13(1), 59-69.
- Wurtzel, O., Yoder-Himes, D. R., Han, K., Dandekar, A. A., Edelheit, S., Greenberg, E. P., ... & Lory, S. (2012). The single-nucleotide resolution transcriptome of Pseudomonas aeruginosa grown in body temperature. *Plos pathogens*, 8(9), e1002945.



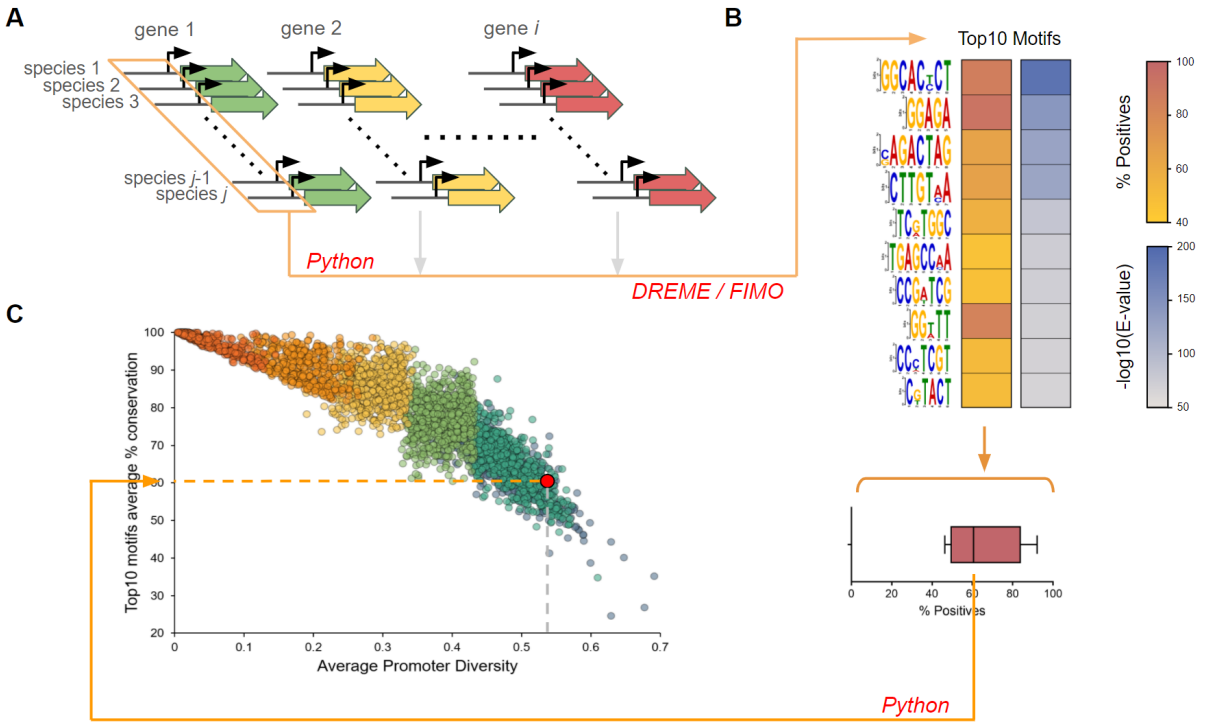
Supplementary Figure 1: Maximum-Likelihood phylogenetic tree of 503 *Pseudomonas* complete genomes. The tree was generated from the multiple alignment of the concatenated sequences of 66 core genes for each strain with 100 bootstraps.



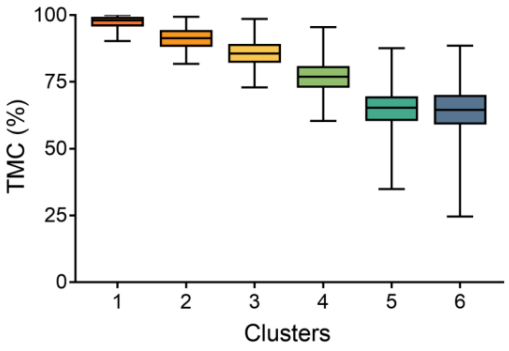
Supplementary Figure 2: Analysis pipeline for measuring sequence diversity. (**A**) Both coding and promoter sequences were retrieved in each ortholog group using Python 3.7. (**B**) Sequences were aligned using MAFFT resulting in 6,894 multiple alignment. FastTree was used for visualization when necessary. (**C**) A phylogenetic pairwise distance matrix was generated for each alignment using dist.seqs from Mothur. (**D**) Distances values were averaged for each ortholog group and used to quantify sequence diversity.



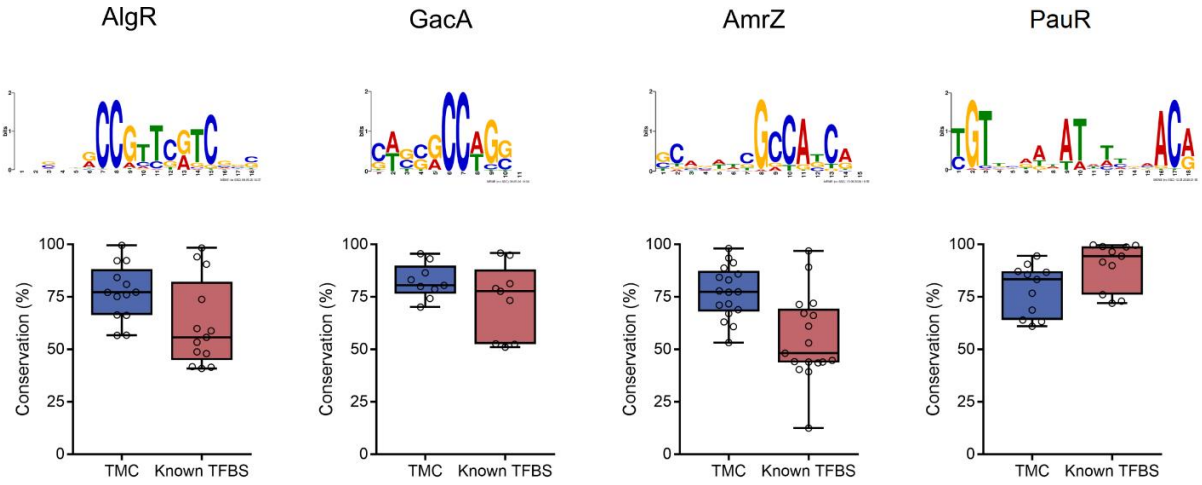
Supplementary Figure 3: Distribution of ortholog groups by number of strains. (**A**) Clusters and number of strains found in ortholog groups. (**B**) Distribution of ortholog groups by gene and promoter diversity.



Supplementary Figure 4: Analysis pipeline for TMC calculation. (**A**) Promoter sequences were retrieved in each ortholog group using Python 3.7. (**B**) *de novo* motif search was performed in each ortholog group using DREME (or FIMO for known TFBSs). The percentage of strains containing each of the 10 motifs with highest *E*-values were computed and averaged, resulting in a TMC value that was then used in further analyses (**C**).



Supplementary Figure 5 : TMC values sorted by clusters. TMC values obtained for all promoter ortholog groups are shown for each cluser, colored as previously.



Supplementary Figure 6: Known TFBSs conservation analysis. For each TF: name (top), motif found when reanalyzing published data (middle), conservation of *de novo*-generated (TMC) and known motifs in promoter ortholog groups of target genes (bottom). A total of 50 known TFBSs are shown. AlgR and GacA regulons were obtained from Huang *et al*., 2019, AmrZ regulon from Jones *et al*., 2014 and PauR regulon from Trouillon *et al*., 2020b. AlgR, GacA and AmrZ are considered broad global regulator while PauR is a specialized TF.

Chapter 5

Post-transcriptional regulation

Contents

5.1 The Hfq regulatory RBP

5.1.1 The multiple roles of Hfq across *P. aeruginosa* lineages

Julian Trouillon✉, Ina Attrée & Stephen Lory✉

Manuscript in preparation

While I mostly worked on DNA-binding proteins during my PhD, I also wanted to work on RNA-binding proteins as they represent an important layer of regulation. To that aim, I went to work in the lab of Stephen Lory in Boston, USA for 3 months during the summer of 2019. During that time, I worked on the major regulatory RNA-binding protein Hfq and characterized its sRNAs interactomes by RIP-seq (RNA-co-Immunoprecipitation and sequencing). I used RIP-seq on Hfq at three different time points during growth in PAO1, PA14 and IHMA87 to compare the roles of Hfq in these three strains. This two-variable comparative analysis allowed the identification of several important targets of Hfq and revealed numerous examples of strain-specific regulations notably involving global TFs, the T3SS or CRISPR systems.
In this work, I performed all experiments and analyses.

The core and accessory Hfq interactomes across *P. aeruginosa* lineages

Julian Trouillon[1*], Kook Han[2], Ina Attrée[1] & Stephen Lory[2*]

[1]Université Grenoble Alpes, CNRS ERL5261, CEA BIG-BCI, INSERM UMR1036, Grenoble, France
[2]Department of Microbiology, Harvard Medical School, Boston, Massachusetts, USA
* stephen_lory@hms.harvard.edu
* julian.trouillon@cea.fr

Abstract

Post-transcriptional regulation is a central mechanism for bacterial response to environmental changes. The major RNA-binding protein Hfq mediates interactions between regulatory small noncoding RNAs (sRNAs) and their target mRNAs in order to tune their translation. Here, in the light of the important variability of sRNA content across *Pseudomonas aeruginosa* genomes, we used RIP-seq to determine the Hfq interactome in representative strains of the three major phylogenetic lineages. The two-variable comparative approach identified Hfq core targets and new sRNAs across growth phases as well as strain-specific regulations. Key metabolic pathways and processes involved in transition from fast growth to starvation were found as conserved Hfq targets. Accessory Hfq interactome notably included most of mRNAs of the Type III Secretion System (T3SS) components as well as cognate secreted toxins in the two T3SS$^+$ lineages, while a cluster of CRISPR guide RNAs was enriched exclusively in one. As most differences observed between strains were not due to different RNA abundances, we propose that novel context-dependent regulatory mechanisms dictate the interactions of RNAs and Hfq across strains. Overall, our work reveals an important intra-species variability in post-transcriptional regulatory networks and highlights the necessity of such approach to apprehend the induced phenotypic diversity.

Introduction

Post-transcriptional regulation is a key mechanism allowing bacteria to control protein synthesis in response to rapid environmental changes. The stability or accessibility of a messenger RNA (mRNA) is often modified through interaction of a RNA-binding protein (RBP) alone or in complex with a small non-coding RNA (sRNA) (Holmqvist & Vogel, 2018). The recent advances in deep sequencing allowed the discovery of thousands of sRNAs suggesting that bacterial post-transcriptional regulatory networks are far more complex than previously thought (Wagner & Romby, 2015). Widely distributed across the bacterial kingdom, Hfq is one of the major regulatory RBPs that has a prime importance for the regulation of key biological processes (Sobrero & Valverde, 2012). Hfq drives post-transcriptional regulation by facilitating the interactions between regulatory sRNAs and their mRNA targets (Vogel & Luisi, 2011). In the major human pathogen *Pseudomonas aeruginosa*, Hfq notably regulates the synthesis of virulence factors (Sorger-Domenigg *et al*., 2007), carbon metabolism (Sonnleitner *et al*., 2003), antibiotic resistance (Zhang *et al*., 2017), as well as the switch between planktonic and biofilm lifestyle (Chihara *et al*., 2019).

While most of the current knowledge on bacterial regulatory networks usually derives from work on one reference strain and is assumed to apply to the entire species, we are on the other hand starting to grasp the vast intra-species genetic diversity that can be found in some bacteria. *Pseudomonas aeruginosa* embodies that fact with now more than 5,000 sequenced genomes spanning various intra-species lineages with specific characteristics (Freschi *et al*., 2019). Numerous studies have focused on the genomic characterization and identification of these different lineages and found various important differences in most central cellular processes (Selezska *et al*., 201; Jeukins *et al*., 2017; Freschi *et al*., 2018; Freschi *et al*., 2019). Indeed, *P. aeruginosa* core genes represents only 1% of the entire species pan-genome which encompasses and explains large differences in antimicrobial and virulence related phenotypes (Freschi *et al*., 2019). For example, strains belonging to three different *P. aeruginosa* lineages

48 harbor mutually exclusive virulence-related secretion systems (Reboud *et al*., 2017) and different
49 antibiotic resistance gene content (Roy *et al*., 2010). While the discovery of these differences calls for
50 further molecular investigation, very few studies have focused on identifying potential intra-species
51 regulatory network specificities.

52 Here, we established the growth-dependent as well as strain-dependent RNA interactomes of Hfq
53 in the three *P. aeruginosa* lineages by RNA co-Immunoprecipitation and sequencing (RIP-seq). We
54 described the core and accessory interactomes of Hfq and highlighted important targeted cellular
55 mechanisms. Our approach identified two new major strain-dependent Hfq targets; most of the operons
56 of the Type III Secretion System (T3SS) apparatus in two lineages, and a cluster of CRISPR RNAs in the
57 third. The untargeted analysis of the obtained datasets resulted in the discovery of potential new sRNAs
58 across lineages. Altogether, our comparative study reveals a wide functional intra-species diversity of Hfq
59 targets.

60

61 # Results

62 **Comparison of RBPs and sRNAs across *P. aeruginosa* lineages**

63 The phylogenetic analysis of core genes in 192 complete *P. aeruginosa* genomes revealed three distinct
64 major lineages (Figure 1A) in accordance with previous reports (Freschi *et al*., 2019). The two most
65 populated groups are represented by reference laboratory strains PAO1 and PA14 (Stover *et al*., 2020; He
66 *et al*., 2004). The third group, recently delineated into two subgroups (Freschi *et al*., 2019), includes its
67 first fully sequenced member, PA7 (Roy *et al*., 2010), and IHMA879472 (IHMA87; Kos *et al*., 2015), a strain
68 characterized for its Exolysin A-dependent hypervirulence (Reboud *et al*., 2016). At the whole genome

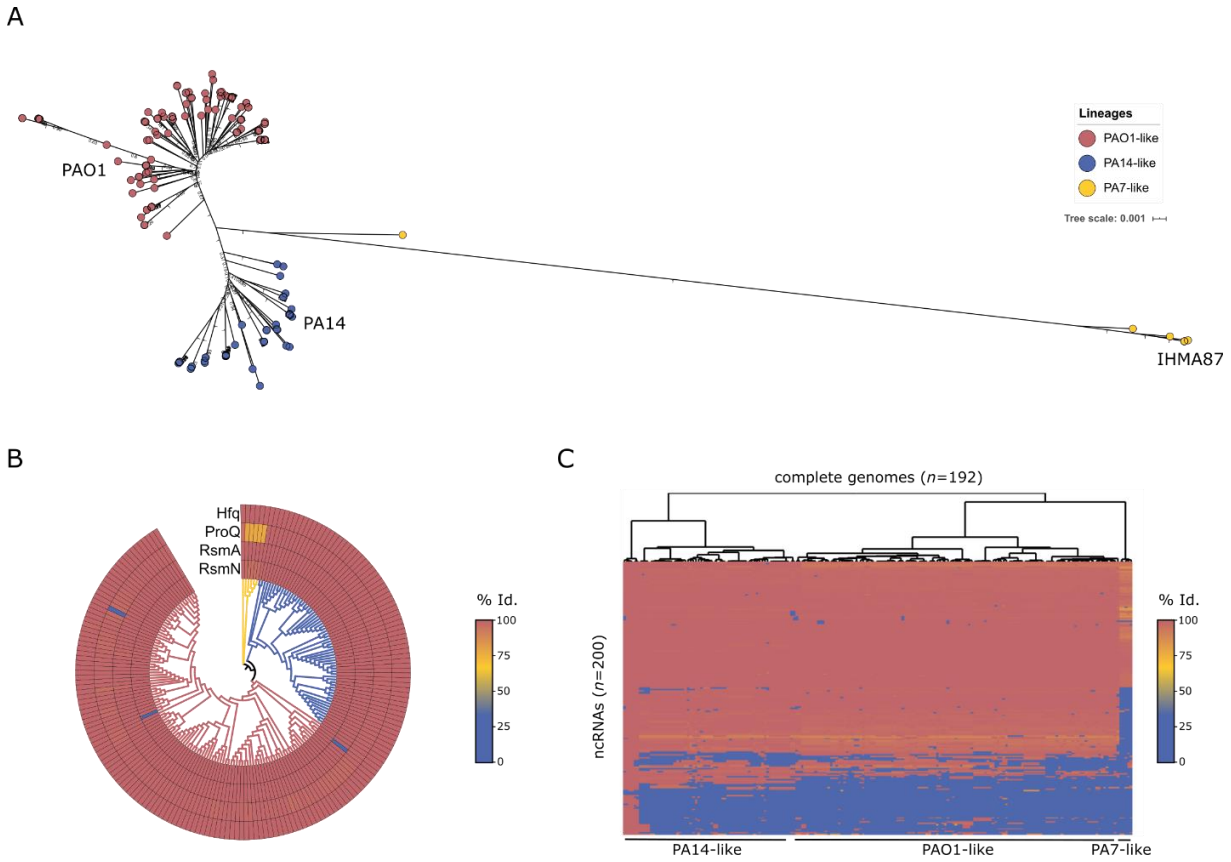


Figure 1: Conservation of *P. aeruginosa* post-transcriptional regulatory networks. (**A**) Maximum-Likelihood phylogenetic tree of 192 *P. aeruginosa* complete genomes. The tree was generated from the multiple alignment of the concatenated sequences of 66 core genes for each strain with 100 bootstraps. PAO1-like strains are depicted in red, PA14-like in blue and IHMA87-like in yellow. (**B**) Heatmap showing % protein sequence identity with *P. aeruginosa* PA14 sequences for 4 major regulatory RBPs. Strains are arranged on the rounded phylogenetic tree from (A) with leafs colors corresponding to the three different strain lineages. (**C**) Hierarchical clustering analysis of the conservation of 200 sRNAs in the 192 *P. aeruginosa* genomes. The clustering was performed on both rows and columns with generation of a dendogram based on Manhattan distances for strains. The heatmap shows % DNA sequence identity with PA14 sRNA sequences.

level, the 3rd group of strains share a significantly lower average nucleotide identity (ANI) (~93-98%) with the two other groups and are devoid of T3SS. While the three lineages differ greatly in their accessory genetic content, the major regulatory RBPs Hfq, RsmA and RsmN, known for their roles in various important processes (Chihara *et al*., 2019; Gebhardt *et al*., 2020; Romero *et al*., 2018), are conserved

between all *P. aeruginosa* strains (Figure 1B), raising the question of the functional conservation of their cognate regulatory sRNAs and target mRNAs. Hfq was found especially well conserved, with 100% of sequence identity between PAO1, PA14 and IHMA87. Interestingly, ProQ, originally identified as a major sRNA-binding regulator in *Salmonella enterica* (Smirnov *et al*., 2016), exhibits a five amino acids insertion at the N-terminus in the PA7/IHMA87-like lineage (Figure 1B). To further investigate potential differences in post-transcriptional regulatory networks, we examined the conservation of 200 sRNAs previously identified in PA14 (Wurtzel *et al*., 2012) in all 192 complete genomes of *P. aeruginosa* (Winsor *et al*., 2016; Figure 1C). Strikingly, only six strains phylogenetically close to PA14 possess more than 90% of these sRNAs, ~70% of them were detected across the PA14 and PAO1 groups, and only half is present in the 3rd group, giving a first glance at the diversity of sRNAs found between and inside *P. aeruginosa* groups of strains. Therefore, the sRNA genomic content is strain-dependent and reveals the need of inter-strain comparison studies.

The Hfq interactome in *P. aeruginosa* lineages during growth

In the light of the observed differences in predicted sRNA content, we investigated the interactome of the major regulatory RBP Hfq in the representative strains from each *P. aeruginosa* lineages (Figure 1A). To that aim, we created C-terminal 3xFLAG tagged Hfq from *P. aeruginosa* PAO1, PA14 and IHMA87. We demonstrated that the tagging of Hfq with the Flag peptide did not affect the growth of the corresponding *P. aeruginosa* strain, and that they were functional as each tagged Hfq immunoprecipitated Prrf1, an sRNA conserved in all three strains known to bind Hfq (Supplemental Figure 1). Using 3xFLAG-Hfq expressing strains of PAO1, PA14 and IHMA87, we performed RNA-co-Immunoprecipitation followed by deep sequencing (RIP-seq) (Saliba *et al*., 2017) at exponential, transition and stationary growth phases, and analyzed enriched RNAs compared to control untagged samples (Figure 2A-C; Supplementary Table 1).

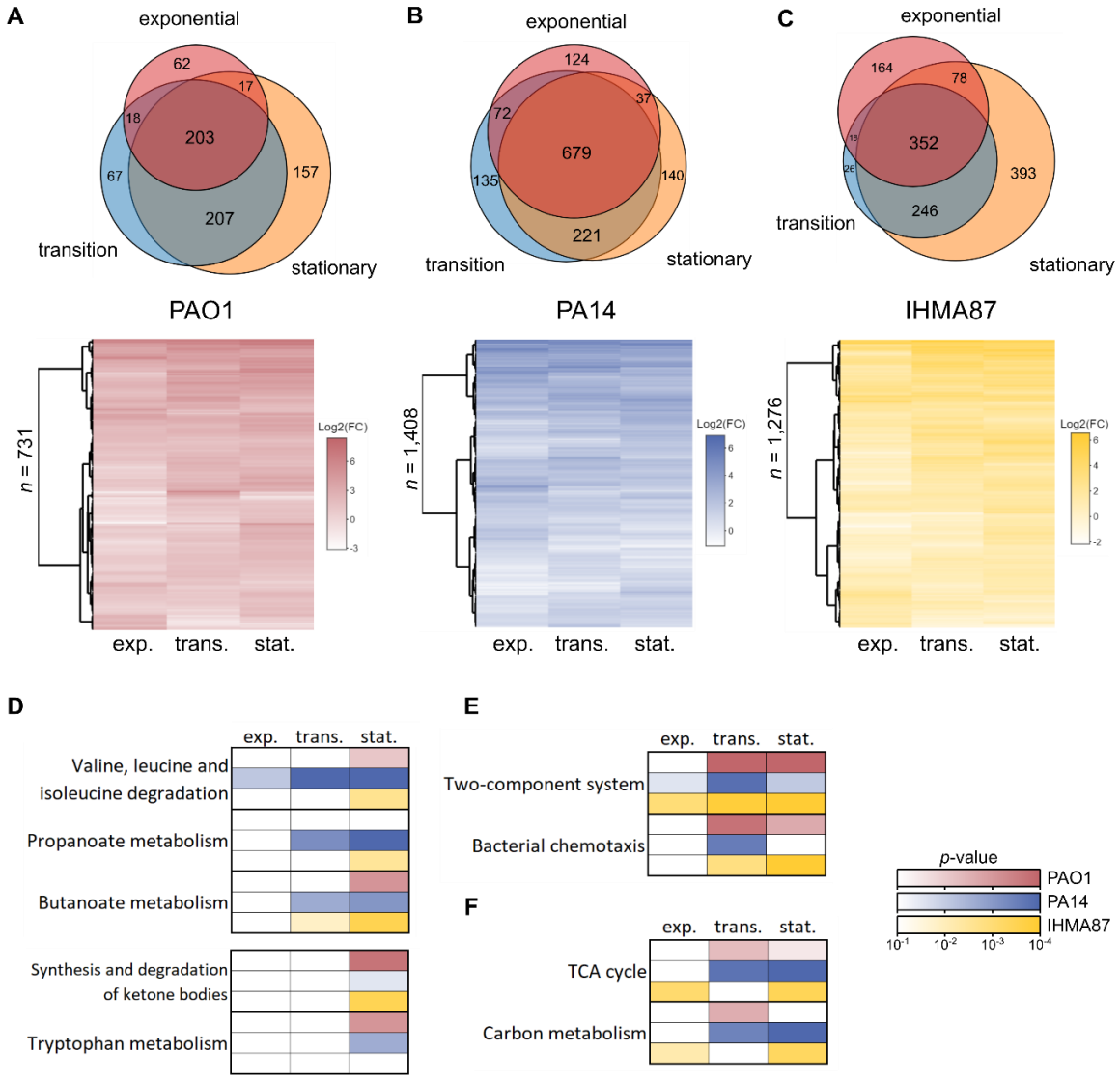


Figure 2: Determination of Hfq interactome during bacterial growth. (**A** - **C**) Venn diagram showing the comparative analysis of number of RNA targets between growth phases for PAO1 (**A**), PA14 (**B**) and IHMA87 (**C**). Lower panels: Hierarchical clustering of all significantly enriched targets between growth phases for each strain. The heatmaps show fold enrichment in Hfq-3xFLAG samples compared to WT results from DESeq2 analysis. (**D** - **F**) Heatmaps showing results of functional enrichment analysis for selected KEGG pathways and GO Terms for each strain and growth phase. Only genes shared with PAO1 homologs were used on the DAVID website (Huang *et al.*, 2007).

Similar patterns of growth phase clustering were observed for the three strains. Namely, (i) a 30-50% core set of targets were enriched in all three growth phases, (ii) there was a 3- to 10-fold stronger overlap in transcripts between transition and stationary phases and (iii) specific sets of targets identified in each

growth phase, with a higher number of specific targets during stationary phase. Functional enrichment analysis allowed the identification of major biological processes enriched in all nine combinations (strains and growth phases). Notably, we found that the branched-chain amino acids (BCAA) degradation functional group was enriched in a growth phase dependent manner (Figure 2D), in concordance with their known regulatory role in response to amino acid starvation through the DNA-binding protein Lrp (Kaiser & Heinrichs, 2018). This result reveals a new conserved role for Hfq in the modulation of BCAA-dependent growth inhibition, as it was already suggested in *Vibrio alginolyticus* (Deng *et al*., 2016). Other growth-dependent metabolic pathways, such as the synthesis and degradation of ketone bodies (Magnuson *et al*., 1993), were enriched in the late phases of growth, emphasizing the central role of Hfq in growth rate determination and starvation response. Similarly, two-component systems and bacterial chemotaxis related targets were also enriched in the later phases of growth (Figure 2E; Supplementary Table 2), in agreement with the known role of Hfq in community-based behaviors found in high cell density conditions (Sonnleitner *et al*., 2006). Finally, the two KEGG pathways Carbon metabolism and TCA cycle were enriched during growth, which corroborates the well-established function of Hfq in central metabolism regulation through binding to the catabolite repression control protein Crc (Sonnleitner *et al*., 2018; Kambara *et al*., 2018). Altogether, our approach clearly delineates the growth-dependent major roles of Hfq and show that these impact globally the core and accessory metabolisms as well as bacterial communication systems. The Hfq-dependent regulation of these key mechanisms seems overall well conserved between the three tested strains.

The global comparison revealed 427 common targets between strains. A number of targets were shared between each pair and Hfq bound transcripts unique to individual strains were also identified: 133 for PAO1, 562 for PA14 and 542 for IHMA87 (Figure 3A). Among the 427 common targets, 109 were found associated with Hfq in all three samplings of growth stages (Figure 3D). These 427 growth-dependent (GD) and 109 growth-independent (GI) core targets of Hfq, and span major key cellular functions (Figure 3E).

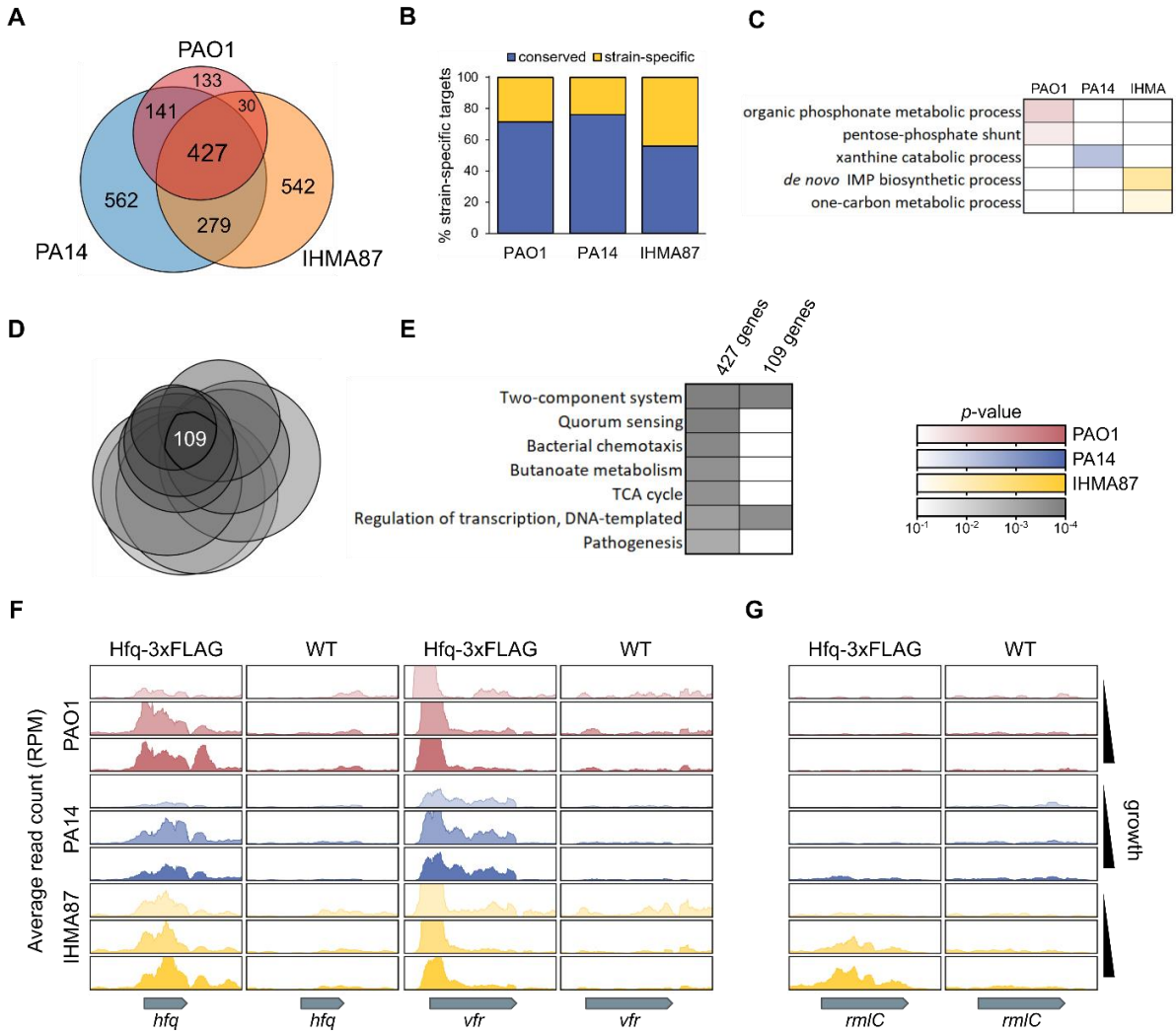


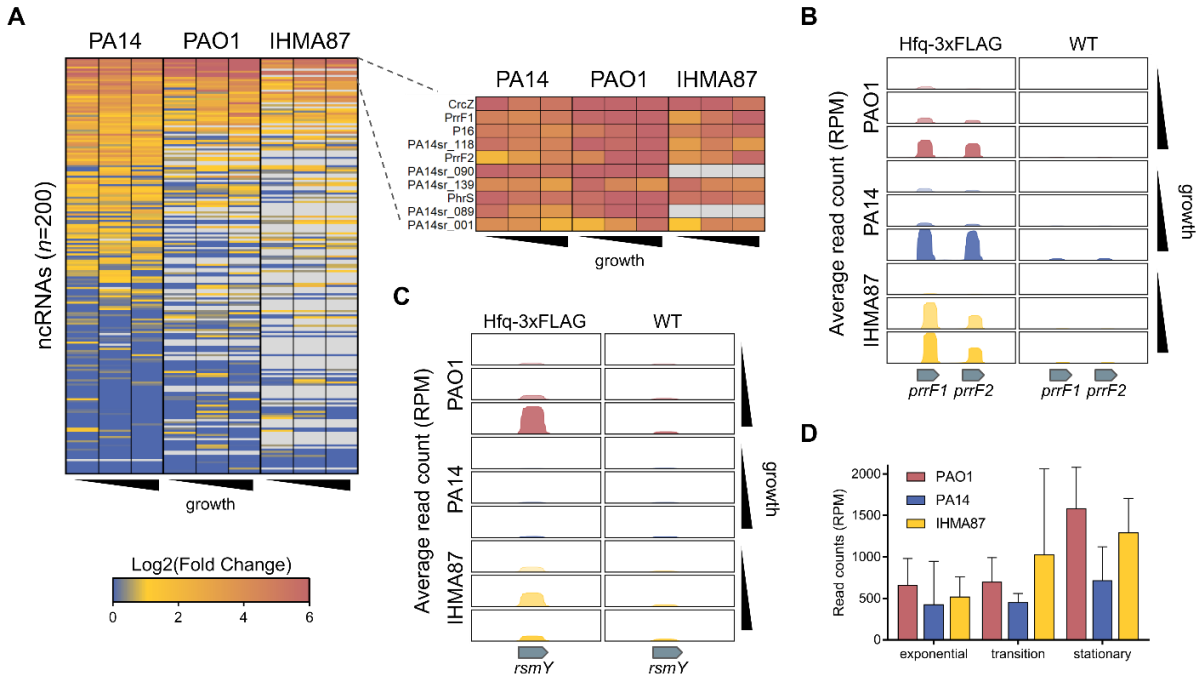
Figure 3: Comparative analysis of Hfq targets across strains. (**A**) Venn diagram showing overall Hfq target overlaps between all three strains. (**B**) Proportion of strain-specific Hfq targets that are shared between genomes (blue) or that are genes specific to the corresponding strain (yellow). (**C**) Results of functional enrichment analysis of strain-specific Hfq targets. (**D**) Venn diagram showing Hfq targets overlaps between all combination of strains and growth phases. (**E**) Results of functional enrichment analysis of core Hfq targets using the group of 427 growth-dependent core targets identified in (A) and the group of 109 growth-independent core targets identified in (D). Only genes shared with PAO1 homologs were used on the DAVID website (Huang *et al.*, 2007). (**F**) Read coverages, averaged between biological replicates, at the *hfq* and *vfr* genes for Hfq-tagged and control samples in all strains and growth phases. (**G**) Same as in (F) for the *rmlC* gene. Read counts are expressed in read per million reads (RPM).

The set of GI core targets includes two-component systems and transcription regulators, demonstrating the important role of Hfq as a master regulator of transcriptional as well as post-transcriptional regulatory networks. Notable among these targets are *hfq* mRNA, confirming the observation in *E. coli* that it is subject to autoregulation (Morita & Aila, 2019) and the global virulence factor regulator *vfr*, previously reported to be regulated, at the post-transcriptional level, by simultaneous binding of the Hfq and RsmA RNA-binding proteins (Irie *et al*., 2018) (Figure 3F).

Interestingly, among strain-specific Hfq targets, only ~30% of them are transcribed from the accessory genome (Figure 3B), while the rest are conserved in at least two strains, suggesting that most of the differences between strains arises from differential binding of Hfq to the same conserved target. Among the strain-specific targets, there were few secondary metabolic pathways (Figure 3C), which could potentially correspond to the strains respective specific niches. Interestingly, one of the accessory targets identified in IHMA87 was *rmlC*, encoding an isomerase responsible for core oligosaccharide and O polysaccharide assembly (Rahim *et al*., 2000) (Figure 3G). This result suggests a role of Hfq in regulation of LPS synthesis in IHMA87 which is O11_O12 serotype (Reboud *et al*., 2016), and not in PAO1 or PA14 which belong to serotypes O5 and O10, respectively.

Annotated and novel Hfq-bound sRNAs

To account for all Hfq-bound sRNAs, we integrated in the analysis 200 sRNAs from PA14 (Wurtzel *et al*., 2012) and the corresponding 150 and 101 homologs identified computationally in PAO1 and IHMA87 respectively (Supplementary Table 3). We found approximately 40% bound to Hfq in at least one growth phase in each strains (Figure 4A), highlighting again the pivotal role of Hfq in sRNA-driven post-transcriptional regulation. Interestingly, most of the differences between strains stemmed from the conservation of sRNAs themselves and not different binding to Hfq as the three strains displayed a globally



181

182 **Figure 4: sRNAs interactions with Hfq.** (**A**) Heatmap of fold enrichments in Hfq-3xFLAG samples compared
183 to WT for a set of 200 sRNAs in all strains and growth phases. (**B** - **C**) Read coverages, averaged between
184 biological replicates, at the *prrf1-2* (B) and *rsmY* (C) loci for Hfq-tagged and control samples in all strains
185 and growth phases. (**D**) Normalized read counts for the sRNA RsmY in RIP-seq control samples. Read
186 counts are expressed in read per million reads (RPM).

187 similar pattern of sRNA enrichment. This suggests that sRNA-Hfq interactions are relatively conserved,
188 and that strain-specific sRNA-Hfq interactions mainly arose from different genomic sRNA pools. Most
189 previously studied sRNA-Hfq interactions were conserved across strains, such as the interaction with PrrF1
190 and Prrf2 which was found increasingly stronger with growth (Figure 4B). Surprisingly, the major
191 regulatory sRNA RsmY was found significantly enriched in PAO1 and IHMA87, but not in PA14 (Figure 4C).
192 RsmY was slightly less abundant in PA14 as assessed in RIP-seq control samples (Figure 4D), which may be
193 due to a mutation in the upstream regulator histidine kinase LadS in PA14 (Mikkelsen *et al.*, 2011) and
194 could explain the difference in RIP-seq enrichment.

195 Most specificities in sRNA-Hfq interactions were due to differences in sRNAs content, thus raising the need
196 of identifying all strain-specific sRNAs. To that aim, we performed an annotation-independent peak-calling

197 analysis in order to rescue missed targets and unannotated sRNAs using the CLIP-seq and RIP-seq peak
198 caller PEAKachu (Li *et al*., 2018) (Supplementary Table 4). 332 intergenic peaks were identified across all
199 strains and growth phases (Supplementary Table 5). Among these, we identified 7 sRNA predicted
200 structures matching the StructRNAfinder bacterial sRNA database (Arias-Carrasco *et al*., 2018)
201 (Supplementary Figure 2 and Supplementary Table 6), which could correspond to unannotated sRNAs
202 from diverse known families of sRNAs. Among these, two were actually the same sRNA detected in both
203 PA14 and IHMA87 genomes, and its sequence can also be found in PAO1. Moreover, more than half of
204 intergenic peaks showed high enrichment (>10-fold) and could correspond to even more new regulatory
205 sRNAs. Also, antisense intragenic peaks were identified in 286 genes across strains (Supplementary Table
206 7), revealing a wide array of potential antisense regulatory sRNAs which are common across the bacterial
207 kingdom and often regulators of their cognate gene (Saberi *et al*., 2016). Notably, a new antisense
208 transcript, encompassing a predicted sRNA secondary structure close to the *Mycobacterium* sRNA
209 Ms_IGR-4, was detected in PAO1 and PA14 within the *dnaA* gene which encodes a replication initiation
210 protein essential for bacterial life (Katayama *et al*., 2010) (Supplementary Figure 3). The detected
211 transcript is found directly downstream of a transcription start site that was previously detected in PA14
212 but not associated with any sRNA (Wurtzel *et al*., 2012). The binding of Hfq to a *dnaA*-antisense sRNA
213 reveals a potential new regulatory mechanism for a key cellular process; i.e. DNA replication, that would
214 need further investigation. Overall, the untargeted analysis of our RIP-seq experiment allowed the
215 identification of numerous potential new sRNAs that could be involved in the observed differences in Hfq
216 interactomes between strains.

217 **The accessory Hfq interactome comprises T3SS mRNAs and a cluster of CRISPR guide RNAs**

218 The Type III Secretion System is the major bacterial virulence-related machinery allowing the injection of
219 toxins inside of host cells (Hauser, 2009). T3SS is present in PAO1- and PA14-lineages and its expression is
220 tightly regulated at both the transcriptional and post-transcriptional levels (for a review see McMackin *et*

al., 2019). The Hfq enrichment analysis resulted in the identification of nearly all T3SS- genes being bound to Hfq throughout all growth phases in both strains (Figure 5A-C). The *exsA* mRNA, encoding for a major transcriptional activator of T3SS gene, was strongly enriched (Figure 5C), in agreement with our previous findings on the role of Hfq in base pairing between sRNA sr0161 and *exsA* mRNA (Zhang *et al*., 2017). The T3SS genetic determinants comprise more than 30 genes encoding the T3SS machinery and its regulatory proteins and at least 7 transcriptional units resulting in the transcription of long polycistronic mRNAs. The observaton that Hfq binds to some of these mRNAs reveals the existence of a novel direct regulatory mechanism potentially globally affecting T3SS-related translation. Moreover, all T3SS toxins, that are encoded at different genetic loci, were found among the most enriched RNAs in our Hfq RIP-seq experiment in both strains, including the shared ExoT and ExoY, as well as strain-specific toxins, ExoS and ExoU (Figure 5D). Therefore, in addition to the known Hfq influence on T3SS expression through *fis* mRNA or RsmY (Lu et al., 2016), Hfq governs multiple interactions with most T3SS mRNAs, potentially involving regulatory sRNAs yet to be identified.

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) systems are widely used prokaryotic antiviral defense mechanisms. Bacteria acquire resistance to bacteriophages by chromosomic integration of short fragments of viral nucleic acids serving later as RNA templates (crRNA) to interfere with phage proliferation (Brouns *et al*., 2008). Several CRISPR systems are present in different *P. aeruginosa* strains (van Belkum *et al*., 2015) suggesting that they were acquired though horizontal gene transfer. While the PAO1 strain does not encode any, PA14 and IHM87 each carries a CRISPR type I-F locus (*PA14_33300-PA14_33350* and *IHMA87_02728-IHMA87_02733*, respectively). The PA14 CRISPR system has been shown to be regulated by Quorum sensing (Høyland-Kroghsbo *et al*., 2017) and by the PhrS sRNA (Lin *et al*., 2019). PhrS represses the Rho-dependent termination of CRISPR RNAs (crRNA) in the absence of Hfq, promoting CRISPR adaptive immunity against bacteriophage infection and represents the first example of sRNA-

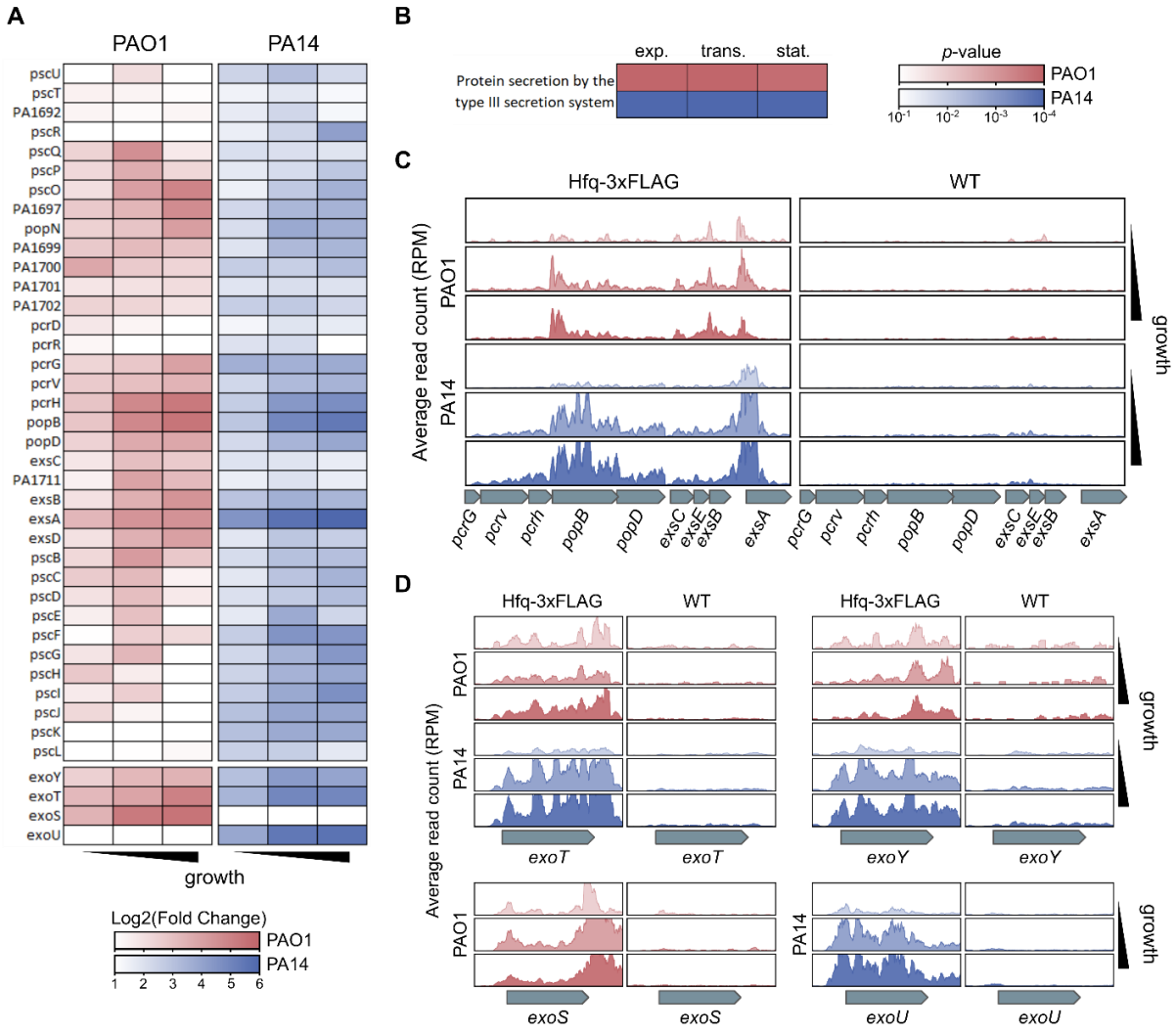


Figure 5: Hfq binds to most of the T3SS-related RNAs. (**A**) Heatmap of fold enrichments in Hfq-3xFLAG samples compared to WT for all T3SS-related genes in PAO1 and PA14 in all growth phases. (**B**) Heatmap showing results of functional enrichment analysis for the T3SS-related GO Term for PAO1 and PA14 in all growth phases. Only genes shared with PAO1 homologs were used on the DAVID website (Huang *et al.*, 2007). (**C** - **D**) Read coverages, averaged between biological replicates, at the *pcrG-to-exsA* (C) and all exotoxins (C) loci for Hfq-tagged and control samples in both strains in all growth phases. Read counts are expressed in read per million reads (RPM).

mediated regulation of CRISPR systems. In bacterial genomes, the crRNAs clusters are usually found linked to the loci encoding CRISPR functional proteins, and PA14 and IHMA87 possess 2 and 3 crRNAs clusters,

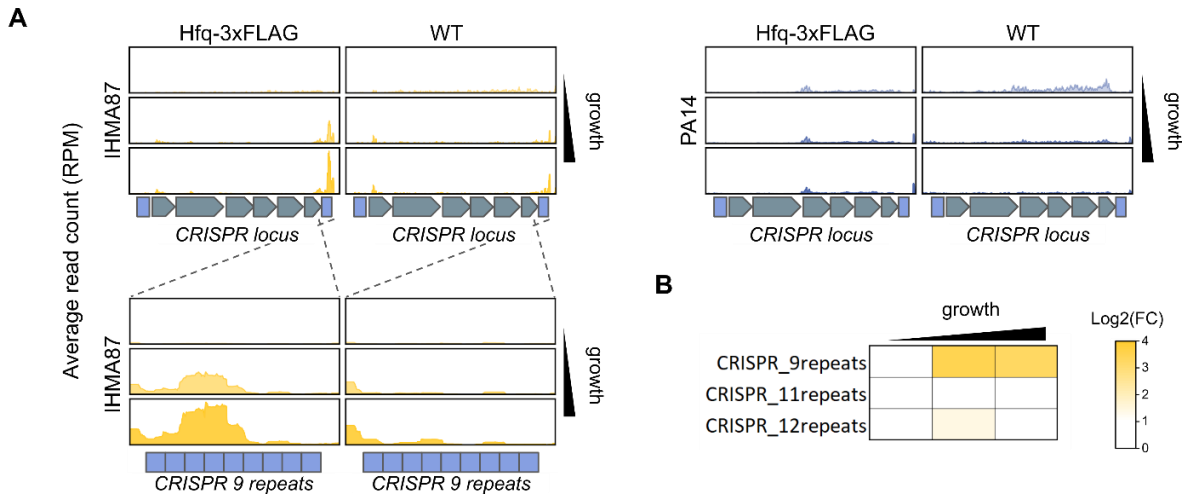


Figure 6: Hfq binds to a cluster of CRISPR RNA specifically in IHMA87. (**A**) Read coverages, averaged between biological replicates, at the major CRISPR loci for Hfq-tagged and control samples in IHMA87 and PA14 in all growth phases. (**B**) Heatmap of fold enrichments in Hfq-3xFLAG samples compared to WT for the annotated CRISPR clusters in IHMA87 in all growth phases. Read counts are expressed in read per million reads (RPM).

respectively. Here, the Hfq RIP-seq experiment revealed one cluster of crRNAs enriched up to 12-fold in late growth phases, specifically in IHMA87 (Figure 6A). This cluster encodes crRNAs predicted to target various phages and unknown sequences (Supplementary Table 8). Interestingly, none of the two other crRNA clusters in IHMA87, nor any PA14 crRNA clusters were found bound to Hfq (Figure 6B). PA14 crRNAs were not enriched in our RIP-seq experiment, confirming the Hfq-independent nature of PhrS CRISPR regulation in this strain (Lin *et al.*, 2019). This result suggests a novel Hfq-dependent regulatory mechanism for expression of CRISPR guide RNAs in a strain- and crRNA-specific way.

Altogether our results revealed the accessory interactomes of Hfq and suggest completely novel types of post-transcriptional regulations of important bacterial features.

Discussion

RBP-driven post-transcriptional regulatory networks of key bacterial processes have been the topic of numerous studies in the last decade (Han *et al*., 2016; Zhang *et al*., 2017; Kambara *et al*., 2018; Chihara *et al*., 2019; Gebhardt *et al*., 2020). The major human pathogen *P. aeruginosa* is represented by three major phylogenetic lineages (Freschi *et al*., 2019), populated by strains displaying different genotypic and phenotypic features. In this study, we provide a first glance at the intra-species post-transcriptional regulatory diversity by comparative RIP-seq analysis using Hfq. Our approach showed that Hfq interacts with 13 to 21% of all annotated RNAs, spanning conserved major biological processes and hundreds of strain-specific targets.

The accessory Hfq interactome arises on one hand from specific Hfq binding to common mRNAs and on the other hand from the differential sRNA contents, highlighting the need of both discovery and characterization of pan sRNAs pools. One could argue that the differences in the detected interactions can stem from differences in RNA abundance between strains. To verify this hypothesis, we analyzed RNA abundance using read counts from RIP-seq control samples. Indeed, as RIP-seq usually identifies full-length transcripts (Li *et al*., 2018), and no specific enrichment is done in the control samples, the read counts reflect RNA abundance in the exact same condition. The comparison between the tested strains and growth phases revealed an overall good correlation of RNA expression between strains, and between growth phases (Supplementary Figure 3). For instance, the overall Hfq interactome differs the most between PAO1 and IHMA87 in the stationary phase, however the vast majority of RNAs exhibit similar expression patterns between these two, showing that most of the differences in Hfq interactomes cannot be explained by different RNA abundances. Concomitantly, this analysis also revealed which strain-specific targets are indeed due to differential RNA abundance, as it is the case for the IHMA87-specific target *pvdA*, encoding a pyoverdine biosynthetic enzyme (Visca *et al*., 1994), which was found significantly more

297 abundant in IHMA87 (Supplementary Table 1 and Supplementary Figure 4). This confirms the previously
298 reported higher expression of the *pvd* cluster in IHMA87 in comparison to PAO1 (Trouillon *et al*., 2020),
299 and probably explains the difference in RIP-seq enrichment observed. Overall, RIP-seq control data can
300 be used to infer RNA expression patterns and define the origin of binding differences observed across
301 strains or conditions, which revealed that most of the strain-specific targets identified here are not due
302 difference in RNA expression but potentially to differences in affinity to Hfq or unknown context-specific
303 binding regulatory mechanisms.

304 Hfq regulates bacterial virulence through multiple direct and indirect mechanisms (Sonnleitner *et al*.,
305 2003; Sonnleitner *et al*., 2006; Sorger-Domenigg *et al*., 2007). Our analysis identified both known and
306 novel virulence-related regulatory targets for Hfq in the different strains. First, the *vfr* mRNA, encoding
307 for a global virulence regulator (Fuchs *et al*., 2010), was strongly bound to Hfq in all three strains.
308 Interestingly, while Vfr is known to regulate the T3SS (Marsden *et al*., 2016), it has been recently shown
309 to directly regulate the two-partner secretion system ExlBA (Berry *et al*., 2018), which is the main virulence
310 factor in the T3SS-lacking IHMA87-like lineage (Elsen *et al*., 2014; Reboud *et al*., 2017). This reveals an
311 interesting feature of regulatory plasticity, as the conserved interaction between Hfq and *vfr* mRNA results
312 in different virulence-related regulatory outputs due to the diversity found in Vfr targets across *P.*
313 *aeruginosa* lineages. The T3SS is regulated by Hfq in direct manner through stabilizing the interactions
314 between *exsA* mRNA, encoding the T3SS transcription regulator, and the sRNA sr0161 (Zhang *et al*., 2017).
315 Here, we show that in addition, Hfq interacts with most of the T3SS-related mRNAs, including those of all
316 four T3SS toxins. For now, it is not possible to infer on the regulatory mechanism of Hfq on T3SS targets.
317 Indeed, while the most studied function of Hfq is to facilitate regulatory mRNA-sRNA interactions, it can
318 also play other sRNA-independent roles, including modulation of mRNA stability and translation (dos
319 Santos *et al*., 2019).

Our study also revealed a cluster of CRISPR RNAs as Hfq accessory targets, specifically in the IHMA87 strain. Although some crRNAs were found regulated by a sRNA in PA14 (Lin *et al*., 2019), our study showed that this interaction was Hfq-independent suggesting diverse mechanisms governing crRNA transcription and stability. The crRNA enrichment seems to be centered on three repeats inside of the crRNA cluster, raising the possibility that Hfq interaction might occur after crRNA processing. In light of the importance that CRIPSR systems have taken in many biotechnology fields over the last few years (Pickar-Oliver *et al*., 2019), it would be important to further investigate the underlying mechanism of crRNAs stabilization revealed here.

In conclusion, we showed that the RBP-driven post-transcriptional regulatory networks exhibit intra-species diversity which probably reflects phenotypic readouts. We identified several possible underlining mechanisms including i) differences in gene or sRNA content or expression, ii) differences in affinities to Hfq which could result from mRNA sequence differences and iii) strain-specific regulatory role of the Hfq target, as illustrated for Vfr. All three mechanisms were found responsible for major differences in Hfq role across *P. aeruginosa* lineages and should be investigated in more details in the context of regulatory networks plasticity. Altogether, this work highlights the importance of considering regulatory and genomic content strain specificities when studying regulatory networks, and calls for more comparative approaches in the fields of transcriptional and post-transcriptional regulation.

Material & Methods

Bacterial strains. *P. aeruginosa* and *E. coli* strains were grown in Lysogeny Broth (LB) at 37°C under agitation. Gentamicin was added when needed at the following concentrations: 75 µg/ml for *P. aeruginosa* and 25 µg/ml for *E. coli*.

Plasmids and genetic manipulations. For Hfq C-terminal tagging, approximately 500 bp of upstream and downstream sequences flanking *hfq* stop codon were amplified using primers carrying the sequence coding for a 3xFLAG tag (DYKDHDGDYKDHDIDYKDDDDK) and a 3-Glycine linker. The two resulting, overlapping fragments were then cloned into *SmaI*-cut pEXG2 by SLIC. After cloning, all plasmids were transformed into competent TOP10 *E. coli* cells and verified by sequencing. The pEXG2-derived vectors were transferred into *P. aeruginosa* strains by triparental mating using pRK600 or pRK2013 as helper plasmids. Merodiploids resulting from cointegration events were selected on LB plates containing 25µg/ml irgasan and gentamicin. Single colonies were then plated on NaCl-free LB agar plates containing 10% (wt/vol) sucrose to select for the loss of plasmid, and the resulting sucrose-resistant strains were checked for mutant genotype and gentamicin sensitivity.

Homolog genes/sRNAs identification and phylogeny. All *P. aeruginosa* complete genomes (*n*=192) were retrieved from the *Pseudomonas* Genome Database (Winsor *et al*., 2016). Homolog identification was performed by Reciprocal Best Blast Hit (RBBH) analysis (Cock *et al*., 2015) on the European Galaxy server (Jalili *et al*., 2020) using all protein sequences from *P. aeruginosa* PA14 against all protein sequences from the other 191 genomes with minimum percentage alignment coverages of 90 and sequence identity of 50. The sequences from 66 core genes were concatenated for each genome and a multiple alignment was performed with MAFFT Galaxy version 7.221.3 (Katoh & Standley, 2013) using default settings. The resulting alignment was used to build a Maximum-Likelihood phylogenetic tree using MEGA X with 100 bootstraps which was visualized and annotated using iTOL v5 (Kumar *et al*., 2018; Letunic & Bork, 2019). Annotations for 200 PA14 sRNAs were obtained from Wurtzel *et al*., 2012 and used to retrieve the corresponding sRNAs sequences. sRNA homologs were identified in *P. aeruginosa* complete genomes using megablast (blastn Galaxy version 0.3.3) with minimum percentage alignment coverages of 90 and sequence identity of 80. All newly identified sRNAs were added to the PAO1 and IHMA87 genome

annotations that were used in this work. Hierarchical clustering analysis of the 200 sRNAs conservation in *P. aeruginosa* genomes was performed with BioVinci v1.1.5 (BioTuring Inc.) using Ward's minimum variance and Manhattan distances.

Western Blot. Proteins amount in total bacteria was evaluated from the equivalent of 1 ml of liquid culture grown to OD_{600}=1 which was pelleted, resuspended in 100 µl of loading buffer and incubated at 100°C for 10 min. Samples were separated on denaturing 12 % PAGE and transferred onto polyvinyl difluoride (PVDF) membranes. Primary antibodies used for immunodetection anti-FLAG M2 mouse monoclonal antibodies (Sigma). Secondary antibodies were anti-mouse-HRP (Sigma) diluted in PBS-Tween. The membranes were developed with Luminata Classico Western HRP (Millipore) substrate.

RT-qPCR. RNAs were extracted as described for RIP-seq and treated with the TURBO DNA-free kit following manufacturer's instructions (Thermofisher). For RT-qPCR, 4 µl of DNase-treated RNA were used with the Luna Universal One-Step RT-qPCR kit (NEB) and primers specific to PrrF1 or 6S RNA. Experiments were performed in biological duplicates for each strain, and the ΔΔCq method was used to obtain fold change values between WT and Hfq-tagged strains normalized to the 6S RNA abundance.

Hfq-RNA co-Immunoprecipitation. Bacterial cultures were started in 200ml of LB medium at OD_{600}=0.1 from overnight cultures. At time of sampling (after 200, 400 or 600 minutes of incubation), the equivalent of 50 ml of OD_{600}=1 were pelleted by centrifugation and snap frozen in liquid nitrogen. Pellets were then resuspended in 700 µl of Lysis buffer (20 mM Tris-HCl pH 8, 150 mM KCl, 1 mM MgCl2, 1 mM DTT) and, after addition of 700 µl of acid-washed, <106 µm glass beads (Sigma), vortexed by cycles of 30 seconds

vortexing and 30 seconds on ice for a total of 10 min. Lysates were centrifugated for 30 min at 16,000 g at 4°C and supernatant were then incubated with 40 µl of washed Anti-FLAG M2 Magnetic beads (Sigma) for 1.5 hour at 4°C. The beads were then washed five times by resuspension in 500 µl of Lysis buffer. After the last wash, 500 µl of Phenol:Chlorophorm:Isoamyl alcohol was added to the resuspended beads and RNA was extracted following standard Phenol:Chlorophorm extraction followed by overnight ethanol precipitation. RIP-seq was also conducted on WT strains in each condition as negative controls. All RIP-seq experiments were performed in biological duplicates for a total of 36 samples.

Library construction and sequencing. Precipitated RNAs were treated with TURBO DNAse following manufacturer's instructions (Thermofisher) and further cleaned up by Phenol:Chlorophorm extraction followed by overnight ethanol precipitation. The quantity and quality of purified RNAs were then assessed on an Agilent Bioanalyzer and 1-5 ng of RNA were then used for library construction using the NEBNext Ultra II Directional Library Prep kit (NEB). Sequencing was performed at the Biopolymers Facility at Harvard Medical School (https://genome.med.harvard.edu/) on an Illumina NextSeq500 for the 24 PAO1 and IHMA87 samples, and at the sequencing core facility of I2BC (http://www.i2bc.paris-saclay.fr) on an Illumina NextSeq500 for the 12 PA14 samples with a total average of ~8.4 million reads per sample.

RIP-seq data analysis. Raw reads were trimmed for adapter sequences and quality using Trimmomatic (Galaxy version 0.38.0) (Bolger *et al.*, 2014) and aligned to their corresponding genome with Bowtie2 (Galaxy version 2.3.4.3) (Langmead *et al.*, 2012; Jalili *et al.*, 2020) with default parameters. sRNAs from our genome screen (Figure 1C) were added to the annotation files used in the analysis (Supplementary Table 3). Read counts per feature were then obtained with htseq-count (Galaxy version 0.9.1) (Anders *et al.*, 2015). Differential RNA enrichment between Hfq-tagged and wild-type strains was assessed using

DESeq2 (Galaxy version 2.11.40.6) (Love *et al.*, 2014). Enriched group comparisons between growth phases were done by hierarchical clustering analysis using BioVinci with Ward's minimum variance and Euclidian distances. For coverage visualization, Bowtie2 alignment output files were converted to bedgraphs with the Genome Coverage tool from BEDTools (Quinlan *et* al., 2010), counts were transformed to reads per million (RPM), and averaged between biological replicates (for figure space reasons). Average local RPM counts were used to generate coverage plots using matplotlib in Python 3.7. RNA abundance comparisons were done using averages of DESeq2 normalized read counts from WT samples transformed to RPM counts among the 4776 genes shared by all three strains. Genes were ranked on their RNA abundance based on their RPM values and a rank matching analysis was performed for all pairs of strains in all growth phases using Pearson's correlation test in BioVinci. Ratios of ranks between pairs of strains in each growth phase were used to identify outlier genes with significantly different RNA abundance between strains. For complementary analysis, peak calling was done using PEAKachu version 0.1.0.2 (Li *et al*., 2018) with maximum insert sizes of 60. The closest gene and distance from it were identified for each peak using the ClosestBed tool from BEDTools (Quinlan *et al*., 2010). Peaks that did not correspond to a gene found enriched with the DESeq2 method either corresponded to mRNA targets that were missed or to new unannotated sRNAs. For detection of potential new intergenic sRNAs, enriched sequences, extended by 10 bp in both directions using BEDTools SlopBed, were obtained from strictly intergenic peaks using BEDTools GetFastaBed (Quinlan *et al*., 2010). Prediction of sRNAs was performed using StructRNAfinder (Arias-Carrasco *et al*., 2018) on the obtained sequences. Potential new antisense sRNAs were manually curated from intragenic peaks that were detected on the opposite strand of their corresponding gene found with ClosestBed.

Inter-strain comparison and functional enrichment analysis. Results from the RBBH analysis were used to generate lists of corresponding gene IDs between the PAO1, PA14 and IHMA87 genomes. For functional

436 enrichment analysis, only genes with PAO1 homologs could be used for PA14 and IHMA87. The resulting
437 lists were analyzed with DAVID v6.8 (Huang *et al*., 2007) against each corresponding strain functional
438 background using PAO1 KEGG Pathways and GOTerm Biological Processes annotations with default
439 parameters.

440

441 **CRISPR spacer prediction and annotation.** Clusters of CRISPR crRNAs were detected using
442 CRISPRCasFinder with default parameters (Couvin *et al*., 2018). Spacer target prediction was done with
443 CRISPRTarget with a score cutoff of 28 (Biswas *et al*., 2013).

References

- Arias-Carrasco, R., Vásquez-Morán, Y., Nakaya, H. I., & Maracaja-Coutinho, V. (2018). StructRNAfinder: an automated pipeline and web server for RNA families prediction. BMC bioinformatics, 19(1), 55.
- Anders, S., Pyl, P.T. and Huber, W. (2015) HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics, 31, 166-169.
- Berry, A., Han, K., Trouillon, J., Robert-Genthon, M., Ragno, M., Lory, S., ... & Elsen, S. (2018). cAMP and Vfr control exolysin expression and cytotoxicity of Pseudomonas aeruginosa taxonomic outliers. Journal of bacteriology, 200(12), e00135-18.
- Biswas, A., Gagnon, J. N., Brouns, S. J., Fineran, P. C., & Brown, C. M. (2013). CRISPRTarget: bioinformatic prediction and analysis of crRNA targets. RNA biology, 10(5), 817-827.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114-2120.
- Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., ... & Van Der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. Science, 321(5891), 960-964.
- Chihara, K., Bischler, T., Barquist, L., Monzon, V. A., Noda, N., Vogel, J., & Tsuneda, S. (2019). Conditional Hfq Association with Small Noncoding RNAs in Pseudomonas aeruginosa Revealed through Comparative UV Cross-Linking Immunoprecipitation Followed by High-Throughput Sequencing. mSystems, 4(6).
- Cock, P. J., Chilton, J. M., Grüning, B., Johnson, J. E., & Soranzo, N. (2015). NCBI BLAST+ integrated into Galaxy. Gigascience, 4(1), s13742-015.
- Couvin, D., Bernheim, A., Toffano-Nioche, C., Touchon, M., Michalik, J., Néron, B., ... & Pourcel, C. (2018). CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic acids research, 46(W1), W246-W251.
- Deng, Y., Chen, C., Zhao, Z., Zhao, J., Jacq, A., Huang, X., & Yang, Y. (2016). The RNA chaperone Hfq is involved in colony morphology, nutrient utilization and oxidative and envelope stress response in Vibrio alginolyticus. PloS one, 11(9).
- dos Santos, R. F., Arraiano, C. M., & Andrade, J. M. (2019). New molecular interactions broaden the functions of the RNA chaperone Hfq. Current genetics, 65(6), 1313-1319.
- Fuchs, E. L., Brutinel, E. D., Jones, A. K., Fulcher, N. B., Urbanowski, M. L., Yahr, T. L., & Wolfgang, M. C. (2010). The Pseudomonas aeruginosa Vfr regulator controls global virulence factor expression through cyclic AMP-dependent and-independent mechanisms. Journal of bacteriology, 192(14), 3553-3564.
- Gebhardt, M. J., Kambara, T. K., Ramsey, K. M., & Dove, S. L. (2020). Widespread targeting of nascent transcripts by RsmA in Pseudomonas aeruginosa. Proceedings of the National Academy of Sciences.
- Han, K., Tjaden, B., & Lory, S. (2016). GRIL-seq provides a method for identifying direct targets of bacterial small regulatory RNA by in vivo proximity ligation. Nature microbiology, 2(3), 1-10.
- Hauser, A. R. (2009). The type III secretion system of Pseudomonas aeruginosa: infection by injection. Nature Reviews Microbiology, 7(9), 654-665.

- He, J., Baldini, R. L., Déziel, E., Saucier, M., Zhang, Q., Liberati, N. T., ... & Rahme, L. G. (2004). The broad host range pathogen Pseudomonas aeruginosa strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. Proceedings of the National Academy of Sciences, 101(8), 2530-2535.
- Høyland-Kroghsbo, N. M., Paczkowski, J., Mukherjee, S., Broniewski, J., Westra, E., Bondy-Denomy, J., & Bassler, B. L. (2017). Quorum sensing controls the Pseudomonas aeruginosa CRISPR-Cas adaptive immune system. Proceedings of the National Academy of Sciences, 114(1), 131-135.
- Huang, D. W., Sherman, B. T., Tan, Q., Kir, J., Liu, D., Bryant, D., ... & Lempicki, R. A. (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic acids research, 35(suppl_2), W169-W175.
- Irie, Y., Geyer, J. E., & Shingler, V. (2018). Hfq-assisted RsmA regulation is central to Pseudomonas aeruginosa biofilm and motility. bioRxiv, 434555.
- Jalili, V., Afgan, E., Gu, Q., Clements, D., Blankenberg, D., Goecks, J., ... & Nekrutenko, A. (2020). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. Nucleic Acids Research.
- Kaiser, J. C., & Heinrichs, D. E. (2018). Branching out: alterations in bacterial physiology and virulence due to branched-chain amino acid deprivation. MBio, 9(5), e01188-18.
- Kambara, T. K., Ramsey, K. M., & Dove, S. L. (2018). Pervasive targeting of nascent transcripts by Hfq. Cell reports, 23(5), 1543-1552.
- Katayama, T., Ozaki, S., Keyamura, K., & Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. Nature Reviews Microbiology, 8(3), 163-170.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution, 30(4), 772-780.
- Kos, V. N., Deraspe, M., McLaughlin, R. E., Whiteaker, J. D., Roy, P. H., Alm, R. A., ... & Gardner, H. (2015). The resistome of Pseudomonas aeruginosa in relationship to phenotypic susceptibility. Antimicrobial agents and chemotherapy, 59(1), 427-436.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular biology and evolution, 35(6), 1547-1549.
- Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods, 9, 357-359.
- Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic acids research, 47(W1), W256-W259.
- Li, L., Förstner, K. U., & Chao, Y. (2018). Computational Analysis of RNA–Protein Interactions via Deep Sequencing. In Transcriptome Data Analysis (pp. 171-182). Humana Press, New York, NY.
- Lin, P., Pu, Q., Wu, Q., Zhou, C., Wang, B., Schettler, J., ... & Li, G. (2019). High-throughput screen reveals sRNAs regulating crRNA biogenesis by targeting CRISPR leader to repress Rho termination. Nature communications, 10(1), 1-12.
- Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol, 15, 550

- Lu P, Wang Y, Zhang Y, Hu Y, Thompson KM, Chen S. 2016. RpoS-dependent sRNA RgsA regulates Fis and AcpP in Pseudomonas aeruginosa. Mol Microbiol 102:244–259. doi:10.1111/mmi.13458.
- Magnuson, K., Jackowski, S., Rock, C. O., & Cronan, J. E. (1993). Regulation of fatty acid biosynthesis in Escherichia coli. Microbiology and Molecular Biology Reviews, 57(3), 522-542.
- Marsden, A. E., Intile, P. J., Schulmeyer, K. H., Simmons-Patterson, E. R., Urbanowski, M. L., Wolfgang, M. C., & Yahr, T. L. (2016). Vfr directly activates exsA transcription to regulate expression of the Pseudomonas aeruginosa type III secretion system. Journal of bacteriology, 198(9), 1442-1450.
- McMackin, E. A. W., Djapgne, L., Corley, J. M., & Yahr, T. L. (2019). Fitting pieces into the puzzle of Pseudomonas aeruginosa type III secretion system gene expression. Journal of bacteriology, 201(13), e00209-19.
- Mikkelsen, H., McMullan, R., & Filloux, A. (2011). The Pseudomonas aeruginosa reference strain PA14 displays increased virulence due to a mutation in ladS. PloS one, 6(12).
- Morita, T., & Aiba, H. (2019). Mechanism and physiological significance of autoregulation of the Escherichia coli hfq gene. RNA, 25(2), 264-276.
- Pickar-Oliver, A., & Gersbach, C. A. (2019). The next generation of CRISPR–Cas technologies and applications. Nature Reviews Molecular Cell Biology, 20(8), 490-507.
- Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics, 26, 841-842.
- Rahim, R., Burrows, L. L., Monteiro, M. A., Perry, M. B., & Lam, J. S. (2000). Involvement of the rml locus in core oligosaccharide and O polysaccharide assembly in Pseudomonas aeruginosa. Microbiology, 146(11), 2803-2814.
- Reboud, E., Elsen, S., Bouillot, S., Golovkine, G., Basso, P., Jeannot, K., ... & Huber, P. (2016). Phenotype and toxicity of the recently discovered exlA-positive Pseudomonas aeruginosa strains collected worldwide. Environmental microbiology, 18(10), 3425-3439.
- Reboud, E., Basso, P., Maillard, A. P., Huber, P., & Attrée, I. (2017). Exolysin shapes the virulence of Pseudomonas aeruginosa clonal outliers. Toxins, 9(11), 364.
- Romero, M., Silistre, H., Lovelock, L., Wright, V. J., Chan, K. G., Hong, K. W., ... & Heeb, S. (2018). Genome-wide mapping of the RNA targets of the Pseudomonas aeruginosa riboregulatory protein RsmN. Nucleic acids research, 46(13), 6823-6840.
- Roy, P. H., Tetu, S. G., Larouche, A., Elbourne, L., Tremblay, S., Ren, Q., ... & Mahamoud, Y. (2010). Complete genome sequence of the multiresistant taxonomic outlier Pseudomonas aeruginosa PA7. PloS one, 5(1), e8842.
- Saberi, F., Kamali, M., Najafi, A., Yazdanparast, A., & Moghaddam, M. M. (2016). Natural antisense RNAs as mRNA regulatory elements in bacteria: a review on function and applications. Cellular & molecular biology letters, 21(1), 6.
- Saliba, A. E., Santos, S. C., & Vogel, J. (2017). New RNA-seq approaches for the study of bacterial pathogens. Current opinion in microbiology, 35, 78-87.
- Sonnleitner, E., Hagens, S., Rosenau, F., Wilhelm, S., Habel, A., Jäger, K. E., & Bläsi, U. (2003). Reduced virulence of a hfq mutant of Pseudomonas aeruginosa O1. Microbial pathogenesis, 35(5), 217-228.

- Sonnleitner, E., Schuster, M., Sorger-Domenigg, T., Greenberg, E. P., & Bläsi, U. (2006). Hfq-dependent alterations of the transcriptome profile and effects on quorum sensing in Pseudomonas aeruginosa. Molecular microbiology, 59(5), 1542-1558.
- Sonnleitner, E., Wulf, A., Campagne, S., Pei, X. Y., Wolfinger, M. T., Forlani, G., ... & Luisi, B. F. (2018). Interplay between the catabolite repression control protein Crc, Hfq and RNA in Hfq-dependent translational regulation in Pseudomonas aeruginosa. Nucleic acids research, 46(3), 1470-1485.
- Sorger-Domenigg, T., Sonnleitner, E., Kaberdin, V. R., & Bläsi, U. (2007). Distinct and overlapping binding sites of Pseudomonas aeruginosa Hfq and RsmA proteins on the non-coding RNA RsmY. Biochemical and biophysical research communications, 352(3), 769-773.
- Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warrener, P., Hickey, M. J., ... & Garber, R. L. (2000). Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen. Nature, 406(6799), 959-964.
- Trouillon, J., Sentausa, E., Ragno, M., Robert-Genthon, M., Lory, S., Attrée, I., & Elsen, S. (2020). Species-specific recruitment of transcription factors dictates toxin expression. Nucleic Acids Research.
- van Belkum, A., Soriaga, L. B., LaFave, M. C., Akella, S., Veyrieras, J. B., Barbu, E. M., ... & Miller, K. (2015). Phylogenetic distribution of CRISPR-Cas systems in antibiotic-resistant Pseudomonas aeruginosa. MBio, 6(6).
- Visca, P., Ciervo, A., & Orsi, N. (1994). Cloning and nucleotide sequence of the pvdA gene encoding the pyoverdin biosynthetic enzyme L-ornithine N5-oxygenase in Pseudomonas aeruginosa. Journal of bacteriology, 176(4), 1128-1140.
- Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A., & Brinkman, F. S. (2016). Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. Nucleic acids research, 44(D1), D646-D653.
- Zhang, Y. F., Han, K., Chandler, C. E., Tjaden, B., Ernst, R. K., & Lory, S. (2017). Probing the sRNA regulatory landscape of P. aeruginosa: post-transcriptional control of determinants of pathogenicity and antibiotic susceptibility. Molecular microbiology, 106(6), 919-937.

5.1.2 Additional results

While I was setting up and assessing the suitability of RIP-seq to determine Hfq interactomes in the three tested strains, I was also exploring other experimental options. Since I had just been working with DAP-seq at that time, I tried to develop an *in vitro* alternative to RIP-seq in a similar way as DAP-seq is the *in vitro* alternative to ChIP-seq. To that aim, I expressed and purified a His-tagged Hfq protein that I then used in *in vitro* RNA binding assays. I purified total RNAs from *P. aeruginosa* PAO1, treated them with DNase and incubated them with beads-bound Hfq protein, similarly to the DAP-seq protocol for DNA binding incubation. I tested several quantities of protein (0, 50, 500, 2000 and 5000 ng), several quantities of RNA (50, 100 and 500 ng), several times of incubation (15, 45 and 120 min), several numbers of washes after incubation (2, 4 or 6) and either untreated total RNA samples or RNAs that have been heat-denatured and renatured before use but I could never see any enrichment of known Hfq targets in the supposedly "bound" RNA samples after incubation and washing.
I unfortunately had to stop trying to make this work as RIP-seq proved efficient at enriching Hfq RNA targets and I thus focused on it. However, even though it would probably require some work optimizing it, I feel like this RAP-seq (DAP-seq equivalent for "R"NA Affinity purificaiton and sequencing) approach could work.

Additional Materials & Methods

Protein purification. Hfq purification was performed as previously described (Sonnleitner et al., 2006).

RNA purification. RNA were purified and DNase-treated as previously described (Trouillon et al., 2020a).

***in vitro* RNA binding assay.** Purified Hfq proteins were diluted in 280 µl of binding buffer (10 mM Tris-Cl, 50 mM NH_4Cl, 2 % Glycerol, pH 7.5) to which 20 µL of Dynabeads His-Tag Isolation and Pulldown magnetic beads(Invitrogen) previously washed three times with 500 µl of binding buffer were added. Hfq protein (0-5 µg) was incubated with washed beads for 20 min at room temperature on a rotating wheel. Protein-bound beads were then washed 4 times in 500 µl of binding buffer and resuspended in 100 µl of binding buffer containing the purified RNAs (50-500 ng). Binding incubation was for 15-120 min at room temperature on a rotating wheel. Beads were then washed 2-6 times in 200 µl of binding buffer

and then resuspended in 200 µl of binding buffer. 20 µl of 3M Na acetate and 220µl Phenol:Chlorophorm:Isoamyl alcohol were added to the beads and RNA was then extracted following standard Phenol:Chlorophorm extraction followed by overnight ethanol precipitation. To check for proper protein binding to the beads, the phenol phase from RNA extraction was added to 600 µl of ice-cold acetone and proteins were then precipitated by standard acetone purification and protein presence was checked by western blotting using anti-His antibodies as described in the manuscript in section 5.1.1.

RT-qPCR. RT-qPCR to check for PrrF1 enrichment related to 6S RNA abundance in samples containing Hfq protein compared to no protein samples were performed as described in the manuscript in section 5.1.1.

5.2 The ProQ regulatory RBP

While in the Lory lab, I actually worked on two RBPs, Hfq, for which the results were presented above, and ProQ. ProQ is a recently discovered regulatory RBP (Holmqvist et al., 2020) which has not been studied in *P. aeruginosa* yet although it possesses a putative ProQ homolog (PA2582 in PAO1, IHMA87_02527 in IHMA87). In an attempt to determine its role in *P. aeruginosa*, I did the exact same RIP-seq analysis described for Hfq above with ProQ in PAO1 and IHMA87. ProQ seemed less abundant than Hfq in both strains but was still detected along growth during the RIP-seq experiment (Figure 5.1A). Additionally, ProQ was correctly recovered after immoprecipitation during RIP-seq (Figure 5.1B). However, the analysis of ProQ RIP-seq experiments yielded relatively unexpected results. Even though the role of ProQ is not known in *P. aeruginosa*, it has been shown to both regulate mRNA stability and mediate sRNA-mRNA interactions in several other species, often impacting hundreds of target RNAs (Smirnov et al., 2016; Holmqvist et al., 2020). In this RIP-seq experiment, only about 100 RNAs on average were enriched in PAO1, including only about 10 with strong fold enrichments (> 5-fold) and 0 sRNA (Figure

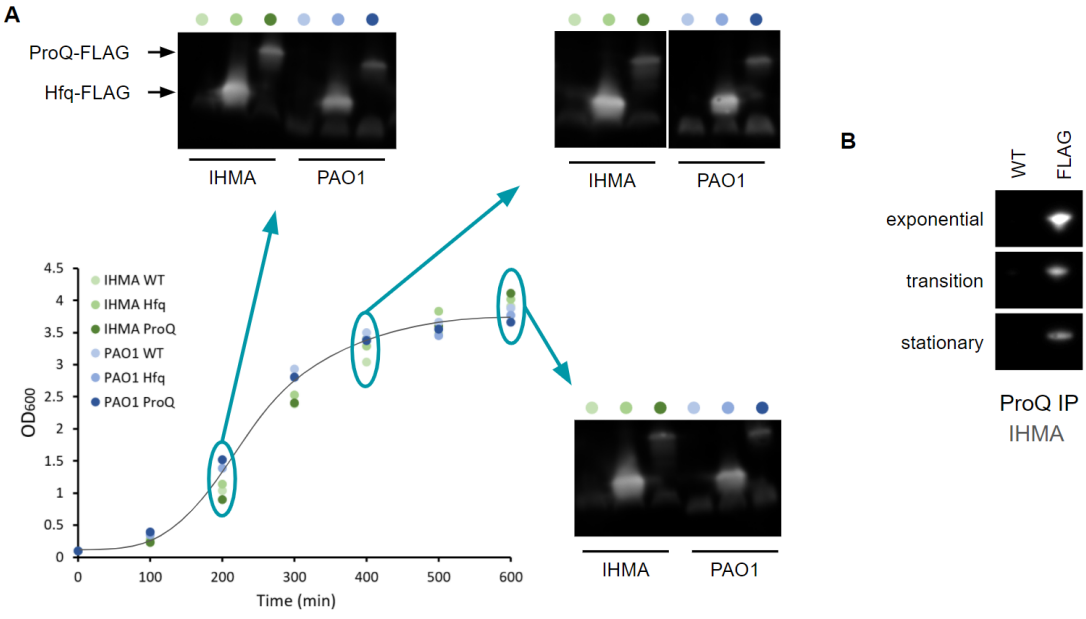


Figure 5.1: RIP-seq experimental procedure. (**A**) Growth curves of strains used in this experiment are shown during growth in LB at 37°C. At each sampling time point (200, 400, 600 min), whole bacteria were sampled and analyzed by SDS-PAGE followed by anti-FLAG antibody immunodetection. (**B**) Protein extracts from phenol phases after RIP-seq experiments at three different time points were concentrated by acetone precipitation and analyzed by SDS-PAGE followed by anti-FLAG antibody immunodetection.

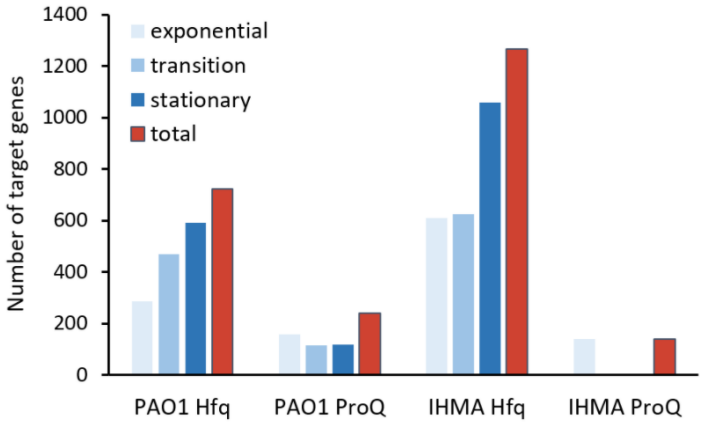


Figure 5.2: Number of significantly enriched RNA in the RIP-seq experiment. All significantly enriched RNAs (p-value < 0.05) in Hfq and ProQ RIP-seq experiments in PAO1 and IHMA87.

5.2). In IHMA87, no RNA was enriched in two time points and no sRNA in the third.

Additionally, only few common targets were found between PAO1 and IHMA87, including none of the most enriched RNAs in PAO1 (Figure 5.3). In an attempt to confirm enrichment of some targets in IHMA87, RT-qPCR did not match the RIP-seq enrichment patterns (Figure 5.4) and enrichment signals were always diffuse along entire genes as opposed to what could be expected from ProQ binding to RNAs. Overall, these results were not convincing enough to know whether the experiment did work or not, and as my time was limited, I stopped working on ProQ and focused on Hfq. As nothing is known on *P. aeruginosa* ProQ, it is still hard to confirm that this RIP-seq experiment failed and does not actually reflect ProQ interactome. However, the above-mentioned limitations in these results suggest that ProQ might not have been active in our conditions. Indeed, it is possible that the addition of the tag to the protein impaired its function. Further work would be

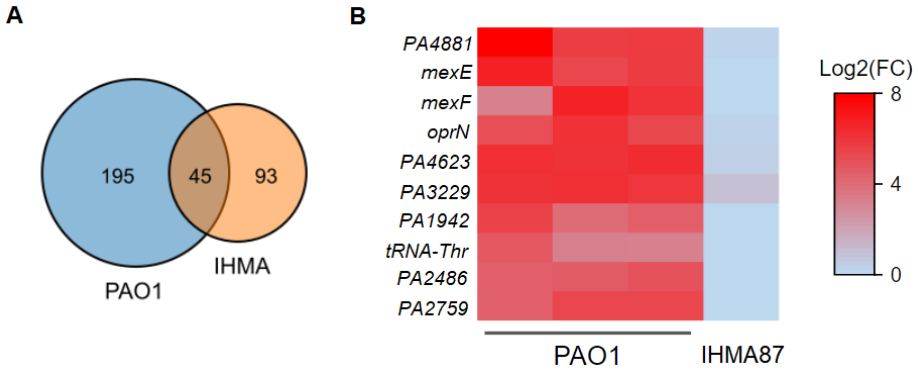


Figure 5.3: Comparison of PAO1 and IHMA87 ProQ RIP-seq results. (**A**) Venn diagram showing all enriched RNAs at all time points in PAO1 and IHMA87. (**B**) RIP-seq fold enrichment of selected RNAs (top 10 in PAO1) in all three PAO1 RIP-seq time points and first IHMA87 time point.

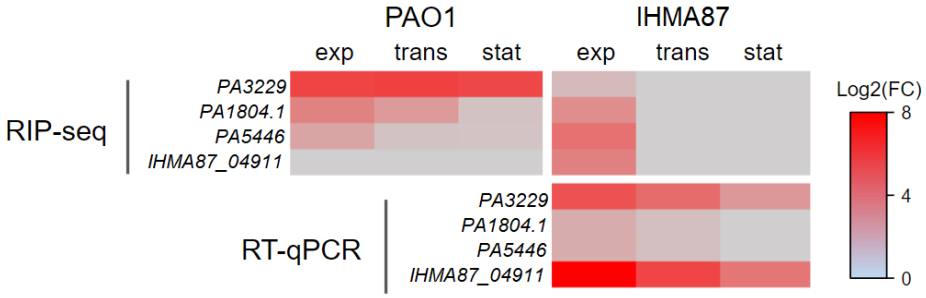


Figure 5.4: Comparison of RIP-seq and RT-qPCR results. Several targets were selected and enrichment was checked by RT-qPCR in IHMA87. Fold enrichment between ProQ-3xFLAG and WT strains were normalized to the abundance of 6S RNA. exp: exponential phase, trans: transition phase, stat: stationary phase.

needed to explain the obtained results.

Materials & Methods

All experiments were performed as described in the manuscript in section 5.1.1.

5.3 The *erfA-ergAB* region

During my stay in the Lory lab, I also worked on *erfA* and *ergAB* in regards to post-transcriptional regulation for several reasons. Indeed, the *erfA* transcription start site is located 253 bp upstream of its start codon, representing a relatively long 5′ untranslated region as the average length is 69 bp in PA14 (Wurtzel et al., 2012). This suggests the implication of post-transcriptional regulation at this region, which could include sRNA-mediated regulation. On the other hand, a new sRNA was discovered inside of the *ergB* gene (Han *et al*., unpublished). This sRNA, called s0223, is at the 3′ end of *ergB* (Figure 5.5) but it is still unclear whether it arises from processing of *ergB* mRNA or from a specific promoter inside of the *ergB* gene.

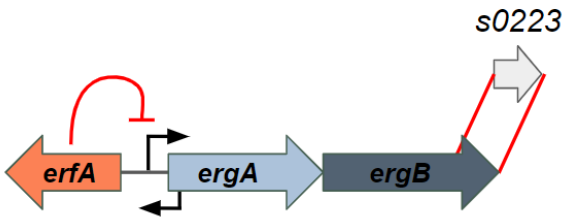


Figure 5.5: The *erfA-ergAB* genomic region. The s0223 sRNA is encoded at the 3′-end of the *ergB* gene. Transcription start sites are shown as black arrows, as predicted with BPROM (Salamov and Solovyevand, 2011).

Overall, these two putative sRNA-mediated input and output of the ErfA regulatory pathway represented new insights to explore on ErfA and ErgAB function. Consequently, I aimed at identifying the potential sRNA regulating *erfA* as well as the role and targets of the s0223 sRNA using GRIL-seq (Global sRNA target Identification by Ligation and Sequencing), a method recently developed in the Lory lab that allows the genome-wide identification of RNA-RNA interactions through *in vivo* ligation and capture of the obtained RNA chimeras (Han et al., 2016). Briefly, both a T4 RNA ligase, which induces ligation of interacting RNAs together as a new chimeric RNA, and the RNA bait (either a sRNA in order to identify its targets or, in reverse GRIL-seq, a mRNA to identify its sRNA regulators) are overexpressed *in vivo*, followed by RNA purification and further enrichment of RNA chimeras through bait-specific capture bead-bound ssDNA primers. I thus overexpressed either s0223 or *erfA* along with T4 RNA ligase in order to identify their specific RNA partners. These overexpressions worked well, were specific, and induced the detection of known RNA-RNA chimeras, as expected (Figure 5.6).
In this work, I performed all experiments and Kook Han analyzed the GRIL-seq results.

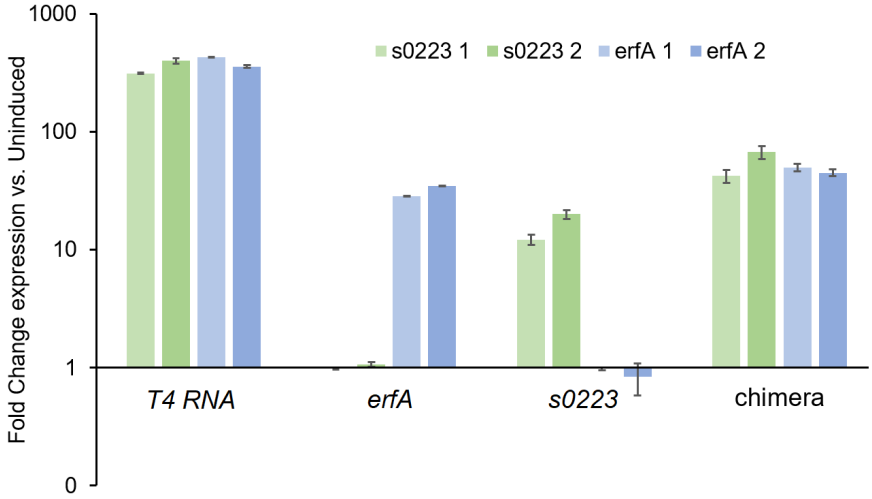


Figure 5.6: GRIL-seq expression controls. RT-qPCR analysis of T4 RNA ligase, *erfA*, s0223 and a known chimera control (PrrF1-*sodB*), normalized to *rpsL* abundance and compared between induced and uninduced cultures grown from OD_{600}=0.01 in LB containing gentamicin and carbenicillin at 37°C. For induced cultures, 1mM IPTG was added at OD_{600}=0.5 (for T4 RNA ligase induction). After 1h of incubation, 0.2% L-Arabinose weas then added (for *erfA* or s0223 induction) and cells were harvested after 20 min of further incubation. 2 biological replicates were done in technical triplicates.

5.3.1 Potential sRNAs regulating *erfA*

The reverse GRIL-seq analysis of *erfA* resulted in three sRNAs forming RNA-RNA chimeras with *erfA* mRNA: RsmZ and two new sRNAs that were previously identified in the Lory lab (Han *et al.*, unpublished), called s3661 and s2790. Notably, the two latter formed chimeras at the 5'-end of *erfA*, suggesting a binding in this region of the mRNA (Figure 5.7). Additionally, both sRNAs have predicted interfaces matching with *erfA* ribosome binding site (Figure 5.8), strongly suggesting a role of these sRNAs in *erfA* translation regulation. In order to confirm the role of these sRNAs in *erfA* regulation, I monitored *ergAB* expression through a *lacZ* fusion, which reflects ErfA activity and abundance, in strains overexpressing or not s2790, s3661, RsmZ or RsmY (as a control). However, no difference in *ergAB* expression was found in any of the strains, suggesting that none of these sRNAs are acting on ErfA translation. This was relatively unexpected and might require more work to be confirmed. Unfortunately, I did not have time to go further into this project and since it seemed that no detected sRNAs had an impact on ErfA, I stopped working on this.

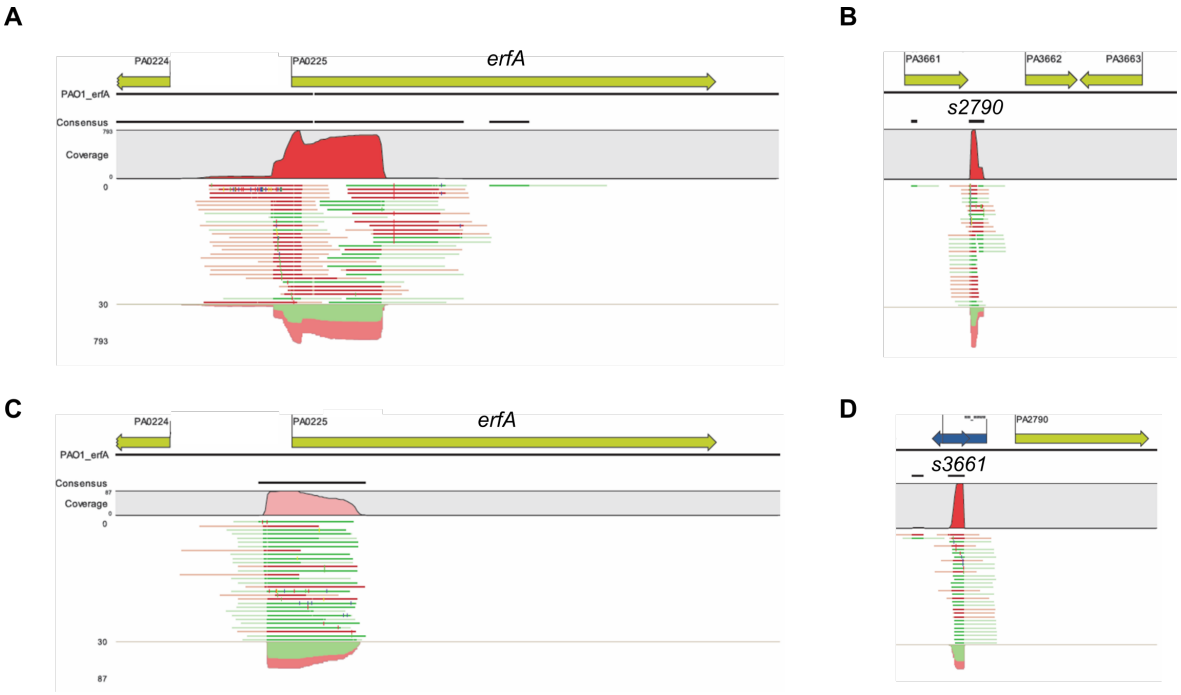


Figure 5.7: GRIL-seq alignement of *erfA* chimeras. Genomic coverage of chimeric reads that align to both *erfA* and another location on the genome for *erfA*-s2790 chimeras on *erfA* (**A**) and s2790 (**B**) and for *erfA*-s3661 chimeras on *erfA* (**C**) and s3661 (**D**).

5.3.2 The *ergB* intragenic s0223 sRNA

The GRIL-seq analysis of s0223 targets resulted in the identification of 33 putative interacting RNAs (Table 5.1). Among them was *rpoS* which is a confirmed target of s0223 (Han *et al.*, unpublished), and other new interesting targets including five genes encoding oxidoreductases (*PA2247*, *PA4613*, *PA2009*, *PA0872* and *PA1759*) and other genes involved in metabolism, suggesting a role of s0223 in stress response (Table 5.1). Interestingly, *pauR*, encoding the PauR TF which shares ErfA architecture (see Section 4.2), was found as interacting with s0223, potentially revealing

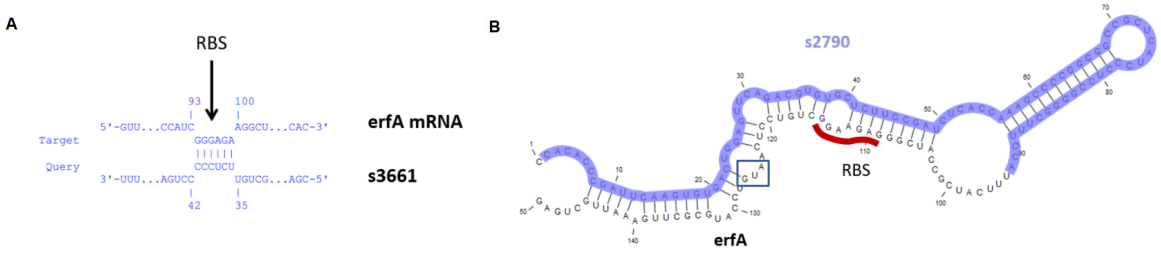


Figure 5.8: Predicted *erfA*-sRNAs interactions. Prediction of interfaces involved in *erfA*-sRNA interactions as predicted by CopraRNA (Wright et al., 2014) for s3661 (**A**) and s2790 (**B**).

Table 5.1: RNAs found in s0223 chimeric reads. All genes to which s0223 chimeric reads aligned with an average max coverage of at least 40 reads are shown, with the coverage values for each duplicates and average.

Rank	Locus_tag	Name	Product	Max coverage		
				rep1	rep2	Average
1	*crcZ*		sRNA	1878	1694	1786
2	*PA2009*	*hmgA*	homogentisate 1,2-dioxygenase	544	649	597
3	*PA2247*	*bkdA1*	2-oxoisovalerate dehydrogenase (alpha subunit)	412	554	483
4	*PA2231*	*pslA*	biofilm initiation	500	380	440
5	*PA3622*	*rpoS*	sigma S transcription factor	300	322	311
6	*PA1092*	*fliC*	flagellin type B	271	271	271
7	*PA2114*		major facilitator superfamily transcription factor	230	175	203
8	*sr003*		sRNA, 5′ utr of PA0049	196	165	181
9	*PA2776*	*pauB3*	FAD-dependent oxidoreductase	123	142	133
10	*PA2760*	*oprQ*	outer membrane protein OprQ	165	97	131
11	*PA4944*	*hfq*	RNA chaperon Hfq	120	124	122
12	*PA3145*	*wbpL*	glycosyltransferase WbpL	107	104	106
13	*PA1288*		long-chain fatty acid transporter	104	102	103
14	*PA4746.1*	*tRNA-met*	tRNA	103	85	94
15	*PA4466*	*ptsH*	phosphocarrier protein HPr ptsH	97	90	94
16	*PA0922.1*	*tRNA-Met*	tRNA	86	83	85
17	*PA2736.1*	*tRNA-pro*	tRNA	78	89	84
18	*PA1338*	*ggt*	gamma-glutamyltranspeptidase precursor	71	67	69
19	*PA4613*	*katB*	catalase	75	59	67
20	*PA4500*	*dppA3*	dipeptide ABC transporter substrate-binding protein DppA3	83	49	66
21	*PA5177*	*yrfG*	probable hydrolase	64	65	65
22	*PA5153*		amino acid ABC transporter periplasmic binding protein	63	65	64
23	*PA0872*	*phhA*	phenylalanine-4-hydroxylase	50	74	62
24	*sr067*		sRNA, 5′ utr of PA2634	53	70	62
25	*PA3027*		AraC family transcriptional regulator	61	57	59
26	*PA1770*	*ppsA*	phosphoenolpyruvate synthase	65	50	58
27	*PA1551*		ferredoxin	55	54	55
28	*PA3190*	*gltB*	sugar ABC transporter substrate-binding protein	58	48	53
29	*PA0866*	*aroP2*	aromatic amino acid transport protein AroP2	68	35	52
30	*PA5301*	*pauR*	transcriptional regulator PauR	53	49	51
31	*PA1126*		hypothetical protein	37	64	51
32	*PA1756*	*cysH*	3′-phosphoadenosine-5′-phosphosulfate reductase	59	35	47
33	*PA2127*	*cgrA*	cupA gene regulator A, CgrA	25	57	41

some interplay regulatory mechanisms between TFs of this family as ErfA activity probably impacts s0223 expression due to its location in *ergB*. Altogether, several known and new targets of s0223 were identified, revealing interesting potential regulatory roles for this sRNA. However, I unfortunately did not have time to continue with this project and try to confirm these interactions and study their phenotypic implications.

5.3.3 Materials & Methods

All experiments were performed as previously described (Han et al., 2016).

Discussion

During my PhD, I have worked on several aspects of regulation in *Pseudomonas aeruginosa* and related species. Most of my work focused on the characterization of transcription factors and their role in transcriptional regulation. I started with the study of the regulation of *exlBA* and identified a new regulator, ErfA, of this operon encoding a major virulence factor. I then found that the main regulatory target of ErfA was not *exlBA* but an operon involved in an undetermined metabolic pathway which is neighboring the *erfA* gene. As 7 other TFs in *P. aeruginosa*, ErfA possesses a XRE DNA-binding domain as well as a cupin domain predicted to act as a signal-sensing regulatory domain. The study of this family of mostly uncharacterized TFs revealed that they all regulate their neighbor genes, which are in most cases predicted to be involved in metabolism. XRE-cupin regulators were found in nearly all *Pseudomonas* species and are mostly found around metabolic genes. To go further into the study of uncharacterized TFs, I also studied the entire family of *P. aerugionsa* DNA-binding response regulators. This ongoing project already yielded a first view of PAO1 TCS regulatory network showing that more than half of all genes might be regulated by RRs.

While studying TFs, I stumbled upon several interesting features of regulatory evolution. Some were very basic, such as differences in either TF or target genes content, which obviously prevent the conservation of direct regulatory interaction. But some differences happened even when both TF and target were conserved. One difference that we thoroughly characterized is the non-conserved regulation of *exlBA* by ErfA between *exlBA*$^+$ *Pseudomonas* species. Indeed, we found that the *exlBA* promoter sequence was highly diverse between *Pseudomonas* species and that the ErfA binding site was found in *exlBA* promoter only in *P. aeruginosa*. Other *exlBA*$^+$ species had different promoter sequences for the operon, which carry different putative TF binding site. This result illustrated a mechanism of evolution where two closely related species that need the same genes but in different conditions adapt their regulation according to their species-specific needs by modification of regulatory sequences inside of gene promoters. Later on, to assess how common

this mechanism is and to obtain a global view of promoter conservation, I studied promoter diversity *in silico* across the *Pseudomonas* genus and found that many genes exhibit different promoters across species. Strikingly, very few genes were found to carry the same TF binding sites in all strains.

The study of DNA-binding proteins and of their regulatory networks inevitably made me curious about regulatory RNA-binding proteins in an attempt to grasp a more global view of regulation mechanisms in bacteria. To that aim and while trying to better understand the regulation of virulence and of ExlBA, I followed some clues suggesting the implication of sRNA-driven post-transcriptional regulation in the *erfA-ergAB* region that could potentially induce indirect *exlBA* regulation. However, I was not able to confirm any post-transcriptional regulation of *erfA* mRNA, and could not continue the other work on the characterization of a sRNA encoded inside of the *ergB* gene due to lack of time. On the other hand, I also studied the major regulatory RBP Hfq. This analysis across *P. aeruginosa* strains and growth phases allowed the complete delineation of the Hfq interactome and of its strain- or -growth specific targets. In this work, I identified hundreds of Hfq targets, including known ones but many new ones. Among these were genes involved in important processes such as virulence and adaptive immunity that were differentially impacted by Hfq between strains.

Across my different projects, I always tried to bring comparison and insights about the conservation and relevance of the observation that I made at higher scales than just a single strain. Indeed, I studied ErfA, all DNA-binding RRs, Hfq and promoter diversity across at least several *P. aeruginosa* strains or even different *Pseudomonas* species. For XRE-cupin regulators, although I only studied them experimentally in one strain, I still brought some insights about their conservation and the conservation of their targets. In my opinion, such comparative approaches are not done often enough. As scientists working in perfectly defined environments with well-characterized easily-cultivable strains which probably reflect only a minute proportion of all that is found in nature, we should at least try to assess diversity and conservation of our results in the small range we have. Especially since we know that there are many differences in genetic content between strains of the same species, which inevitably induce changes in regulatory networks. While this is the exact purpose of several fields of phylogenetics and evolution, the study and characterization of molecular mechanisms often omit that aspect. In the case of evolution or regulatory networks through TF binding sites modification, our

knowledge is still sparse and needs both more global binding sites mapping and conservation analysis. The results presented here illustrate such approaches, which will hopefully guide more studies in the future.

1. *exlBA* regulation

With Alice Berry's thesis work and mine, we have identified two regulators of *exlBA* in *P. aeruginosa*, coding for the major virulence factor of PA7-like strains (Berry et al., 2018; Trouillon et al., 2020a). Both TFs - Vfr and ErfA - bind at two different locations on the promoter of *exlBA*, allowing the independent binding of either the activator (Vfr) or the repressor (ErfA). This relatively common promoter architecture (Mejia-Almonte et al., 2020; Ireland et al., 2020) allows the integration of both regulatory pathways and thus provides a combinational embedding of both signals transduced by the two TFs. While Vfr is known to be activated by cAMP, which is synthesized in respond to different stimuli including calcium depletion, there has been no signal identified yet that modulates ErfA activity. Interestingly, although Vfr was shown to activate *exlBA* expression in certain conditions, the upregulation was much higher in a *erfA* mutant showing that ErfA counteracts the Vfr-dependent positive effect on *exlBA* transcription and that ErfA has to be inactivated for full expression of the operon. While ExlBA is known to be important for virulence, the fact that the *erfA* mutant is much more virulent than the wild-type strain in cytotoxicity assays or during *G. mellonella* larvae infection shows that ErfA inhibition is not alleviated in these conditions. Other XRE-cupin inhibitors have already been shown to detach from DNA upon sensing of a signal molecule, including the putrescine-sensing PuuR TF in *E. coli* (Nemoto et al., 2012a), thus relieving the repression of their target genes, suggesting that ErfA might act similarly. In this case and others (Chou et al., 2013; Kiely et al., 2008), the sensed molecule was always linked to the function of the TF target genes. As *ergAB* is the major target of ErfA, determining its precise function will probably help for the identification of ErfA signal. The characterization of ErfA signal would bring the missing information about *exlBA* regulation and reveal the conditions in which *exlBA*, and thus virulence, is fully active, i.e. where both Vfr and ErfA signals are present.

As shown in different studies from our lab (Basso et al., 2017a; Job et al., 2019; Trouillon et al., 2020a), the ExlBA TPS system is not only present in the human opportunistic pathogen *P. aeruginosa*, but also in other *Pseudomonas* species that are

considered as environmental species. Among them are species known for their capacity to protect plants against fungi or insects and it was shown that ExlBA contributes to insect killing in this context, notably in *P. chlororaphis* or *P. protegens* (Job et al., 2019). However, we recently showed that the regulation of *exlBA* by Vfr and ErfA was specific to *P. aeruginosa* and that other *Pseudomonas* species do not carry their two respective binding sites in the promoter of *exlBA*. Instead, they possess putative binding sites for other TFs but the regulation of *exlBA* has not yet been studied outside of *P. aeruginosa*. While Vfr and ErfA signals might correspond to *P. aeruginosa* needs as a human pathogen to regulate and express ExlBA, it is probable that in other species, which live in different environments and infect different types of hosts, these signals are not suited for *exlBA* regulation. For that reason, evolution has shaped different regulatory sequences in the promoters of *exlBA* in each of these species. To prove that and go further into the study of signal integration and regulatory rewiring between closely related species, it would now be interesting to characterize *exlBA* regulation in these other species.

2. The XRE-cupin regulators

To better understand ErfA and further extend our knowledge on *P. aeruginosa* regulatory network, I also studied 7 other TFs that share ErfA protein architecture. These TFs, called here the XRE-cupin regulators, possess a XRE DNA-binding domain and a cupin domain predicted to act as a signal-sensing domain. The characterization of this family of TFs defined them as local, specialized inhibitors involved in metabolism regulation. However, while most of their targets are indeed predicted to be involved in metabolism, in most cases (for ErfA, PA1359, PA1884 and PA4987, and PA0535 to a lesser extent), there is no experimental confirmation available and the precise metabolic pathway is not precisely determined. This is partly due to the large number of predicted enzyme-coding genes found in *P. aeruginosa*, which can originate from gene duplication, sometimes resulting in very similar genes having divergent functions. This is notably illustrated in the γ-glutamylation pathway, which is partly regulated by two of the XRE-cupin regulators (PA0535 and PauR), and for which there have been six PauA and four PauB homologs identified in *P. aeruginosa*, all with either partly overlapping or different functions (Chou et al., 2013; Luengo and Olivera, 2020). Consequently, the characterization of the targets of the XRE-cupin regulators is still needed in order to fully apprehend the role of these TFs.

One major feature of the XRE-cupin regulators is their putative cupin signal-sensing domain. While this signal-sensing mechanism has been proved and even used for biosensor engineering in *E. coli* allowing the monitoring of putrescine levels (Nemoto et al., 2012a; Chen et al., 2018), it has only been suggested in *P. aeruginosa*, for PauR and PsdR (Kiely et al., 2008; Chou et al., 2013), and further work is still needed to understand it. *In vivo* effects on target genes have been reported in presence of putrescine or dipeptides for PauR and PsdR, respectively, but the direct interactions between signal molecules and TFs still need to be proved. Indeed, it still cannot be excluded that while there is an effect on the TF, the actual interaction with the TF is mediated by another molecule such as a byproduct resulting from the presence of the added molecule. Both *in vivo* and *in vitro* analyses are now needed to identify and confirm these potential metabolite-TF interactions, which will probably depend on the function of the target genes. Additionally, while the cupin domain is predicted as a signal-sensing domain, this has not been experimentally confirmed and the actual interaction interface is not identified. The structure-function analysis of this uncharacterized family of potential signal-sensing domains, which could be guided by the only resolved structure of a XRE-cupin TF (unpublished, PDB ID: 1Y9Q), is among the next steps needed to understand their mechanisms. Notably, there are several relatively conserved amino acids in the cupin domains of these regulators, including a Glycine-Aspartic acid motif found in the cupin domains of all 8 *P. aeruginosa* XRE-cupin regulators and several inward-oriented phenylalanine residues found inside of the putative binding site of the domain. The investigation of the roles of specific residues and mechanism of signal sensing by the cupin domain would bring valuable information concerning the regulatory pathways relying on these regulators.

3. The response regulators

As only a small proportion of *P. aeruginosa* transcriptional regulatory network is characterized, we aimed at partly filling this knowledge gap by studying all DNA-binding RRs. To the best of my knowledge, this ongoing project already represents the largest bacterial TF binding dataset ever reported in a single study, with 312 DAP-seq experiments performed and probably about 50-100 more to do, and will bring valuable information on this pathogen. This project is still ongoing and much work is still needed to refine data analysis and obtain more binding results which will be needed to draw any conclusions. However, we can already see that our approach was largely successful and that recent methodological development, such

as DAP-seq, are now making possible such large-scale analyses. Indeed, the first, and still largest, DAP-seq project reported about 2,300 DAP-seq experiments carried on 1,812 *A. thaliana* TFs (O'Malley et al., 2016). This kind of scales, completely impossible in the past, will become more and more common and I believe that all most-studied organisms will have all their TFs characterized in similar ways in the near future, including *P. aeruginosa* and other bacterial pathogens. In this first DAP-seq report, they achieved DNA motif determination for 529 TFs, or about 29% of the tested TFs (O'Malley et al., 2016). In our project, and only with the first batch of experiments (as some have to be redone due to technical issues), we were able to determine DNA motifs for more than 74% of the tested TFs, illustrating the success of our approach. This difference can be explained in several ways: (i) we tailored our protocol to only one TF family while their dataset represented 80 TF families, which could potentially each need different conditions to be active, (ii) bacterial genomes are smaller and thus require smaller sequencing capacities to obtain sufficient depth and (iii) data quality is often inversely proportional to scale of study (Cai and Zhu, 2015).

There is still a relatively large amount of work to do for this project, as we only obtained the first results recently. While I delineated a first global view of the binding landscape of RRs in PAO1, more work is now required to compare these new results to what is already known for the ones that have been studied. The manual comparison of known binding sites and motifs already suggests that our data recover known interactions quite well, but we still need to comprehensively gather all known DNA motifs for all RRs that have been sufficiently studied and properly compare them with what we obtained.

In my different projects, I have experimentally illustrated several examples of regulatory differences between *P. aeruginosa* strains, involving either a TF (ErfA) or a RBP (Hfq), using two or three of the reference strains PAO1, PA14 and IHMA87 which represent the three main *P. aeruginosa* phylogenetic lineages. In each cases, the phenotypic implication of the regulatory differences was major, such as with virulence regulation, and reflects divergent evolutionary pathways between the studied strains or species. To assess this at a larger scale, we decided to study all RRs in all three representative strains. As many RRs are implicated in major regulatory pathways involved in all important biological processes of this pathogen, we expect that this approach will reveal new, and explain known phenotypic strain- or lineage-specific characteristics, bringing a global view at the

regulatory network of *P. aeruginosa*, as a species. First, among the core set of 48 RRs conserved in all three strains, I anticipate that some targets will not be directly regulated by the same RRs in all three strains. Since we will have a DNA motifs for most of them, we will also be able to explain these differences by looking at motif conservation in the corresponding targets promoters and then try to associate these differences with known phenotypic differences between the strains. Other differences could occur from lack or presence of different target or HK genes, such as with TtsR, which is an orphan RR in PAO1 and PA14 but is associated with TtsS, its cognate HK, and a cluster of neighbor target genes only present in PA7-like strains (Varga et al., 2015). Additionally, 7 RRs are present in only one or two of the three tested strains which also provides opportunities for regulatory differences.

4. Promoter diversity

To better understand regulatory rewiring driven by promoter evolution, I set up an *in silico* analysis for the determination of promoter diversity in all ortholog groups across the *Pseudomonas* genus. This approach suggested that most genes are not regulated in the same way in all species that possess them, revealing high intra-genus promoter diversity levels. This is also an ongoing project that still requires more work to improve the analysis methods and go into more details. Notably, I want to redo the analysis with different length of promoter regions used in order to assess its effect on the results, and to increase the number of TFs analyzed with experimental data as for now there are only 4. Also, even though preliminary results showed that there was no enrichment of functional groups in any of the identified clusters, I plan on going into more details on the distribution of functional groups to assess if some functions tend to be more or less prone to regulatory diversification. As it is, this analysis already shows large diversity in upstream promoter regions of ortholog genes, illustrated with examples of different levels of conservation of experimentally-derived TF binding sites.

In order to fully apprehend this mechanism in bacteria, we need more comparative analyses such as this one and most importantly more TF binding data. The size comparison of TF binding databases between eukaryotes, with for instance JASPAR and CIS-BP which each regroup more than 1,500 and 4,500 TF motifs, respectively (Fornes et al., 2020; Weirauch et al., 2014), and prokaryotes, with CollecTF and Prodoric with 253 and 160 motifs (Kılıç et al., 2014; Münch et al., 2003), quickly illustrates the vast gap between the two and how little we still know

about bacterial TFs. Increasing our experimental evidences will inevitably allow a better characterization of regulatory diversification.

5. The Hfq regulatory RBP

The characterization of Hfq by RIP-seq in the three *P. aeruginosa* reference strains at different growth phases allowed the complete delienation of Hfq core and accessory interactomes. In this work, I show that Hfq interacts with 13 to 21% of all annotated mRNAs and sRNAs in *P. aeruginosa*, spanning nearly all major biological processes. A core set of 427 RNAs (mRNAs or sRNAs) was found bound to Hfq in all three strains, including 109 that were also bound to Hfq during all three tested growth phases. This Hfq core interactome is enriched in mRNAs encoding TCSs and TFs, illustrating the role of Hfq as a master regulator. This result also show a massive interplay between transcriptional and post-transcriptional regulation which should be investigated in more details.

An interesting newly discovered interaction was that of Hfq with the *vfr* mRNA. Indeed, we showed that Vfr has different regulatory targets between PA7-like strains like IHMA87 (*exlBA*) and the *T3SS*$^{+}$ strains (T3SS-related genes). Consequently, a potential direct regulation of Vfr by Hfq might reveal a new common regulatory pathway between the T3SS and ExlBA. Interestingly, Vfr is known to be indirectly regulated by RsmA, which has recently been shown to be regulated by the sRNA 179 and Hfq (Janssen et al., 2020); this reveals in addition to the direct potential post-transcriptional regulation of Vfr that we found, numerous different levels of regulation of this master TF and the complexity of regulatory networks. Apart from common targets, numerous interesting regulatory features were found among strain- or growth-specific targets. These include many T3SS-related mRNAs as well as CRISPR RNAs. Here again, these new interactions could have major implications in virulence-related phenotypes. While this analysis did identify numerous new potential regulatory interactions, much work is now needed to confirm and study them. Notably, while RIP-seq identifies all RNAs bound to a RBP, it does not identify interacting RNAs in the case of Hfq-mediated sRNA-mRNA interactions for instance. Consequently, the need is now to identify the sRNAs involved in the regulation of all mRNAs found enriched here in order to understand these interactions.

Appendices

5.4 Conference presentations

Flash presentation & poster

Gordon Conference on Microbial Transcription, USA. 08.2019

Genetic rearrangement drives diversification of virulence regulation across *Pseudomonas* species

Julian Trouillon[1]*, Erwin Sentausa[1†], Michel Ragno[1], Mylène Robert-Genthon[1], Stephen Lory[2], Ina Attree[1], Sylvie Elsen[1]

[1]Université Grenoble Alpes, CNRS ERL5261, CEA BIG-BCI, INSERM UMR1036, Grenoble, France

[2]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, USA

*julian.trouillon@cea.fr

[†]Present address: Evotec ID (Lyon) SAS, Marcy l'Étoile, France

Abstract

Tight and coordinate regulation of virulence determinants is essential for bacterial biology and involves dynamic transcriptional regulatory networks shaped during evolution. *Pseudomonas aeruginosa* taxonomic outliers constitute an emerging group of infectious bacteria endowed with ExlBA, a two-partner pore-forming toxin secretion system responsible for their pathogenicity. Here, by screening a transposon mutant library, we identify a Cro-like repressor named ErfA, maintaining *exlBA* expression at basal levels. In addition to *exlBA*, RNA-seq and ChIP-seq revealed that ErfA directly represses an operon adjacent to its gene, *ergAB,* that encodes two enzymes with functions unrelated to ExlBA. By characterizing the pan-cistrome of ErfA in several *exlBA*$^+$ *Pseudomonas* species using DAP-seq, we show that *ergAB* is the unique conserved target of ErfA. The analysis of 451 *exlBA* promoter sequences from all *exlBA*$^+$ species unraveled a wide variety of promoters for *exlBA* across *Pseudomonas*. Indeed, only in *P. aeruginosa* is *exlBA* under the direct control of the two regulators, ErfA and the global virulence regulator Vfr. In these strains, the activation of *exlBA* transcription by Vfr in a cAMP-responsive manner is hindered by ErfA activity. Overall, we propose that the emergence of new regulatory traits specific to *P. aeruginosa exlBA* is an example of plasticity of virulence regulatory networks that evolved to provide an adapted response in a given niche.

Oral presentation

BactoGre Congress, France. 04.2019

Species-specific recruitment of transcription factors dictates toxin expression

Julian Trouillon[1]*, Erwin Sentausa[1], Michel Ragno[1], Mylène Robert-Genthon[1], Stephen Lory[2], Ina Attree[1], Sylvie Elsen[1]*

[1]Université Grenoble Alpes, CNRS ERL5261, CEA-IRIG-BCI, INSERM UMR1036, Grenoble, France
[2]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, USA

Present Address: Erwin Sentausa, Evotec ID (Lyon) SAS, Marcy l'Étoile, France

ABSTRACT

Tight and coordinate regulation of virulence determinants is essential for bacterial biology and involves dynamic shaping of transcriptional regulatory networks during evolution. The horizontally transferred two-partner secretion system ExlB-ExlA is instrumental in the virulence of different *Pseudomonas* species, ranging from soil- and plant-dwelling biocontrol agents to the major human pathogen *P. aeruginosa*. Here, we identify a Cro/CI-like repressor, named ErfA, which together with Vfr, a CRP-like activator, controls *exlBA* expression in *P. aeruginosa*. The characterization of ErfA regulon across *P. aeruginosa* subfamilies revealed a second conserved target, the *ergAB* operon, with functions unrelated to virulence. To gain insights into this functional dichotomy, we defined the pan-regulon of ErfA in several *Pseudomonas* species and found *ergAB* as the sole conserved target of ErfA. The analysis of 446 *exlBA* promoter sequences from all *exlBA*$^+$ genomes revealed a wide variety of regulatory sequences, as ErfA and Vfr binding sites were found to have evolved specifically in *P. aeruginosa* and nearly each species carries different regulatory sequences for this operon. We propose that the emergence of different regulatory *cis*-elements in the promoters of horizontally transferred genes is an example of plasticity of regulatory networks evolving to provide an adapted response in each individual niche.

Oral presentation

Pseudomonas French Congress, France. 09.2018

Regulation of the Two-Partner Secretion system ExlBA of *Pseudomonas aeruginosa*

J. Trouillon, A. Berry, I. Attrée & S. Elsen.

Biology of Cancer and Infection, Bacterial Pathogenesis & Cellular Responses – CEA Grenoble

ExlBA is a newly discovered Two-Partner Secretion system responsible for the virulence of PA7-like *Pseudomonas aeruginosa* strains. While the Type 3 Secretion System (T3SS) is known as the major virulence factor of *P. aeruginosa*, it is absent in *exlBA*$^+$ strains and both secretion systems are mutually exclusive. Here we investigated the regulation of ExlBA, which is used for host cell membrane disruption through the pore-forming activity of the secreted toxin ExlA. We first demonstrated that the *exlBA* genes are under the control of the cAMP/Vfr regulatory pathway. We showed that upon interaction with the secondary messenger cAMP, Vfr binds directly to *exlBA* promoter (P*exlBA*) with high affinity and induces transcription. In agreement with its known function as the main source of cAMP, the adenylate cyclase CyaB was found to be required for Vfr-dependent upregulation of *exlBA*. Furthermore, a library of 100,000 transposon mutants was built in a strain carrying an *in situ exlA::lacZ* transcriptional fusion and screened to identify new regulators of *exlBA*. This allowed the identification of an uncharacterized transcription factor as a strong inhibitor of *exlBA* expression that we renamed ErfA (*exlBA* regulatory factor A). We showed that ErfA binds to P*exlBA* slightly downstream of its transcription start site and that its binding accounts for a 40-fold decrease of *exlBA* expression in comparison to a Δ*erfA* mutant. Consequently, the Δ*erfA* mutant secretes higher quantities of ExlA and was found to have increased toxicity towards epithelial cells as well as in a *Galleria mellonella* model of infection. To go further with the investigation of ErfA and as its gene is highly conserved across *P. aeruginosa* strains, including in *exlBA*$^-$ strains, we determined its regulon by RNA-seq and ChIP-seq in both *T3SS*$^+$ and *exlBA*$^+$ strains. This approach led to the identification of other targets of ErfA and will improve our understanding of its function in *P. aeruginosa*.

5.5 Posters

Gordon Conference on Microbial Transcription, USA. 08.2019

Exolysin regulation showcases evolution of virulence regulatory networks across *Pseudomonas* species

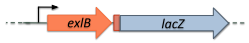
Julian Trouillon[1]*, Erwin Sentausa[1†], Michel Ragno[1], Mylène Robert-Genthon[1], Stephen Lory[2], Ina Attrée[1], Sylvie Elsen[1]

[1] Bacterial Pathogenicity and Cellular Responses, ERL5261 CNRS, U1036 INSERM, CEA, University of Grenoble Alpes, Grenoble, France.
[2] Department of Microbiology, Harvard Medical School, Boston, Massachusetts, USA. [†] Present address: Evotec ID, Marcy L'Étoile, France.

Introduction

Exolysin (ExlA) is a 172 kDa pore-forming toxin secreted by *Pseudomonas aeruginosa* strains lacking the Type III Secretion System (T3SS). ExlA is part of a Two-Partner Secretion System along with its cognate outer membrane transporter ExlB. We recently found that *exlBA* is under the positive regulation of the global transcriptional regulator Vfr (Berry *et al.*, 2018). However, even when fully activated by Vfr, *exlBA* expression levels remain extremely low. We thus hypothesized that a transcriptional repressor might be involved in *exlBA* regulation and looked for it by screening a transposon library for mutants exhibiting *exlBA* upregulation.

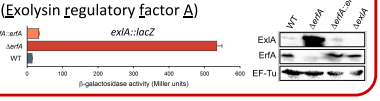
1 Screen of 100.000 transposon mutants carrying a chromosomal *exlBA::lacZ* transcriptional fusion



➡ Identification of an uncharacterized repressor

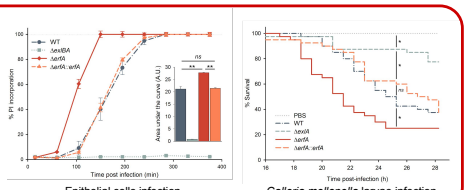
2 Role confirmation of the identified repressor, named **ErfA** (Exolysin regulatory factor A)

➡ **ErfA strongly represses *exlBA* transcription and ExlBA synthesis**



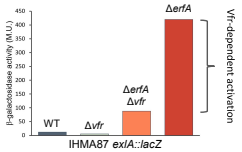
3 Effect on virulence

➡ Increased ExlBA synthesis in the *erfA* mutant induces higher cytotoxicity towards epithelial cells as well as virulence during infection of *G. mellonella* larvae



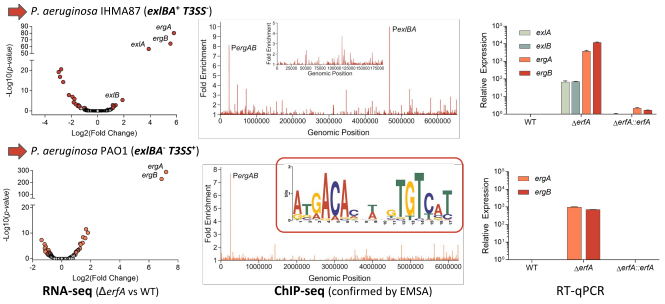
Epithelial cells infection — *Galleria mellonella* larvae infection

4 Vfr activation of *exlBA* is hindered by ErfA



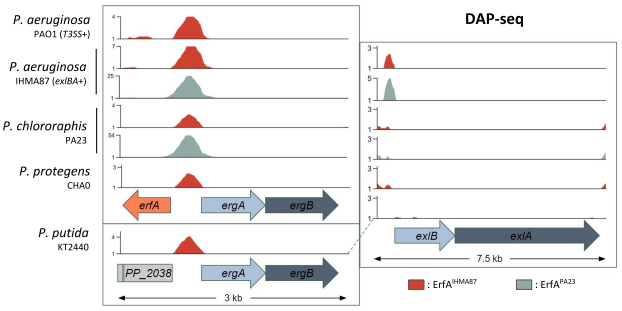
5 ErfA regulon consists of *ergAB*, in addition to *exlBA* in *T3SS*- strains.

ergAB encodes two putative enzymes with functions unrelated to ExlBA



RNA-seq (Δ*erfA* vs WT) — **ChIP-seq** (confirmed by EMSA) — RT-qPCR

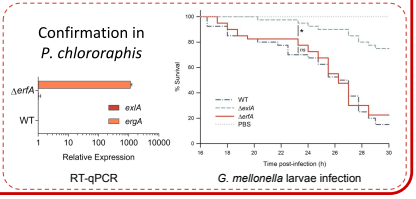
6 Determination of ErfA regulon in 4 different *exlBA*+ *Pseudomonas* species by DAP-seq



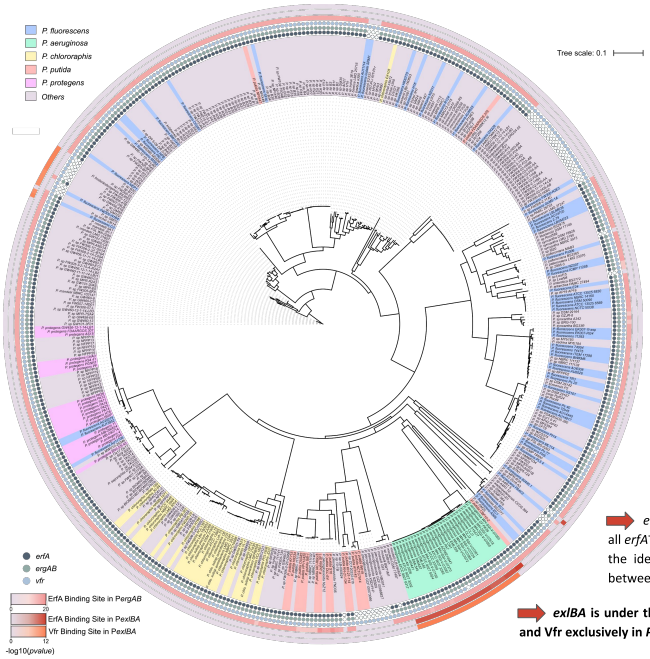
➡ ***exlBA* is not regulated by ErfA in these other species**

➡ *ergAB* is the only target of ErfA across the tested *Pseudomonas* species

Confirmation in *P. chlororaphis*



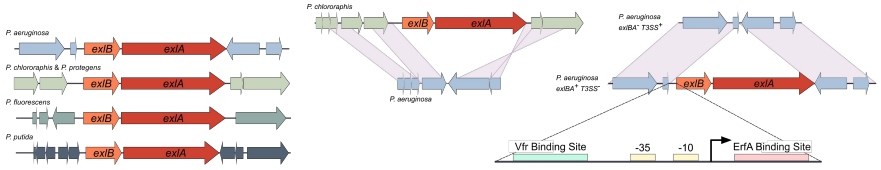
7 Genomic comparative analysis of 451 *exlBA* promoter sequences from all *exlBA*+ *Pseudomonas* strains reveals **diversity in cis-regulatory elements**



➡ *erfA* regulates *ergAB* in all *erfA*+ strains, strengthening the idea of a functional link between them

➡ ***exlBA* is under the regulation of ErfA and Vfr exclusively in *P. aeruginosa***

8 Both *exlBA* promoter sequence and genetic environment vary between *Pseudomonas* species



➡ **The different regulation is due to the presence of a *P. aeruginosa*-specific sequence upstream of *exlBA* containing ErfA and Vfr binding sites**

➡ Many different *exlBA* promoter sequences are found across *Pseudomonas* species, suggesting as many different regulations

Conclusions

➡ **ErfA locks *exlBA* expression**, and counteracts Vfr cAMP-dependent activation by binding downstream of *exlBA* Transcription Start Site.

➡ ErfA also regulates *ergAB*, encoding two putative enzymes with functions unrelated to *exlBA*. *ergAB* forms the **core regulon of ErfA**.

➡ ***exlBA* grafted on ErfA and Vfr regulons specifically in *P. aeruginosa*** by acquisition of the corresponding binding sites in its promoter region. In other species, the variety of promoters suggest several different *exlBA* regulatory networks, each one being probably adapted to give an advantage in a given niche.

* email: julian.trouillon@cea.fr
@JulianTrouillon

Université Grenoble Alpes — CNRS — CEA — Inserm — HARVARD MEDICAL SCHOOL

Bibliography

Afgan, Enis et al. (2018). "The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update". In: *Nucleic acids research* 46.W1, W537–W544.

Almblad, Henrik et al. (2015). "The cyclic AMP-Vfr signaling pathway in Pseudomonas aeruginosa is inhibited by cyclic di-GMP". In: *Journal of bacteriology* 197.13, pp. 2190–2200.

Almblad, Henrik et al. (2019). "High levels of cAMP inhibit Pseudomonas aeruginosa biofilm formation through reduction of the c-di-GMP content". In: *Microbiology* 165.3, pp. 324–333.

Anderson, Anne J and Young Cheol Kim (2018). "Biopesticides produced by plant-probiotic Pseudomonas chlororaphis isolates". In: *Crop Protection* 105, pp. 62–69.

Arora, Meenakshi (2013). "Cell culture media: a review". In: *Mater methods* 3.175, p. 24.

Arzanlou, Mohsen, Wern Chern Chai, and Henrietta Venter (2017). "Intrinsic, adaptive and acquired antimicrobial resistance in Gram-negative bacteria". In: *Essays in biochemistry* 61.1, pp. 49–59.

Attaiech, Laetitia et al. (2016). "Silencing of natural transformation by an RNA chaperone and a multitarget small RNA". In: *Proceedings of the National Academy of Sciences* 113.31, pp. 8813–8818.

Babitzke, Paul and Tony Romeo (2007). "CsrB sRNA family: sequestration of RNA-binding regulatory proteins". In: *Current opinion in microbiology* 10.2, pp. 156–163.

Backofen, Rolf and Wolfgang R Hess (2010). "Computational prediction of sRNAs and their targets in bacteria". In: *RNA biology* 7.1, pp. 33–42.

Bai, Fang et al. (2018). "Bacterial type III secretion system as a protein delivery tool for a broad range of biomedical applications". In: *Biotechnology advances* 36.2, pp. 482–493.

Balleza, Enrique et al. (2008). "Regulation by transcription factors in bacteria: beyond description". In: *FEMS microbiology reviews* 33.1, pp. 133–151.

Bandyra, Katarzyna J et al. (2012). "The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E". In: *Molecular cell* 47.6, pp. 943–953.

Barken, Kim B et al. (2008). "Roles of type IV pili, flagellum-mediated motility and extracellular DNA in the formation of mature multicellular structures in Pseudomonas aeruginosa biofilms". In: *Environmental microbiology* 10.9, pp. 2331–2343.

Barr, Helen L et al. (2015). "Pseudomonas aeruginosa quorum sensing molecules correlate with clinical status in cystic fibrosis". In: *European Respiratory Journal* 46.4, pp. 1046–1054.

Bartlett, Anna et al. (2017). "Mapping genome-wide transcription-factor binding sites using DAP-seq". In: *Nature protocols* 12.8, p. 1659.

Bassetti, Matteo et al. (2018). "How to manage Pseudomonas aeruginosa infections". In: *Drugs in context* 7.

Basso, Pauline et al. (2017a). "Multiple Pseudomonas species secrete exolysin-like toxins and provoke Caspase-1-dependent macrophage death". In: *Environmental microbiology* 19.10, pp. 4045–4064.

Basso, Pauline et al. (2017b). "Pseudomonas aeruginosa pore-forming exolysin and type IV pili cooperate to induce host cell lysis". In: *MBio* 8.1.

Batrich, Mary et al. (2019). "Pseudomonas diversity within urban freshwaters". In: *Frontiers in microbiology* 10, p. 195.

Beckert, Urike et al. (2014). "ExoY from Pseudomonas aeruginosa is a nucleotidyl cyclase with preference for cGMP and cUMP formation". In: *Biochemical and biophysical research communications* 450.1, pp. 870–874.

Belyy, Alexander et al. (2018). "ExoY, an actin-activated nucleotidyl cyclase toxin from P. aeruginosa: A minireview". In: *Toxicon* 149, pp. 65–71.

Berry, Alice (2019). "Premiers mécanismes de régulation d'exlBA, le facteur de virulence des souches de Pseudomonas aeruginosa de type PA7: First regulatory mechanisms of exlBA, virulence factor of Pseudomonas aeruginosa PA7-like strains". PhD thesis. Université Grenoble Alpes.

Berry, Alice et al. (2018). "cAMP and Vfr control exolysin expression and cytotoxicity of Pseudomonas aeruginosa taxonomic outliers". In: *Journal of bacteriology* 200.12.

Bertrand, Quentin et al. (2020). "Exolysin (ExlA) from Pseudomonas aeruginosa Punctures Holes into Target Membranes Using a Molten Globule Domain". In: *Journal of Molecular Biology*.

Bhagirath, Anjali Y et al. (2017). "Characterization of the direct interaction between hybrid sensor kinases PA1611 and RetS that controls biofilm formation and the

type III secretion system in Pseudomonas aeruginosa". In: *ACS infectious diseases* 3.2, pp. 162–175.

Bielecki, Piotr et al. (2015). "Cross talk between the response regulators PhoB and TctD allows for the integration of diverse environmental signals in Pseudomonas aeruginosa". In: *Nucleic acids research* 43.13, pp. 6413–6425.

Blair, Jessica MA et al. (2015). "Molecular mechanisms of antibiotic resistance". In: *Nature reviews microbiology* 13.1, pp. 42–51.

Bleves, Sophie et al. (2010). "Protein secretion systems in Pseudomonas aeruginosa: a wealth of pathogenic weapons". In: *International Journal of Medical Microbiology* 300.8, pp. 534–543.

Bodey, Gerald P et al. (1983). "Infections caused by Pseudomonas aeruginosa". In: *Reviews of infectious diseases* 5.2, pp. 279–313.

Bordi, Christophe et al. (2010). "Regulatory RNAs and the HptB/RetS signalling pathways fine-tune Pseudomonas aeruginosa pathogenesis". In: *Molecular microbiology* 76.6, pp. 1427–1443.

Botelho, João, Filipa Grosso, and Luísa Peixe (2019). "Antibiotic resistance in Pseudomonas aeruginosa–Mechanisms, epidemiology and evolution". In: *Drug Resistance Updates* 44, p. 100640.

Bouillet, Sophie et al. (2019). "Connected partner-switches control the life style of Pseudomonas aeruginosa through RpoS regulation". In: *Scientific reports* 9.1, pp. 1–11.

Bouillot, Stéphanie et al. (2020). "Inflammasome activation by Pseudomonas aeruginosa's ExlA pore-forming toxin is detrimental for the host". In: *Cellular Microbiology*, e13251.

Bourret, Robert B (2010). "Receiver domain structure and function in response regulator proteins". In: *Current opinion in microbiology* 13.2, pp. 142–149.

Bouvier, Marie et al. (2008). "Small RNA binding to 5 mRNA coding region inhibits translational initiation". In: *Molecular cell* 32.6, pp. 827–837.

Brantl, Sabine and Reinhold Brückner (2014). "Small regulatory RNAs from low-GC Gram-positive bacteria". In: *RNA biology* 11.5, pp. 443–456.

Breidenstein, Elena BM, César de la Fuente-Núñez, and Robert EW Hancock (2011). "Pseudomonas aeruginosa: all roads lead to resistance". In: *Trends in microbiology* 19.8, pp. 419–426.

Brencic, Anja and Stephen Lory (2009). "Determination of the regulon and identification of novel mRNA targets of Pseudomonas aeruginosa RsmA". In: *Molecular microbiology* 72.3, pp. 612–632.

Brencic, Anja et al. (2009). "The GacS/GacA signal transduction system of Pseudomonas aeruginosa acts exclusively through its control over the transcription of the RsmY and RsmZ regulatory small RNAs". In: *Molecular microbiology* 73.3, pp. 434–445.

Brinkman, Fiona SL, Robert EW Hancock, and C Kendal Stover (2000). "Sequencing solution: use volunteer annotators organized via Internet". In: *Nature* 406.6799, pp. 933–933.

Britten, Roy J and Eric H Davidson (1969). "Gene regulation for higher cells: a theory". In: *Science* 165.3891, pp. 349–357.

Broder, Ursula N, Tina Jaeger, and Urs Jenal (2016). "LadS is a calcium-responsive kinase that induces acute-to-chronic virulence switch in Pseudomonas aeruginosa". In: *Nature microbiology* 2.1, pp. 1–11.

Brodsky, Igor E et al. (2010). "A Yersinia effector protein promotes virulence by preventing inflammasome recognition of the type III secretion system". In: *Cell host & microbe* 7.5, pp. 376–387.

Browning, Douglas F and Stephen JW Busby (2004). "The regulation of bacterial transcription initiation". In: *Nature Reviews Microbiology* 2.1, pp. 57–65.

— (2016). "Local and global regulation of transcription initiation in bacteria". In: *Nature Reviews Microbiology* 14.10, p. 638.

Browning, Douglas F, Matej Butala, and Stephen JW Busby (2019). "Bacterial transcription factors: regulation by pick "N" mix". In: *Journal of molecular biology* 431.20, pp. 4067–4077.

Browning, Douglas F, David C Grainger, and Stephen JW Busby (2010). "Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression". In: *Current opinion in microbiology* 13.6, pp. 773–780.

Bull, Carolee Theresa et al. (2010). "Comprehensive list of names of plant pathogenic bacteria, 1980-2007". In: *Journal of plant pathology*, pp. 551–592.

Cabot, Gabriel et al. (2011). "Overexpression of AmpC and efflux pumps in Pseudomonas aeruginosa isolates from bloodstream infections: prevalence and impact on resistance in a Spanish multicenter study". In: *Antimicrobial agents and chemotherapy* 55.5, pp. 1906–1911.

Cadoret, Frédéric et al. (2014). "Txc, a new type II secretion system of Pseudomonas aeruginosa strain PA7, is regulated by the TtsS/TtsR two-component system and directs specific secretion of the CbpE chitin-binding protein". In: *Journal of bacteriology* 196.13, pp. 2376–2386.

Cai, Li and Yangyong Zhu (2015). "The challenges of data quality and data quality assessment in the big data era". In: *Data science journal* 14.

Carlier, Evelin et al. (2008). "Improvement of growth, under field conditions, of wheat inoculated with Pseudomonas chlororaphis subsp. aurantiaca SR1". In: *World Journal of Microbiology and Biotechnology* 24.11, pp. 2653–2658.

Carroll, Sean B (2008). "Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution". In: *Cell* 134.1, pp. 25–36.

Cases, Ildefonso, Victor De Lorenzo, and Christos A Ouzounis (2003). "Transcription regulation and environmental adaptation in bacteria". In: *Trends in microbiology* 11.6, pp. 248–253.

Casilag, Fiordiligie et al. (2016). "The LasB elastase of Pseudomonas aeruginosa acts in concert with alkaline protease AprA to prevent flagellin-mediated immune recognition". In: *Infection and immunity* 84.1, pp. 162–171.

Casino, Patricia, Vicente Rubio, and Alberto Marina (2010). "The mechanism of signal transduction by two-component systems". In: *Current opinion in structural biology* 20.6, pp. 763–771.

Castellanos, Milagros, Nivin Mothi, and Victor Muñoz (2020). "Eukaryotic transcription factors can track and control their target genes using DNA antennas". In: *Nature communications* 11.1, pp. 1–13.

Cézard, C, N Farvacques, and P Sonnet (2015). "Chemistry and biology of pyoverdines, Pseudomonas primary siderophores". In: *Current medicinal chemistry* 22.2, pp. 165–186.

Chakraborty, Smarajit et al. (2017). "Non-canonical activation of OmpR drives acid and osmotic stress responses in single bacterial cells". In: *Nature communications* 8.1, pp. 1–14.

Chandrangsu, Pete, Christopher Rensing, and John D Helmann (2017). "Metal homeostasis and resistance in bacteria". In: *Nature Reviews Microbiology* 15.6, p. 338.

Chao, Yanjie and Jörg Vogel (2010). "The role of Hfq in bacterial pathogens". In: *Current opinion in microbiology* 13.1, pp. 24–33.

Chen, I-Chen Kimberly, Gregory J Velicer, and Yuen-Tsu Nicco Yu (2017). "Divergence of functional effects among bacterial sRNA paralogs". In: *BMC evolutionary biology* 17.1, p. 199.

Chen, Xue-Feng et al. (2018). "Engineering tunable biosensors for monitoring putrescine in Escherichia coli". In: *Biotechnology and bioengineering* 115.4, pp. 1014–1027.

Chevalier, Sylvie et al. (2017). "Structure, function and regulation of Pseudomonas aeruginosa porins". In: *FEMS Microbiology Reviews* 41.5, pp. 698–722.

Chihara, Kotaro et al. (2019). "Conditional Hfq Association with Small Noncoding RNAs in Pseudomonas aeruginosa Revealed through Comparative UV Cross-Linking Immunoprecipitation Followed by High-Throughput Sequencing". In: *Msystems* 4.6.

Choi, Yusang et al. (2011). "Growth phase-differential quorum sensing regulation of anthranilate metabolism in Pseudomonas aeruginosa". In: *Molecules and cells* 32.1, pp. 57–65.

Chou, Han Ting et al. (2013). "Molecular characterization of PauR and its role in control of putrescine and cadaverine catabolism through the γ-glutamylation pathway in Pseudomonas aeruginosa PAO1". In: *Journal of bacteriology* 195.17, pp. 3906–3913.

Chubukov, Victor et al. (2014). "Coordination of microbial metabolism". In: *Nature Reviews Microbiology* 12.5, pp. 327–340.

Conlon, Brian P et al. (2016). "Persister formation in Staphylococcus aureus is associated with ATP depletion". In: *Nature microbiology* 1.5, pp. 1–7.

Connolly, James PR, Nicky O'Boyle, and Andrew J Roe (2020). "Widespread strain-specific distinctions in chromosomal binding dynamics of a highly conserved Escherichia coli transcription factor". In: *Mbio* 11.3.

Conrad, Jacinta C et al. (2011). "Flagella and pili-mediated near-surface single-cell motility mechanisms in P. aeruginosa". In: *Biophysical journal* 100.7, pp. 1608–1616.

Cornelis, Pierre (2020). "Putting an end to the Pseudomonas aeruginosa IQS controversy". In: *MicrobiologyOpen* 9.2, e962.

Crack, Jason, Jeffrey Green, and Andrew J Thomson (2004). "Mechanism of oxygen sensing by the bacterial transcription factor fumarate-nitrate reduction (FNR)". In: *Journal of Biological Chemistry* 279.10, pp. 9278–9286.

Darst, Seth A (2001). "Bacterial RNA polymerase". In: *Current opinion in structural biology* 11.2, pp. 155–162.

David, Baliah V, Govindan Chandrasehar, and Pamila N Selvam (2018). "Pseudomonas fluorescens: a plant-growth-promoting rhizobacterium (PGPR) with potential role in biocontrol of pests of crops". In: *Crop improvement through microbial biotechnology*. Elsevier, pp. 221–243.

Davis, Pamela B (2006). "Cystic fibrosis since 1938". In: *American journal of respiratory and critical care medicine* 173.5, pp. 475–482.

De, Nabanita et al. (2009). "Determinants for the activation and autoinhibition of the diguanylate cyclase response regulator WspR". In: *Journal of molecular biology* 393.3, pp. 619–633.

De Bentzmann, Sophie and Patrick Plésiat (2011). "The Pseudomonas aeruginosa opportunistic pathogen and human infections". In: *Environmental microbiology* 13.7, pp. 1655–1665.

Deutscher, Josef (2008). "The mechanisms of carbon catabolite repression in bacteria". In: *Current opinion in microbiology* 11.2, pp. 87–93.

Di Iulio, Julia et al. (2018). "The human noncoding genome defined by genetic diversity". In: *Nature genetics* 50.3, pp. 333–337.

Diggle, Stephen P and Marvin Whiteley (2020). "Microbe Profile: Pseudomonas aeruginosa: opportunistic pathogen and lab rat". In: *Microbiology* 166.1, pp. 30–33.

Dillon, Shane C and Charles J Dorman (2010). "Bacterial nucleoid-associated proteins, nucleoid structure and gene expression". In: *Nature Reviews Microbiology* 8.3, pp. 185–195.

Dolan, Stephen K et al. (2020). "Transcriptional regulation of central carbon metabolism in Pseudomonas aeruginosa". In: *Microbial biotechnology* 13.1, pp. 285–289.

Driscoll, James A, Steven L Brody, and Marin H Kollef (2007). "The epidemiology, pathogenesis and treatment of Pseudomonas aeruginosa infections". In: *Drugs* 67.3, pp. 351–368.

Du, Dijun et al. (2018). "Multidrug efflux pumps: structure, function and regulation". In: *Nature reviews microbiology* 16.9, pp. 523–539.

Elías-Arnanz, Montserrat, S Padmanabhan, and Francisco J Murillo (2011). "Light-dependent gene regulation in nonphototrophic bacteria". In: *Current opinion in microbiology* 14.2, pp. 128–135.

Eljounaidi, Kaouthar, Seung Kyu Lee, and Hanhong Bae (2016). "Bacterial endophytes as potential biocontrol agents of vascular wilt diseases–review and future prospects". In: *Biological Control* 103, pp. 62–68.

Elsen, Sylvie et al. (2014). "A type III secretion negative clinical strain of Pseudomonas aeruginosa employs a two-partner secreted exolysin to induce hemorrhagic pneumonia". In: *Cell host & microbe* 15.2, pp. 164–176.

Falcone, Marilena et al. (2018). "The small RNA ErsA of Pseudomonas aeruginosa contributes to biofilm development and motility through post-transcriptional modulation of AmrZ". In: *Frontiers in microbiology* 9, p. 238.

Feklistov, Andrey et al. (2014). "Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective". In: *Annual Review of Microbiology* 68.1. PMID: 25002089, pp. 357–376. DOI: 10.1146/annurev-micro-092412-155737.

Feng, Lei et al. (2016). "Effects of quorum sensing systems on regulatory T cells in catheter-related Pseudomonas aeruginosa biofilm infection rat models". In: *Mediators of inflammation* 2016.

Fernández, Lucía et al. (2010). "Adaptive resistance to the "last hope" antibiotics polymyxin B and colistin in Pseudomonas aeruginosa is mediated by the novel two-component regulatory system ParR-ParS". In: *Antimicrobial agents and chemotherapy* 54.8, pp. 3372–3382.

Ferrara, Silvia et al. (2015). "Post-transcriptional regulation of the virulence-associated enzyme A lg C by the σ22-dependent small RNA E rs A of P seudomonas aeruginosa". In: *Environmental Microbiology* 17.1, pp. 199–214.

Filipowicz, Witold, Suvendra N Bhattacharyya, and Nahum Sonenberg (2008). "Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?" In: *Nature reviews genetics* 9.2, pp. 102–114.

Filloux, Alain (2011). "Protein secretion systems in Pseudomonas aeruginosa: an essay on diversity, evolution, and function". In: *Frontiers in microbiology* 2, p. 155.

Folkesson, Anders et al. (2012). "Adaptation of Pseudomonas aeruginosa to the cystic fibrosis airway: an evolutionary perspective". In: *Nature Reviews Microbiology* 10.12, pp. 841–851.

Fornes, Oriol et al. (2020). "JASPAR 2020: update of the open-access database of transcription factor binding profiles". In: *Nucleic acids research* 48.D1, pp. D87–D92.

Francis, Vanessa I, Emma C Stevenson, and Steven L Porter (2017). "Two-component systems required for virulence in Pseudomonas aeruginosa". In: *FEMS microbiology letters* 364.11, fnx104.

Francis, Vanessa I et al. (2018). "Multiple communication mechanisms between sensor kinases are crucial for virulence in Pseudomonas aeruginosa". In: *Nature communications* 9.1, pp. 1–11.

Freschi, Luca et al. (2015). "Clinical utilization of genomics data produced by the international Pseudomonas aeruginosa consortium". In: *Frontiers in microbiology* 6, p. 1036.

Freschi, Luca et al. (2019). "The Pseudomonas aeruginosa pan-genome provides new insights on its population structure, horizontal gene transfer, and pathogenicity". In: *Genome biology and evolution* 11.1, pp. 109–120.

Frimmersdorf, Eliane et al. (2010). "How Pseudomonas aeruginosa adapts to various environments: a metabolomic approach". In: *Environmental microbiology* 12.6, pp. 1734–1747.

Fuchs, Erin L et al. (2010). "The Pseudomonas aeruginosa Vfr regulator controls global virulence factor expression through cyclic AMP-dependent and-independent mechanisms". In: *Journal of bacteriology* 192.14, pp. 3553–3564.

Gajiwala, Ketan S and Stephen K Burley (2000). "Winged helix proteins". In: *Current opinion in structural biology* 10.1, pp. 110–116.

Galán-Vásquez, Edgardo, Beatriz Luna, and Agustino Martínez-Antonio (2011). "The regulatory network of Pseudomonas aeruginosa". In: *Microbial informatics and experimentation* 1.1, p. 3.

Galperin, Michael Y (2006). "Structural classification of bacterial response regulators: diversity of output domains and domain combinations". In: *Journal of bacteriology* 188.12, pp. 4169–4182.

Gao, Jinxin et al. (2005). "Binding of the global response regulator protein CovR to the sag promoter of Streptococcus pyogenes reveals a new mode of CovR-DNA interaction". In: *Journal of Biological Chemistry* 280.47, pp. 38948–38956.

Gao, Rong, Timothy R Mack, and Ann M Stock (2007). "Bacterial response regulators: versatile regulatory strategies from common domains". In: *Trends in biochemical sciences* 32.5, pp. 225–234.

Garber, Megan E et al. (2018). "Multiple signaling systems target a core set of transition metal homeostasis genes using similar binding motifs". In: *Molecular microbiology* 107.6, pp. 704–717.

García-Contreras, Rodolfo (2016). "Is quorum sensing interference a viable alternative to treat Pseudomonas aeruginosa infections?" In: *Frontiers in Microbiology* 7, p. 1454.

Gaspar, MC et al. (2013). "Pseudomonas aeruginosa infection in cystic fibrosis lung disease and new perspectives of treatment: a review". In: *European Journal of Clinical Microbiology & Infectious Diseases* 32.10, pp. 1231–1252.

Gebhardt, Michael J et al. (2020). "Widespread targeting of nascent transcripts by RsmA in Pseudomonas aeruginosa". In: *Proceedings of the National Academy of Sciences* 117.19, pp. 10520–10529.

Gelfand, Mikhail S (2006). "Evolution of transcriptional regulatory networks in microbial genomes". In: *Current opinion in structural biology* 16.3, pp. 420–429.

Gellatly, Shaan L and Robert EW Hancock (2013). "Pseudomonas aeruginosa: new insights into pathogenesis and host defenses". In: *Pathogens and disease* 67.3, pp. 159–173.

Georg, Jens et al. (2020). "The power of cooperation: Experimental and computational approaches in the functional characterization of bacterial sRNAs". In: *Molecular microbiology* 113.3, pp. 603–612.

Germain, Elsa et al. (2015). "Stochastic induction of persister cells by HipA through (p) ppGpp-mediated activation of mRNA endonucleases". In: *Proceedings of the National Academy of Sciences* 112.16, pp. 5171–5176.

Gerosa, Luca et al. (2013). "Dissecting specific and global transcriptional regulation of bacterial gene expression". In: *Molecular systems biology* 9.1, p. 658.

Gessard, Carle (1882). "Sur les colorations bleue et verte des linges a pansements". In: *CR Acad. Sci. Hebd. Seances Acad. Sci* 94, pp. 536–538.

Glick, Bernard R (2012). "Plant growth-promoting bacteria: mechanisms and applications". In: *Scientifica* 2012.

Glover, JN Mark et al. (2015). "The FinO family of bacterial RNA chaperones". In: *Plasmid* 78, pp. 79–87.

Gómez-Lozano, María et al. (2012). "Genome-wide identification of novel small RNAs in Pseudomonas aeruginosa". In: *Environmental microbiology* 14.8, pp. 2006–2016.

Gomila, Margarita et al. (2015). "Phylogenomics and systematics in Pseudomonas". In: *Frontiers in microbiology* 6, p. 214.

Gooderham, W James and Robert EW Hancock (2009). "Regulation of virulence and antibiotic resistance by two-component regulatory systems in Pseudomonas aeruginosa". In: *FEMS microbiology reviews* 33.2, pp. 279–294.

Gorgani, Neda et al. (2009). "Detection of point mutations associated with antibiotic resistance in Pseudomonas aeruginosa". In: *International journal of antimicrobial agents* 34.5, pp. 414–418.

Grainger, David C et al. (2006). "Association of nucleoid proteins with coding and non-coding segments of the Escherichia coli genome". In: *Nucleic acids research* 34.16, pp. 4642–4652.

Grosso-Becera, María Victoria, Luis Servín-González, and Gloria Soberón-Chávez (2015). "RNA structures are involved in the thermoregulation of bacterial virulence-associated traits". In: *Trends in microbiology* 23.8, pp. 509–518.

Grosso-Becerra, María Victoria et al. (2014). "Regulation of Pseudomonas aeruginosa virulence factors by two novel RNA thermometers". In: *Proceedings of the National Academy of Sciences* 111.43, pp. 15562–15567.

Haas, Dieter and Geneviève Défago (2005). "Biological control of soil-borne pathogens by fluorescent pseudomonads". In: *Nature reviews microbiology* 3.4, pp. 307–319.

Han, Kook, Brian Tjaden, and Stephen Lory (2016). "GRIL-seq provides a method for identifying direct targets of bacterial small regulatory RNA by in vivo proximity ligation". In: *Nature microbiology* 2.3, pp. 1–10.

Hauser, Alan R (2009). "The type III secretion system of Pseudomonas aeruginosa: infection by injection". In: *Nature Reviews Microbiology* 7.9, pp. 654–665.

Hauser, Alan R et al. (2002). "Type III protein secretion is associated with poor clinical outcomes in patients with ventilator-associated pneumonia caused by Pseudomonas aeruginosa". In: *Critical care medicine* 30.3, pp. 521–528.

Haynes, WC (1951). "Pseudomonas aeruginosa—its Characterization and Identification". In: *Microbiology* 5.5, pp. 939–950.

Heimann, John D (2002). "The extracytoplasmic function (ECF) sigma factors". In.

Helmann, John D (2019). "Where to begin? Sigma factors and the selectivity of transcription initiation in bacteria". In: *Molecular microbiology* 112.2, pp. 335–347.

Hengge, Regine (2009). "Principles of c-di-GMP signalling in bacteria". In: *Nature Reviews Microbiology* 7.4, pp. 263–273.

Hesse, Cedar et al. (2018). "Genome-based evolutionary history of Pseudomonas spp". In: *Environmental microbiology* 20.6, pp. 2142–2159.

Hickman, Jason W and Caroline S Harwood (2008). "Identification of FleQ from Pseudomonas aeruginosa as ac-di-GMP-responsive transcription factor". In: *Molecular microbiology* 69.2, pp. 376–389.

Hickman, Jason W, Delia F Tifrea, and Caroline S Harwood (2005). "A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels". In: *Proceedings of the National Academy of Sciences* 102.40, pp. 14422–14427.

Hoe, Chee-Hock et al. (2013). "Bacterial sRNAs: regulation in stress". In: *International Journal of Medical Microbiology* 303.5, pp. 217–229.

Hoffman, Lucas R et al. (2009). "Pseudomonas aeruginosa lasR mutants are associated with cystic fibrosis lung disease progression". In: *Journal of Cystic Fibrosis* 8.1, pp. 66–70.

Hoffmann, Nils and Bernd HA Rehm (2004). "Regulation of polyhydroxyalkanoate biosynthesis in Pseudomonas putida and Pseudomonas aeruginosa". In: *FEMS microbiology letters* 237.1, pp. 1–7.

Høiby, Niels et al. (2010). "Antibiotic resistance of bacterial biofilms". In: *International journal of antimicrobial agents* 35.4, pp. 322–332.

Holmqvist, Erik, Sofia Berggren, and Alisa Rizvanovic (2020). "RNA-binding activity and regulatory functions of the emerging sRNA-binding protein ProQ". In: *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, p. 194596.

Holmqvist, Erik and Jörg Vogel (2018). "RNA-binding proteins in bacteria". In: *Nature Reviews Microbiology* 16.10, pp. 601–615.

Holmqvist, Erik and E Gerhart H Wagner (2017). "Impact of bacterial sRNAs in stress responses". In: *Biochemical Society Transactions* 45.6, pp. 1203–1212.

Holmqvist, Erik et al. (2016). "Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo". In: *The EMBO journal* 35.9, pp. 991–1011.

Holmqvist, Erik et al. (2018). "Global maps of ProQ binding in vivo reveal target recognition via RNA structure and stability control at mRNA 3 ends". In: *Molecular cell* 70.5, pp. 971–982.

Hołówka, Joanna and Jolanta Zakrzewska-Czerwńska (2020). "Nucleoid Associated Proteins: The Small Organizers That Help to Cope With Stress". In: *Frontiers in Microbiology* 11, p. 590.

Hong, Duck Jin et al. (2015). "Epidemiology and characteristics of metallo-β-lactamase-producing Pseudomonas aeruginosa". In: *Infection & chemotherapy* 47.2, pp. 81–97.

Hör, Jens et al. (2020). "Grad-seq in a Gram-positive bacterium reveals exonucleolytic sRNA activation in competence control". In: *The EMBO Journal* 39.9, e103852.

Houot, Laetitia et al. (2012). "A bacterial two-hybrid genome fragment library for deciphering regulatory networks of the opportunistic pathogen Pseudomonas aeruginosa". In: *Microbiology* 158.8, pp. 1964–1971.

Huang, Hao et al. (2019). "An integrated genomic regulatory network of virulence-related transcriptional factors in Pseudomonas aeruginosa". In: *Nature communications* 10.1, pp. 1–13.

Huber, Philippe et al. (2016). "Pseudomonas aeruginosa renews its virulence factors". In: *Environmental microbiology reports* 8.5, pp. 564–571.

Ichinose, Yuki, Fumiko Taguchi, and Takafumi Mukaihara (2013). "Pathogenicity and virulence factors of Pseudomonas syringae". In: *Journal of general plant pathology* 79.5, pp. 285–296.

Impey, Rachael E et al. (2020). "Identification of two dihydrodipicolinate synthase isoforms from Pseudomonas aeruginosa that differ in allosteric regulation". In: *The FEBS Journal* 287.2, pp. 386–400.

Indurthi, Sai Madhuri, Han-Ting Chou, and Chung-Dar Lu (2016). "Molecular characterization of lysR-lysXE, gcdR-gcdHG and amaR-amaAB operons for lysine export and catabolism: a comprehensive lysine catabolic network in Pseudomonas aeruginosa PAO1". In: *Microbiology* 162.5, pp. 876–888.

Ireland, William T et al. (2020). "Deciphering the regulatory genome of Escherichia coli, one hundred promoters at a time". In: *arXiv preprint arXiv:2001.07396*.

Jacob-Dubuisson, Françoise et al. (2018). "Structural insights into the signalling mechanisms of two-component systems". In: *Nature Reviews Microbiology* 16.10, pp. 585–593.

Janssen, Kayley H et al. (2018). "RsmV, a small noncoding regulatory RNA in Pseudomonas aeruginosa that sequesters RsmA and RsmF from target mRNAs". In: *Journal of bacteriology* 200.16.

Janssen, Kayley H et al. (2020). "Hfq and sRNA 179 Inhibit Expression of the Pseudomonas aeruginosa cAMP-Vfr and Type III Secretion Regulons". In: *Mbio* 11.3.

Jin, Qiaoling et al. (2003). "Type III protein secretion in Pseudomonas syringae". In: *Microbes and Infection* 5.4, pp. 301–310.

Job, Viviana et al. (2019). "Pseudomonas two-partner secretion toxin Exolysin contributes to insect killing". In: *bioRxiv*, p. 807867.

Joly, Muriel et al. (2013). "Ice nucleation activity of bacteria isolated from cloud water". In: *Atmospheric environment* 70, pp. 392–400.

Jones, Allison M et al. (2005). "Role of BvgA phosphorylation and DNA binding affinity in control of Bvg-mediated phenotypic phase transition in Bordetella pertussis". In: *Molecular microbiology* 58.3, pp. 700–713.

Jones, Christopher J et al. (2014). "ChIP-Seq and RNA-Seq reveal an AmrZ-mediated mechanism for cyclic di-GMP synthesis and biofilm development by Pseudomonas aeruginosa". In: *PLoS Pathog* 10.3, e1003984.

Kambara, Tracy K, Kathryn M Ramsey, and Simon L Dove (2018). "Pervasive targeting of nascent transcripts by Hfq". In: *Cell reports* 23.5, pp. 1543–1552.

Kanack, Kristen J et al. (2006). "Characterization of DNA-binding specificity and analysis of binding sites of the Pseudomonas aeruginosa global regulator, Vfr, a homologue of the Escherichia coli cAMP receptor protein". In: *Microbiology* 152.12, pp. 3485–3496.

Karimova, Gouzel, Agnes Ullmann, and Daniel Ladant (2000). "[5] A bacterial two-hybrid system that exploits a cAMP signaling cascade in Escherichia coli". In: *Methods in enzymology*. Vol. 328. Elsevier, pp. 59–73.

Khademi, SM Hossein, Pavelas Sazinas, and Lars Jelsbak (2019). "Within-host adaptation mediated by intergenic evolution in Pseudomonas aeruginosa". In: *Genome biology and evolution* 11.5, pp. 1385–1397.

Khaledi, Ariane et al. (2016). "Transcriptome profiling of antimicrobial resistance in Pseudomonas aeruginosa". In: *Antimicrobial agents and chemotherapy* 60.8, pp. 4722–4733.

Khorasanizadeh, Sepideh (2004). "The nucleosome: from genomic organization to genomic regulation". In: *Cell* 116.2, pp. 259–272.

Kiely, Patrick D et al. (2008). "Genetic analysis of genes involved in dipeptide metabolism and cytotoxicity in Pseudomonas aeruginosa PAO1". In: *Microbiology* 154.8, pp. 2209–2218.

Kılıç, Sefa et al. (2014). "CollecTF: a database of experimentally validated transcription factor-binding sites in Bacteria". In: *Nucleic acids research* 42.D1, pp. D156–D160.

Kohler, R et al. (2017). "Architecture of a transcribing-translating expressome". In: *Science* 356.6334, pp. 194–197.

Kong, Weina et al. (2013). "Hybrid sensor kinase PA 1611 in P seudomonas aeruginosa regulates transitions between acute and chronic infection through direct interaction with RetS". In: *Molecular microbiology* 88.4, pp. 784–797.

Kong, Weina et al. (2015). "ChIP-seq reveals the global regulator AlgR mediating cyclic di-GMP synthesis in Pseudomonas aeruginosa". In: *Nucleic acids research* 43.17, pp. 8268–8282.

Kortmann, Jens and Franz Narberhaus (2012). "Bacterial RNA thermometers: molecular zippers and switches". In: *Nature Reviews Microbiology* 10.4, p. 255.

Kos, Veronica N et al. (2015). "The resistome of Pseudomonas aeruginosa in relationship to phenotypic susceptibility". In: *Antimicrobial agents and chemotherapy* 59.1, pp. 427–436.

Kulasekara, Bridget R et al. (2006). "Acquisition and evolution of the exoU locus in Pseudomonas aeruginosa". In: *Journal of bacteriology* 188.11, pp. 4037–4050.

Kulasekara, Hemantha D et al. (2005). "A novel two-component system controls the expression of Pseudomonas aeruginosa fimbrial cup genes". In: *Molecular microbiology* 55.2, pp. 368–380.

Kung, Vanderlene L, Egon A Ozer, and Alan R Hauser (2010). "The accessory genome of Pseudomonas aeruginosa". In: *Microbiology and molecular biology reviews* 74.4, pp. 621–641.

Lagares Jr, Antonio, Indra Roux, and Claudio Valverde (2016). "Phylogenetic distribution and evolutionary pattern of an α-proteobacterial small RNA gene that controls polyhydroxybutyrate accumulation in Sinorhizobium meliloti". In: *Molecular phylogenetics and evolution* 99, pp. 182–193.

Lai, Xuelei et al. (2020). "Genome-wide binding of SEPALLATA3 and AGAMOUS complexes determined by sequential DNA-affinity purification sequencing". In: *Nucleic Acids Research*.

Lambert, PA (2002). "Mechanisms of antibiotic resistance in Pseudomonas aeruginosa." In: *Journal of the royal society of medicine* 95.Suppl 41, p. 22.

Langmead, Ben and Steven L Salzberg (2012). "Fast gapped-read alignment with Bowtie 2". In: *Nature methods* 9.4, p. 357.

Le Berre, Rozenn et al. (2011). "Relative contribution of three main virulence factors in Pseudomonas aeruginosa pneumonia". In: *Critical care medicine* 39.9, pp. 2113–2120.

Lee, Daniel G et al. (2006). "Genomic analysis reveals that Pseudomonas aeruginosa virulence is combinatorial". In: *Genome biology* 7.10, R90.

Lee, Jasmine and Lianhui Zhang (2015). "The hierarchy quorum sensing network in Pseudomonas aeruginosa". In: *Protein & cell* 6.1, pp. 26–41.

Lee, Ji-Young and Kwan Soo Ko (2014). "Mutations and expression of PmrAB and PhoPQ related with colistin resistance in Pseudomonas aeruginosa clinical isolates". In: *Diagnostic microbiology and infectious disease* 78.3, pp. 271–276.

Lee, Ji-Young et al. (2016). "Evolved resistance to colistin and its loss due to genetic reversion in Pseudomonas aeruginosa". In: *Scientific reports* 6.1, pp. 1–13.

Lee, Vincent T et al. (2005). "Activities of Pseudomonas aeruginosa effectors secreted by the type III secretion system in vitro and during infection". In: *Infection and immunity* 73.3, pp. 1695–1705.

Leid, Jeff G et al. (2005). "The exopolysaccharide alginate protects Pseudomonas aeruginosa biofilm bacteria from IFN-γ-mediated macrophage killing". In: *The Journal of Immunology* 175.11, pp. 7512–7518.

Lemay, Jean-François et al. (2006). "Folding of the adenine riboswitch". In: *Chemistry & biology* 13.8, pp. 857–868.

Lhospice, Sébastien et al. (2017). "Pseudomonas aeruginosa zinc uptake in chelating environment is primarily mediated by the metallophore pseudopaline". In: *Scientific reports* 7.1, pp. 1–10.

Li, Qunhua et al. (2011). "Measuring reproducibility of high-throughput experiments". In: *The annals of applied statistics* 5.3, pp. 1752–1779.

Li, Wuju et al. (2012). "Predicting sRNAs and their targets in bacteria". In: *Genomics, proteomics & bioinformatics* 10.5, pp. 276–284.

Lin, Ping et al. (2019). "High-throughput screen reveals sRNAs regulating crRNA biogenesis by targeting CRISPR leader to repress Rho termination". In: *Nature communications* 10.1, pp. 1–12.

Linares, Juan F et al. (2010). "The global regulator Crc modulates metabolism, susceptibility to antibiotics and virulence in Pseudomonas aeruginosa". In: *Environmental microbiology* 12.12, pp. 3196–3212.

Lindeberg, Magdalen, Sébastien Cunnac, and Alan Collmer (2012). "Pseudomonas syringae type III effector repertoires: last words in endless arguments". In: *Trends in microbiology* 20.4, pp. 199–208.

Liu, Yi-Yun et al. (2016). "Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study". In: *The Lancet infectious diseases* 16.2, pp. 161–168.

Livny, Jonathan and Matthew K Waldor (2007). "Identification of small RNAs in diverse bacterial species". In: *Current opinion in microbiology* 10.2, pp. 96–101.

Lozada-Chavez, Irma, Sarath Chandra Janga, and Julio Collado-Vides (2006). "Bacterial regulatory networks are extremely flexible in evolution". In: *Nucleic acids research* 34.12, pp. 3434–3445.

Lu, Pei et al. (2018). "RgsA, an RpoS-dependent sRNA, negatively regulates rpoS expression in Pseudomonas aeruginosa". In: *Microbiology* 164.4, pp. 716–724.

Luengo, José M and Elías R Olivera (2020). "Catabolism of biogenic amines in Pseudomonas species". In: *Environmental Microbiology* 22.4, pp. 1174–1192.

Lyczak, Jeffrey B, Carolyn L Cannon, and Gerald B Pier (2000). "Establishment of Pseudomonas aeruginosa infection: lessons from a versatile opportunist". In: *Microbes and infection* 2.9, pp. 1051–1060.

Machanick, Philip and Timothy L Bailey (2011). "MEME-ChIP: motif analysis of large DNA datasets". In: *Bioinformatics* 27.12, pp. 1696–1697.

MacKenzie, Todd et al. (2014). "Longevity of patients with cystic fibrosis in 2000 to 2010 and beyond: survival analysis of the Cystic Fibrosis Foundation patient registry". In: *Annals of internal medicine* 161.4, pp. 233–241.

Maddocks, Sarah E and Petra CF Oyston (2008). "Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins". In: *Microbiology* 154.12, pp. 3609–3623.

Madigan, MT et al. (2017). "Brock Biology of Microorganisms 15 edn". In.

Maisonneuve, Etienne and Kenn Gerdes (2014). "Molecular mechanisms underlying bacterial persisters". In: *Cell* 157.3, pp. 539–548.

Mandal, Maumita et al. (2004). "A glycine-dependent riboswitch that uses cooperative binding to control gene expression". In: *Science* 306.5694, pp. 275–279.

Mandin, Pierre and Susan Gottesman (2010). "Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA". In: *The EMBO journal* 29.18, pp. 3094–3107.

Mannervik, Mattias (1999). "Target genes of homeodomain proteins". In: *BioEssays* 21.4, pp. 267–270.

Marden, Jeremiah N et al. (2013). "An unusual CsrA family member operates in series with RsmA to amplify posttranscriptional responses in Pseudomonas aeruginosa". In: *Proceedings of the National Academy of Sciences* 110.37, pp. 15055–15060.

Martínez, Eriel et al. (2019). "Oxylipins mediate cell-to-cell communication in Pseudomonas aeruginosa". In: *Communications biology* 2.1, pp. 1–10.

Martínez-García, Esteban and Víctor de Lorenzo (2019). "Pseudomonas putida in the quest of programmable chemistry". In: *Current opinion in biotechnology* 59, pp. 111–121.

Marvig, Rasmus Lykke et al. (2015). "Convergent evolution and adaptation of Pseudomonas aeruginosa within patients with cystic fibrosis". In: *Nature genetics* 47.1, p. 57.

Massé, Eric, Freddy E Escorcia, and Susan Gottesman (2003). "Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli". In: *Genes & development* 17.19, pp. 2374–2383.

Massé, Eric, Carin K Vanderpool, and Susan Gottesman (2005). "Effect of RyhB small RNA on global iron use in Escherichia coli". In: *Journal of bacteriology* 187.20, pp. 6962–6971.

Mathee, Kalai et al. (2008). "Dynamics of Pseudomonas aeruginosa genome evolution". In: *Proceedings of the National Academy of Sciences* 105.8, pp. 3100–3105.

Matsuo, Yasuhiro et al. (2004). "MexZ-mediated regulation of mexXY multidrug efflux pump expression in Pseudomonas aeruginosa by binding on the mexZ-mexX intergenic DNA". In: *FEMS microbiology letters* 238.1, pp. 23–28.

McCown, Phillip J et al. (2017). "Riboswitch diversity and distribution". In: *Rna* 23.7, pp. 995–1011.

Medina, Eva and Dietmar Helmut Pieper (2016). "Tackling threats and future problems of multidrug-resistant bacteria". In: *How to overcome the antibiotic crisis*. Springer, pp. 3–33.

Mejia-Almonte, Citlalli et al. (2020). "Redefining fundamental concepts of transcription initiation in bacteria". In: *Nature Reviews Genetics*, pp. 1–16.

Melamed, Sahar et al. (2020). "RNA-RNA interactomes of ProQ and Hfq reveal overlapping and competing roles". In: *Molecular Cell* 77.2, pp. 411–425.

Michalska, Marta and Philipp Wolf (2015). "Pseudomonas Exotoxin A: optimized by evolution for effective killing". In: *Frontiers in microbiology* 6, p. 963.

Mikkelsen, Helga, Rachel McMullan, and Alain Filloux (2011a). "The Pseudomonas aeruginosa reference strain PA14 displays increased virulence due to a mutation in ladS". In: *PloS one* 6.12, e29113.

Mikkelsen, Helga, Melissa Sivaneson, and Alain Filloux (2011b). "Key two-component regulatory systems that control biofilm formation in Pseudomonas aeruginosa". In: *Environmental microbiology* 13.7, pp. 1666–1681.

Miller, Christine L et al. (2016). "RsmW, Pseudomonas aeruginosa small non-coding RsmA-binding RNA upregulated in biofilm versus planktonic growth conditions". In: *BMC microbiology* 16.1, pp. 1–16.

Mills, Erez et al. (2008). "Real-time analysis of effector translocation by the type III secretion system of enteropathogenic Escherichia coli". In: *Cell host & microbe* 3.2, pp. 104–113.

Minezaki, Yoshiaki, Keiichi Homma, and Ken Nishikawa (2005). "Genome-wide survey of transcription factors in prokaryotes reveals many bacteria-specific families not found in archaea". In: *DNA Research* 12.5, pp. 269–280.

Mishra, Jitendra and Naveen Kumar Arora (2018). "Secondary metabolites of fluorescent pseudomonads in biocontrol of phytopathogens for sustainable agriculture". In: *Applied Soil Ecology* 125, pp. 35–45.

Moradali, M Fata, Shirin Ghods, and Bernd HA Rehm (2017). "Pseudomonas aeruginosa lifestyle: a paradigm for adaptation, survival, and persistence". In: *Frontiers in cellular and infection microbiology* 7, p. 39.

Morita, Teppei, Kimika Maki, and Hiroji Aiba (2005). "RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs". In: *Genes & development* 19.18, pp. 2176–2186.

Mulcahy, Heidi, Laetitia Charron-Mazenod, and Shawn Lewenza (2008). "Extracellular DNA chelates cations and induces antibiotic resistance in Pseudomonas aeruginosa biofilms". In: *PLoS Pathog* 4.11, e1000213.

Münch, Richard et al. (2003). "PRODORIC: prokaryotic database of gene regulation". In: *Nucleic acids research* 31.1, pp. 266–269.

Narberhaus, Franz, Torsten Waldminghaus, and Saheli Chowdhury (2006). "RNA thermometers". In: *FEMS microbiology reviews* 30.1, pp. 3–16.

Nemoto, Naoki et al. (2012a). "Mechanism for regulation of the putrescine utilization pathway by the transcription factor PuuR in Escherichia coli K-12". In: *Journal of bacteriology* 194.13, pp. 3437–3447.

— (2012b). "Regulation mechanism of the putrescine utilization pathway by the transcription factor PuuR in Escherichia coli K-12". In: *Journal of Bacteriology*.

Ngo, Tuan Dung et al. (2020). "The PopN gate-keeper complex acts on the ATPase PscN to regulate the T3SS secretion switch from early to middle substrates in Pseudomonas aeruginosa". In: *bioRxiv*.

Nikel, Pablo I, Esteban Martínez-García, and Víctor De Lorenzo (2014). "Biotechnological domestication of pseudomonads using synthetic biology". In: *Nature Reviews Microbiology* 12.5, pp. 368–379.

Nudler, Evgeny and Alexander S Mironov (2004). "The riboswitch control of bacterial metabolism". In: *Trends in biochemical sciences* 29.1, pp. 11–17.

Obritsch, Marilee D et al. (2005). "Nosocomial infections due to multidrug-resistant Pseudomonas aeruginosa: epidemiology and treatment options". In: *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* 25.10, pp. 1353–1364.

O'Loughlin, Colleen T et al. (2013). "A quorum-sensing inhibitor blocks Pseudomonas aeruginosa virulence and biofilm formation". In: *Proceedings of the National Academy of Sciences* 110.44, pp. 17981–17986.

O'Malley, Ronan C et al. (2016). "Cistrome and epicistrome features shape the regulatory DNA landscape". In: *Cell* 165.5, pp. 1280–1292.

Ortet, Philippe et al. (2012). "P2TF: a comprehensive resource for analysis of prokaryotic transcription factors". In: *BMC genomics* 13.1, pp. 1–8.

O'Toole, George A and Gerard CL Wong (2016). "Sensational biofilms: surface sensing in bacteria". In: *Current opinion in microbiology* 30, pp. 139–146.

Ozer, Egon A et al. (2019). "The population structure of pseudomonas aeruginosa is characterized by genetic isolation of exoU+ and exoS+ Lineages". In: *Genome biology and evolution* 11.7, pp. 1780–1796.

Page, Rebecca and Wolfgang Peti (2016). "Toxin-antitoxin systems in bacterial growth arrest and persistence". In: *Nature chemical biology* 12.4, pp. 208–214.

Paget, Mark SB and John D Helmann (2003). "The σ 70 family of sigma factors". In: *Genome biology* 4.1, pp. 1–6.

Pannen, Derk et al. (2016). "Interaction of the RcsB response regulator with auxiliary transcription regulators in Escherichia coli". In: *Journal of Biological Chemistry* 291.5, pp. 2357–2370.

Panpatte, Deepak G et al. (2016). "Pseudomonas fluorescens: a promising biocontrol agent and PGPR for sustainable agriculture". In: *Microbial inoculants in sustainable agricultural productivity*. Springer, pp. 257–270.

Paramythiotou, Elisabeth et al. (2004). "Acquisition of multidrug-resistant Pseudomonas aeruginosa in patients in intensive care units: role of antibiotics with antipseudomonal activity". In: *Clinical infectious diseases* 38.5, pp. 670–677.

Park, Se Chang et al. (2000). "Isolation of bacteriophages specific to a fish pathogen, Pseudomonas plecoglossicida, as a candidate for disease control". In: *Applied and environmental microbiology* 66.4, pp. 1416–1422.

Pastor, Alexandrine et al. (2005). "PscF is a major component of the Pseudomonas aeruginosa type III secretion needle". In: *FEMS microbiology letters* 253.1, pp. 95–101.

Patten, Cheryl L and Bernard R Glick (2002). "Role of Pseudomonas putida indoleacetic acid in development of the host plant root system". In: *Applied and environmental microbiology* 68.8, pp. 3795–3801.

Pavlova, Nikolet and Robert Penchovsky (2019). "Genome-wide bioinformatics analysis of FMN, SAM-I, glmS, TPP, lysine, purine, cobalamin, and SAH riboswitches for their applications as allosteric antibacterial drug targets in human pathogenic bacteria". In: *Expert opinion on therapeutic targets* 23.7, pp. 631–643.

Pazos, Michael A et al. (2017). "Pseudomonas aeruginosa ExoU augments neutrophil transepithelial migration". In: *PLoS pathogens* 13.8, e1006548.

Pei, Xue Yuan et al. (2019). "Architectural principles for Hfq/Crc-mediated regulation of gene expression". In: *Elife* 8, e43158.

Peix, Alvaro, Martha-Helena Ramírez-Bahena, and Encarna Velázquez (2009). "Historical evolution and current status of the taxonomy of genus Pseudomonas". In: *Infection, Genetics and Evolution* 9.6, pp. 1132–1147.

Pelosi, Ludovic et al. (2019). "Ubiquinone Biosynthesis over the Entire O2 Range: Characterization of a Conserved O2-Independent Pathway". In: *Mbio* 10.4, e01319–19.

Perez, J Christian and Eduardo A Groisman (2009a). "Evolution of transcriptional regulatory circuits in bacteria". In: *Cell* 138.2, pp. 233–244.

— (2009b). "Transcription factor function and promoter architecture govern the evolution of bacterial regulons". In: *Proceedings of the National Academy of Sciences* 106.11, pp. 4319–4324.

Perez-Rueda, Ernesto and Mario Alberto Martinez-Nuñez (2012). "The repertoire of DNA-binding transcription factors in prokaryotes: functional and evolutionary lessons". In: *Science Progress* 95.3, pp. 315–329.

Pesavento, Christina and Regine Hengge (2009). "Bacterial nucleotide-based second messengers". In: *Current opinion in microbiology* 12.2, pp. 170–176.

Petrova, Olga E et al. (2017). "Divide and conquer: the P seudomonas aeruginosa two-component hybrid S ag S enables biofilm formation and recalcitrance of biofilm cells to antimicrobial agents via distinct regulatory circuits". In: *Environmental microbiology* 19.5, pp. 2005–2024.

Pham, Van TK et al. (2017). "The plant growth-promoting effect of the nitrogen-fixing endophyte Pseudomonas stutzeri A15". In: *Archives of microbiology* 199.3, pp. 513–517.

Pieterse, Corné MJ et al. (2014). "Induced systemic resistance by beneficial microbes". In: *Annual review of phytopathology* 52.

Pirnay, Jean-Paul et al. (2002). "Analysis of the Pseudomonas aeruginosa oprD gene from clinical and environmental isolates". In: *Environmental Microbiology* 4.12, pp. 872–882.

Pita, Tiago, Joana R Feliciano, and Jorge H Leitão (2018). "Small noncoding regulatory RNAs from Pseudomonas aeruginosa and Burkholderia cepacia complex". In: *International journal of molecular sciences* 19.12, p. 3759.

Pohl, Carolina H and Johan LF Kock (2014). "Oxidized fatty acids as inter-kingdom signaling molecules". In: *Molecules* 19.1, pp. 1273–1285.

Poirel, Laurent et al. (2001). "OXA-28, an extended-spectrum variant of OXA-10 β-lactamase from Pseudomonas aeruginosa and its plasmid-and integron-located gene". In: *Antimicrobial Agents and Chemotherapy* 45.2, pp. 447–453.

Poole, Keith (2001). "Multidrug efflux pumps and antimicrobial resistance in Pseudomonas aeruginosa and related organisms". In: *Journal of molecular microbiology and biotechnology* 3.2, pp. 255–264.

— (2011). "Pseudomonas aeruginosa: resistance to the max". In: *Frontiers in microbiology* 2, p. 65.

Potvin, Eric, François Sanschagrin, and Roger C Levesque (2008). "Sigma factors in Pseudomonas aeruginosa". In: *FEMS microbiology reviews* 32.1, pp. 38–55.

Potvin, Eric et al. (2003). "In vivo functional genomics of Pseudomonas aeruginosa for high-throughput screening of new virulence factors and antibacterial targets". In: *Environmental microbiology* 5.12, pp. 1294–1308.

Preston, Gail M (2004). "Plant perceptions of plant growth-promoting Pseudomonas". In: *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 359.1446, pp. 907–918.

Price, Morgan N, Paramvir S Dehal, and Adam P Arkin (2008). "Horizontal gene transfer and the evolution of transcriptional regulation in Escherichia coli". In: *Genome biology* 9.1, R4.

Pukatzki, Stefan, Richard H Kessin, and John J Mekalanos (2002). "The human pathogen Pseudomonas aeruginosa utilizes conserved virulence pathways to infect the social amoeba Dictyostelium discoideum". In: *Proceedings of the National Academy of Sciences* 99.5, pp. 3159–3164.

Qiu, X, BR Kulasekara, and S Lory (2009). "Role of horizontal gene transfer in the evolution of Pseudomonas aeruginosa virulence". In: *Microbial Pathogenomics*. Vol. 6. Karger Publishers, pp. 126–139.

Quinlan, Aaron R and Ira M Hall (2010). "BEDTools: a flexible suite of utilities for comparing genomic features". In: *Bioinformatics* 26.6, pp. 841–842.

Rahme, Laurence G et al. (2000). "Plants and animals share functionally common bacterial virulence factors". In: *Proceedings of the National Academy of Sciences* 97.16, pp. 8815–8821.

Rajeev, Lara, Megan E Garber, and Aindrila Mukhopadhyay (2020). "Tools to map target genes of bacterial two-component system response regulators". In: *Environmental microbiology reports* 12.3, pp. 267–276.

Raoust, Eloïse et al. (2009). "Pseudomonas aeruginosa LPS or flagellin are sufficient to activate TLR-dependent signaling in murine alveolar macrophages and airway epithelial cells". In: *PloS one* 4.10, e7259.

Reboud, Emeline et al. (2016). "Phenotype and toxicity of the recently discovered exlA-positive Pseudomonas aeruginosa strains collected worldwide". In: *Environmental microbiology* 18.10, pp. 3425–3439.

Reboud, Emeline et al. (2017a). "Exolysin shapes the virulence of Pseudomonas aeruginosa clonal outliers". In: *Toxins* 9.11, p. 364.

Reboud, Emeline et al. (2017b). "Pseudomonas aeruginosa ExlA and Serratia marcescens ShlA trigger cadherin cleavage by promoting calcium influx and ADAM10 activation". In: *PLoS pathogens* 13.8, e1006579.

Reen, F Jerry et al. (2013). "Molecular evolution of LysR-type transcriptional regulation in Pseudomonas aeruginosa". In: *Molecular phylogenetics and evolution* 66.3, pp. 1041–1049.

Reiss, Caroline W and Scott A Strobel (2017). "Structural basis for ligand binding to the guanidine-II riboswitch". In: *RNA* 23.9, pp. 1338–1343.

Rivera-Gómez, Nancy et al. (2017). "Dissecting the protein architecture of DNA-binding transcription factors in bacteria and archaea". In: *Microbiology* 163.8, pp. 1167–1178.

Romeo, Tony, Christopher A Vakulskas, and Paul Babitzke (2013). "Post-transcriptional regulation on a global scale: form and function of Csr/Rsm systems". In: *Environmental microbiology* 15.2, pp. 313–324.

Romero, Pedro and Peter Karp (2003). "PseudoCyc, a pathway-genome database for Pseudomonas aeruginosa". In: *Journal of molecular microbiology and biotechnology* 5.4, pp. 230–239.

Roy, Paul H et al. (2010). "Complete genome sequence of the multiresistant taxonomic outlier Pseudomonas aeruginosa PA7". In: *PloS one* 5.1, e8842.

Salamov, Victor Solovyevand Asaf and A Solovyevand (2011). "Automatic annotation of microbial genomes and metagenomic sequences". In: *Metagenomics and its*

applications in agriculture, biomedicine and environmental studies. Hauppauge: Nova Science Publishers, pp. 61–78.

Saliba, Antoine-Emmanuel, Sara C Santos, and Jörg Vogel (2017). "New RNA-seq approaches for the study of bacterial pathogens". In: *Current opinion in microbiology* 35, pp. 78–87.

Sall, Khady Mayebine et al. (2014). "A gacS deletion in Pseudomonas aeruginosa cystic fibrosis isolate CHA shapes its virulence". In: *PLoS One* 9.4, e95936.

Sato, Hiromi et al. (2003). "The mechanism of action of the Pseudomonas aeruginosa-encoded type III cytotoxin, ExoU". In: *The EMBO journal* 22.12, pp. 2959–2969.

Sawa, Teiji et al. (2014a). "Anti-PcrV antibody strategies against virulent Pseudomonas aeruginosa". In: *Human vaccines & immunotherapeutics* 10.10, pp. 2843–2852.

Sawa, Teiji et al. (2014b). "Association between Pseudomonas aeruginosa type III secretion, antibiotic resistance, and clinical outcome: a review". In: *Critical Care* 18.6, p. 668.

Schalk, Isabelle J and Laurent Guillon (2013). "Pyoverdine biosynthesis and secretion in P seudomonas aeruginosa: implications for metal homeostasis". In: *Environmental microbiology* 15.6, pp. 1661–1673.

Schnell, Robert et al. (2012). "Tetrahydrodipicolinate N-succinyltransferase and dihydrodipicolinate synthase from Pseudomonas aeruginosa: structure analysis and gene deletion". In: *PLoS One* 7.2, e31133.

Schoehn, Guy et al. (2003). "Oligomerization of type III secretion proteins PopB and PopD precedes pore formation in Pseudomonas". In: *The EMBO journal* 22.19, pp. 4957–4967.

Schuster, Martin and E Peter Greenberg (2006). "A network of networks: quorum-sensing gene regulation in Pseudomonas aeruginosa". In: *International journal of medical microbiology* 296.2-3, pp. 73–81.

Schuster, Martin et al. (2003). "Identification, timing, and signal specificity of Pseudomonas aeruginosa quorum-controlled genes: a transcriptome analysis". In: *Journal of bacteriology* 185.7, pp. 2066–2079.

Schweizer, Herbert P and Cecilia Po (1996). "Regulation of glycerol metabolism in Pseudomonas aeruginosa: characterization of the glpR repressor gene." In: *Journal of bacteriology* 178.17, pp. 5215–5221.

Sciuto, Alessandra Lo et al. (2020). "Effect of lipid A aminoarabinosylation on Pseudomonas aeruginosa colistin resistance and fitness". In: *International Journal of Antimicrobial Agents*, p. 105957.

Scortichini, M and FG Troplano (1994). "Severe outbreak of Pseudomonas syringae pv. avellanae on hazelnut in Italy". In: *Journal of Phytopathology* 140.1, pp. 65–70.

Serganov, Alexander et al. (2006). "Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch". In: *Nature* 441.7097, pp. 1167–1171.

Shafikhani, Sasha H and Joanne Engel (2006). "Pseudomonas aeruginosa type III-secreted toxin ExoT inhibits host-cell division by targeting cytokinesis at multiple steps". In: *Proceedings of the National Academy of Sciences* 103.42, pp. 15605–15610.

Shahrezaei, Vahid and Samuel Marguerat (2015). "Connecting growth with gene expression: of noise and numbers". In: *Current opinion in microbiology* 25, pp. 127–135.

Shen-Orr, Shai S et al. (2002). "Network motifs in the transcriptional regulation network of Escherichia coli". In: *Nature genetics* 31.1, pp. 64–68.

Shivaji, S and Jogadhenu SS Prakash (2010). "How do bacteria sense and respond to low temperature?" In: *Archives of microbiology* 192.2, pp. 85–95.

Siam, Rania and Gregory T Marczynski (2000). "Cell cycle regulator phosphorylation stimulates two distinct modes of binding at a chromosome replication origin". In: *The EMBO journal* 19.5, pp. 1138–1147.

Sineva, Elena, Maria Savkina, and Sarah E Ades (2017). "Themes and variations in gene regulation by extracytoplasmic function (ECF) sigma factors". In: *Current opinion in microbiology* 36, pp. 128–137.

Skurnik, David et al. (2013). "A comprehensive analysis of in vitro and in vivo genetic fitness of Pseudomonas aeruginosa using high-throughput sequencing of transposon libraries". In: *PLoS Pathog* 9.9, e1003582.

Smirnov, Alexandre et al. (2016). "Grad-seq guides the discovery of ProQ as a major small RNA-binding protein". In: *Proceedings of the National Academy of Sciences* 113.41, pp. 11591–11596.

Smirnov, Alexandre et al. (2017). "Molecular mechanism of mRNA repression in trans by a ProQ-dependent small RNA". In: *The EMBO journal* 36.8, pp. 1029–1045.

Sonnleitner, Elisabeth and Udo Bläsi (2014). "Regulation of Hfq by the RNA CrcZ in Pseudomonas aeruginosa carbon catabolite repression". In: *PLoS Genet* 10.6, e1004440.

Sonnleitner, Elisabeth, Isabella Moll, and Udo Bläsi (2002). "Functional replacement of the Escherichia coli hfq gene by the homologue of Pseudomonas aeruginosa". In: *Microbiology* 148.3, pp. 883–891.

Sonnleitner, Elisabeth et al. (2003). "Reduced virulence of a hfq mutant of Pseudomonas aeruginosa O1". In: *Microbial pathogenesis* 35.5, pp. 217–228.

Sonnleitner, Elisabeth et al. (2006). "Hfq-dependent alterations of the transcriptome profile and effects on quorum sensing in Pseudomonas aeruginosa". In: *Molecular microbiology* 59.5, pp. 1542–1558.

Sonnleitner, Elisabeth et al. (2011). "The small RNA PhrS stimulates synthesis of the Pseudomonas aeruginosa quinolone signal". In: *Molecular microbiology* 80.4, pp. 868–885.

Sonnleitner, Elisabeth et al. (2012). "Novel targets of the CbrAB/Crc carbon catabolite control system revealed by transcript abundance in Pseudomonas aeruginosa". In: *PloS one* 7.10, e44637.

Sonnleitner, Elisabeth et al. (2018). "Interplay between the catabolite repression control protein Crc, Hfq and RNA in Hfq-dependent translational regulation in Pseudomonas aeruginosa". In: *Nucleic acids research* 46.3, pp. 1470–1485.

Sorger-Domenigg, Theresa et al. (2007). "Distinct and overlapping binding sites of Pseudomonas aeruginosa Hfq and RsmA proteins on the non-coding RNA RsmY". In: *Biochemical and biophysical research communications* 352.3, pp. 769–773.

Sosinsky, Alona et al. (2007). "Discovering transcriptional regulatory regions in Drosophila by a nonalignment method for phylogenetic footprinting". In: *Proceedings of the National Academy of Sciences* 104.15, pp. 6305–6310.

Storz, Gisela (1999). "An RNA thermometer". In: *Genes & development* 13.6, pp. 633–636.

Stover, CK et al. (2000). "Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen". In: *Nature* 406.6799, p. 959.

Strateva, Tanya and Daniel Yordanov (2009). "Pseudomonas aeruginosa–a phenomenon of bacterial resistance". In: *Journal of medical microbiology* 58.9, pp. 1133–1148.

Streeter, Klrissa and Mohammad Katouli (2016). "Pseudomonas aeruginosa: A review of their Pathogenesis and Prevalence in Clinical Settings and the Environment". In: *Infection, Epidemiology and Microbiology*.

Sundin, Charlotta et al. (2004). "Polarisation of type III translocation by Pseudomonas aeruginosa requires PcrG, PcrV and PopN". In: *Microbial pathogenesis* 37.6, pp. 313–322.

Tamber, Sandeep and RE Hancock (2003). "On the mechanism of solute uptake in Pseudomonas." In: *Frontiers in bioscience: a journal and virtual library* 8, s472.

Thi Bach Nguyen, Hue et al. (2018). "Negative control of RpoS synthesis by the sRNA ReaL in Pseudomonas aeruginosa". In: *Frontiers in microbiology* 9, p. 2488.

Tjaden, Brian (2012). "Computational identification of sRNA targets". In: *Bacterial Regulatory RNA*. Springer, pp. 227–234.

Toffano-Nioche, Claire et al. (2012). "Transcriptomic profiling of the oyster pathogen Vibrio splendidus opens a window on the evolutionary dynamics of the small RNA repertoire in the Vibrio genus". In: *Rna* 18.12, pp. 2201–2219.

Touzain, Fabrice et al. (2011). "DNA motifs that sculpt the bacterial chromosome". In: *Nature Reviews Microbiology* 9.1, pp. 15–26.

Tran, Anh M, Kelly M Chambers, and Katherine E Berry (2020). "Identifying species-specific Hfq-RNA Interactions using a bacterial three-hybrid assay". In: *The FASEB Journal* 34.S1, pp. 1–1.

Tree, Jai J et al. (2014). "Identification of bacteriophage-encoded anti-sRNAs in pathogenic Escherichia coli". In: *Molecular cell* 55.2, pp. 199–213.

Trouillon, Julian et al. (2020a). "Species-specific recruitment of transcription factors dictates toxin expression". In: *Nucleic acids research* 48.5, pp. 2388–2400.

Trouillon, Julian et al. (2020b). "Transcription inhibitors with XRE DNA-binding and cupin signal-sensing domains drive metabolic diversification in Pseudomonas". In: *bioRxiv*.

Trunk, Katharina et al. (2010). "Anaerobic adaptation in Pseudomonas aeruginosa: definition of the Anr and Dnr regulons". In: *Environmental microbiology* 12.6, pp. 1719–1733.

Turner, Keith H et al. (2014). "Requirements for Pseudomonas aeruginosa acute burn and chronic surgical wound infection". In: *PLoS Genet* 10.7, e1004518.

Uğur, Aysel, Özgür Ceylan, Belma Aslım, et al. (2012). "Characterization of Pseudomonas spp. from seawater of the southwest coast of Turkey." In: *Journal of Biological & Environmental Sciences* 6.16, pp. 25–33.

Ulrich, Luke E, Eugene V Koonin, and Igor B Zhulin (2005). "One-component systems dominate signal transduction in prokaryotes". In: *Trends in microbiology* 13.2, pp. 52–56.

Upadhyay, Supriya, Malay Ranjan Sen, and Amitabha Bhattacharjee (2010). "Presence of different beta-lactamase classes among clinical isolates of Pseudomonas aeruginosa expressing AmpC beta-lactamase enzyme". In: *The Journal of Infection in Developing Countries* 4.04, pp. 239–242.

Valentini, Martina and Alain Filloux (2016). "Biofilms and cyclic di-GMP (c-di-GMP) signaling: lessons from Pseudomonas aeruginosa and other bacteria". In: *Journal of Biological Chemistry* 291.24, pp. 12547–12555.

Valentini, Martina et al. (2018). "Lifestyle transitions and adaptive pathogenesis of Pseudomonas aeruginosa". In: *Current opinion in microbiology* 41, pp. 15–20.

Vallet, Isabelle et al. (2001). "The chaperone/usher pathways of Pseudomonas aeruginosa: identification of fimbrial gene clusters (cup) and their involvement in biofilm formation". In: *Proceedings of the National Academy of Sciences* 98.12, pp. 6911–6916.

Valot, Benoît et al. (2015). "What it takes to be a Pseudomonas aeruginosa? The core genome of the opportunistic pathogen updated". In: *PloS one* 10.5, e0126468.

Van Opijnen, Tim, Kip L Bodi, and Andrew Camilli (2009). "Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms". In: *Nature methods* 6.10, pp. 767–772.

Vanneste, Joel L (2017). "The scientific, economic, and social impacts of the New Zealand outbreak of bacterial canker of kiwifruit (Pseudomonas syringae pv. actinidiae)". In: *Annual Review of Phytopathology* 55, pp. 377–399.

Varga, John J et al. (2015). "Genotypic and phenotypic analyses of a Pseudomonas aeruginosa chronic bronchiectasis isolate reveal differences from cystic fibrosis and laboratory strains". In: *BMC genomics* 16.1, p. 883.

Veesenmeyer, Jeffrey L et al. (2009). "Pseudomonas aeruginosa virulence and therapy: evolving translational strategies". In: *Critical care medicine* 37.5, p. 1777.

Verma, Subhash C, Zhong Qian, and Sankar L Adhya (2019). "Architecture of the Escherichia coli nucleoid". In: *PLoS genetics* 15.12, e1008456.

Vettoretti, Lucie et al. (2009). "Efflux unbalance in Pseudomonas aeruginosa isolates from cystic fibrosis patients". In: *Antimicrobial agents and chemotherapy* 53.5, pp. 1987–1997.

Vodovar, Nicolas et al. (2006). "Complete genome sequence of the entomopathogenic and metabolically versatile soil bacterium Pseudomonas entomophila". In: *Nature biotechnology* 24.6, pp. 673–679.

Vogel, Jörg and Ben F Luisi (2011). "Hfq and its constellation of RNA". In: *Nature Reviews Microbiology* 9.8, pp. 578–589.

Vogel, Jörg et al. (2003). "RNomics in Escherichia coli detects new sRNA species and indicates parallel transcriptional output in bacteria". In: *Nucleic acids research* 31.22, pp. 6435–6443.

Wagner, E Gerhart H and Pascale Romby (2015). "Small RNAs in bacteria and archaea: who they are, what they do, and how they do it". In: *Advances in genetics*. Vol. 90. Elsevier, pp. 133–208.

Wang, H, F Tu, and Z Gui (2013). "Virulence factors in Pseudo-monas aeruginosa: mechanisms and modes of regulation". In: *Indian J Microbiol* 2, pp. 163–167.

Wasi, Samina, Shams Tabrez, and Masood Ahmad (2013). "Use of Pseudomonas spp. for the bioremediation of environmental pollutants: a review". In: *Environmental monitoring and assessment* 185.10, pp. 8147–8155.

Wassarman, Karen M (2018). "6S RNA, a global regulator of transcription". In: *Regulating with RNA in Bacteria and Archaea*, pp. 355–367.

Wassarman, Karen Montzka and Gisela Storz (2000). "6S RNA regulates E. coli RNA polymerase activity". In: *Cell* 101.6, pp. 613–623.

Waterman, Scott R and David W Holden (2003). "Functions and effectors of the Salmonella pathogenicity island 2 type III secretion system". In: *Cellular microbiology* 5.8, pp. 501–511.

Waters, Christopher M and Bonnie L Bassler (2005). "Quorum sensing: cell-to-cell communication in bacteria". In: *Annu. Rev. Cell Dev. Biol.* 21, pp. 319–346.

Waters, Elaine M et al. (2017). "Phage therapy is highly effective against chronic lung infections with Pseudomonas aeruginosa". In: *Thorax* 72.7, pp. 666–667.

Waters, Lauren S and Gisela Storz (2009). "Regulatory RNAs in bacteria". In: *Cell* 136.4, pp. 615–628.

Weirauch, Matthew T et al. (2014). "Determination and inference of eukaryotic transcription factor sequence specificity". In: *Cell* 158.6, pp. 1431–1443.

Widmaier, Daniel M et al. (2009). "Engineering the Salmonella type III secretion system to export spider silk monomers". In: *Molecular systems biology* 5.1, p. 309.

Wienken, Christoph J et al. (2010). "Protein-binding assays in biological liquids using microscale thermophoresis". In: *Nature communications* 1, p. 100.

Wilderman, Paula J et al. (2004). "Identification of tandem duplicate regulatory small RNAs in Pseudomonas aeruginosa involved in iron homeostasis". In: *Proceedings of the National Academy of Sciences* 101.26, pp. 9792–9797.

Wilhelm, Susanne et al. (2007). "The autotransporter esterase EstA of Pseudomonas aeruginosa is required for rhamnolipid production, cell motility, and biofilm formation". In: *Journal of bacteriology* 189.18, pp. 6695–6703.

Winkler, Wade C and Ronald R Breaker (2005). "Regulation of bacterial gene expression by riboswitches". In: *Annu. Rev. Microbiol.* 59, pp. 487–517.

Winsor, Geoffrey L et al. (2016). "Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database". In: *Nucleic acids research* 44.D1, pp. D646–D653.

Winstanley, Craig, Siobhan O'Brien, and Michael A Brockhurst (2016). "Pseudomonas aeruginosa evolutionary adaptation and diversification in cystic fibrosis chronic lung infections". In: *Trends in microbiology* 24.5, pp. 327–337.

Wittkopp, Patricia J and Gizem Kalay (2012). "Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence". In: *Nature Reviews Genetics* 13.1, pp. 59–69.

Wolf, Philipp and Ursula Elsässer-Beile (2009). "Pseudomonas exotoxin A: from virulence factor to anti-cancer agent". In: *International Journal of Medical Microbiology* 299.3, pp. 161–176.

Wray, Gregory A et al. (2003). "The evolution of transcriptional regulation in eukaryotes". In: *Molecular biology and evolution* 20.9, pp. 1377–1419.

Wright, Patrick R et al. (2014). "CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains". In: *Nucleic acids research* 42.W1, W119–W123.

Wu, Hao et al. (2012). "On the molecular mechanism of GC content variation among eubacterial genomes". In: *Biology direct* 7.1, p. 2.

Würtele, Martin et al. (2001). "How the Pseudomonas aeruginosa ExoS toxin downregulates Rac". In: *Nature structural biology* 8.1, pp. 23–26.

Wurtzel, Omri et al. (2010). "A single-base resolution map of an archaeal transcriptome". In: *Genome research* 20.1, pp. 133–141.

Wurtzel, Omri et al. (2012). "The single-nucleotide resolution transcriptome of Pseudomonas aeruginosa grown in body temperature". In: *Plos pathog* 8.9, e1002945.

Xavier, Danilo E et al. (2010). "Efflux pumps expression and its association with porin down-regulation and β-lactamase production among Pseudomonas aeruginosa causing bloodstream infections in Brazil". In: *BMC microbiology* 10.1, p. 217.

Xin, Xiu-Fang, Brian Kvitko, and Sheng Yang He (2018). "Pseudomonas syringae: what it takes to be a pathogen". In: *Nature Reviews Microbiology* 16.5, p. 316.

Yahr, Timothy L and Matthew C Wolfgang (2006). "Transcriptional regulation of the Pseudomonas aeruginosa type III secretion system". In: *Molecular microbiology* 62.3, pp. 631–640.

Yakhnin, Alexander V et al. (2013). "CsrA activates flhDC expression by protecting flhDC m RNA from RN ase E-mediated cleavage". In: *Molecular microbiology* 87.4, pp. 851–866.

Yamamoto, Kaneyoshi et al. (2005). "Functional characterization in vitro of all two-component signal transduction systems from Escherichia coli". In: *Journal of Biological Chemistry* 280.2, pp. 1448–1456.

Yang, Lei et al. (2011). "Evolutionary dynamics of bacteria in a human host environment". In: *Proceedings of the National Academy of Sciences* 108.18, pp. 7481–7486.

Yi, Buqing and Alexander H. Dalpke (2020). "Revisiting the Genomic Structure of the Genus Pseudomonas with Whole Genome Data: Insights into Diversity and Host-related Genetic Determinants". In: *bioRxiv*.

Zhang, Qiuhong et al. (2020). "Identification of genes regulated by the two-component system response regulator NarP of Actinobacillus pleuropneumoniae via DNA-affinity-purified sequencing". In: *Microbiological research* 230, p. 126343.

Zhang, Yi-Fan et al. (2017). "Probing the sRNA regulatory landscape of P. aeruginosa: post-transcriptional control of determinants of pathogenicity and antibiotic susceptibility". In: *Molecular microbiology* 106.6, pp. 919–937.

Zhang, Yong et al. (2004). "Conservation analysis of small RNA genes in Escherichia coli". In: *Bioinformatics* 20.5, pp. 599–603.

Zhang, Yong et al. (2008). "Model-based analysis of ChIP-Seq (MACS)". In: *Genome biology* 9.9, pp. 1–9.

Zhang, Yue, Qing Deng, and Joseph T Barbieri (2007). "Intracellular localization of type III-delivered Pseudomonas ExoS with endosome vesicles". In: *Journal of Biological Chemistry* 282.17, pp. 13022–13032.

Zheng, Ming and Gisela Storz (2000). "Redox sensing by prokaryotic transcription factors". In: *Biochemical pharmacology* 59.1, pp. 1–6.

Insights into the regulatory networks of *Pseudomonas aeruginosa*

Pseudomonas aeruginosa is an opportunistic human pathogen and a leading cause of nosocomial infections. This Gram-negative bacterium possesses one of the most complex regulatory networks, which allows it to sense and adapt to a wide variety of environmental conditions. In this work, several aspects of *P. aeruginosa* regulation were investigated. A new transcription factor (TF), ErfA, was found involved in the regulation of the ExlBA-dependent virulence in the PA7-like lineage of *P. aeruginosa* strains. The study of *exlBA* regulation in several *Pseudomonas* species revealed a diversity of regulatory mechanisms for this virulence factor due to differences in promoter *cis*-regulatory elements, which illustrated a mechanism of regulatory network evolution between closely related species. This type of promoter diversity was further investigated at a genome-wide scale and was found to be very common across all genes. Additionally, the characterization of all regulators sharing ErfA domain architecture allowed the definition of this family of TFs as comprising local, specialized regulators involved in the inhibition of small metabolic pathways. To go further into *P. aeruginosa* transcriptional regulation, the genome-wide determination of the regulons of all DNA-binding response regulators allowed the delineation of the two-component systems regulatory network, which comprises more than half of all genes of the bacterium. Finally, the investigation of post-transcriptional regulation through the comparative determination of RNA interactomes of Hfq between strains and growth phases identified numerous new common or specific regulatory interactions.

Etude des réseaux de régulation chez *Pseudomonas aeruginosa*

Pseudomonas aeruginosa est un pathogène opportuniste pour l'humain et une des premières causes d'infections nosocomiales. Cette bactérie Gram-négative possède un des réseaux de régulation les plus complexes, ce qui lui permet de s'adapter à un grand nombre d'environnements différents. Durant cette thèse, différents aspects de la régulation chez *P. aeruginosa* ont été investigués. Un nouveau facteur de transcription (FT), ErfA, a été identifié comme impliqué dans la régulation de la virulence dépendante d'ExlBA dans le groupe phylogénétique des souches PA7-like de *P. aeruginosa*. L'étude de la régulation d'*exlBA* dans plusieurs espèces de *Pseudomonas* a révélé une diversité de mécanismes de régulation pour ce facteur de virulence due à des différences de séquences régulatrices dans les promoteurs, ce qui illustre un mécanisme d'évolution de la régulation entre espèce proches. La diversité de promoteurs a été étudiée à l'échelle du génome et s'est révélée être relativement commune parmi tous les gènes. De plus, la caractérisation de tous les régulateurs qui possèdent la même architecture protéique que ErfA a permis d'identifier ces FTs comme des régulateurs locaux et spécialisés dans l'inhibition de voies métaboliques. Pour aller plus loin dans l'exploration de la régulation transcriptionnelle, la détermination des régulons de tous les régulateurs de réponses qui sont des FTs a permis la définition du réseau de régulation des systèmes à deux composantes, qui comprend plus de la moitié des gènes de *P. aeruginosa*. Enfin, l'étude de la régulation post-transcriptionnelle au travers de la caractérisation des interactomes de Hfq dans trois phases de croissance de différentes souches a permis d'identifier de nombreuses nouvelles interactions régulatrices communes ou spécifiques.