

Natural Transformation and Gene Transfer Agents: New Perspectives on Horizontal Gene Transfer in environmental bacteria

A dissertation submitted to the University of Neuchâtel

for the degree of Docteur.e ès sciences

by

Camille Tinguely

Thesis committee members:

Dr Diego Gonzalez, thesis co-director, University of Neuchâtel

Prof. Pilar Junier, thesis co-director, University of Neuchâtel

Prof. Josephus Vermeer, internal expert, University of Neuchâtel

Prof. Christoph Dehio, external expert, University of Basel

Defended on the 10th of June 2025

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise
l'impression de la présente thèse soutenue par

Madame Camille TINGUELY

Titre :

**“Natural Transformation and Gene Transfer Agents:
New Perspectives on Horizontal Gene Transfer in
environmental bacteria”**

sur le rapport des membres du jury composé comme suit :

- **Dr Diego Gonzalez**, co-directeur de thèse, Université de Neuchâtel
- **Prof. Pilar Junier**, co-directrice de thèse, Université de Neuchâtel
- **Prof. Josephus Vermeer**, Université de Neuchâtel
- **Prof. Christoph Dehio**, Université de Bâle

Neuchâtel, le 13 août 2025

Le Doyen, Prof. P. Brunner



A mon grand-papa Emile,
pour son amour de la nature et son application dans tout travail.

A ma grand-maman Marie,
pour son sens du détail en jardinage, en couture et en cuisine.

A mon grand-papa Arsène,
pour sa persévérance lors de la cueillette de champignons.

A ma grand-maman Eveline,
pour son sens de la conversation et l'importance qu'elle porte aux autres.

Acknowledgements

First and foremost, I would like to express my gratitude to Diego Gonzalez, my supervisor, for offering me a position in his dynamic subgroup and proposing such exciting projects. He not only shared his knowledge but also instilled in me a passion for research. Our stimulating discussions on project ideas, hypotheses, and philosophical questions have profoundly shaped the person I am today. I am indebted to him for his supervision, support, and the time he dedicated to reviewing my thesis, which were key in its completion.

My sincere appreciation extends to Pilar Junier, my co-supervisor, for welcoming me into her group. Her mentorship and unwavering support throughout my thesis have been invaluable.

Moreover, I feel privileged to have been part of the LAMUN group. To all present and former members, I extend my heartfelt thanks for the constant good mood and friendly working atmosphere. A special acknowledgment goes to the former apprentices Naïma Mangia, Matys Constantino, and Lorenzo Bisanti. Their contributions have been valuable to this thesis. It was a pleasure to introduce them to my work, share my bench with them, and mentor them. The teaching experience was reciprocal, as they helped me become a better instructor. I am also grateful to Valentin Voegeli for his excellent work on the GTA project. Sharing both the GTA topic and office space with him was truly enjoyable. Special thanks go to Thierry Kuhn and Céline Terrettaz for their support and the countless moments of laughter and camaraderie, both within and outside the laboratory.

My thanks go to Prof. Jan-Willem Veening for his insightful comments and productive discussions during my mid-thesis committee meeting. I am also grateful to Prof. Josephus Vermeer and Prof. Christoph Dehio, for agreeing to be part of my thesis committee and for their time in reviewing this work.

Finally, I want to thank my family and friends for their unwavering support and interest in my thesis. I am grateful for their patience in listening to my explanations and their encouragement during challenging times. A playful thanks to my cats, who always reminded me to take breaks and recharge by playing with them. Finally, I am truly thankful to my partner, Didier, for his patience and endless support, especially during the thesis writing phase. His understanding of the long hours and occasional frustrations has been a cornerstone of my thesis.

Thesis summary

Horizontal gene transfer (HGT) is a key driver of evolution and adaptation in prokaryotes. This process enables the acquisition of genes from non-related organisms through various mechanisms: conjugation, natural transformation, transduction, as well as less known systems like gene transfer agents (GTAs), membrane vesicles, or intercellular nanotubes.

While HGT increases genetic diversity, it may also confer other advantages. For example, studies suggest that natural transformation may confer advantages to recipient cells by enabling them to utilize internalized DNA as a template for genome repair or as a substrate for nutrient acquisition. However, the HGT carries risks for the recipient: improper integration may cause chromosomal breaks, acquisition of incompatible or deleterious genes can reduce bacterial fitness, and the process itself might incur an energy cost. Thus, HGT presents a complex trade-off between costs and benefits.

The main goal of this thesis was to deepen our understanding of HGT by addressing two fundamental questions: What specific advantages do natural transformation and HGT mediated by GTAs provide to bacteria? Under what conditions are they most beneficial?

The first part focused on natural transformation in *Pseudomonas stutzeri*. This process involves extracellular DNA uptake by the recipient bacterium, followed by its integration into the genome via homologous recombination. To achieve this, the bacterium must develop a competence state, characterized especially by the production of a specialized apparatus. We first evaluated the role of competence in using DNA as a nutrient source. Our results indicate that while *P. stutzeri* can metabolize DNA for carbon and nitrogen, this ability does not primarily depend on the DNA import machinery, suggesting an alternative extracellular degradation mechanism, likely mediated by nucleases. We then identified conditions favoring competence, demonstrating that genes required for natural transformation are upregulated in response to genotoxic stress (e.g., UV) and peak when bacterial populations exit exponential growth. Finally, we adapted a method to globally identify genes essential for natural transformation in donors and recipients. Though further optimization is needed, this approach revealed essential genes for *P. stutzeri* growth and survival.

The second part investigated the ability of *Paracoccus versutus* to perform gene transfer mediated by GTAs. GTAs are virus-like particles produced by host bacteria. They encapsulate donor DNA fragments and inject them into recipient cells, where the DNA fragments may integrate via homologous recombination. Our work demonstrated the ability of *P. versutus* to perform GTA-mediated transfer for the first time. We also showed that a deletion in the gene *cpdB* increases cyclic diguanylate concentration, rendering *P. versutus* resistant to the phage PvEP4A but reducing its ability for GTA-mediated transfer. This highlights a possible trade-off between phage resistance and gene transfer. Additionally, *P. versutus* GTA exhibits unique features compared to the one of *Rhodobacter capsulatus*, particularly in terms of the regulation of its production and DNA encapsulation, despite genetic proximity between both bacterial models.

In conclusion, this thesis provides new insights into bacterial regulation of natural transformation and HGT mediated by GTAs and highlights species-specific adaptations. Though HGT mechanisms are common across multiple species, each species optimizes these processes to maximize benefits while minimizing risks and costs. Broadly, this work advances our understanding of the diversity of bacterial adaptive strategies in terms of gene exchange, highlighting how HGT functions as a tool in microbial evolutionary arsenal.

Keywords: horizontal gene transfer | natural transformation | *Pseudomonas stutzeri* | gene transfer agents | *Paracoccus versutus* | phages resistance | cyclic di-GMP | cost-benefits | genes regulation.

Résumé

Le transfert horizontal de gènes (THG) est un moteur essentiel de l'évolution et de l'adaptation chez les procaryotes. Ce processus permet l'acquisition de gènes issus d'organismes non apparentés via des mécanismes bien connus comme la conjugaison, la transformation naturelle ou la transduction, mais aussi via des systèmes moins conventionnels tels que les agents de transfert de gènes (ATG), les vésicules membranaires ou les nanotubes intercellulaires.

Si le THG contribue à accroître la diversité génétique, il pourrait offrir également d'autres avantages. Par exemple, des études suggèrent que la transformation naturelle pourrait conférer des avantages aux cellules réceptrices en leur permettant d'utiliser l'ADN internalisé comme matrice pour la réparation de leur génome ou comme substrat pour l'acquisition de nutriments. Cependant, le THG n'est pas sans risque pour le receveur : une intégration inadéquate peut provoquer des cassures chromosomiques, l'acquisition de gènes incompatibles ou délétères peut réduire la valeur sélective d'un individu, et le processus lui-même engendre un coût énergétique. Ainsi, l'utilisation du THG doit répondre à un équilibre complexe entre avantages et inconvénients.

L'objectif principal de cette thèse est d'approfondir notre compréhension du THG en répondant à deux questions fondamentales : quels sont les avantages spécifiques que la transformation naturelle et le THG via les ATG confèrent aux bactéries ? Dans quelles conditions ces mécanismes confèrent-ils le plus d'avantages ?

La première partie de ce travail s'est concentrée sur le mécanisme de transformation naturelle chez l'espèce bactérienne *Pseudomonas stutzeri*. Ce processus implique l'absorption d'ADN extracellulaire par la bactérie réceptrice, suivie de son intégration dans le génome via recombinaison homologue. Pour ce faire, la bactérie doit développer un état de compétence, caractérisé notamment par la production d'un appendice spécialisé. Nous avons d'abord évalué le rôle de la compétence dans l'utilisation de l'ADN comme source de nutriments. Nos résultats indiquent que, bien que *P. stutzeri* puisse effectivement métaboliser l'ADN comme source de carbone et d'azote, cette capacité ne dépend pas principalement de l'appareil d'importation d'ADN, suggérant l'existence d'un mécanisme alternatif de dégradation

extracellulaire, probablement réalisé par des nucléases. Nous avons ensuite identifié les conditions favorisant le développement de la compétence, démontrant que l'expression des gènes nécessaires à la transformation naturelle est augmentée en réponse à un stress génotoxique (comme les UV) et atteint son maximum lorsque la population bactérienne sort de la phase exponentielle de croissance. Enfin, nous avons mis au point une approche pour identifier globalement les gènes essentiels à la transformation naturelle chez les donneurs et les receveurs. Bien que cette méthode nécessite encore des optimisations, elle nous a permis d'identifier des gènes essentiels pour la croissance et la survie de *P. stutzeri*.

La seconde partie de cette thèse s'est intéressée à *Paracoccus versutus* et au transfert de gènes via les ATG. Les ATG sont des particules similaires à des virus produites par une bactérie donneuse. Ils encapsulent des fragments du génome de celle-ci et les injectent dans des cellules réceptrices, où ces fragments peuvent s'intégrer par recombinaison homologe. Nos travaux ont permis de démontrer pour la première fois la capacité de *P. versutus* à effectuer du transfert de gènes via des ATG. Nous avons également montré qu'une délétion dans le gène *cpdB* a pour conséquence une concentration plus élevée de diguanylate cyclique, un messenger intracellulaire utilisé par de nombreuses bactéries. Cette délétion a rendu notre souche de *P. versutus* résistante au phage PvEP4A mais a également diminué sa capacité à transmettre des gènes via les ATG, mettant en lumière un compromis auquel cette bactérie fait face : la résistance à ce phage se fait au détriment du transfert génétique. Par ailleurs, nous avons établi que les ATG de *P. versutus* présentent des caractéristiques uniques par rapport à ceux de *Rhodobacter capsulatus*, notamment en ce qui concerne la régulation de leur production et l'encapsulation de l'ADN, malgré la proximité génétique entre ces deux modèles bactériens.

En conclusion, cette thèse apporte un nouvel éclairage sur les mécanismes par lesquels les bactéries régulent la transformation naturelle et le transfert de gènes via les ATG, et met en évidence des adaptations spécifiques à chaque espèce étudiée, en contraste avec les modèles bactériens plus classiques. Bien que ces mécanismes de transfert de gènes soient communs à plusieurs espèces, chaque espèce adapte ces processus pour maximiser leurs bénéfices tout en minimisant les risques et les coûts associés. Plus globalement, ces travaux ouvrent de nouvelles perspectives sur les stratégies d'adaptation bactériennes en matière d'échanges

génétiques, où le THG apparaît comme un outil clé dans l'arsenal évolutif des microorganismes.

Mots-clés : transfert horizontal de gènes | transformation naturelle | *Pseudomonas stutzeri* | agents de transfert de gènes | *Paracoccus versutus* | résistance aux phages | diguanylate cyclique | rapport coût-bénéfice | régulation des gènes.

Table of Contents

Chapter 1 - General introduction	19
1.1 Overview of HGT mechanisms in bacteria	19
1.2 Bacterial conjugation	21
1.3 Transduction	23
1.4 Natural transformation	25
1.5 HGT mediated by Gene Transfer Agents	32
1.6 HGT mechanisms: specificities and implications.....	39
1.7 Study models	42
1.8 Scope of the thesis	44
1.9 References	47
Part I: Natural transformation of <i>Pseudomonas stutzeri</i>	59
Chapter 2 - Utilization of DNA as a nutrient source by <i>P. stutzeri</i>	61
2.1 Foreword	61
2.2 Abstract	63
2.3 Introduction	64
2.4 Materials & Methods	68
2.5 Results	72
2.6 Discussion	80
2.7 References	84
2.8 Supplementary Material.....	89
Chapter 3 - Regulation of natural competence in <i>P. stutzeri</i>.....	101
3.1 Foreword	101
3.2 Abstract	103

3.3	Introduction	104
3.4	Materials and Methods	107
3.5	Results and Discussion	113
3.6	Conclusion	128
3.7	References	130
3.8	Supplementary Material.....	136
Chapter 4 - Broad-scale study of essential genes and natural transformation genetics in <i>P. stutzeri</i>		
		147
4.1	Foreword	147
4.2	Abstract	149
4.3	Introduction.....	150
4.4	Materials and Methods	155
4.5	Results and Discussion	161
4.6	Conclusion	169
4.7	References	170
4.8	Supplementary Material.....	175
Part II: Phage and Phage-like elements of <i>Paracoccus versutus</i>		
		195
Chapter 5 - Characterization of a novel bacteriophage, PvEP4A, with lytic activity against <i>P. versutus</i>		
		197
5.1	Foreword	197
5.2	Abstract	199
5.3	Introduction	200
5.4	Materials & Methods	201
5.5	Results and Discussion	209
5.6	Conclusion	222

5.7	References	224
5.8	Supplementary Material.....	232
Chapter 6 - Study of a conserved Gene Transfer Agent in <i>Paracoccus</i> species.....		261
6.1	Foreword	261
6.2	Abstract	263
6.3	Introduction	264
6.4	Materials & Methods	266
6.5	Results	272
6.6	Discussion	295
6.7	Conclusion	301
6.8	References	303
6.9	Supplementary Material.....	309
Chapter 7 - Synthesis, general discussion, perspectives and conclusion.....		329
7.1	Research background	329
7.2	Synthesis.....	329
7.3	Discussion	332
7.4	Perspectives.....	338
7.5	Conclusion	339
7.6	References	341
Chapter 8 - Appendices		345
8.1	Conference presentations	347
8.2	Other projects.....	351
8.3	Curriculum Vitae	353

Chapter 1 - General introduction

Horizontal gene transfer (HGT) is a key process that significantly influences bacterial evolution and adaptability. It can be advantageous in multiple ways. For example, HGT might allow bacteria to quickly adapt to new environments by acquiring new genes that enable them to utilize novel compounds or resources from their surroundings. Furthermore, HGT can also be beneficial by replacing deleterious alleles or genomic parasites such as mobile genetic elements, thereby improving the overall fitness of the organism. Another advantage of HGT is its ability to combine beneficial mutations that have arisen in different individuals into a single organism, accelerating the process of adaptation. Lastly, the incoming DNA transferred through HGT might also serve as a template for DNA repair, contributing to the maintenance of genomic integrity. The specific benefits largely depend on the length and content of the incoming DNA.

However, while HGT offers numerous advantages, it also presents potential drawbacks, such as the risk of acquiring detrimental genes and the energy costs associated with the process. To address these challenges, bacteria have evolved strategies to optimize HGT mechanisms, striving to maximize benefits while minimizing associated risks and costs. This optimization raises several questions: What specific benefits do different HGT mechanisms provide to bacteria? Under what conditions are these mechanisms most beneficial? How do bacteria minimize the costs associated with these mechanisms while maximizing their benefits?

In this PhD thesis, I investigated some of these questions using two environmental bacterial species, *Pseudomonas stutzeri* and *Paracoccus versutus*, as models. The focus was on two HGT mechanisms: natural transformation and HGT mediated by gene transfer agents (GTAs). The overall goal of this research is to provide deeper insights into the evolutionary and biological functions of these HGT mechanisms, contributing to our understanding of bacterial adaptation and evolution.

1.1 Overview of HGT mechanisms in bacteria

Prokaryotes exhibit a great diversity in metabolisms, lifestyles, and ecological niches. Indeed, they are found associated with plants and animals, in water and soil, and they can use a broad

range of substrates as energy sources. Bacteria can undergo evolution through various mechanisms, including single nucleotide mutations and gene duplications. While these genetic changes can account for significant genetic diversity over long periods, HGT plays a crucial role in accelerating bacterial evolution. HGT involves the transfer of genetic material between two unrelated organisms, rather than the vertical transmission from parent to offspring. Horizontal gene transfer has been even observed between species from different kingdoms (Gladyshev *et al.*, 2008). Similarly to spontaneous mutations, the genetic changes acquired through HGT can be beneficial, neutral, or deleterious and are subject to natural selection. HGT is a crucial driving force behind rapid bacterial evolution, enabling bacteria to swiftly adapt to new environmental niches. Understanding the mechanisms and significance of HGT is an important task of modern microbiology because HGT is a key element of bacterial evolutionary success and has impacts beyond microorganisms, such as on human health and agriculture (Boto *et al.*, 2019; Emamalipour *et al.*, 2020; Ghaly *et al.*, 2024).

The process of horizontal gene transfer requires that DNA gets released from a donor cell and internalized by a recipient cell through various mechanisms. HGT is typically thought to involve three major mechanisms (Figure 1.1): conjugation, transduction, and natural transformation. Conjugation consists in the transfer of genetic material between two bacteria through a structure called a sex pilus; it requires cell-to-cell contact between a living donor and recipient bacteria. In transduction, the transfer of DNA is mediated by a bacteriophage (hereafter phage) to a recipient bacterium; the donor bacterium can be spatially and temporarily separated from the recipient because phages can persist in the environment (Jończyk *et al.*, 2011). Natural transformation consists of the internalization of naked extracellular DNA. Gene transfer by transformation does not require a living donor cell; in nature, the transforming DNA has often been released by lysing cells. Donor and recipient cells can be separated spatially and temporally, but the persistence and spread of DNA in the environment are limited by degradation and adsorption (Lorenz and Wackernagel, 1994). In addition to these modes of transfer, another less traditional system exists in which transfer is mediated by gene transfer agents (GTAs) (Solioz *et al.*, 1975). GTAs are defective phages, incapable of self-replication, that package fragments of host DNA. Although all these HGT mechanisms allow the transfer of DNA from a donor to a recipient, each one of them has its own characteristics that will be discussed in the next sections.

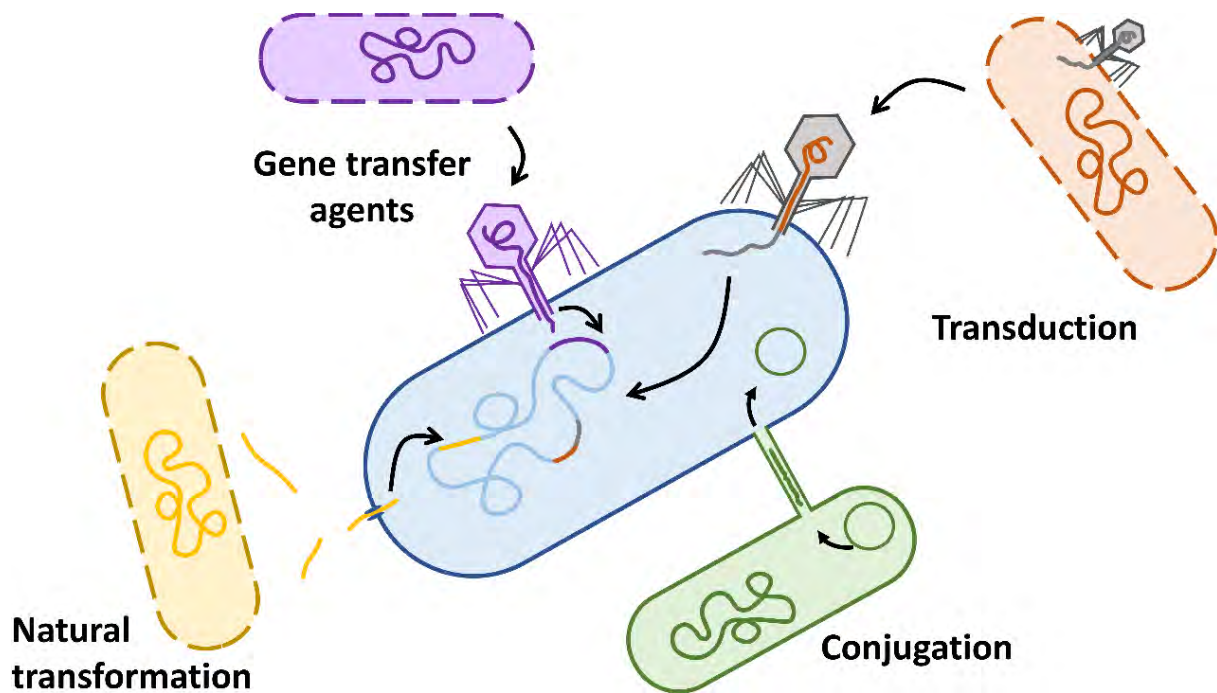


Figure 1.1 Horizontal gene transfer mechanisms in bacteria. Natural transformation is the process by which free environmental DNA is internalized by a recipient cell (blue). The DNA can be actively secreted by a donor cell or released after cell lysis (yellow). Transduction is a process by which genetic material from a donor cell is transferred to a recipient bacterium by a phage (grey). The donor DNA is in the phage capsid due to aberrant excision of a prophage from its host DNA or due to random picking of bacteria DNA during the packaging. Conjugation refers to the transfer of DNA from a donor cell (green) to a recipient cell via a sex pilus. Besides the three canonical HGT mechanisms, gene transfer agents mediate a fourth mechanism of DNA transfer, acting as defective phage elements that package random fragments of genomic DNA (purple).

1.2 Bacterial conjugation

Conjugation is a process by which genetic material is transferred from a donor cell to a recipient cell via direct cell-to-cell contact. The transfer requires a specific machinery that mobilizes the DNA and allows the formation of a mating pair. These genes can be encoded either by an autonomously replicating plasmid or by integrative conjugative elements (ICEs), which are integrated into the chromosome (Cabezón *et al.*, 2015; Johnson and Grossman, 2015).

1.2.1 Mechanism

The first discovered and most extensively studied conjugation system involves the transfer of the F plasmid in *Escherichia coli* from a donor cell (traditionally called “male”) possessing the plasmid to a recipient cell (traditionally called “female”) lacking it (Villa *et al.*, 2019; Virolle *et al.*, 2020). Conjugation starts with the expression of the *tra* genes. Certain Tra proteins are involved in forming the Type IV Secretion System (T4SS) and the conjugative pilus, which helps attachment to the recipient cell and stabilization of the mating pair. Other Tra proteins

constitute the relaxosome, which, along with the integration host factor, binds to the origin of transfer (*oriT*) and prepares the plasmid for transfer by nicking one strand of the double-stranded DNA (dsDNA). Then, the single-stranded DNA (ssDNA) end of the replicating plasmid, bound to the complex, is transferred into the recipient cell via the conjugative pilus while rolling circle replication occurs in the donor, converting ssDNA into dsDNA. Once inside the recipient, the ssDNA is recircularized and converted back into a dsDNA plasmid. Consequently, the recipient cell has now the potential to become a future donor cell.

Bacterial conjugation is a tightly regulated process (Koraimann and Wagner, 2014; Cabezón *et al.*, 2015; Kohler *et al.*, 2019; Virolle *et al.*, 2020). The expression of transfer *tra* genes is crucial for conjugation but is typically repressed to avoid fitness costs. Regulation mechanisms vary across conjugation systems, including quorum sensing in some Gram-negative and positive bacteria and specific activators like TraR, which can be triggered by environmental signals such as specific nutrients in *Agrobacterium tumefaciens*. Environmental conditions, physiological status of donor cells, and even certain antibiotics can influence conjugative behavior through complex regulatory networks (Headd and Bradford, 2018; Virolle *et al.*, 2020). Despite diverse molecular mechanisms, a common principle emerges across conjugation systems: conjugation-related gene products are typically regulated either by remaining in a default “OFF” state until induced by signaling molecules from potential recipients or the environment, or by being constitutively produced at low levels. This regulation balances the potential benefits of gene transfer with the need to minimize metabolic burden, ensuring that conjugation occurs optimally in response to specific environmental cues and in a small fraction of the population.

1.2.2 Implications of conjugation

Conjugation can introduce new genes into the recipient cell. During this process, any genes present on the conjugative plasmid, including those not directly involved in conjugation, can be transferred. As a result, the recipient cell not only acquires the genes necessary for the conjugative machinery but also any additional genes carried by the plasmid. Typically, many conjugative plasmids have been identified to carry pathogenicity islands, which are regions containing gene clusters that encode various virulence factors and antibiotic-resistance genes (Norman *et al.*, 2009; Koraimann, 2018; Virolle *et al.*, 2020). Genes might also “hitchhike” on conjugative plasmids to enter recipient cells and may either remain on the plasmid or

integrate into the host chromosome through mechanisms such as homologous recombination or transposon-mediated integration (Setlow *et al.*, 1984; Norman *et al.*, 2009).

A distinct conjugation mechanism involves the chromosomal integration of conjugative plasmids. The most famous example concerns *Hfr* (High-frequency recombination) strains, in which F plasmid becomes integrated into the host chromosome (Villa *et al.*, 2019).

Transfer originates from the F plasmid *oriT* site but extends beyond the plasmid coding region into adjacent chromosomal DNA, enabling the processive transfer of genomic sequences. Researchers exploited this phenomenon to map chromosomal gene order by tracking the timing of gene transfer to recipient cells (Villa *et al.*, 2019). However, due to the bacterial chromosome length, full transfer rarely occurs, leaving recipients with partial chromosomal DNA and an incomplete F plasmid that cannot confer donor capability.

As mentioned above, besides conjugative plasmids, ICEs are also capable of transfer via conjugation. They usually get integrated into particular genomic regions. Upon stimulation, they excise from the genome, circularize into a plasmid, and are transferred through a standard conjugative process (Johnson and Grossman, 2015). Once inside the recipient cell, the ICE re-integrates itself into the host genome. The excision and integration events can also lead to the incorporation of chromosomally adjacent genes into the plasmid during incorrect excision. This phenomenon allows for the transfer of additional genetic material beyond the ICE itself (Bedzyk *et al.*, 1992; Johnson and Grossman, 2015).

Overall, conjugation is a tightly regulated and coordinated process that enables the precise unidirectional transfer of a specific set of genes between spatially adjacent bacteria.

1.3 Transduction

Transduction is the process by which genes are transferred from one bacterial cell to another through the action of a phage. This mechanism involves the accidental inclusion of host DNA within the virion during the phage life cycle. There are two primary types of transduction: generalized and specialized transduction.

1.3.1 Mechanism

Phage infection is initiated by the specific adsorption of viral particles to receptors on the bacterial cell surface. This receptor-mediated interaction is followed by the injection of phage

genetic material into the host cell. Once internalized, the phage genome can proceed through either a lytic or lysogenic cycle (Villa *et al.*, 2019).

In the lytic cycle, phage genetic material undergoes replication and is packaged into newly formed virions, which are subsequently released via host cell lysis. While these progeny virions typically encapsulate the phage genome, errors in the packaging process can result in the incorporation of host DNA, leading to the formation of non-infective particles (Ikeda and Tomizawa, 1965). Upon infection of a new host by such a virion, the encapsulated host DNA neither replicates nor contributes to phage production. Instead, this genetic material may be integrated into the host genome through a process termed generalized transduction. This mechanism allows for the transfer of any genetic marker, provided it falls within the size constraints of the phage capsid.

In contrast, temperate phages integrate their genetic material into the host genome at specific sites and remain dormant. When the lysogenic state is induced, the phage DNA is excised from the chromosome, and adjacent host DNA may also be excised, replicated, and packaged into virions. Since this extra DNA is locally adjacent to the integration site, specific genes are consistently transferred through this mechanism, named specialized transduction. In some cases, a considerable portion of the viral DNA may be left behind during DNA excision, resulting in defective phage particles that cannot complete their life cycle (Villa *et al.*, 2019).

The regulation of transduction is intricately linked to the life cycle of phages, particularly their ability to switch between lytic and lysogenic states. Upon infection, phages can enter either the lytic or lysogenic cycle, with the decision influenced by multiple factors. For example, the choice of lambda phage to enter lysogeny is influenced by several parameters, including genetic compatibility (such as the presence of host *attB* integration sites), the host physiological state (for example, nutrient scarcity promotes lysogeny), and phage concentration (with higher multiplicity of infection increasing the likelihood of lysogeny) (Casjens and Hendrix, 2015; Howard-Varona *et al.*, 2017). Once lysogeny is established, prophages integrate into the bacterial chromosome and replicate alongside it. In the lambda phage, this lysogenic state is maintained by the CI repressor through a complex genetic feedback circuit (Bednarz *et al.*, 2014). The transition from lysogenic to lytic state can occur spontaneously or in response to external stressors that lead to CI inactivation (Zhang *et al.*, 2022). These stressors include changes in nutrients, salinity, temperature, and UV radiation,

although their effects may vary among different phage species (Zhang *et al.*, 2022). It is important to note that our understanding of the molecular mechanisms governing life cycle transitions in phages remains limited (Zhang *et al.*, 2022). Understanding the molecular mechanisms governing phage life cycles is crucial not only for elucidating their ecological roles but also for shedding light on transduction, as this process is directly influenced by phage life cycle decisions. Further research is needed to fully grasp these complex interactions.

1.3.2 Evolutionary functions of transduction

Overall, transduction serves as a mechanism for horizontal gene transfer. However, it remains uncertain whether transduction is merely an incidental result of the DNA packaging process or an evolutionary adaptation. Typically, the relationship between phages and bacteria is often viewed as a predator-prey relationship, but some studies suggest possible mutualistic interactions between transducing phages and their host, where both the phages and the hosts may benefit from transduction (Fillol-Salom *et al.*, 2019).

1.4 Natural transformation

Natural transformation is a biological process through which bacterial cells actively acquire extracellular DNA from their environment and integrate it into their own genome. For transformation to occur, bacteria must enter a specialized physiological state called “competence”, during which they assemble intricate cellular structures to facilitate DNA uptake (Claverys *et al.*, 2006; Dubnau and Blokesch, 2019). Natural transformation is a widespread HGT mechanism documented in more than 80 bacterial species (Johnston *et al.*, 2014). However, our understanding of this process primarily comes from studies on a limited number of model organisms. Among Gram-negative bacteria, *Pseudomonas stutzeri*, *Vibrio cholerae*, *Neisseria gonorrhoeae*, *Haemophilus influenzae*, *Helicobacter pylori*, and *Acinetobacter* spp. are the main models. For Gram-positive bacteria, *Bacillus subtilis* and *Streptococcus pneumoniae* have received the most attention.

1.4.1 Mechanism

Natural transformation relies on a specialized cellular machinery, in most cases encoded by the *com* genes, which functions to actively transport dsDNA into the cell (Johnston *et al.*, 2014). The DNA uptake system exhibits similarities between Gram-positive and Gram-negative

bacteria except for the transport across the outer membrane which is specific to Gram-negative bacteria (Figure 1.2) (Johnston *et al.*, 2014).

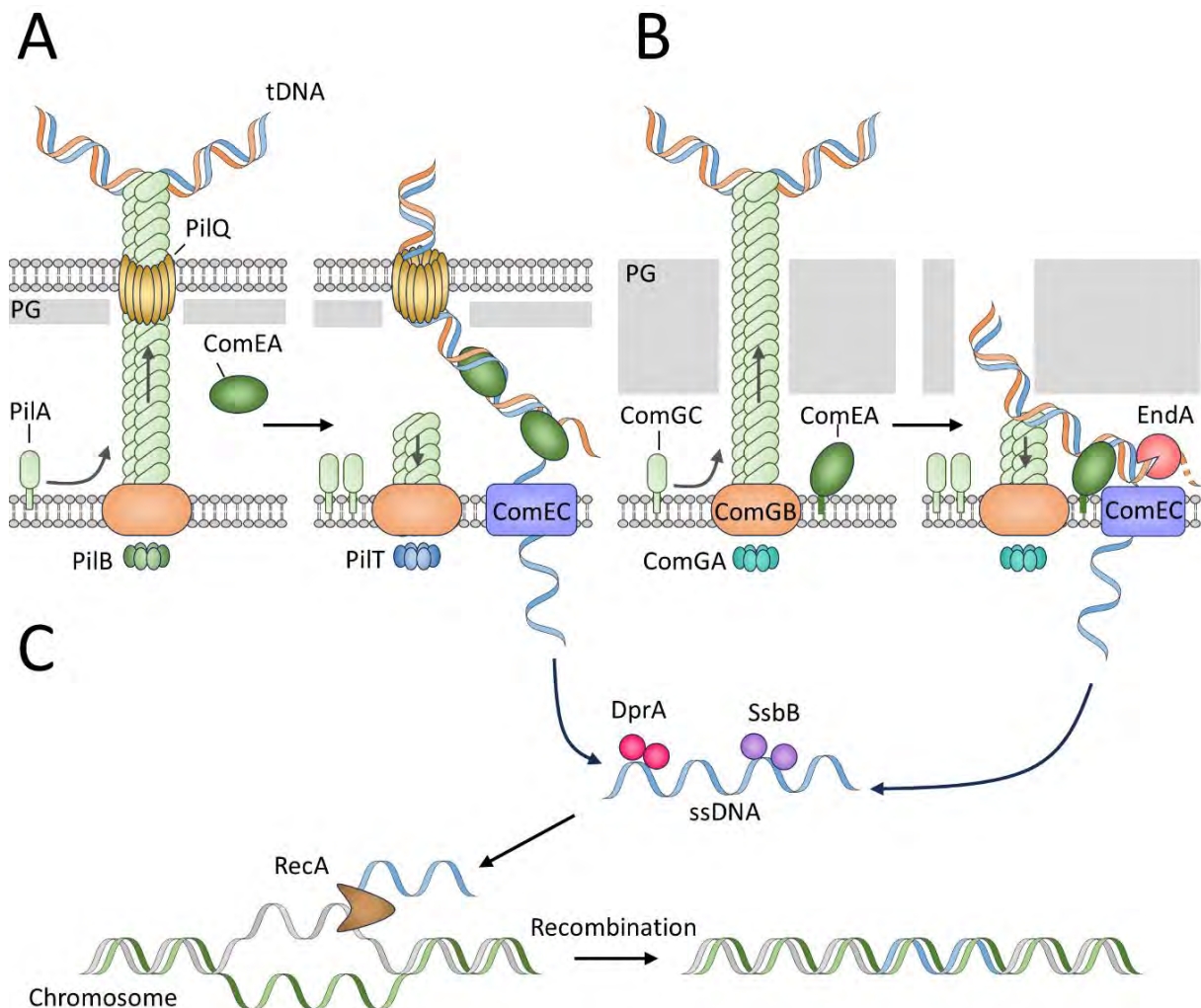


Figure 1.2 Schematic representation of natural transformation in Gram-negative bacteria like *V. cholerae* and Gram-positive bacteria like *B. subtilis* and *S. pneumoniae* (modified from (Dubnau and Blokesch, 2019; Liu *et al.*, 2020)). (A) In Gram-negative bacteria, extracellular DNA attaches to a pilus and is transported across the outer membrane with the help of the periplasmic protein ComEA. The pilus is a complex assembly of multiple proteins including PilA as the primary structural fiber protein, PilB and PilT ATPases that provide energy for pilus extension and retraction, PilQ forming a secretin pore in the outer membrane, PilC acting as the motor for pilus movement, and various other auxiliary proteins. (B) In Gram-positive bacteria, the pseudopilus, primarily composed of ComGC (major pilin), ComGB (membrane protein), and ComGA (ATPase), binds DNA and transports it through the cell wall with help of ComEA. The EndA nuclease degrades one DNA strand in *S. pneumoniae* but not in *B. subtilis*, which does not seem to have an EndA homolog (Dubnau and Blokesch, 2019), while ComEC transports the other into the cytoplasm. (C) Once in the cytoplasm, the ssDNA is bound by SsbB and DprA before being inserted in the host genome by the recombinase RecA. Abbreviations: PG, peptidoglycan; tDNA, transforming DNA.

The process of natural transformation involves several steps, with the machinery components often having species-specific nomenclature (Table 1.1). First, transforming DNA (tDNA) binds to cell surface and is transported to the periplasm by a type IV pilus or pilus-like structure except for *H. pylori* which uses type IV secretion system (Hofreuter *et al.*, 2001). In

Gram-negative bacteria, the translocation of tDNA is mediated by the retraction of the DNA-uptake pilus and binding of competence protein ComEA. The tDNA uptake mechanism is less clear in Gram-positive bacteria (Johnston *et al.*, 2014). Then, the ssDNA is translocated into the cytoplasm through the (inner) membrane channel protein ComEC, the other strand being typically degraded (Chen and Dubnau, 2004). It is not known exactly why DNA is transported as a single strand, but it might be that ssDNA is less sensitive to restriction enzymes, or that the ssDNA is a better substrate for recombination (Johnston *et al.*, 2014; Dubnau and Blokesch, 2019). Once ssDNA has entered the cytoplasm, the ssDNA-binding protein B (SsbB) and the DNA processing protein A (DprA) protect it from nuclease action, and with the help of RecA, incoming DNA is integrated into the host genome via recombination (Sharma *et al.*, 2023).

While the DNA-uptake apparatus is relatively conserved among bacteria, the regulation of competence varies considerably between species (Claverys *et al.*, 2006; Seitz and Blokesch, 2012). This diversity in regulatory mechanisms reflects the adaptation of bacterial species to their specific environmental niches and evolutionary pressures. Integration of competence regulation with other cellular regulators presumably enables these species to finely tune competence development in response to the overall physiological state of the cell.

Table 1.1 Homologues of proteins involved in natural transformation in various bacteria. This table is a modified version from Chen and Dubnau (Chen and Dubnau, 2004) and from Nolan and colleagues (Nolan *et al.*, 2020). n.a., not applicable.

	<i>B. subtilis</i> ¹	<i>S. pneumoniae</i> ²	<i>H. influenzae</i> ³	<i>N. gonorrhoeae</i> ⁴	<i>V. cholerae</i> ⁵	<i>P. stutzeri</i> ⁶
Competence pilus						
Traffic NTPase(s)	ComGA	ComGA	PilB	PilF, PilT	PilB, PilT	PilT, PilU
Polytopic membrane protein	ComGB	ComGB	PilC	PilG	PilC, -O, -M, -N, -P	PilC
Pilins or pseudopilins	ComGC, -GD, -GE, -GG	CglC, CglD	PilA	PilE, ComP	PilA	PilA
Prepilin peptidase	ComC	CilC	PilD	PilD	PilD	
Secretin/pilot	n.a.	n.a.	ComE	PilQ/PilP	PilQ	
DNA translocation machinery						
DNA receptor	ComEA	ComEA		ComE	ComEA	
Membrane channel	ComEC	ComEC	Rec-2	ComA	ComEC	ComA
ATP-binding protein	ComFA	ComFA				ExbB

References: 1. (Dubnau and Provvedi, 2000), 2. (Dubnau and Blokesch, 2019), 3. (Dougherty and Smith, 1999), 4. (Kooimey, 1998), 5. (Seitz and Blokesch, 2013; Matthey and Blokesch, 2016), 6. (Graupner *et al.*, 2000; Graupner and Wackernagel, 2001; Graupner *et al.*, 2001).

Competence is typically induced in response to specific cellular or environmental signals, with the exception of *H. pylori* and *N. gonorrhoeae*, which are considered to be constitutively competent (Sparling, 1966; Baltrus and Guillemin, 2006; Corbinais *et al.*, 2016). For instance, *H. influenzae* develops competence in response to nutrient's reduction, such as during stationary phase or when transferred to a nutrient-poor medium (Herriott *et al.*, 1970; Redfield, 1991). This process requires the expression of regulatory genes *sxy* and *crp*, along with elevated intracellular levels of cyclic AMP (Wise *et al.*, 1973; Redfield, 1991; Redfield *et al.*, 2005). In *V. cholerae*, natural transformation is regulated through multiple environmental cues, including chitin availability, quorum sensing, and carbon catabolite repression (Metzger and Blokesch, 2016). Chitin induces the expression of the TfoX transcriptional regulator, which in turn activates competence genes.

The gram-positive species *B. subtilis* and *S. pneumoniae* share similarities in their competence regulatory cascades (Claverys *et al.*, 2006). Both species initiate competence through a small

peptide pheromone, competence-stimulating peptide (CSP) for *S. pneumoniae* and ComX for *B. subtilis*, detected by a specialized two-component regulatory systems. This detection triggers a signaling cascade that ultimately activates the transcription of genes in the competence regulon, but the effectors have not evolved from a common ancestor (Claverys and Martin, 2003). In *S. pneumoniae*, competence is induced by various stressors, including certain antibiotics as aminoglycoside and quinolone, nutrient availability, and cell wall integrity (Claverys and Havarstein, 2002; Prudhomme *et al.*, 2006; Fontaine *et al.*, 2015). *B. subtilis* competence development is controlled by a complex network involving cell-density, nutritional signals, and cell cycle cues (Hamoen *et al.*, 2003; Claverys *et al.*, 2006).

The control of competence by the host allows bacteria to adapt their DNA uptake capabilities to their specific environmental conditions and physiological needs, potentially maximizing the benefits of natural transformation.

1.4.2 Benefits of natural transformation

Natural transformation is a widespread mechanism of HGT that has been documented in more than 80 bacterial species (Johnston *et al.*, 2014). However, the evolutionary forces driving the development and maintenance of natural competence remain debated. Three primary, non-mutually exclusive hypotheses suggest that natural transformation might be beneficial due to the potential use of DNA for nutrition, genomic repair, and evolution (Redfield, 1993b; Ambur *et al.*, 2016).

1.4.2.1 DNA as a nutrient source

The nutritional hypothesis proposes that bacteria benefit from natural competence by using the tDNA as a source of nutrients, particularly when resources are scarce. This process involves breaking down the tDNA into smaller components, such as oligonucleotides or single nucleotides, which can then serve as building blocks for DNA synthesis or as, for instance, a carbon source for cellular processes (Redfield, 1993b; Finkel and Kolter, 2001). Supporting evidence includes the observation that some species, like *H. influenzae*, require nutritional downshifts to induce competence, and purine depletion triggers the competence activator Sxy translation (Herriott *et al.*, 1970; Sinha *et al.*, 2013). Moreover, experiments with *E. coli* suggest that this bacterial species can grow using DNA as a source of nutrients with the help of the competence apparatus (Finkel and Kolter, 2001; Palchevskiy and Finkel, 2006).

However, this hypothesis faces several challenges. First, exogenous DNA does not induce competence in any known transformable species (Claverys *et al.*, 2006; Seitz and Blokesch, 2012). Secondly, only one strand of DNA is degraded while the other enters the cytoplasm (Dubnau, 1999; Chen and Dubnau, 2004). This selective degradation process is inconsistent with an optimal nutritional strategy, which would likely involve complete breakdown of both DNA strands to maximize nutrient availability. Thirdly, some species preferentially incorporate genetic material from closely related organisms due to specific DNA uptake sequences (Graves *et al.*, 1982; Redfield *et al.*, 2006), despite the potential nutritional value of non-homologous DNA. Finally, it remains unclear whether the metabolic benefits derived from exogenous DNA compensate for the energy expenditure associated with DNA transport (Seitz and Blokesch, 2012). An alternative option for feeding on DNA rely on the secretion of nucleases such as EddB for *Pseudomonas aeruginosa* (Mulcahy *et al.*, 2010), EndA and ExeM for *Shewanella oneidensis* (Gödeke *et al.*, 2011; Heun *et al.*, 2012), and Dns and Xds for *V. cholerae* (Seper *et al.*, 2011). These nucleases have been shown to break down DNA for growth benefits. Interestingly, *dns* transcription is downregulated before the onset of natural transformation, indicating that nuclease production and natural competence can coexist in the same strain, as long as they are not simultaneous (Blokesch and Schoolnik, 2008). The presence of this more energy-efficient alternative does not necessarily invalidate the potential nutritional role of competence. However, it suggests that the evolutionary advantages of maintaining competence machinery might extend beyond mere nutritional benefits.

1.4.2.2 DNA as a Template for Genome Repair

The concept of transformation as a beneficial mechanism for genome repair has been a longstanding hypothesis in the field. This idea suggests that acquired DNA serves as a substrate for repairing genomic damage. Early experiments in *B. subtilis* indicated an immediate survival benefit for transformants after DNA-damaging conditions (Michod and Wojciechowski, 1988; Wojciechowski *et al.*, 1989; Hoelzer and Michod, 1991). In addition, during competence induction, genes used for DNA repair are also upregulated in *B. subtilis*, supporting this potential connection between the two pathways (Berka *et al.*, 2002). Nevertheless, other studies in the same species challenged this notion by demonstrating that genotoxic stress did not induce competence as predicted (Redfield, 1993a).

A more recent extension of this hypothesis suggests that natural transformation serves as a general stress response mechanism (Claverys *et al.*, 2006). This idea is supported by observations in some naturally transformable human pathogens, such as *S. pneumoniae*, *Legionella pneumophila*, and *H. pylori*, which lack a general SOS response but exhibit competence induction in response to DNA-damaging agents, including certain antibiotics (Claverys *et al.*, 2006; Dorer *et al.*, 2010; Charpentier *et al.*, 2011). However, it is noteworthy that competence induction is not exclusively linked to genotoxic stress responses (Lorenz and Wackernagel, 1990; Palmen *et al.*, 1993; Kim *et al.*, 2008; Seitz and Blokesch, 2012; Johnston *et al.*, 2014). In numerous bacterial species, competence is primarily regulated by quorum sensing mechanisms, which operate independently of external stressors (Seitz and Blokesch, 2012; Johnston *et al.*, 2014). Furthermore, a diverse array of stress factors can modulate natural transformation processes.

1.4.2.3 DNA as Genetic Material for Genome Evolution

Natural transformation, like other HGT mechanisms, is recognized as a significant contributor to bacterial evolution. Mathematical modelling supports the notion that natural competence has evolved and persists as a mechanism promoting genome plasticity, thereby enhancing rates of environmental adaptation (Levin and Cornejo, 2009; Moradigaravand and Engelstädter, 2013). Indeed, natural transformation has the potential to shuffle alleles within a population, allowing natural selection to act on this diverse population. Some experimental results confirm the benefits of natural transformation as accelerating adaptation (Baltrus *et al.*, 2007). However, this reshuffling might also break up co-adapted genes and be detrimental (Ambur *et al.*, 2016).

Conversely, some authors propose that natural transformation may function to maintain genome stability and facilitate the elimination of detrimental mobile genetic elements from the bacterial chromosome (Treangen *et al.*, 2008; Ambur *et al.*, 2016). In *S. pneumoniae*, competent populations showed less fitness improvement compared to non-competent populations when exposed to mild stress conditions, but the competent populations accumulated fewer new mutations over time, suggesting that competence helps in maintaining genome stability during evolution in benign conditions (Engelmoer *et al.*, 2013). A whole-genome study by Croucher *et al.* revealed that transformable streptococci contained fewer prophages compared to non-transformable strains despite tending to have larger

genomes (Croucher *et al.*, 2016), suggesting that natural transformation may help remove harmful mobile genetic elements from the bacterial genome.

The effects of natural transformation on bacterial genomes are multifaceted, promoting either genomic conservation or evolution. These outcomes are influenced by species-specific factors, the origin of the genetic material, and the experimental conditions employed. Further research is necessary to elucidate the specific circumstances under which natural transformation facilitates genome conservation versus evolution. This research should involve three complementary facets: studies of natural populations *in situ*, simplified evolution experiments conducted under controlled conditions (e.g. liquid cultures with constant parameters or single variable manipulation), and more complex studies simulating environmental conditions. The latter could include experiments incorporating fluctuating parameters, structured habitats, or multi-species interactions. By integrating data from both simplified and ecologically relevant experiments, alongside studies of natural populations, we can develop a more comprehensive understanding of the evolutionary forces driving the maintenance of natural transformation and clarify the ecological significance of this HGT mechanism.

In summary, while naturally competent bacteria may derive nutritional advantages from their ability to take up exogenous DNA, it remains unclear whether these benefits outweigh the associated metabolic costs. The potential benefits of natural transformation for DNA repair and evolutionary adaptation are less predictable, as they depend on factors such as sequence homology and ecological or evolutionary opportunities. It is important to note that these hypotheses regarding the function of natural competence are not mutually exclusive. Rather, competence may serve multiple purposes for a given bacterial species, analogous to the versatility of a Swiss Army knife. This multifunctional nature of natural transformation highlights its complex and significant role in bacterial genetics and evolution.

1.5 HGT mediated by Gene Transfer Agents

A fourth, less studied mechanism of gene transfer involves gene transfer agents (GTAs). GTAs are phage-like particles encoded by gene clusters within a bacterial or archaeal host genome, which encapsulate fragments of host dsDNA (Lang and Beatty, 2000). Since their first discovery in the alphaproteobacterium *Rhodobacter capsulatus* (Marrs, 1974), GTAs have garnered

increasing interest due to their role in bacterial gene flow, ecological interactions, and evolutionary dynamics, especially in the marine ecosystem (McDaniel *et al.*, 2010, 2012; Von Wintersdorff *et al.*, 2016). Nowadays, multiple models have been identified (Table 1.2) (Banks and Le, 2024; Craske *et al.*, 2024).

Although the various GTA models are encoded by phage-derived genes with similar functions, they exhibit evidence of independent evolutionary trajectories, with five genetically unrelated GTAs known: RcGTA, BaGTA, DdGTA, BhGTA, and MvGTA (Lang *et al.*, 2017; Redfield *et al.*, 2019). Nevertheless, the main RcGTA genes cluster, encoding the structural elements of GTA particles, is widely conserved throughout the Rhodobacterales order, with multiple occurrences in other alphaproteobacterial orders (Lang and Beatty, 2007; Lang *et al.*, 2017). Besides genetic homology, experimental data indicate that some species other than *R. capsulatus* produce GTAs.

Table 1.2 GTA models and their characteristics.

GTA	Species	Size of main GTA cluster [kb]	DNA packaged [kb]	Tail morphology ⁸	Ref.
RcGTA	<i>Rhodobacter capsulatus</i>	15	4	Siphovirus	1
CcGTA	<i>Caulobacter crescentus</i>	15.6	8.3	Siphovirus	2
DsGTA	<i>Dinoroseobacter shibae</i>	15	4.2	Siphovirus	3
BaGTA	<i>Bartonella</i> spp.	80	14	Unknown	4
DdGTA (formerly Dd1)	<i>Desulfovibrio desulfuricans</i>	17.8	13.5	Podovirus	5
BhGTA (formerly VSH-1)	<i>Brachyspira hyodysenteriae</i>	16.3	7.5	Tailed	6
MvGTA (formerly VTA)	<i>Methanococcus voltae</i>	14.2	4.4	Tailed	7

1: (Yen *et al.*, 1979; Westbye *et al.*, 2013; Hynes *et al.*, 2016; Bárdy *et al.*, 2020), 2 : (Gozzi *et al.*, 2022), 3 : (Tomasch *et al.*, 2018), 4 : (Berglund *et al.*, 2009; Québatte *et al.*, 2017), 5: (Rapp and Wall, 1987), 6: (Humphrey *et al.*, 1997; Matson *et al.*, 2005), 7: (Bertani, 1999; Eiserling *et al.*, 1999), 8: (Craske *et al.*, 2024).

For instance, *Ruegeria pomeroyi* possesses a complete RcGTA-like gene cluster and demonstrates gene transfer capabilities (Biers *et al.*, 2008). This species produces tail-less particles ranging from 50 to 70 nm in size, with its genome apparently free of other prophage sequences. Additionally, *Roseovarius nubinhibens*, one other marine Rhodobacterales species, has been shown to produce particles classified as GTAs (McDaniel *et al.*, 2010). Moreover, the Rhodobacterales *Dinoroseobacter shibae* produces RcGTA-like GTA particles, called DsGTA, and performs HGT mediated by GTAs (Tomasch *et al.*, 2018). Outside of Rhodobacterales, *Caulobacter crescentus*, which belongs to the Caulobacterales order, performed HGT with

CcGTA, which shares several characteristics with RcGTA (Gozzi *et al.*, 2022). While CcGTA and DsGTA share similarities with the canonical RcGTA model, they also exhibit some differences. This highlights the importance of studying diverse GTA systems to identify common and variable features across different bacterial species.

1.5.1 Mechanism

GTA are believed to have evolved from defective prophages during a host-driven domestication process (Lang *et al.*, 2017; Xu *et al.*, 2024). Unlike typical phages, which primarily package their own DNA, GTAs mainly package fragments of the bacterial genome (Lang *et al.*, 2012). Additionally, the length of the DNA fragments packaged into GTAs is insufficient to encompass the entire GTA locus (Table 1.2), meaning that GTAs cannot transfer themselves horizontally in one transfer event (Lang *et al.*, 2012). In *R. capsulatus* and other GTA-producing species, several genes essential for GTA production are located far away from the operon that encodes the main structural components of the GTA (Figure 1.3) (Craske *et al.*, 2024). Thus, GTAs are not infectious phage-like particles.

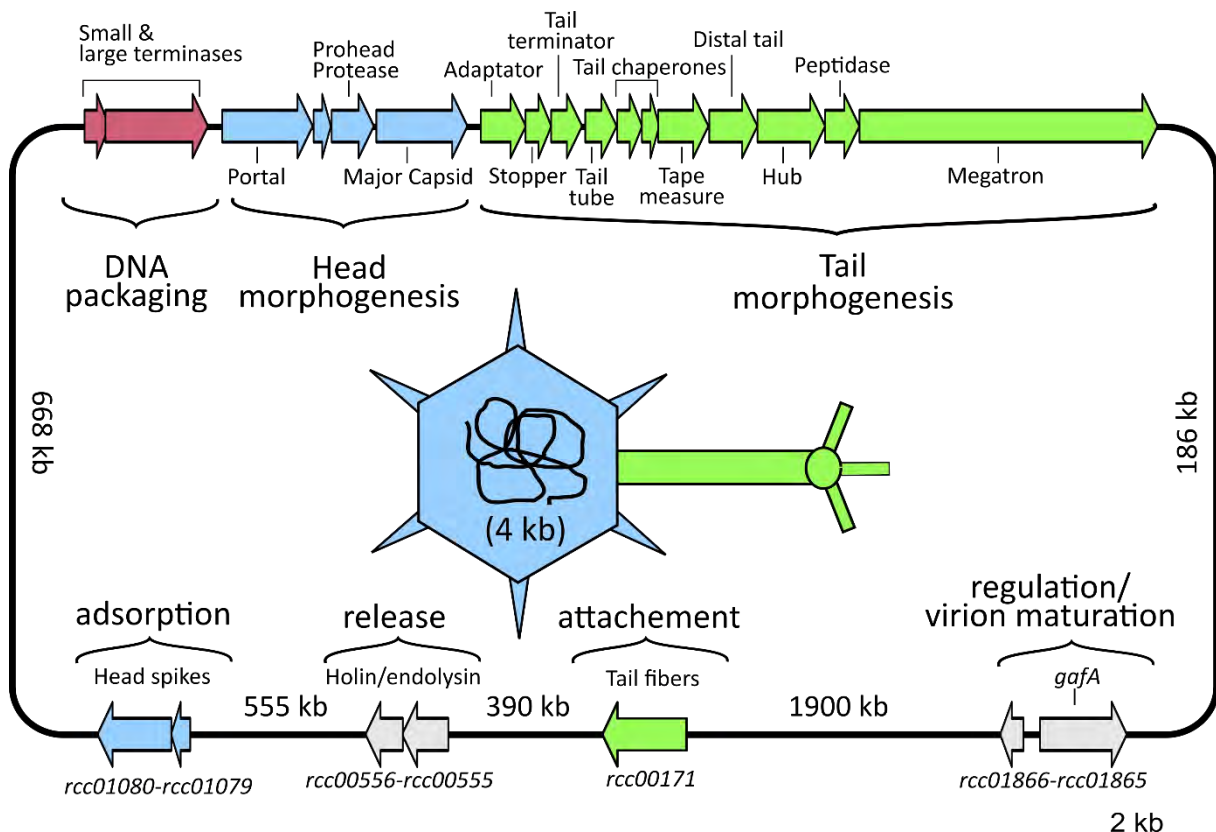


Figure 1.3 Genomic organization of GTA-related genes in *R. capsulatus*. The *R. capsulatus* genome is represented by a black line, with distances between the main GTA gene cluster and auxiliary operons shown to highlight the fact that GTA-associated genes are scattered along the genome. Arrows denote predicted ORF linked to RcGTA production, color-coded by broad functional categories: DNA packaging (red), head formation (blue), tail assembly (green), and GTA regulation and release (grey). The genome and particle are not drawn to scale.

HGT mediated by GTAs involves several steps (Figure 1.4) (Redfield *et al.*, 2019). In the donor cell, GTA production begins with the expression of the phage-like structural and DNA packaging genes (Figure 1.3). These components assemble to form the GTA particles, which package host DNA fragments. Once GTAs are assembled, holins and endolysins trigger cell lysis, which releases the GTA particles. GTA particles then attach to the surface of the recipient cell through specific interactions between tail fibers, head spikes, and the recipient's surface-associated proteins and capsule components. Next, DNA is translocated into the cytoplasm, thanks to natural competence-associated genes. Non-homologous DNA is degraded while homologous DNA then undergoes recombination via the RecA-mediated pathway, integrating into the recipient's genome. The cost of GTA production for the population is substantial since it results in the death of the donor cell. This raises intriguing questions about the evolutionary persistence of such a mechanism, suggesting that the benefits to the population must outweigh the cost of losing the producers.

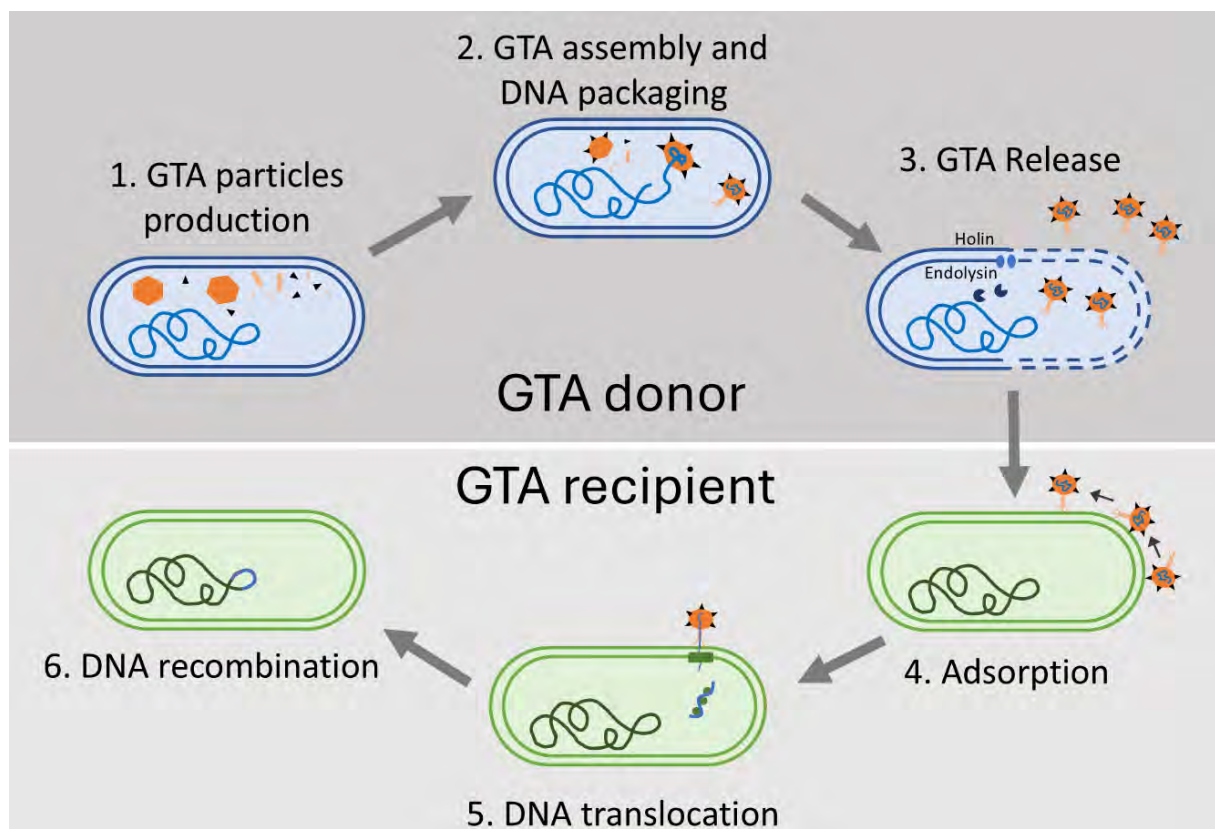


Figure 1.4 Typical steps in the production and reception of bacterial GTAs. (1) Transcription and translation of the GTA genes. (2) Assembly of GTA structural proteins and packaging of DNA fragments into heads and attachment of tails. (3) Lysis of the cell. (4) GTA particles attach to the recipient cell. (5) GTAs inject their DNA fragments and cellular proteins translocate the DNA across the inner membrane. (6) In the case of homology with the host genome, internalized DNA undergoes homologous recombination.

Given that the production of GTA particles requires cell lysis, the production of GTAs must be maintained under strict regulatory control to prevent total population extinction. Consequently, only a small proportion of cells within a population produce GTA particles with 0.1–3% of *R. capsulatus* cells and up to 6% of *Bartonella henselae* cells (Fogg *et al.*, 2012; Hynes *et al.*, 2012; Québatte *et al.*, 2017). On a population level, RcGTA production is triggered when the bacterial culture enters the stationary phase and is influenced by multiple factors (Solioz *et al.*, 1975). This includes nutrient scarcity, particularly carbon or amino acid deficiency (Westbye *et al.*, 2017). Additionally, quorum sensing plays also a role in regulating RcGTA (Schaefer *et al.*, 2002). Interestingly, BaGTA follows a different regulatory pattern, with production occurring exclusively in a subset of the most rapidly dividing cells (Québatte *et al.*, 2017).

Our understanding of GTA regulation on the cellular-level is mainly based on a RcGTA-centric model, given that the underlying regulatory mechanisms remain largely unidentified for the majority of known GTA-producing organisms (Banks and Le, 2024). In *R. capsulatus*, GTA production is intricately regulated by a complex network involving quorum sensing, the CtrA phosphorelay system, and the direct activator GafA (Figure 1.5) (Fogg, 2019; Sherlock and Fogg, 2022). Key to this regulation is the transcriptional activator GafA, which binds to the promoter region of the GTA operon, initiating the production of structural components essential for GTA assembly (Fogg, 2019).

Population density directly influences *gafA* expression through the LuxIR-type quorum sensing system (Schaefer *et al.*, 2002). The transcriptional repressor GtaR is de-repressed by binding to N-acyl-homoserine lactones (HSLs) produced by the HSL synthase GtaI. This de-repression allows *gafA* expression (Leung *et al.*, 2012; Fogg, 2019; Sherlock and Fogg, 2022). Quorum sensing is also important for RcGTA reception, via the induction of *dprA* expression (Brimacombe *et al.*, 2014).

CtrA plays a dual role in RcGTA regulation: unphosphorylated CtrA is required for *gafA* transcription, while phosphorylated CtrA is necessary for GTA maturation and lysis (Fogg, 2019). The activity of CtrA is modulated by cyclic-dimeric-GMP (c-di-GMP) levels through a phosphorelay signaling cascade involving CckA-ChpT-CtrA (Farrera-Calderon *et al.*, 2020). High intracellular c-di-GMP levels maintain CtrA in an unphosphorylated form, enhancing *gafA* induction through positive autoregulation (Farrera-Calderon *et al.*, 2020; Sherlock and Fogg,

2022). The concentration of c-di-GMP in *R. capsulatus* is controlled by diguanylate cyclase (DGC) enzymes, which produce c-di-GMP from GTP, and phosphodiesterase (PDE) enzymes, which degrade c-di-GMP, respectively (Pallegar *et al.*, 2020). This regulatory system ensures that RcGTA production is induced under specific conditions, such as high cell density or nutrient limitation, promoting gene transfer within bacterial communities (Marrs, 1974; Lang and Beatty, 2000; Westbye *et al.*, 2017).

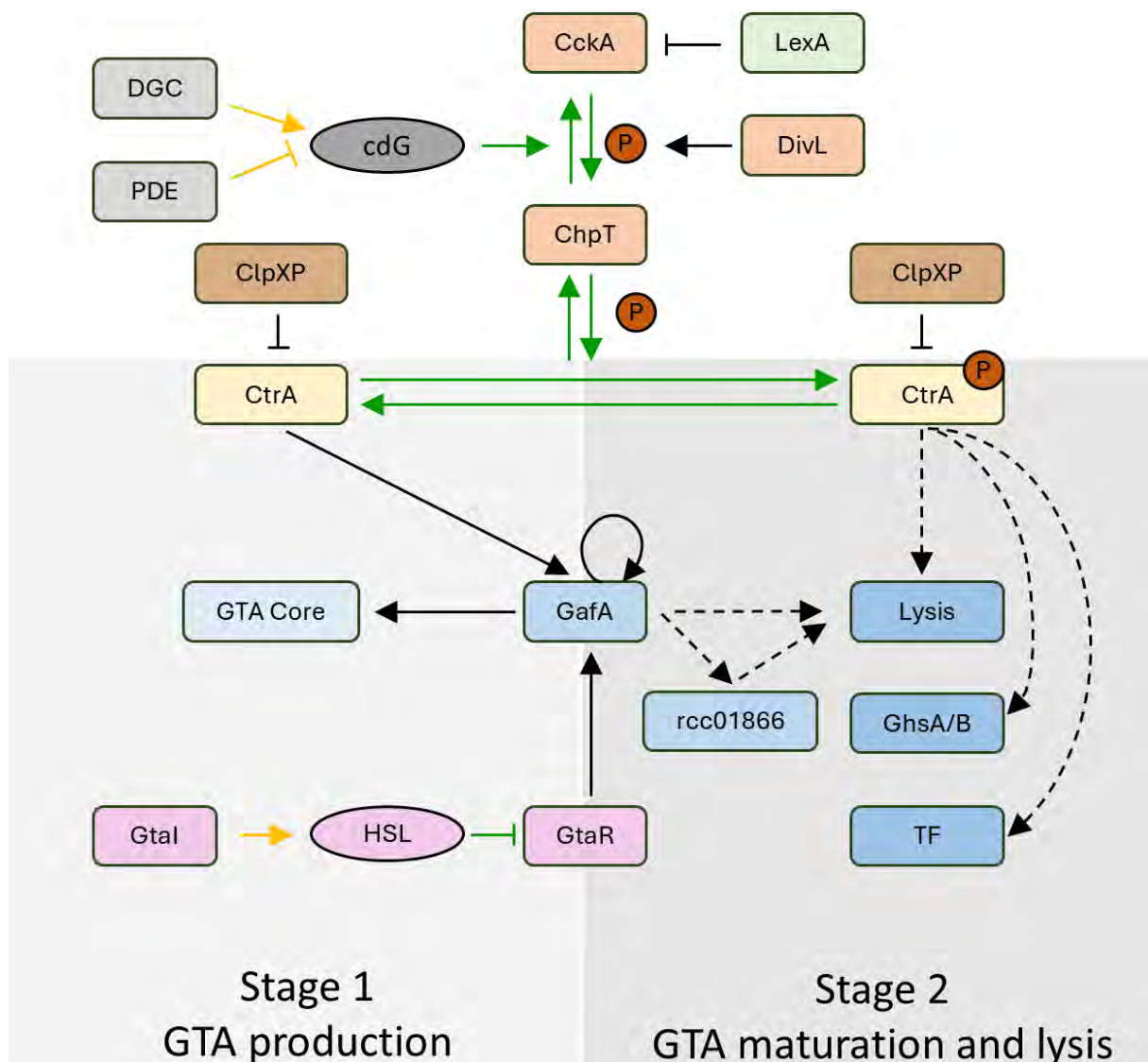


Figure 1.5 Simplified model of the regulatory network for RcGTAs production (Sherlock and Fogg, 2022). Contributors to this network are categorized according to their primary influence: those affecting the early production of structural proteins (stage 1) and those involved in late-stage maturation and lysis (stage 2). Arrows in the diagram signify different types of regulation; standard arrows indicate positive regulation, while flat-headed arrows denote repression. Black arrows represent transcriptional regulation, green arrows illustrate post-translational or ligand-protein interactions, orange arrows depict biosynthesis or degradation processes, and dashed arrows indicate mechanisms that remain uncertain. Box colours denote: grey for c-di-GMP (cdG) synthesis/degradation, beige/orange for phosphorelay system, blue for GTA specific genes and pink for quorum sensing system.

Despite these advances, important aspects of RcGTA regulation remain unresolved. Recent findings in *C. crescentus* and *Bartonella* reveal that transcriptional anti-termination plays a central role in GTA regulation (Tran and Le, 2024; Korotaev *et al.*, 2025). Because RcGTA/CcGTA and BaGTA originate from independent phage domestication events (Lang *et al.*, 2017; Tamarit *et al.*, 2018; Kogay *et al.*, 2022), their shared reliance on anti-termination suggests this mechanism may represent a common regulatory strategy across diverse GTAs, most likely inherited from phages given the shared structural similarity of the antiterminators GafZ (*C. crescentus*) and BrrG (*Bartonella*) with phage antiterminators (Korotaev *et al.*, 2025). Altogether, these insights underscore that GTA regulation remains an active area of research, with important discoveries continuing to reshape our understanding of horizontal gene transfer in bacteria.

1.5.2 Evolutionary functions of GTAs

Transfer mediated by GTAs stands out because of the high fitness cost imposed on the donor. Despite this cost, there is an increasing body of research suggesting that RcGTA-like elements have been maintained in bacterial genomes for an estimated period exceeding 700 million years (Shakya *et al.*, 2017). Interestingly, evidence suggests that evolutionary forces are actively preserving the functionality of these GTA genes across various species within the Rhodobacterales order (Lang *et al.*, 2012). Nevertheless, the specific advantages of GTA production and reception are not yet clear.

Researchers have proposed several potential reasons for the maintenance of GTAs in bacterial populations, including facilitating beneficial gene exchange, enhancing resilience to environmental stressors like DNA damage, and improving adaptation to nutrient scarcity (Shakya *et al.*, 2017). These potential benefits are similar to those hypothesized for natural transformation (Redfield, 1993b; Seitz and Blokesch, 2012). Unrelated to the transfer of DNA by GTAs, the GTA cluster might also be beneficial by providing protection against viral infections through superinfection immunity (Díaz-Muñoz, 2017; Shakya *et al.*, 2017).

Alternatively, the natural transformation mechanism was also hypothesized to confer a fitness advantage in changing environments through a process termed “genetic time travel”: by transforming with old genetic material kept in the environment, bacteria might revert to a previous genotype that could be beneficial in current or future environmental conditions (Engelstädter and Moradigaravand, 2013). Computer simulations indicate that cells capable

of transformation generally outcompete non-transforming cells across a broad spectrum of variables, including the rate at which DNA degrades in the environment, the frequency of transformation, and how often the environment changes (Engelstädter and Moradigaravand, 2013). Given their phage-like structure, GTAs likely provide a protective environment for encapsulated DNA, as suggested by nuclease resistance (Marrs, 1974). This could allow recipient bacteria to access preserved genetic material—effectively enabling genetic time travel. However, the long-term stability of GTA particles remains unstudied.

Limited experimental data exists regarding the potential benefits of GTAs. In *C. crescentus*, transfer mediated by GTAs enhances bacterial survival during the stationary phase and after DNA damage, by supplying recipient cells with genetic material for homologous recombination-based repair (Gozzi *et al.*, 2022). In *B. henselae*, the most rapidly dividing cells both produce and receive BaGTA, suggesting a strategy for genetic exchange among the fittest individuals in the population (Québatte *et al.*, 2017).

Despite these insights, there remains a significant gap in experimental evidence exploring GTAs benefits across different species. It is noteworthy that GTA-like elements may not confer a universal advantage across all bacterial species. Instead, diverse bacterial lineages might experience distinct selection pressures based on their specific ecological niches, resulting in varied benefits for maintaining these elements in different groups.

1.6 HGT mechanisms: specificities and implications

While all four HGT mechanisms mentioned previously—conjugation, transformation, transduction, and transfer mediated by GTAs—facilitate the exchange of genetic material between bacteria, they differ significantly in their mechanisms and requirements. Table 1.3 summarizes both convergent and divergent aspects of these systems.

The source of transferred DNA and the vector responsible for its delivery vary among these processes. Conjugation typically involves the transfer of plasmid or chromosomal DNA via a conjugative plasmid or ICE (Setlow *et al.*, 1984; Norman *et al.*, 2009). Transformation relies on the uptake of extracellular DNA, originating from plasmids or chromosomes without the involvement of a specific vector. Transduction uses bacteriophages to transfer chromosomal or plasmid DNA (Mašlačňová *et al.*, 2016; Valero-Rello *et al.*, 2017), whereas GTAs predominantly package chromosomal fragments into specialized particles (Québatte *et al.*,

2017; Tomasch *et al.*, 2018). The size of transferred DNA also varies across mechanisms: GTA-mediated transfer typically involves short fragments ranging from 4 kb to 14 kb, while natural transformation can accommodate fragments up to 150 kb in length (Matthey *et al.*, 2019; Banks and Le, 2024). Transfer of longer fragments and DNA from extracellular elements may provide greater opportunities for introducing genetic novelty.

Table 1.3 Comparative elements of bacterial conjugation, natural transformation, transduction, and transfer mediated by GTAs

	Conjugation	Transformation	Transduction	HGT mediated by GTAs
DNA Source	Chromosome or plasmid	Extracellular DNA (chromosome or plasmid)	Chromosome or plasmid	Chromosome fragments
Vector	Conj. plasmid or ICE	None	Bacteriophage	GTAs
DNA	F pilus or conj. pore	None	Bacteriophage	GTAs
Protection Constraints	- Induction of conjugation - Cell-to-cell contact - Donor-recipient compatibility	- Induction of competence - DNA homology	- Host range - DNA homology	- Host range - DNA homology
Controlled by	Vector	Host	Vector	Host

The type of DNA transfer also impacts the level of protection afforded to the genetic material, which has critical implications for its integrity and dissemination. Bacteriophages and GTAs offer robust protection by encapsulating DNA within proteinaceous particles, which shield the genetic material from environmental factors and nucleases (Marrs, 1974; Jończyk *et al.*, 2011). Conjugation provides also protection, as DNA is transferred through specialized structures such as the F pilus or a conjugative pore (Virolle *et al.*, 2020; Goldlust *et al.*, 2023), which safeguards genetic material during direct cell-to-cell contact. In contrast, extracellular DNA used for natural transformation is exposed to the environment, leaving it vulnerable to degradation by bacterial DNases, radiation, and temperature fluctuations (Levy-Booth *et al.*, 2007; Nielsen *et al.*, 2007). However, studies have demonstrated that free DNA can remain stable and retain transformability for extended periods under certain conditions, such as up to 10 weeks in soil environments (Romanowski *et al.*, 1992; Nielsen *et al.*, 2007). This differential protection influences the persistence and dissemination of genetic material across temporal and spatial scales, shaping bacterial evolutionary trajectories. Interestingly, some bacteria have strategies to mitigate the risks associated with transformation with degraded DNA. For example, *V. cholerae* and *S. pneumoniae* engage in predatory behavior towards

neighboring cells before performing natural transformation (Veening and Blokesch, 2017), potentially ensuring the quality and integrity of internalized genetic material.

The four HGT mechanisms exhibit varying constraints for successful genetic material transfer between donor and recipient cells, particularly during the initial stages. Natural transformation requires the recipient cell to enter a state of competence. Once competence is achieved, natural transformation demonstrates the highest level of permissiveness among HGT mechanisms regarding DNA internalization. Generally, the DNA uptake machinery in naturally competent bacteria lacks specificity for same-species DNA (Palmen *et al.*, 1993; Dubnau, 1999; Suckow *et al.*, 2011), allowing for a broad range of genetic material to be internalized. However, exceptions exist, such as *H. influenzae* and *N. gonorrhoeae*, which preferentially uptake species-specific DNA sequences (Sisco and Smith, 1979; Graves *et al.*, 1982). Nevertheless, some competent bacteria have strategies to enhance the probability of integrating species-specific DNA, for example by linking the regulation of natural competence and transformation to species-specific autoinducer molecules (Seitz and Blokesch, 2012). Upon release into the environment, phages and GTA particles face their initial challenge in genetic transfer: compatibility between host cell receptors and the transferring particles. Conjugation imposes additional constraints, by necessitating physical contact between donor and recipient cells in addition to the structural compatibility for stabilization of the mating pair (Frankel *et al.*, 2023). While the type of HGT can influence the initial transfer process, once the genetic material enters the recipient cell, its fate within the recipient cell and its long-term persistence in the population, are primarily determined by several key factors. These include the genetic content and methylation pattern of the transferred DNA, its interactions with the host genome, and the impact of the transferred genes on the recipient fitness across various environmental conditions.

The evolutionary drivers behind these mechanisms further reflect distinct selective pressures. Conjugation and transduction are primarily driven by the self-propagation interests of mobile genetic elements and phages. In contrast, natural transformation and GTA-mediated transfer appear to serve the genetic interests of the host bacterium. In this section, we have highlighted a few specific features of these four HGT mechanisms. Despite enabling a similar function—gene transfer—, these mechanisms exhibit vastly different requirements. These

differences underscore how each mechanism has evolved to address distinct ecological challenges.

1.7 Study models

1.7.1 *Pseudomonas stutzeri*, an old workhorse for new studies on natural transformation

Pseudomonas stutzeri is renowned for its metabolic adaptability and ecological resilience. This ubiquitous species has been extensively studied for its capabilities in bioremediation, nitrogen fixation, and pollutant degradation, highlighting its adaptability in diverse environmental niches (Lalucat *et al.*, 2006). Moreover, *P. stutzeri* was one of the early models of environmental bacteria to study natural transformation (Carlson *et al.*, 1983), with about one third of population isolates capable of natural transformation (Lorenz and Sikorski, 2000; Sikorski *et al.*, 2002).

In recent years, some studies have led to the reclassification of many *Pseudomonas* species that fall outside the core *Pseudomonas* group into several new genera. In particular, the group known as the “Stutzeri cluster” has been reassigned to the genus *Stutzerimonas* (Gomila *et al.*, 2022; Lalucat *et al.*, 2022; Rudra and Gupta, 2024). However, in this thesis, we use the former nomenclature, *Pseudomonas stutzeri*, for consistency with the majority of the existing literature and to facilitate comparison with previous studies.

Analysis of genetic diversity using multi-gene sequencing revealed that *P. stutzeri* exhibits a very high level of genetic variability among bacterial species (Cladera *et al.*, 2004). This remarkable genetic diversity and the presence of *P. stutzeri* in diverse environments may be attributed to two complementary mechanisms of genetic adaptation. Evolutionary adaptation allows a versatile ancestral strain to respond to divergent selection pressures in different ecological niches, leading to specialized variants. Simultaneously, HGT mechanisms enable the bacterium to acquire specific genes from other organisms, such as likely the catechol 2,3 dioxygenase genes (Cladera *et al.*, 2004), enhancing its metabolic capabilities and enabling its ability to thrive in a variety of habitats. These mechanisms collectively might have contributed to the species remarkable genetic plasticity and environmental versatility.

The ability of *P. stutzeri* to engage in HGT, particularly through natural transformation, has been a subject of significant scientific interest. Research on *P. stutzeri* natural transformation

has primarily focused on two key aspects: the mechanism of DNA uptake and the factors that influence the development of competence. The DNA uptake apparatus and DNA processing steps are known to be well conserved among the transformable species, and *P. stutzeri* does not stand as an exception (Table 1.1, Figure 1.2) (Johnston *et al.*, 2014).

Previous work into the regulatory mechanisms governing natural genetic transformation in *P. stutzeri* has essentially examined the influence of diverse environmental and physiological parameters on its transformation efficiency. Like in other species, the competence of *P. stutzeri* varies during growth with a maximum during the transition from exponential to stationary growth phase in minimal medium and a peak of competence at the beginning of the stationary phase in rich medium (Lorenz and Wackernagel, 1990, 1991). Nutrient availability plays a crucial role: transformation frequencies increase when phosphorus, nitrogen, or carbon limit growth, but severe nutrient restrictions may prevent competence development (Lorenz and Wackernagel, 1991, 1992). In addition, temperatures below 20°C or above 40°C and pH levels below 6.5 or above 8 highly reduce natural transformation (Lorenz and Wackernagel, 1990, 1992). Importantly, transformation has been also observed in soil microcosms for *P. stutzeri* (Sikorski *et al.*, 1998; Kittredge *et al.*, 2022). Interestingly, natural transformation of *P. fluorescens* was successfully demonstrated in soil, but not in various other conditions including two soil-based media (Demanèche *et al.*, 2001), suggesting that specific environmental factors in soil may trigger competence development. This observation aligns with findings in other species, such as *V. cholerae*, where competence induction involves multiple environmental cues and key regulators (Seitz and Blokesch, 2012). The regulatory mechanisms governing natural transformation in *P. stutzeri* remain largely unknown, presenting an opportunity for future research, particularly using next-generation sequencing methods, which have not yet been extensively applied to *P. stutzeri*.

1.7.2 The genus *Paracoccus*: the next generation models for studying GTAs

The genus *Paracoccus*, belonging to the alphaproteobacteria class, and more precisely to the Rhodobacterales order, comprises a diverse group of metabolically versatile bacteria (Baker *et al.*, 1998; Kelly *et al.*, 2006). These organisms have demonstrated a remarkable adaptability, thriving in a wide range of environmental niches such as soil, rhizosphere, groundwater, sea water, marine sediments, and sewage treatment plants; some species are even known to cause infections in humans (Daneshvar *et al.*, 2003; Lee *et al.*, 2004; Kelly *et al.*, 2006; Liu *et*

al., 2008; Lasek *et al.*, 2018). This widespread distribution across diverse habitats underscores their exceptional capacity for adaptation.

The identification of the RcGTA genes cluster in *R. capsulatus* led to the realization that similar genes are prevalent among alphaproteobacteria (Lang and Beatty, 2007). RcGTA-like genes have been found in all taxonomic orders of alphaproteobacteria with sequenced genomes (Lang *et al.*, 2012). The order Rhodobacterales shows particularly high conservation, with nearly every sequenced member containing a GTA gene cluster similar to that of *R. capsulatus*. This widespread distribution suggests that a GTA gene cluster existed in the common ancestor of this bacterial class (Shakya *et al.*, 2017), although some lineages may have lost it over time.

In addition to the mere presence of the GTA gene cluster in Rhodobacterales, there are also indications of GTAs production for *R. pomeroyi*, *R. nubinhibens* and *D. shibae* (Biers *et al.*, 2008; McDaniel *et al.*, 2010; Tomasch *et al.*, 2018). These findings indicate that multiples species from the Rhodobacterales order are capable of HGT mediated by GTAs.

Exploring the potential of *Paracoccus* species to produce GTAs and mediate HGT via GTAs represents a promising research direction for several reasons. As a member of the Rhodobacterales order, *Paracoccus* species are closely related to other species known to produce GTAs and likely contain the full RcGTA operon (Lang and Beatty, 2007). Additionally, numerous genetic tools are available given the extensive research on *Paracoccus* (Baker *et al.*, 1998; Pal *et al.*, 2024), allowing to dissect in details for instance GTAs production, reception and functions. Finally, the widespread distribution of *Paracoccus* across diverse ecological niches allow to study the complex interplay between environmental conditions and GTAs dynamics, where environmental pressures potentially modulate GTAs maintenance, while GTAs simultaneously facilitate bacterial adaptation.

1.8 Scope of the thesis

1.8.1 Motivations

This thesis explores two important mechanisms of HGT in environmental bacterial species: natural transformation in *P. stutzeri* and GTA-mediated transfer in *Paracoccus* genus.

The investigation of natural transformation in *P. stutzeri* was motivated by the opportunity to apply new techniques to this well-established model organism and to address fundamental

questions about the potential benefits of natural transformation. This includes examining whether *P. stutzeri* can use DNA as a nutrient source, studying the regulation of competence at both population and single-cell levels, and attempting to adapt transposon insertion sequencing for broad genetic analysis of natural transformation.

For GTA-mediated transfer, the motivation stemmed from the need to expand beyond the current model organism, *R. capsulatus*. By introducing *P. versutus* as a new model for studying GTA-mediated HGT, this research aims to distinguish between shared and species-specific characteristics of this mechanism. This approach allows for a more comprehensive understanding of GTAs biology and their role in bacterial evolution.

The thesis also explores related aspects of bacterial biology, including phage resistance mechanisms in *P. versutus*, which contributes to our understanding of bacteria-phage interactions and their potential impact on HGT.

By investigating these diverse HGT mechanisms, this research contributes to the broader field of evolutionary systems biology.

1.8.2 Research questions and general structure

This thesis investigates two types of HGT mechanisms, presented in two different parts. The first section explores the characteristics of natural transformation in *P. stutzeri*, while the second section focuses on another bacterial species, *P. versutus*. The second part includes the study of a new phage, the description of a resistance mechanism developed by *P. versutus* and introduces *P. versutus* as a new model for studying HGT mediated by GTA.

In Chapter 2, we examined whether *P. stutzeri* can utilize DNA as a nutrient source. After confirming that, we investigated the involvement of the competence apparatus and of an extracellular nuclease in this ability.

Chapter 3 presents studies on the regulation of natural transformation in *P. stutzeri* at both the population and single-cell level, using transcriptional reporter strains. This chapter also explores the role of a homolog of the master regulator of natural transformation TfoX.

In Chapter 4, we tried to adapt a transposon insertion sequencing method for *P. stutzeri* to broadly study the genetics behind natural transformation in this species from both the recipient and donor perspectives.

Chapter 5 introduces a newly isolated phage that infects *P. versutus* and details the investigation into a resistance mechanism against this phage.

In Chapter 6, we identified *P. versutus* as a new model to study HGT mediated by GTAs and explored aspects regarding regulation of this HGT mechanism.

Finally, Chapter 7 provides a comprehensive summary of the main findings, offers a general discussion of the dissertation, outlines future research perspectives, and concludes the study on HGT mechanisms in environmental bacterial species.

1.9 References

- Ambur, O.H., Engelstädter, J., Johnsen, P.J., Miller, E.L., and Rozen, D.E. (2016) Steady at the wheel: Conservative sex and the benefits of bacterial transformation. *Philos Trans R Soc B Biol Sci* **371**: 1–11.
- Baker, S.C., Ferguson, S.J., Ludwig, B., Page, M.D., Richter, O.-M.H., and van Spanning, R.J.M. (1998) Molecular Genetics of the Genus *Paracoccus* : Metabolically Versatile Bacteria with Bioenergetic Flexibility . *Microbiol Mol Biol Rev* **62**: 1046–1078.
- Baltrus, D.A. and Guillemin, K. (2006) Multiple phases of competence occur during the *Helicobacter pylori* growth cycle. *FEMS Microbiol Lett* **255**: 148–155.
- Baltrus, D.A., Guillemin, K., and Phillips, P.C. (2007) Natural transformation increases the rate of adaptation in the human pathogen *Helicobacter pylori*. *Evolution (N Y)* **62**: 39–49.
- Banks, E.J. and Le, T.B.K. (2024) Co-opting bacterial viruses for DNA exchange: structure and regulation of gene transfer agents. *Curr Opin Microbiol* **78**: 1–10.
- Bárdy, P., Füzik, T., Hrebík, D., Pantůček, R., Thomas Beatty, J., and Plevka, P. (2020) Structure and mechanism of DNA delivery of a gene transfer agent. *Nat Commun* **11**: 1–13.
- Bedzyk, L.A., Shoemaker, N.B., Young, K.E., and Salyers, A.A. (1992) Insertion and excision of *Bacteroides* conjugative chromosomal elements. *J Bacteriol* **174**: 166–172.
- Berglund, E.C., Frank, A.C., Calteau, A., Pettersson, O.V., Granberg, F., Eriksson, A.S., et al. (2009) Run-off replication of host-adaptability genes is associated with gene transfer agents in the genome of mouse-infecting *Bartonella grahamii*. *PLoS Genet* **5**: 1–16.
- Berka, R.M., Hahn, J., Albano, M., Draskovic, I., Persuh, M., Cui, X., et al. (2002) Microarray analysis of the *Bacillus subtilis* K-state: Genome-wide expression changes dependent on ComK. *Mol Microbiol* **43**: 1331–1345.
- Bertani, G. (1999) Transduction-like gene transfer in the methanogen *methanococcus voltae*. *J Bacteriol* **181**: 2992–3002.
- Biers, E.J., Wang, K., Pennington, C., Belas, R., Chen, F., and Moran, M.A. (2008) Occurrence and expression of gene transfer agent genes in marine bacterioplankton. *Appl Environ Microbiol* **74**: 2933–2939.
- Blokesch, M. and Schoolnik, G.K. (2008) The extracellular nuclease Dns and its role in natural transformation of *Vibrio cholerae*. *J Bacteriol* **190**: 7232–7240.
- Boto, L., Pineda, M., and Pineda, R. (2019) Potential impacts of horizontal gene transfer on human health and physiology and how anthropogenic activity can affect it. *FEBS J* **286**: 3959–3967.
- Brimacombe, C.A., Ding, H., and Beatty, J.T. (2014) *Rhodobacter capsulatus* DprA is essential for RecA-mediated gene transfer agent (RcGTA) recipient capability regulated by quorum-sensing and the CtrA response regulator. *Mol Microbiol* **92**: 1260–1278.

- Cabezón, E., Ripoll-Rozada, J., Peña, A., de la Cruz, F., and Arechaga, I. (2015) Towards an integrated model of bacterial conjugation. *FEMS Microbiol Rev* **39**: 81–95.
- Carlson, C.A., Pierson, L.S., Rosen, J.J., and Ingraham, J.L. (1983) *Pseudomonas stutzeri* and related species undergo natural transformation. *J Bacteriol* **153**: 93–99.
- Casjens, S.R. and Hendrix, R.W. (2015) Bacteriophage lambda: Early pioneer and still relevant. *Virology* **479–480**: 310–330.
- Charpentier, X., Kay, E., Schneider, D., and Shuman, H.A. (2011) Antibiotics and UV radiation induce competence for natural transformation in *Legionella pneumophila*. *J Bacteriol* **193**: 1114–1121.
- Chen, I. and Dubnau, D. (2004) DNA uptake during bacterial transformation. *Nat Rev Microbiol* **2**: 241–249.
- Cladera, A.M., Bennasar, A., Barceló, M., Lalucat, J., and García-Valdés, E. (2004) Comparative genetic diversity of *Pseudomonas stutzeri* genomovars, clonal structure, and phylogeny of the species. *J Bacteriol* **186**: 5239–5248.
- Claverys, J.-P. and Havarstein, L.S. (2002) Extracellular-peptide control of competence for genetic transformation in *Streptococcus pneumoniae*. *Front Biosci* **7**: 1798–1814.
- Claverys, J.P. and Martin, B. (2003) Bacterial “competence” genes: Signatures of active transformation, or only remnants? *Trends Microbiol* **11**: 161–165.
- Claverys, J.P., Prudhomme, M., and Martin, B. (2006) Induction of competence regulons as a general response to stress in gram-positive bacteria. *Annu Rev Microbiol* **60**: 451–475.
- Corbinais, C., Mathieu, A., Kortulewski, T., Radicella, J.P., and Marsin, S. (2016) Following transforming DNA in *Helicobacter pylori* from uptake to expression. *Mol Microbiol* **101**: 1039–1053.
- Craske, M.W., Wilson, J.S., and Fogg, P.C.M. (2024) Gene transfer agents: structural and functional properties of domesticated viruses. *Trends Microbiol* **32**: 1–12.
- Croucher, N.J., Mostowy, R., Wymant, C., Turner, P., Bentley, S.D., and Fraser, C. (2016) Horizontal DNA Transfer Mechanisms of Bacteria as Weapons of Intragenomic Conflict. *PLoS Biol* **14**: 1–42.
- Daneshvar, M.I., Hollis, D.G., Weyant, R.S., Steigerwalt, A.G., Whitney, A.M., Douglas, M.P., et al. (2003) *Paracoccus yeeii* sp. nov. (Formerly CDC group EO-2), a novel bacterial species associated with human infection. *J Clin Microbiol* **41**: 1289–1294.
- Demanèche, S., Kay, E., Gourbière, F., and Simonet, P. (2001) Natural Transformation of *Pseudomonas fluorescens* and *Agrobacterium tumefaciens* in Soil. *Appl Environ Microbiol* **67**: 2617–2621.
- Díaz-Muñoz, S.L. (2017) Viral coinfection is shaped by host ecology and virus-virus interactions across diverse microbial taxa and environments. *Virus Evol* **3**: 1–14.

- Dorer, M.S., Fero, J., and Salama, N.R. (2010) DNA damage triggers genetic exchange in *Helicobacter pylori*. *PLoS Pathog* **6**: 1–10.
- Dougherty, B.A. and Smith, H.O. (1999) Identification of *Haemophilus influenzae* Rd transformation genes using cassette mutagenesis. *Microbiology* **145**: 401–409.
- Dubnau, D. (1999) DNA Uptake in Bacteria. *Annu Rev Microbiol* **53**: 217–244.
- Dubnau, D. and Blokesch, M. (2019) Mechanisms of DNA Uptake by Naturally Competent Bacteria. *Annu Rev Genet* **53**: 217–237.
- Dubnau, D. and Provvedi, R. (2000) Internalizing DNA. *Res Microbiol* **151**: 475–480.
- Eiserling, F., Pushkin, A., Gingery, M., and Bertani, G. (1999) Bacteriophage-like particles associated with the gene transfer agent of *Methanococcus voltae* PS. *J Gen Virol* **80**: 3305–3308.
- Emamalipour, M., Seidi, K., Zununi Vahed, S., Jahanban-Esfahlan, A., Jaymand, M., Majdi, H., et al. (2020) Horizontal Gene Transfer: From Evolutionary Flexibility to Disease Progression. *Front Cell Dev Biol* **8**.
- Engelmoer, D.J.P., Donaldson, I., and Rozen, D.E. (2013) Conservative Sex and the Benefits of Transformation in *Streptococcus pneumoniae*. *PLoS Pathog* **9**: 3–10.
- Engelstädter, J. and Moradigaravand, D. (2013) Adaptation through genetic time travel? Fluctuating selection can drive the evolution of bacterial transformation. *Proc R Soc B Biol Sci* **281**: 1–7.
- Farrera-Calderon, R.G., Pallegar, P., Westbye, A.B., Wiesmann, C., Lang, A.S., and Beatty, J.T. (2020) The CckA-ChpT-CtrA phosphorelay controlling *Rhodobacter capsulatus* gene transfer agent production is bidirectional and regulated by cyclic di-GMP. *J Bacteriol* **203**: 1–17.
- Fillol-Salom, A., Alsaadi, A., de Sousa, J.A.M., Zhong, L., Foster, K.R., Rocha, E.P.C., et al. (2019) Bacteriophages benefit from generalized transduction. *PLoS Pathog* **15**: 1–22.
- Finkel, S.E. and Kolter, R. (2001) DNA as a nutrient: Novel role for bacterial competence gene homologs. *J Bacteriol* **183**: 6288–6293.
- Fogg, P.C.M. (2019) Identification and characterization of a direct activator of a gene transfer agent. *Nat Commun* **10**: 1–10.
- Fogg, P.C.M., Westbye, A.B., and Beatty, J.T. (2012) One for all or all for one: Heterogeneous expression and host cell lysis are key to gene transfer agent activity in *Rhodobacter capsulatus*. *PLoS One* **7**: 1–11.
- Fontaine, L., Wahl, A., Fléchar, M., Mignolet, J., and Hols, P. (2015) Regulation of competence for natural transformation in streptococci. *Infect Genet Evol* **33**: 343–360.
- Frankel, G., David, S., Low, W.W., Seddon, C., Wong, J.L.C., and Beis, K. (2023) Plasmids pick a bacterial partner before committing to conjugation. *Nucleic Acids Res* **51**: 8925–8933.

- Ghaly, T.M., Gillings, M.R., Rajabal, V., Paulsen, I.T., and Tetu, S.G. (2024) Horizontal gene transfer in plant microbiomes: integrons as hotspots for cross-species gene exchange. *Front Microbiol* **15**: 1–10.
- Gladyshev, E.A., Meselson, M., and Arkhipova, I.R. (2008) Massive horizontal gene transfer in bdelloid rotifers. *Science (80-)* **320**: 1210–1213.
- Gödeke, J., Heun, M., Bubendorfer, S., Paul, K., and Thormann, K.M. (2011) Roles of two *Shewanella oneidensis* MR-1 extracellular endonucleases. *Appl Environ Microbiol* **77**: 5342–5351.
- Goldlust, K., Ducret, A., Halte, M., Dedieu-Berne, A., Erhardt, M., and Lesterlin, C. (2023) The F pilus serves as a conduit for the DNA during conjugation between physically distant bacteria. *Proc Natl Acad Sci* **120**: 1–8.
- Gomila, M., Mulet, M., García-Valdés, E., and Lalucat, J. (2022) Genome-Based Taxonomy of the Genus *Stutzerimonas* and Proposal of *S. frequens* sp. nov. and *S. degradans* sp. nov. and Emended Descriptions of *S. perfectomarina* and *S. chloritidismutans*. *Microorganisms* **10**..
- Gozzi, K., Tran, N.T., Modell, J.W., Le, T.B.K.K., and Laub, M.T. (2022) Prophage-like gene transfer agents promote *Caulobacter crescentus* survival and DNA repair during stationary phase. *PLoS Biol* **20**: 1–34.
- Graupner, S., Frey, V., Hashemi, R., Lorenz, M.G., Brandes, G., and Wackernagel, W. (2000) Type IV pilus genes *pilA* and *pilC* of *Pseudomonas stutzeri* are required for natural genetic transformation, and *pilA* can be replaced by corresponding genes from nontransformable species. *J Bacteriol* **182**: 2184–2190.
- Graupner, S. and Wackernagel, W. (2001) *Pseudomonas stutzeri* has two closely related *pilA* genes (Type IV pilus structural protein) with opposite influences on natural genetic transformation. *J Bacteriol* **183**: 2359–2366.
- Graupner, S., Weger, N., Sohni, M., and Wackernagel, W. (2001) Requirement of novel competence genes *pilT* and *pilU* of *Pseudomonas stutzeri* for natural transformation and suppression of *pilT* deficiency by a hexahistidine tag on the type IV pilus protein *PilA*. *J Bacteriol* **183**: 4694–4701.
- Graves, J.F., Biswas, G.D., and Sparling, P.F. (1982) Sequence-specific DNA uptake in transformation of *Neisseria gonorrhoeae*. *J Bacteriol* **152**: 1071–1077.
- Hamoen, L.W., Venema, G., and Kuipers, O.P. (2003) Controlling competence in *Bacillus subtilis*: Shared use of regulators. *Microbiology* **149**: 9–17.
- Headd, B. and Bradford, S.A. (2018) Physicochemical factors that favor conjugation of an antibiotic resistant plasmid in non-growing bacterial cultures in the absence and presence of antibiotics. *Front Microbiol* **9**: 1–14.

- Herriott, R.M., Meyer, E.M., and Vogt, M. (1970) Defined nongrowth media for stage II development of competence in *Haemophilus influenzae*. *J Bacteriol* **101**: 517–524.
- Heun, M., Binnenkade, L., Kreienbaum, M., and Thormann, K.M. (2012) Functional specificity of extracellular nucleases of *Shewanella oneidensis* MR-1. *Appl Environ Microbiol* **78**: 4400–4411.
- Hoelzer, M.A. and Michod, R.E. (1991) DNA repair and the evolution of transformation in *Bacillus subtilis*. III. Sex with damaged DNA. *Genetics* **128**: 215–223.
- Hofreuter, D., Odenbreit, S., and Haas, R. (2001) Natural transformation competence in *Helicobacter pylori* is mediated by the basic components of a type IV secretion system. *Mol Microbiol* **41**: 379–391.
- Howard-Varona, C., Hargreaves, K.R., Abedon, S.T., and Sullivan, M.B. (2017) Lysogeny in nature: Mechanisms, impact and ecology of temperate phages. *ISME J* **11**: 1511–1520.
- Humphrey, S.B., Stanton, T.B., Jensen, N.S., and Zuerner, R.L. (1997) Purification and characterization of VSH-1, a generalized transducing bacteriophage of *Serpulina hyodysenteriae*. *J Bacteriol* **179**: 323–329.
- Hynes, A.P., Mercer, R.G., Watton, D.E., Buckley, C.B., and Lang, A.S. (2012) DNA packaging bias and differential expression of gene transfer agent genes within a population during production and release of the *Rhodobacter capsulatus* gene transfer agent, RcGTA. *Mol Microbiol* **85**: 314–325.
- Hynes, A.P., Shakya, M., Mercer, R.G., Gröll, M.P., Bown, L., Davidson, F., et al. (2016) Functional and Evolutionary Characterization of a Gene Transfer Agent's Multilocus "genome." *Mol Biol Evol* **33**: 2530–2543.
- Ikeda, H. and Tomizawa, J. (1965) Transducing fragments in generalized transduction by phage P1: I. Molecular origin of the fragments. *J Mol Biol* **14**: 85–109.
- Johnson, C.M. and Grossman, A.D. (2015) Integrative and Conjugative Elements (ICEs): What They Do and How They Work. *Annu Rev Genet* **49**: 577–601.
- Johnston, C., Martin, B., Fichant, G., Polard, P., and Claverys, J.P. (2014) Bacterial transformation: Distribution, shared mechanisms and divergent control. *Nat Rev Microbiol* **12**: 181–196.
- Jończyk, E., Kłak, M., Międzybrodzki, R., and Górski, A. (2011) The influence of external factors on bacteriophages-review. *Folia Microbiol (Praha)* **56**: 191–200.
- Kelly, D.P., Rainey, F.A., and Wood, A.N.N.P. (2006) The Genus *Paracoccus*. *The Prokaryotes* **5**: 232–249.
- Kim, J.-S., Kim, J.-W., and Kathariou, S. (2008) Differential Effects of Temperature on Natural Transformation to Erythromycin and Nalidixic Acid Resistance in *Campylobacter coli*. *Appl Environ Microbiol* **74**: 6121–6125.

- Kittredge, H.A., Dougherty, K.M., and Evans, S.E. (2022) Dead but Not Forgotten: How Extracellular DNA, Moisture, and Space Modulate the Horizontal Transfer of Extracellular Antibiotic Resistance Genes in Soil. *Appl Environ Microbiol* **88**: 1–11.
- Kogay, R., Koppenhöfer, S., Beatty, J.T., Kuhn, J.H., Lang, A.S., and Zhaxybayeva, O. (2022) Formal recognition and classification of gene transfer agents as viriforms. *Virus Evol* **8**: 1–13.
- Kohler, V., Keller, W., and Grohmann, E. (2019) Regulation of gram-positive conjugation. *Front Microbiol* **10**: 1–13.
- Koomey, M. (1998) Competence for natural transformation in *Neisseria gonorrhoeae*: A model system for studies of horizontal gene transfer. *APMIS, Suppl* **106**: 56–61.
- Koraimann, G. (2018) Spread and Persistence of Virulence and Antibiotic Resistance Genes: A Ride on the F Plasmid Conjugation Module. *EcoSal Plus* **8**: 1–23.
- Koraimann, G. and Wagner, M.A. (2014) Social behavior and decision making in bacterial conjugation. *Front Cell Infect Microbiol* **4**: 1–7.
- Korotaev, A., Niggli, Q., Congedi, V., and Dehio, C. (2025) Transcriptional antitermination integrates the expression of loci of diverse phage origin in the chimeric *Bartonella* Gene Transfer Agent BaGTA. *Elife* 1–37.
- Lalucat, J., Bennasar, A., Bosch, R., García-Valdés, E., and Palleroni, N.J. (2006) Biology of *Pseudomonas stutzeri*. *Microbiol Mol Biol Rev* **70**: 510–547.
- Lalucat, J., Gomila, M., Mulet, M., Zaruma, A., and García-Valdés, E. (2022) Past, present and future of the boundaries of the *Pseudomonas* genus: Proposal of *Stutzerimonas* gen. Nov. *Syst Appl Microbiol* **45**: 126289.
- Lang, A.S. and Beatty, J.T. (2000) Genetic analysis of a bacterial genetic exchange element: The gene transfer agent of *Rhodobacter capsulatus*. *Proc Natl Acad Sci U S A* **97**: 859–864.
- Lang, A.S. and Beatty, J.T. (2007) Importance of widespread gene transfer agent genes in alpha-proteobacteria. *Trends Microbiol* **15**: 1–9.
- Lang, A.S., Westbye, A.B., and Beatty, J.T. (2017) The Distribution, Evolution, and Roles of Gene Transfer Agents in Prokaryotic Genetic Exchange. *Annu Rev Virol* **4**: 87–104.
- Lang, A.S., Zhaxybayeva, O., and Beatty, J.T. (2012) Gene transfer agents: Phage-like elements of genetic exchange. *Nat Rev Microbiol* **10**: 472–482.
- Lasek, R., Szuplewska, M., Mitura, M., Decewicz, P., Chmielowska, C., Pawlot, A., et al. (2018) Genome structure of the opportunistic pathogen *Paracoccus yeei* (Alphaproteobacteria) and identification of putative virulence factors. *Front Microbiol* **9**: 1–17.
- Lee, J.H., Kim, Y.S., Choi, T.J., Lee, W.J., and Kim, Y.T. (2004) *Paracoccus haeundaensis* sp. nov., a Gram-negative, halophilic, astaxanthin-producing bacterium. *Int J Syst Evol Microbiol* **54**: 1699–1702.

- Leung, M.M., Brimacombe, C.A., Spiegelman, G.B., and Beatty, J.T. (2012) The GtaR protein negatively regulates transcription of the gtaRI operon and modulates gene transfer agent (RcGTA) expression in *Rhodobacter capsulatus*. *Mol Microbiol* **83**: 759–774.
- Levin, B.R. and Cornejo, O.E. (2009) The population and evolutionary dynamics of homologous gene recombination in bacteria. *PLoS Genet* **5**: 1000601.
- Levy-Booth, D.J., Campbell, R.G., Gulden, R.H., Hart, M.M., Powell, J.R., Klironomos, J.N., et al. (2007) Cycling of extracellular DNA in the soil environment. *Soil Biol Biochem* **39**: 2977–2991.
- Liu, M., Huang, M., Wang, M., Zhu, D., Jia, R., Chen, S., et al. (2020) The Clustered Regularly Interspaced Short Palindromic Repeat System and Argonaute: An Emerging Bacterial Immunity System for Defense Against Natural Transformation? *Front Microbiol* **11**: 1–9.
- Liu, Z.P., Wang, B.J., Liu, X.Y., Dai, X., Liu, Y.H., and Liu, S.J. (2008) *Paracoccus halophilus* sp. nov., isolated from marine sediment of the South China Sea, China and emended description of genus *Paracoccus* Davis 1969. *Int J Syst Evol Microbiol* **58**: 257–261.
- Lorenz, M.G. and Sikorski, J. (2000) The potential for intraspecific horizontal gene exchange by natural genetic transformation: Sexual isolation among genomovars of *Pseudomonas stutzeri*. *Microbiology* **146**: 3081–3090.
- Lorenz, M.G. and Wackernagel, W. (1994) Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol Rev* **58**: 563–602.
- Lorenz, M.G. and Wackernagel, W. (1991) High frequency of natural genetic transformation of *Pseudomonas stutzeri* in soil extract supplemented with a carbon/energy and phosphorus source. *Appl Environ Microbiol* **57**: 1246–1251.
- Lorenz, M.G. and Wackernagel, W. (1990) Natural genetic transformation of *Pseudomonas stutzeri* by sand-adsorbed DNA. *Arch Microbiol* **154**: 380–385.
- Lorenz, M.G. and Wackernagel, W. (1992) Stimulation of natural genetic transformation of *Pseudomonas stutzeri* in extracts of various soils by nitrogen or phosphorus limitation and influence of temperature and pH. *Microb Releases* **1**: 173–176.
- Marrs, B. (1974) Genetic recombination in *Rhodopseudomonas capsulata*. *Proc Natl Acad Sci U S A* **71**: 971–973.
- Mašlaňová, I., Stríbná, S., Doškař, J., and Pantůček, R. (2016) Efficient plasmid transduction to *Staphylococcus aureus* strains insensitive to the lytic action of transducing phage. *FEMS Microbiol Lett* **363**: 1–12.
- Matson, E.G., Thompson, M.G., Humphrey, S.B., Zuerner, R.L., and Stanton, T.B. (2005) Identification of genes of VSH-1, a prophage-like gene transfer agent of *Brachyspira hyodysenteriae*. *J Bacteriol* **187**: 5885–5892.
- Matthey, N. and Blokesch, M. (2016) The DNA-Uptake Process of Naturally Competent *Vibrio cholerae*. *Trends Microbiol* **24**: 98–110.

- Matthey, N., Stutzmann, S., Stoudmann, C., Guex, N., Iseli, C., and Blokesch, M. (2019) Neighbor predation linked to natural competence fosters the transfer of large genomic regions in *Vibrio cholerae*. *Elife* **8**: 1–21.
- McDaniel, L.D., Young, E., Delaney, J., Ruhnau, F., Ritchie, K.B., and Paul, J.H. (2010) High frequency of horizontal gene transfer in the oceans. *Science* (80-) **330**: 50.
- McDaniel, L.D., Young, E.C., Ritchie, K.B., and Paul, J.H. (2012) Environmental factors influencing gene transfer agent (GTA) mediated transduction in the subtropical ocean. *PLoS One* **7**: 1–9.
- Metzger, L.C. and Blokesch, M. (2016) Regulation of competence-mediated horizontal gene transfer in the natural habitat of *Vibrio cholerae*. *Curr Opin Microbiol* **30**: 1–7.
- Michod, R.E. and Wojciechowski, M.F. (1988) DNA repair and the evolution of transformation in the Bacterium *Bacillus subtilis*. *Genetics* **118**: 31–39.
- Moradigaravand, D. and Engelstädter, J. (2013) The evolution of natural competence: Disentangling costs and benefits of sex in bacteria. *Am Nat* **182**: 1–15.
- Mulcahy, H., Charron-Mazenod, L., and Lewenza, S. (2010) *Pseudomonas aeruginosa* produces an extracellular deoxyribonuclease that is required for utilization of DNA as a nutrient source. *Environ Microbiol* **12**: 1621–1629.
- Nielsen, K.M., Johnsen, P.J., Bensasson, D., and Daffonchio, D. (2007) Release and persistence of extracellular DNA in the environment. *Environ Biosafety Res* **6**: 37–53.
- Nolan, L.M., Turnbull, L., Katrib, M., Osvath, S.R., Losa, D., Lazenby, J.J., and Whitchurch, C.B. (2020) *Pseudomonas aeruginosa* is capable of natural transformation in biofilms. *Microbiol (United Kingdom)* **166**: 995–1003.
- Norman, A., Hansen, L.H., and Sørensen, S.J. (2009) Conjugative plasmids: Vessels of the communal gene pool. *Philos Trans R Soc B Biol Sci* **364**: 2275–2289.
- Pal, U., Bachmann, D., Pelzer, C., Christiansen, J., Blank, L.M., and Tiso, T. (2024) A genetic toolbox to empower *Paracoccus pantotrophus* DSM 2944 as a metabolically versatile SynBio chassis. *Microb Cell Fact* **23**: 1–17.
- Palchevskiy, V. and Finkel, S.E. (2006) *Escherichia coli* competence gene homologs are essential for competitive fitness and the use of DNA as a nutrient. *J Bacteriol* **188**: 3902–3910.
- Pallegar, P., Langille, E., Gomelsky, M., and Lang, A.S. (2020) Cyclic di-GMP-Mediated Regulation of Gene Transfer and Motility in *Rhodobacter capsulatus*. **202**: 1–17.
- Palmen, R., Vosman, B., Buijsman, P., Breek, C.K.D., and Hellingwerf, K.J. (1993) Physiological characterization of natural transformation in *Acinetobacter calcoaceticus*. *J Gen Microbiol* **139**: 295–305.

- Prudhomme, M., Attaiech, L., Sanchez, G., Martin, B., and Claverys, J.-P. (2006) Antibiotic Stress Induces Genetic Transformability in the Human Pathogen *Streptococcus pneumoniae*. *Science (80-)* **313**: 89–92.
- Québatte, M., Christen, M., Harms, A., Körner, J., Christen, B., and Dehio, C. (2017) Gene Transfer Agent Promotes Evolvability within the Fittest Subpopulation of a Bacterial Pathogen. *Cell Syst* **4**: 611–621.
- Rapp, B.J. and Wall, J.D. (1987) Genetic transfer in *Desulfovibrio desulfuricans*. *Proc Natl Acad Sci* **84**: 9128–9130.
- Redfield, R.J. (1993a) Evolution of natural transformation: Testing the DNA repair hypothesis in *Bacillus subtilis* and *Haemophilus influenzae*. *Genetics* **133**: 755–761.
- Redfield, R.J. (1993b) Genes for breakfast: The have-your-cake and-eat-it-too of bacterial transformation. *J Hered* **84**: 400–404.
- Redfield, R.J. (1991) *sxy-1*, a *Haemophilus influenzae* Mutation Causing Greatly Enhanced Spontaneous Competence. *J Bacteriol* **173**: 5612–5618.
- Redfield, R.J., Beatty, J.T., and Lang, A.S. (2019) Gene transfer agents. *Encycl Microbiol* **1**: 370–377.
- Redfield, R.J., Cameron, A.D.S., Qian, Q., Hinds, J., Ali, T.R., Kroll, J.S., and Langford, P.R. (2005) A novel CRP-dependent regulon controls expression of competence genes in *Haemophilus influenzae*. *J Mol Biol* **347**: 735–747.
- Redfield, R.J., Findlay, W.A., Bossé, J., Kroll, J.S., Cameron, A.D.S., and Nash, J.H.E. (2006) Evolution of competence and DNA uptake specificity in the Pasteurellaceae. *BMC Evol Biol* **6**: 1–15.
- Romanowski, G., Lorenz, M.G., Sayler, G., and Wackernagel, W. (1992) Persistence of free plasmid DNA in soil monitored by various methods, including a transformation assay. *Appl Environ Microbiol* **58**: 3012–3019.
- Rudra, B. and Gupta, R.S. (2024) Phylogenomics studies and molecular markers reliably demarcate genus *Pseudomonas* sensu stricto and twelve other Pseudomonadaceae species clades representing novel and emended genera. *Front Microbiol* **14**: 1–30.
- Schaefer, A.L., Taylor, T.A., Beatty, J.T., and Greenberg, E.P. (2002) Long-chain acyl-homoserine lactone quorum-sensing regulation of *Rhodobacter capsulatus* gene transfer agent production. *J Bacteriol* **184**: 6515–6521.
- Seitz, P. and Blokesch, M. (2012) Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. *FEMS Microbiol Rev* **37**: 336–363.
- Seitz, P. and Blokesch, M. (2013) DNA-uptake machinery of naturally competent *Vibrio cholerae*. *Proc Natl Acad Sci U S A* **110**: 17987–17992.

- Seper, A., Fengler, V.H.I., Roier, S., Wolinski, H., Kohlwein, S.D., Bishop, A.L., et al. (2011) Extracellular nucleases and extracellular DNA play important roles in *Vibrio cholerae* biofilm formation. *Mol Microbiol* **82**: 1015–1037.
- Setlow, J.K., Cabrera-Juarez, E., and Griffin, K. (1984) Mechanism of acquisition of chromosomal markers by plasmids in *Haemophilus influenzae*. *J Bacteriol* **160**: 662–667.
- Shakya, M., Soucy, S.M., and Zhaxybayeva, O. (2017) Insights into origin and evolution of α -proteobacterial gene transfer agents. *Virus Evol* **3**: 1–13.
- Sharma, D.K., Misra, H.S., Bihani, S.C., and Rajpurohit, Y.S. (2023) Biochemical Properties and Roles of DprA Protein in Bacterial Natural Transformation, Virulence, and Pilin Variation. *J Bacteriol* **205**: 1–16.
- Sherlock, D. and Fogg, P.C.M. (2022) The archetypal gene transfer agent RcGTA is regulated via direct interaction with the enigmatic RNA polymerase omega subunit. *Cell Rep* **40**: 1–19.
- Sikorski, J., Graupner, S., Lorenz, M.G., and Wackernagel, W. (1998) Natural genetic transformation of *Pseudomonas stutzeri* in a non-sterile soil. *Microbiology* **144**: 569–576.
- Sikorski, J., Teschner, N., and Wackernagel, W. (2002) Highly different levels of natural transformation are associated with genomic subgroups within a local population of *Pseudomonas stutzeri* from soil. *Appl Environ Microbiol* **68**: 865–873.
- Sinha, S., Mell, J., and Redfield, R.J. (2013) The availability of purine nucleotides regulates natural competence by controlling translation of the competence activator Sxy. *Mol Microbiol* **88**: 1106–1119.
- Sisco, K.L. and Smith, H.O. (1979) Sequence-specific DNA uptake in *Haemophilus* transformation *Microbiology : Proc Natl Acad Sci* **76**: 972–976.
- Solioz, M., Yen, H.-C.C., and Marrs, B. (1975) Release and uptake of gene transfer agent by *Rhodopseudomonas capsulata*. *J Bacteriol* **123**: 651–657.
- Sparling, P.F. (1966) Genetic transformation of *Neisseria gonorrhoeae* to streptomycin resistance. *J Bacteriol* **92**: 1364–1371.
- Suckow, G., Seitz, P., and Blokesch, M. (2011) Quorum Sensing Contributes to Natural Transformation of *Vibrio cholerae* in a Species-Specific Manner. *J Bacteriol* **193**: 4914–4924.
- Tamarit, D., Neuvonen, M.M., Engel, P., Guy, L., and Andersson, S.G.E. (2018) Origin and evolution of the bartonella gene transfer agent. *Mol Biol Evol* **35**: 451–464.
- Tomasch, J., Wang, H., Hall, A.T.K., Patzelt, D., Preusse, M., Petersen, J., et al. (2018) Packaging of *Dinoroseobacter shibae* DNA into gene transfer agent particles is not random. *Genome Biol Evol* **10**: 356–369.
- Tran, N.T. and Le, T.B.K. (2024) Control of a gene transfer agent cluster in *Caulobacter*

- crescentus by transcriptional activation and anti-termination. *Nat Commun* **15**:
- Treangen, T.J., Ambur, O.H., Tonjum, T., and Rocha, E.P.C. (2008) The impact of the neisserial DNA uptake sequences on genome evolution and stability. *Genome Biol* **9**: 1–17.
- Valero-Rello, A., López-Sanz, M., Quevedo-Olmos, A., Sorokin, A., and Ayora, S. (2017) Molecular mechanisms that contribute to horizontal transfer of plasmids by the bacteriophage SPP1. *Front Microbiol* **8**: 1–13.
- Veening, J.W. and Blokesch, M. (2017) Interbacterial predation as a strategy for DNA acquisition in naturally competent bacteria. *Nat Rev Microbiol* **15**: 621–629.
- Villa, T.G., Feijoo-Siota, L., Sánchez-Pérez, A., Rama, J.R., and Sieiro, C. (2019) Horizontal Gene Transfer in Bacteria, an Overview of the Mechanisms Involved, Cam Springer.
- Virolle, C., Goldlust, K., Djermoun, S., Bigot, S., and Lesterlin, C. (2020) Plasmid transfer by conjugation in gram-negative bacteria: From the cellular to the community level. *Genes (Basel)* **11**: 1–33.
- Westbye, A., O'Neill, Z., Schellenberg-Beaver, T., and Beatty, T.J. (2017) The *Rhodobacter capsulatus* gene transfer agent is induced by nutrient depletion and the RNAP omega subunit. *Microbiology* **163**: 1–9.
- Westbye, A.B., Leung, M.M., Florizone, S.M., Taylor, T.A., Johnson, J.A., Fogg, P.C., and Beatty, J.T. (2013) Phosphate concentration and the putative sensor kinase protein CckA modulate cell lysis and release of the *Rhodobacter capsulatus* gene transfer agent. *J Bacteriol* **195**: 5025–5040.
- Von Wintersdorff, C.J.H., Penders, J., Van Niekerk, J.M., Mills, N.D., Majumder, S., Van Alphen, L.B., et al. (2016) Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Front Microbiol* **7**: 1–11.
- Wise, E.M., Alexander, S.P., and Powers, M. (1973) Adenosine 3':5'-cyclic monophosphate as a regulator of bacterial transformation. *Proc Natl Acad Sci U S A* **70**: 471–474.
- Wojciechowski, M.F., Hoelzer, M.A., and Michod, R.E. (1989) DNA repair and the evolution of transformation in *Bacillus subtilis*. II. Role of Inducible Repair. *Genetics* **128**: 215–223.
- Xu, Y., Liu, B., Jiao, N., Liu, J., and Chen, F. (2024) New evidence supports the prophage origin of RcGTA. *Appl Environ Microbiol* **90**: 1–14.
- Yen, H.C., Hu, N.T., and Marrs, B.L. (1979) Characterization of the gene transfer agent made by an overproducer mutant of *Rhodopseudomonas capsulata*. *J Mol Biol* **131**: 157–168.
- Zhang, M., Zhang, T., Yu, M., Chen, Y.L., and Jin, M. (2022) The Life Cycle Transitions of Temperate Phages: Regulating Factors and Potential Ecological Implications. *Viruses* **14**: 1–20.

Part I:

**Natural transformation of
*Pseudomonas stutzeri***

Chapter 2 - Utilization of DNA as a nutrient source by *P. stutzeri*

2.1 Foreword

In this part of the project, I investigated whether *P. stutzeri* can use DNA as a source of nutrients. In addition, I determined whether DNA was degraded inside the cell after uptake via the natural transformation apparatus or outside of the cell by a homologue of EndA DNase.

The key findings of this chapter are the following:

- *P. stutzeri* can use DNA as a source of carbon or nitrogen, but not phosphorus.
- *P. stutzeri* benefits more from self-DNA than foreign DNA for DNA-based nutrition.
- The ability of *P. stutzeri* to use DNA as a nutrient source does not likely rely on the competence apparatus.
- Deletion of the nuclease EndA does not impair DNA-based growth in *P. stutzeri*.

2.2 Abstract

DNA, beyond its canonical role as a carrier of genetic information, serves multiple other roles, including providing essential nutrients that support bacterial growth. This DNA-based nutrition is typically achieved through the action of nucleases, such as EndA, which degrade extracellular DNA into smaller building blocks components that are then taken up by the cell. Alternatively, a model demonstrated in *Escherichia coli* involves the competence apparatus importing DNA, followed by intracellular degradation. However, evidence supporting this dual function of the competence apparatus is limited, particularly in other species. In this study, we investigated the utilization of DNA as a nutrient source in *Pseudomonas stutzeri*, a versatile bacterium capable of natural transformation. Our results demonstrate that *P. stutzeri* can grow using DNA as the sole source of carbon or nitrogen. Additionally, we observed a distinct preference for self-DNA over DNA derived from other species, indicating a nuanced response to different DNA substrates. Contrary to expectations, deletion of genes involved in natural competence and of a homologue of the extracellular nuclease EndA, which is involved in DNA consumption in *Shewanella oneidensis* and *Vibrio cholerae*, did not impair the ability of *P. stutzeri* to grow on DNA, suggesting that they did not play an essential role in this DNA-based nutritional phenotype. Other nucleases may therefore be involved and should be further investigated, including both extracellular candidates and intracellular nucleases that could be released upon cell lysis.. Overall, our study provides a comprehensive examination of DNA utilization as a nutrient source by *P. stutzeri*, laying the groundwork for future research to uncover the broader ecological and evolutionary implications of this intriguing process.

2.3 Introduction

DNA is one of the fundamental molecules of life, carrying genetic information in all kingdoms of life. DNA is also an environmentally abundant organic substrate in both aquatic and terrestrial ecosystems due to secretion by living bacteria through the type IV secretion system, vesicles, or release by dead cells following autolysis, viral infection, antimicrobial agents, or predation, among others (Hamilton *et al.*, 2005; Nielsen *et al.*, 2007; Liao *et al.*, 2014).

Over the past decades, extracellular DNA (eDNA) has been shown to play multiple roles for bacteria. As a carrier of genetic information, it is involved in gene transfer through natural transformation. It might also provide support for DNA damage repair (Michod and Wojciechowski, 1988). In addition, eDNA plays a role in structuring biofilms (Shak *et al.*, 1990; Whitchurch *et al.*, 2002). As a polymer, eDNA confers viscosity and interacts with various molecules due to the negative charges on its backbone (Shak *et al.*, 1990; Pisetsky, 2017). Biofilms protect bacteria from harsh environmental conditions such as pH and temperature changes, UV light, and high salinity but also from antimicrobial agents (Yin *et al.*, 2019). Moreover, eDNA might serve as a source of nutrients (Finkel and Kolter, 2001; Palchevskiy and Finkel, 2006; Pinchuk *et al.*, 2008; Mulcahy *et al.*, 2010). Considering its chemical structure, it is easily noticeable that, for organisms capable of catabolizing it, DNA is a particularly interesting molecule. It contains mainly carbon, but also nitrogen and phosphorus, elements whose scarcity limits bacterial growth (Aldén *et al.*, 2001; Demoling *et al.*, 2007). These numerous roles highlight the significance of DNA in bacterial pathogenicity, environmental survival, and evolution.

Two strategies allow the use of eDNA as a source of nutrients: intracellular degradation after DNA uptake from the environment or extracellular degradation followed by the import of smaller building blocks such as oligonucleotides or nucleotides (Van Alphen *et al.*, 1978) (Figure 2.1). Extracellular degradation of DNA implies that fragments become readily available to competing bacterial populations. Moreover, when DNA-degrading enzymes like DNases are secreted into the surrounding medium, their concentration becomes significantly diluted, reducing their enzymatic effectiveness. Conversely, internal DNA degradation requires a specialized DNA uptake mechanism, which represents a metabolic investment. Each approach—extracellular and intracellular DNA breakdown—thus involves nuanced energetic and strategic considerations for bacterial survival and are discussed below.

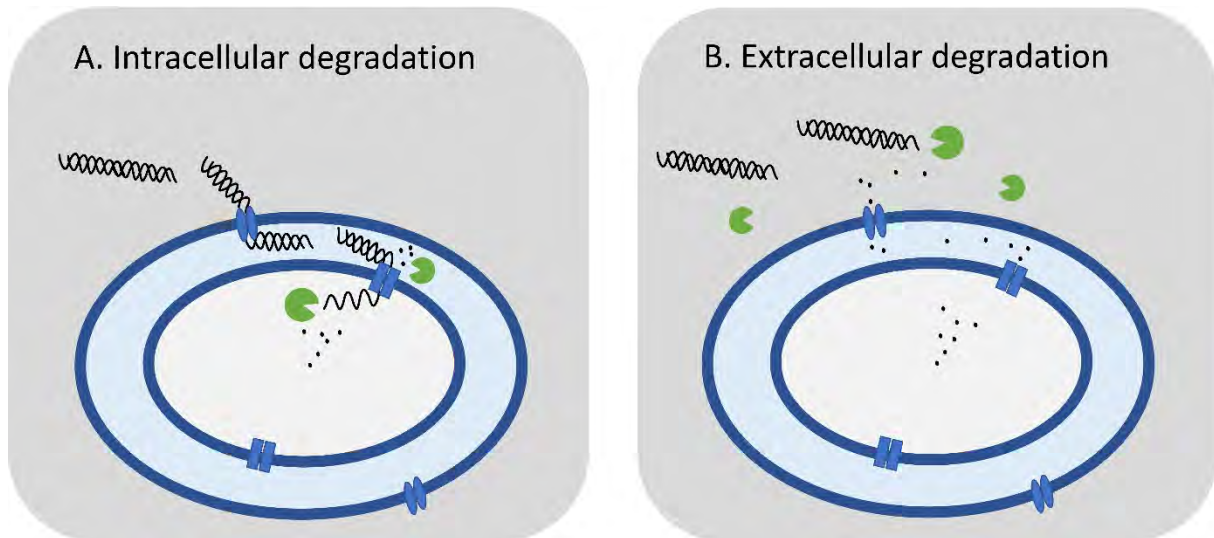


Figure 2.1 Models of DNA degradation by Gram-negative bacteria. (A) During intracellular degradation, double-stranded DNA (dsDNA) is first imported into the periplasm by an extracellular surface receptor through the outer membrane. Then, dsDNA might be processed to single-stranded DNA (ssDNA) in the periplasm or during the transport into the cytoplasm where ssDNA is further fragmented and metabolized. This model was described in *E. coli* (Palchevskiy and Finkel, 2006; Huang et al., 2022). (B) In the case of extracellular degradation, dsDNA degradation occurs extracellularly by secreted nucleases, followed by the import of smaller DNA fragments. Dark blue lines define the outer and inner membranes. The periplasm is represented by the light blue area. The green “Pac-Man” depicts nucleases. Small dots represent degraded DNA subproducts.

The model describing intracellular degradation of DNA relies on the competence apparatus. When the foreign DNA is incorporated into the recipient’s genome, this internalization might lead to natural transformation. The hypothesis that the competence apparatus has evolved as a nutrient-uptake system was proposed decades ago. It is based on the idea that cells always benefit from the nutrients contained in the eDNA they internalize, even when this eDNA integrates into the genome through homologous recombination, as the displaced genomic DNA fragment is subsequently degraded and might be used as a nutrient source (Redfield, 1993). This hypothesis was further emphasized by the fact that some species such as *Haemophilus influenzae* and *Bacillus subtilis* require nutrient-limited conditions to become competent (Herriott *et al.*, 1970; Dubnau, 1991; Sinha *et al.*, 2013).

The role of natural competence genes in DNA-based nutrition has so far been experimentally demonstrated in *E. coli*. It has been shown that *E. coli* can use DNA as a food source and depends for that on homologues of *H. influenzae* competence genes *comA*, *comB*, *comC*, *comD*, *comE*, *comF*, *comG/orfG*, and *comJ/orfJ* (Finkel and Kolter, 2001; Palchevskiy and Finkel, 2006). Mutation of some of these genes in *H. influenzae* leads to impairment of DNA uptake and transformability (Sinha *et al.*, 2012). When the competence gene homologues are mutated in *E. coli*, the mutant growth is highly reduced in a minimal medium supplemented

with DNA, indicating the need for these genes for efficient DNA internalization (Palchevskiy and Finkel, 2006).

Importantly, these competence gene homologues in *E. coli* are not only required for DNA uptake for nutritional purposes, but have also been shown to participate directly in natural transformation. Experimental evidence demonstrates that, under specific conditions—such as induction of the competence regulator Sxy and expression of the chromosomally encoded Red recombinase—*E. coli* cells can undergo natural transformation, yielding transformants (Sinha and Redfield, 2012). This transformation process strictly depends on the integrity of the DNA uptake machinery, as no transformants were observed in mutants lacking the *hofQ/comE* gene (Sinha and Redfield, 2012). These findings confirm that the competence apparatus in *E. coli* is functionally capable of mediating both DNA uptake for nutrition and natural genetic transformation. Interestingly, both competence genes and genes coding for the intracellular nuclease EndA are expressed in the presence of DNA in *E. coli* (Huang *et al.*, 2022). These findings collectively suggest that the DNA-uptake competence apparatus might serve a dual purpose: intracellular DNA consumption and natural transformation. However, data supporting this double function of the competence apparatus is lacking, especially in other species.

The second model for DNA-based nutrition consists of the external degradation of DNA by a nuclease and the subsequent import of DNA building blocks. Several bacteria have been shown to use DNA as a source of nutrients with the help of extracellular nucleases and even, if needed, grow on DNA as their sole source of carbon, nitrogen, and/or phosphorus. For example, *Pseudomonas aeruginosa* secretes an extracellular deoxyribonuclease (DNase), EddB, enabling the degradation of extracellular DNA and the utilization of DNA as a source of carbon, nitrogen, and phosphorus (Mulcahy *et al.*, 2010). *Shewanella oneidensis* produces two extracellular nucleases, EndA and ExeM, which allow it to use DNA as a source of phosphorus (Gödeke *et al.*, 2011; Heun *et al.*, 2012). This ability to feed on DNA is shared by several species of *Shewanella*, which use eDNA as a sole source of carbon, nitrogen, and phosphorus in media lacking these nutrients, under both aerobic and anaerobic conditions (Pinchuk *et al.*, 2008). Furthermore, *V. cholerae* produces two extracellular nucleases, Dns (homologue of EndA^{So}) and Xds, allowing the utilization of eDNA as a phosphorus source (Seper *et al.*, 2011). Interestingly, *dns* transcription is downregulated before the onset of natural transformation,

indicating that nuclease production and natural competence can coexist in the same strain, as long as they are not simultaneous (Blokesch and Schoolnik, 2008). DNA-based nutrition has also been investigated in the naturally competent *Helicobacter pylori* (Liechti and Goldberg, 2013). As a consequence of its auxotrophy for purines, this bacterium needs to acquire them from its environment (Nedenskov, 1994; Doig *et al.*, 1999). For this process, *H. pylori* relies on the outer membrane-associated nuclease NucT, and not the competence apparatus (Liechti and Goldberg, 2013). The reliance of certain naturally competent species on nucleases for DNA consumption underscores the challenge of accurately predicting the preferred strategy for DNA-based nutrition in such organisms.

Pseudomonas stutzeri, a ubiquitous member of the *Pseudomonas* genus renowned for its metabolic versatility and ecological resilience, has drawn attention as a potential candidate for DNA-based nutrition. This bacterium has been extensively studied for its capacities in bioremediation, nitrogen fixation, and pollutant degradation, underscoring its adaptive prowess in diverse environmental niches (Lalucat *et al.*, 2006). However, its ability to retrieve nutrients from DNA remains unexplored. Interestingly, carbon-, nitrogen-, and phosphorous-limited conditions stimulate natural transformation in *P. stutzeri* (Lorenz and Wackernagel, 1991), which means that the competence apparatus is highly functional in these conditions. So, both feeding strategies, extracellular or intracellular DNA degradation, are plausible for *P. stutzeri*. In this context, exploring the utilization of DNA as a source of nutrients by *P. stutzeri* provides not only insights into the metabolic versatility of this bacterium but also broadens our understanding of the ecological significance of DNA as a nutrient source in microbial communities. Furthermore, investigating the involvement of the competence apparatus allows us to shed light on the possible functions of natural competence as supplying nutrients to the cell.

Through a combination of genetic, biochemical, and physiological analyses, we examined the ability of *P. stutzeri* to use DNA as a source of nutrients and investigated the molecular machinery responsible for this phenomenon. We found that *P. stutzeri* can grow using DNA as the sole source of nitrogen or carbon. Interestingly, *P. stutzeri* got the most benefit in terms of growth from self-DNA rather than DNA from another bacterial species or fish DNA. However, this ability does not seem to rely on the competence apparatus. Hence, we investigated whether the nuclease EndA was responsible for DNA-based nutrition. Deletion of

endA did not decrease the growth based on DNA, suggesting that other nucleases are involved in DNA-based nutrition.

2.4 Materials & Methods

2.4.1 Bacterial strains and media

The bacterial strains used in this chapter are listed in Table S2.1. In nutrition experiments, initial cultures of *P. stutzeri* strains were grown overnight in LB medium at 30°C (with 180 rpm agitation). When required, antibiotics were added at the following concentration: gentamicin (Gm), 15 µg/mL for *E. coli* and 30 µg/mL for *P. stutzeri*; kanamycin (Km), 50 µg/mL; rifampicin (Rif), 50 µg/mL. For experiments, we used either LB (tryptone, 10 g/L; yeast extract 5 g/L; NaCl, 5 g/L), M63 medium –the standard recipe (100 mM KH₂PO₄, 15 mM (NH₄)₂SO₄, 0.5 mg/L FeSO₄*7H₂O, 1 mM MgSO₄*7H₂O, 0.2% glucose, vitamins, and metal solution) and a modified version which contains in addition 0.1 M HEPES– or BM2 medium prepared as done previously (Mulcahy *et al.*, 2010), as well for the versions limited in phosphorus, carbon and nitrogen. Assays with DNA as a source of either carbon, nitrogen, or phosphorus were performed in the modified M63 with the “no carbon” medium lacking glucose, the “no phosphorus” medium lacking KH₂PO₄, and the “no nitrogen” medium lacking (NH₄)₂SO₄. The M63 limited in nutrients contains either 8-fold less carbon, 200-fold less nitrogen, or 250-fold less phosphorus than the original recipe. The media were solidified with 1.5% agar.

2.4.2 Strain and plasmid construction

P. stutzeri deletion mutants were constructed by forced recombination between homologous DNA sequences according to the method of Martínez-García and de Lorenzo (Martínez-García and de Lorenzo, 2011). Briefly, the upstream and downstream flanking regions of the genes of interest (*pilA*, *comA*, *endA*) were amplified by PCR separately. Then, both fragments were joined together by PCR. The resulting fragment was purified, digested with EcoRI and BamHI restriction enzymes, and ligated to a EcoRI-BamHI-digested pEMG plasmid. The ligation mix was then inserted in *E. coli* DH5α λ *pir* by heat-shock followed by the selection of complete plasmid and further propagation of the plasmid. Then, *P. stutzeri* was electroporated with the plasmids pEMG-Δ*pilA*, pEMG-Δ*comA*, or pEMG-Δ*endA*. The plasmid inserts itself in the host genome by homologous recombination of the flanking regions insert. To excise the pEMG from the host genome, pSW-2 plasmid encoding the I-SceI restriction enzyme was

electroporated in a *P. stutzeri* clone obtained from the first electroporation. The loss of the pEMG was first evaluated by restreaking Gm-resistant colonies from the second electroporation on LB-Km. The susceptibility to Km indicates the removal of pEMG from the host genome. The deletion of *pilA*, *comA*, and *endA* was finally confirmed by PCR.

2.4.3 Nucleic acids production and purification for nutrition experiments

The DNA used in growth curve experiments was mainly salmon DNA-sodium salt (Amersham Life science, USA). After testing that there were no differences in growth for bacteria fed with ethanol-precipitated DNA or proteinase K-digestion DNA followed by double extraction with phenol/chloroform/isoamyl alcohol (Figure S2.1), we pursued the experiments with ethanol-precipitated fish DNA, resuspended in milli-Q water, except otherwise stated in the legend of the figures. Based on gel electrophoresis, the fragments have a size ranging from 250 bp to 750 bp (data not shown).

Bacterial DNA was extracted with the following protocol. Firstly, the bacterial cells from 500 μ L overnight culture were lysed in 500 μ L lysis buffer (10 mM Tris, 100 mM EDTA, 2% SDS, pH7) supplemented with proteinase K (80 μ g/mL) at 55°C. Then, the RNA was degraded with the help of RNase (40 μ g/mL) for 30 min incubation at 37°C. The proteins were precipitated by the addition of 1 volume of 7.5 M ammonium acetate. A centrifugation step allowed to separate the precipitated proteins from the supernatant. Sodium acetate at 3 M was added to the supernatant in a ratio of 1:10. Then, the DNA was precipitated with the addition of 2.5 volumes of pure ethanol. A centrifugation step allowed to gather the precipitated DNA at the bottom of the tube. Finally, the DNA was washed twice with 70% ethanol before being resuspended in nuclease-free water. The concentrations of DNA were estimated with a Qubit® 2.0 fluorometer (Life Technologies, USA) according to the manufacturer's protocol.

Regarding PCR products for nutrition tests, the gene *rpoB* of *P. stutzeri* ATCC 17587 and *C. crescentus* NA1000 was amplified with primers listed in Table S2.1. The products of 4kb were visualized with gel electrophoresis before purification (Gel and PCR clean-up kit, Macherey-Nagel) and quantification with a Qubit® 2.0 fluorometer (Life Technologies, USA).

2.4.4 Growth curves

Overnight cultures of *P. stutzeri* in LB were washed twice with NaCl 0.9%. Then, optical densities (OD) of cultures were measured at 600nm in cuvettes with a spectrophotometer (Genesys 10-S UV-Vis, Thermo Fisher) before being adjusted to an OD₆₀₀ of 0.2 in NaCl 0.9%.

Bacteria suspensions were then diluted by 100 in the medium supplemented with DNA before being distributed in a 96-well plate with 100µL per well. The OD₆₀₀ was measured over time at 30°C with agitation in various 96-well plate readers Cytation5 (BioTek, USA), Synergy HT (BioTek, USA), SpectraMax i3x (Molecular Devices, USA).

2.4.5 DNA degradation assays on gel and agar plates

For gel electrophoresis, DNA degradation assays were adapted or performed as previously done by Mulcahy (Mulcahy *et al.*, 2010). *P. stutzeri* was grown overnight in M63, BM2, or any variations of these media. BM2 medium was supplemented with 10 mM MgSO₄ and 10 mM CaCl₂ for this assay as these cations were shown to be essential for DNA degradation by *P. aeruginosa* (Mulcahy *et al.*, 2010). Then, overnight cultures were normalized to an OD₆₀₀ of 1 and supernatants were collected by centrifugation at 13'000 rpm for 1min before filtration at 0.22 µm, except otherwise stated in figure legends. Then, the conditioned medium was mixed with *P. stutzeri* genomic DNA and incubated at 30°C. Regularly, aliquots were removed and stored at -20°C until analyzed by DNA gel electrophoresis using a 1% agarose gel.

DNA degradation was also assessed on agar plates similarly as done previously (Waller, Hodel, and Almquist, 1985; Waller, Hodel, and Nuti, 1985). These plates were LB agar plates (pH 7.3) supplemented with fish DNA (2 g/L) and Toluidine Blue O (0.01%). Plates with freshly spread bacteria were incubated at 30°C overnight before assessment of plate colors. In the presence of unhydrolyzed DNA, Toluidine Blue O exhibits a royal blue color. When Toluidine Blue O is complexed with the oligonucleotides and mononucleotides released during DNA hydrolysis, the color spectrum shifts from blue to pink (Waller, Hodel, and Almquist, 1985; Waller, Hodel, and Nuti, 1985).

2.4.6 Natural transformation assays

Recipient cells of *P. stutzeri* strains were grown in LB medium at 30°C overnight (with 180 rpm agitation). After centrifugation for 1 min at 13'000 rpm, the pellet was washed twice with saline (0.9% w/v NaCl) and resuspended in saline at an OD₆₀₀ of 1 measured with spectrophotometer Genesys 10s UV-VIS (Thermo Fisher Scientific). Rifampicin-resistant *P. stutzeri* DNA (final concentration, 20 µg/mL) was added to the bacteria suspension, and 25 µL of the mix was spotted on LB agar plates. After the spots had dried, inoculated agar plates were incubated at 30°C for 24 hours. Then, each bacterial spot was collected with an inoculation loop and resuspended in 1 mL saline in a tube. The tubes were vigorously vortexed

for 15 s to resuspend the bacteria. Samples were diluted and plated on LB-Rif plates for the determination of the number of transformants and total colony-forming units (CFU) on LB plates. Transformation frequencies are given as the ratio of rifampicin-resistant CFU/mL divided by the total number of CFU/mL.

2.4.7 Biofilm assays

Overnight cultures in LB were washed twice and then normalized to an OD₆₀₀ of 0.4 in 5 mL BM2 medium prepared as previously done (Mulcahy *et al.*, 2010). Then, the new cultures were distributed in flat-bottom transparent 96-well plates and incubated without shaking at 30°C for 5 days. To normalize the production of biofilm with the growth of bacterial culture, the absorbance was first measured at 600 nm in a SpectraMax i3X plate-reader (Molecular Devices). Then, for the biofilm staining, liquid cultures were removed by pipetting and wells were washed 3x with 200 µL of MilliQ water. The biofilm was then stained with 150 µL crystal violet 1% (w/v) solubilized in a 30% (v/v) acetic acid solution for 15 min at 30°C. Then, crystal violet was removed, and wells were finally washed 3x with 200 µL MilliQ water. The 150 µL of 30% acetic acid solution was added to the wells to destain the Crystal violet and incubated at room temperature for 15 min with a 300rpm shaking (Heidolf Vibramax 100). Finally, 100 µL were transferred to a new flat-bottom 96-well plate and absorbance was measured at 595 nm in a SpectraMax i3X plate-reader (Molecular Devices). Crystal Violet retention was measured at 595 nm and normalized by culture density measured at 600 nm.

2.4.8 Graphics and data analysis

Statistical analyses and graphical visualizations were performed using *RStudio* software (Core R Team, 2019). Data visualization was accomplished using the *ggplot2* package and base R functions (Wickham, 2008). Growth rate and carrying capacity were estimated by fitting the logistic equation to growth curve data using the *growthcurver* package (Sprouffske and Wagner, 2016). For statistical analyses, normality of data distribution was assessed using the Shapiro-Wilk normality test, while equality of variances was evaluated with either the F-test or Levene's test. Mean comparisons were conducted using either Student's t-test or the Wilcoxon rank-sum test, depending on data characteristics, or via ANOVA followed by post-hoc Tukey or Dunnett tests. Significance levels are indicated on the graphs as follows: *for p-value <0.05; **for p-value <0.01; *** for p-value <0.001; NS for non-significant results

(p-value > 0.05). Multiple sequence alignments of amino acids were generated using the MUSCLE algorithm implemented in Unipro UGENE v45.1 (Okonechnikov et al., 2012).

2.5 Results

2.5.1 *P. stutzeri* can use DNA as a source of carbon or nitrogen.

To investigate the utilization of DNA as a source of nutrients by *P. stutzeri*, we first conducted growth assays using commercial fish DNA. *P. stutzeri* can utilize DNA as the sole carbon source (Figure 2.2). Furthermore, the growth parameters exhibited a direct correlation with the concentration of DNA, suggesting that carbon availability limits the growth of *P. stutzeri*.

To further characterize the utilization of DNA as a source of nutrients, we also verified whether *P. stutzeri* might use DNA as a supplement or as the sole source of carbon, nitrogen, or phosphorus. *P. aeruginosa* has been shown to produce an extracellular DNase under phosphate starvation conditions to use DNA more readily as a source of phosphorous than carbon or nitrogen (Mulcahy et al., 2010). In the absence of DNA supplementation, the media did not support the growth of *P. stutzeri* (Figure 2.3). Furthermore, there was no observed growth when DNA was added as a phosphorus supplement, unlike conditions where DNA served as either a supplement or as the sole source of carbon or nitrogen. Due to the higher total carbon or nitrogen content, nutrient-limited media enriched with DNA supported growth better than media lacking carbon or nitrogen but supplemented with DNA.

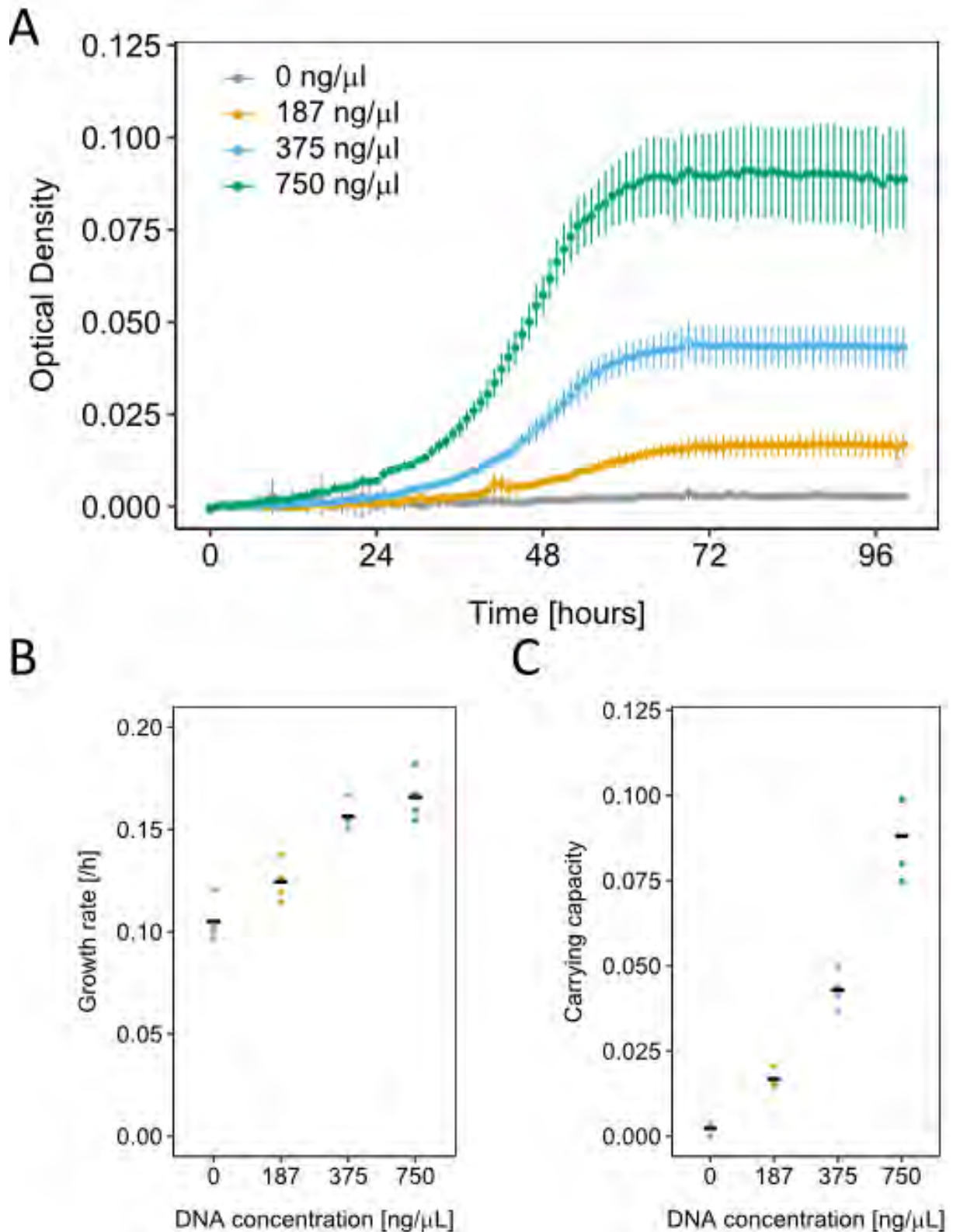
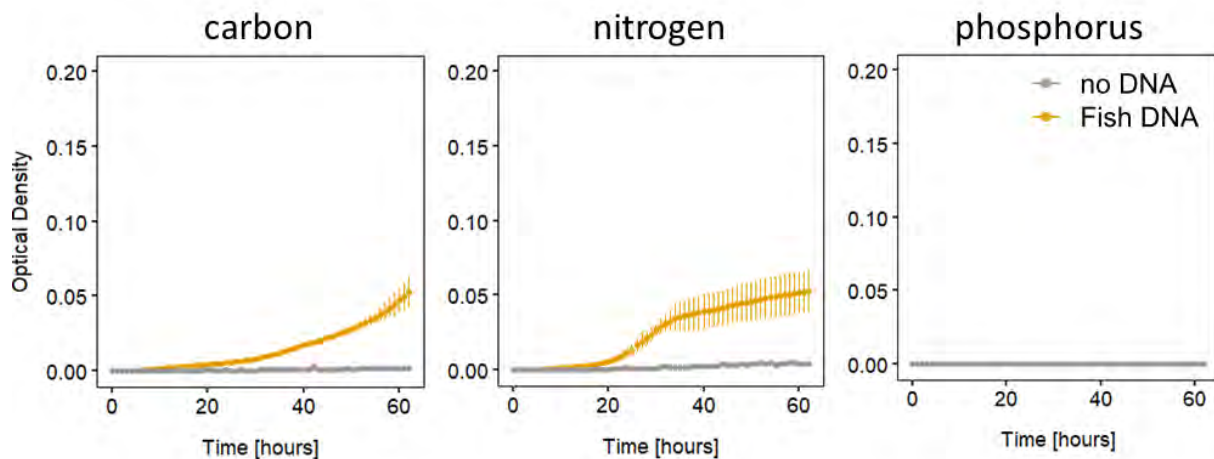


Figure 2.2 *P. stutzeri* can utilize DNA as the sole carbon source. (A) Growth kinetics of *P. stutzeri* using fish DNA as the sole source of carbon in M63. Multiple concentrations of fish DNA, purified by ethanol precipitation, were used: 750 ng/μL (green), 375 ng/μL (blue), 187 ng/μL (orange), and 0 ng/μL (grey) as control. Higher growth measured as optical density was obtained with a higher concentration of DNA. The dots represent the mean of four replicates, the error bars stand for the standard deviation. Growth rate (B) and carrying capacity (C) were obtained by fitting a logistic equation on the data with dots representing four replicates and black lines as the means.

A) DNA as the sole source of



B) DNA as a supplementing source of

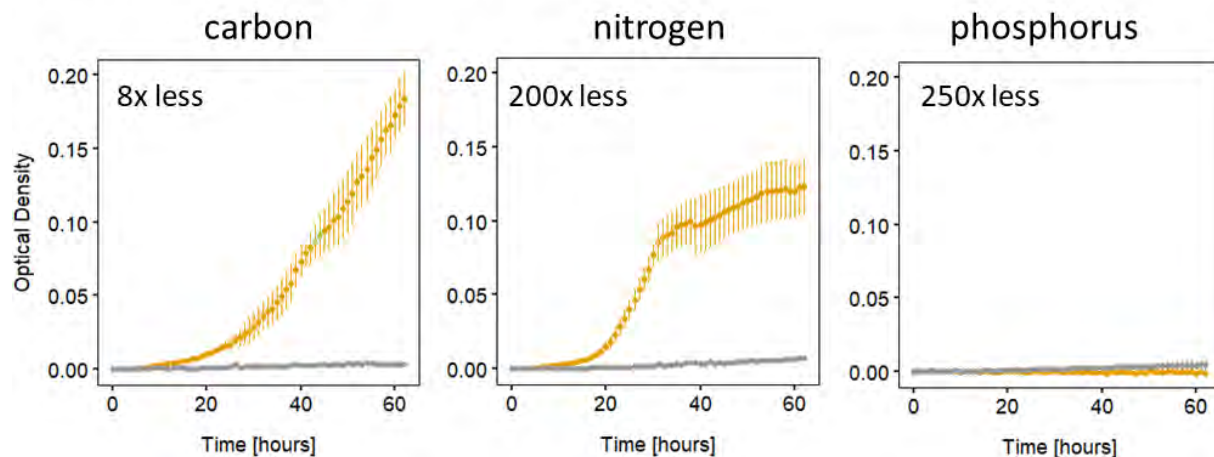


Figure 2.3 DNA supports the growth of *P. stutzeri* as a source of carbon and nitrogen but not as a source of phosphorus. Growth kinetics of *P. stutzeri* in modified M63 using fish DNA (orange) as the sole (top, panel A) or as a supplementing source (bottom, panel B) of carbon, nitrogen, and phosphorus. The control (grey) does not contain any DNA. The limitation in nutrients consists of an 8-fold, 200-fold, or 250-fold reduction of carbon, nitrogen, and phosphorus respectively as done previously for *P. aeruginosa* (Mulcahy *et al.*, 2010). Proteinase K-digested and phenol/chloroform-extracted fish DNA was used as the source of DNA at a concentration of 1.5 mg/mL. The dots represent the mean of four replicates, the error bars stand for the standard deviation.

2.5.2 *P. stutzeri* benefits more from self-DNA than foreign DNA for nutrition.

The experiments performed to this point used commercial fish DNA for practical reasons; it was readily purified and accessible in significant quantities, typically on a multiple-gram scale. Having demonstrated that fish DNA can be used as a nutrient source by *P. stutzeri*, we next asked whether the nucleotide composition of DNA had an impact on its nutritional value, by feeding *P. stutzeri* with genomic DNA from itself, from *Caulobacter crescentus*, and from fish as a control. Interestingly, growth was faster and yield higher with self-DNA than with fish or *C. crescentus* DNA (Figure 2.4).

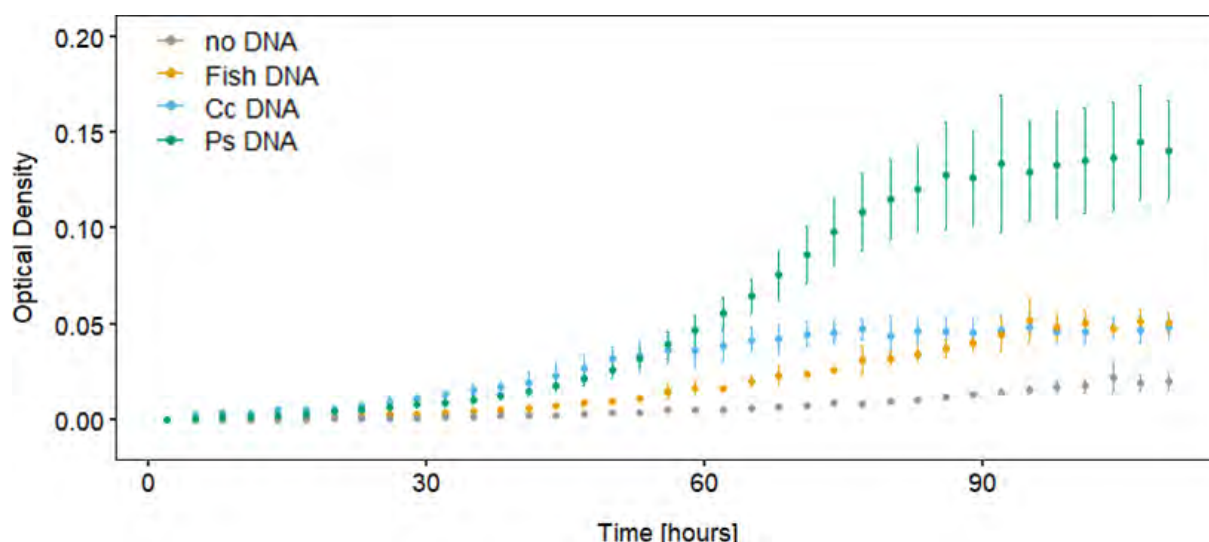


Figure 2.4 *P. stutzeri* benefits more from self-DNA than foreign DNA for nutrition. Growth kinetics of *P. stutzeri* in M63 using genomic DNA from different species as a sole source of carbon. DNA was at a concentration of 50 ng/μL. “Fish DNA” stands for DNA from fish, “Cc DNA” stands for DNA from *C. crescentus* NA1000, and “Ps DNA” stands for DNA from *P. stutzeri* ATCC 17587. The dots represent the mean of 5 replicates, the error bars stand for the standard deviation.

The GC content of *P. stutzeri* ATCC 17587 DNA is approximately 61%, and *C. crescentus* NA1000 has a similar GC content to *P. stutzeri* with 67% (Ely and Scott, 2014). Despite these similarities, the outcomes observed with *C. crescentus* DNA were similar to those obtained with fish DNA, even though fish DNA is AT-rich (Hui *et al.*, 2013). This suggests that factors other than intrinsic GC content influenced the observed growth differences, such as specific DNA sequences.

2.5.3 The ability to use DNA as a nutrient source did not rely on the competence apparatus.

P. stutzeri can use DNA as a source of nutrients, but the process allowing that remains unknown. Natural competence, a mechanism during which extracellular DNA is internalized, might be a good candidate. Indeed, in *E. coli*, competence gene homologues are important for the utilization of DNA as a nutrient: their removal led to a decrease in growth in a minimal medium supplemented with DNA (Finkel and Kolter, 2001; Palchevskiy and Finkel, 2006).

To determine if the ability of *P. stutzeri* to use DNA as a nutrient relied on the competence apparatus, we measured the growth of two mutants, lacking either *pilA* or *comA*, with DNA as the sole source of carbon. These genes are essential for natural transformation (Figure S2.3) (Graupner *et al.*, 2000; Graupner and Wackernagel, 2001). The gene *pilA* encodes a pilin, a structural component of type IV pilus, and the gene *comA* is required for the translocation of DNA from the periplasm into the cytoplasm (Graupner and Wackernagel, 2001), hence important for a later stage of natural transformation. If feeding on DNA relies on the

competence apparatus, the results of these mutants should indicate in which space degradation of DNA occurs. The $\Delta comA$ and $\Delta pilA$ mutants grew like the wild-type strain in LB (Figure S2.4) and in M63 supplemented with *P. stutzeri* DNA as the sole carbon source (Figure 2.5). Our results indicate that the ability to feed on DNA did not rely on a type IV pilus or on the transfer of long DNA fragments to the cytoplasm. Since no reduction in growth was observed for the $\Delta pilA$ mutant, we hypothesized that DNA degradation occurs primarily extracellularly and that small DNA subunits (nucleotides or oligonucleotides) are then imported into the cell.

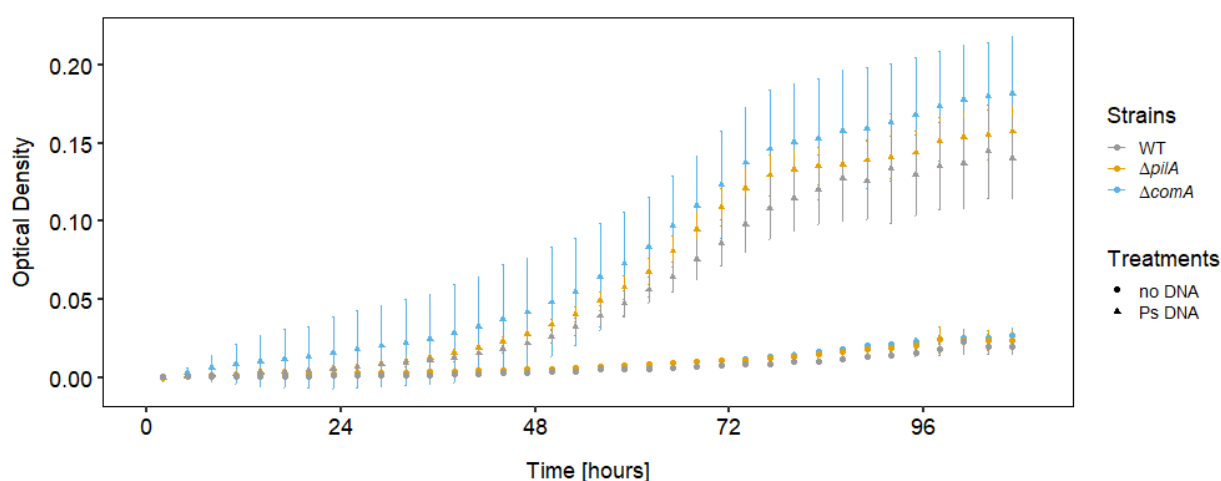


Figure 2.5 The competence apparatus does not play a major role in DNA-based nutrition in *P. stutzeri*. Growth kinetics of *P. stutzeri* wild-type (WT), $\Delta comA$, and $\Delta pilA$ using *P. stutzeri* genomic DNA (at a final concentration of 50 ng/ μ L) as the sole source of carbon in M63. The dots represent the mean of five replicates, the error bars stand for the standard deviation.

2.5.4 The nuclease EndA is not involved in the DNA-based nutrition of *P. stutzeri*, but it reduces the transformability of it.

Since the ability of *P. stutzeri* to feed on DNA does not rely on the competence apparatus and no other system is known to mediate the import of long stretches of DNA in *P. stutzeri*, some DNA degradation should occur extracellularly. To identify potential extracellular enzymes with nucleolytic activity, we screened the genome sequence of *P. stutzeri* ATCC17587 for the presence of genes encoding nucleases using the search term “nuclease.” The screen revealed 41 genes encoding probable or predicted nucleases (Table S2.2). The screen was also performed with the genome assembly of *P. stutzeri* ATCC 17587 (GCF_003410115.1) from the database <https://www.pseudomonas.com/> (Winsor *et al.*, 2009) and gave similar results (Table S2.3). Of all these nucleases, only three presented evidence of a secretory signal peptide according to *signalP* (Almagro Armenteros *et al.*, 2019) (Table 2.1).

Table 2.1 Genes encoding potential and confirmed extracellular nucleases in the genome of *P. stutzeri* ATCC17587. COG: Clusters of Orthologous Genes.

Our genome ID ^a	ps40790	ps29660	ps06830
genome ID from Pseudomonas Genome Database ^b	DOA22_RS20565	DOA22_RS16070	DOA22_RS08380
RefSeq Accession ^b	WP_017244187.1	WP_053527841.1	WP_151320409.1
Gene name ^a	<i>ps40790</i>	<i>endA</i>	<i>ps06830</i>
Description ^a	putative endonuclease	DNA-specific endonuclease I	nuclease
Pfam (Accession) ^a	Staphylococcal nuclease homologue (PF00565)	Endonuclease I (PF04231)	Staphylococcal nuclease homologue (PF00565)
COG ^a	Endonuclease YncB, thermonuclease family (COG1525)	EndA Endonuclease I (COG2356)	Endonuclease YncB, thermonuclease family (COG1525)

^a This info originated from the annotation of our genome from *P. stutzeri* ATCC17587 with the DFAST pipeline (Tanizawa *et al.*, 2018).

^b This info was from <https://www.pseudomonas.com/> webpage.

As we did not find any literature regarding these three potential nucleases of *P. stutzeri*, we blasted them against studied nucleases to find their closest annotated homologues. The protein Ps06830 shared about 35% amino-acid identity with YncB nuclease from *B. subtilis*. This nuclease has been shown to weakly degrade DNA and reduce the transformability of *B. subtilis*, but it is located in the cytoplasm (Sakamoto *et al.*, 2001; Meile *et al.*, 2006). Ps40790 shared a domain with *Staphylococcus* thermostable nuclease according to InterProScan (Zdobnov and Apweiler, 2001). Concerning the last predicted nuclease, EndA^{Ps} (Ps29660), it shared about 40-60% amino-acid identity with the endonuclease EndA^{Pa} of *P. aeruginosa*, with the extracellular nuclease Dns from *V. cholerae*, and with *S. oneidensis* EndA^{So} (Figure 2.6). EndA^{So} allows the use of DNA at least as a sole source of phosphorus (Heun *et al.*, 2012); it has not been tested whether this nuclease allowed also the use of DNA as a source of other nutrients. Regarding Dns, this nuclease is important for the utilization of DNA as a phosphorus source in *V. cholerae* and has been shown to degrade circular and linearized plasmid DNA (Seper *et al.*, 2011). EndA^{Pa} has been mainly studied for its role in biofilm degradation and promoting cell dispersion, rather than its involvement in DNA-based nutrition (Cherny and Sauer, 2019). However, it undoubtedly degrades DNA as shown by gel electrophoresis (Cherny and Sauer, 2019). Another nuclease from *P. aeruginosa*, EddB, has been shown to effectively

allow the utilization of DNA as only source of carbon, nitrogen, and phosphorous (Mulcahy *et al.*, 2010), but no homologue was present in *P. stutzeri*. Out of the three potential extracellular nucleases of *P. stutzeri* and based on homology, EndA^{Ps} seemed to be the most promising candidate for an extracellular nuclease.

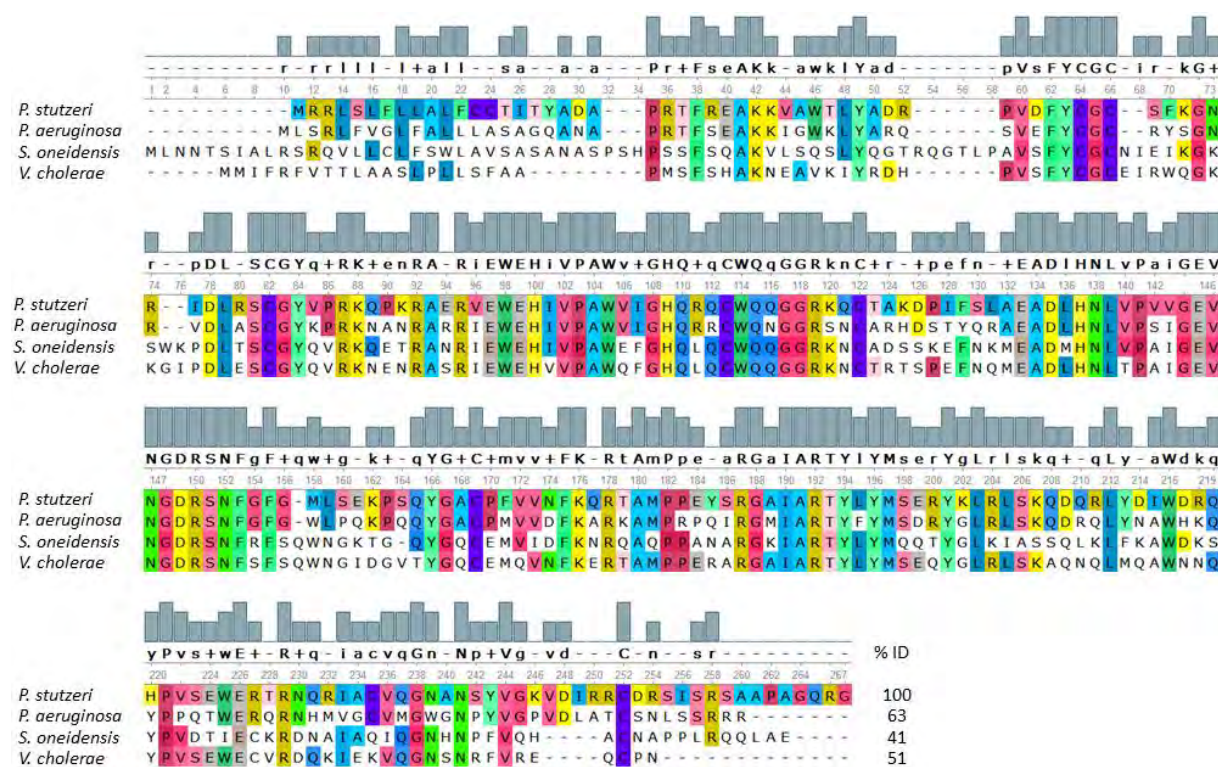


Figure 2.6 Multiple amino acids alignment of *P. stutzeri* EndA (Ps29660) with similar proteins of other bacteria generated by the MUSCLE protein alignment algorithm in Unipro UGENE v45.1 (Okonechnikov *et al.*, 2012). The corresponding genes coded for EndA from *P. aeruginosa* (NP_251439.1), Dns from *V. cholerae* (WP_149560494.1), and EndA from *S. oneidensis* (WP_011071125.1). Gaps introduced for alignment are represented by dashes. Amino acids are colored when they agree with *P. stutzeri* EndA. The consensus sequence is displayed above.

To determine whether the endonuclease EndA is responsible for the nutrition with DNA in *P. stutzeri*, we conducted growth assays using a markerless in-frame deletion mutant (Δ endA). The Δ endA mutant displayed the same ability as the wild-type strain to use DNA as the sole source of carbon (Figure 2.7A), indicating that the nuclease EndA is not an essential contributor to DNA consumption in *P. stutzeri*.

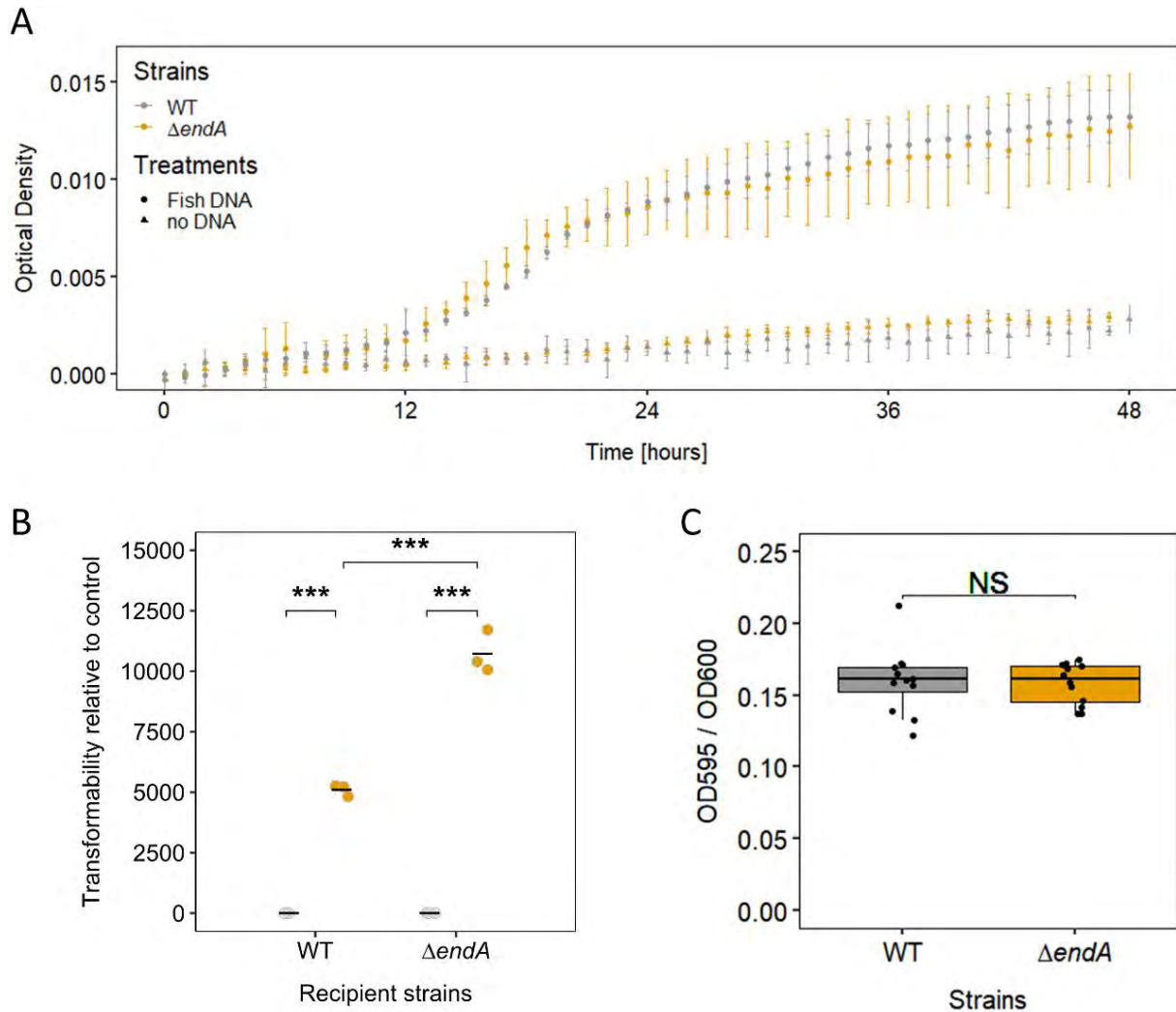


Figure 2.7 The mutant $\Delta endA$ can use DNA as a sole source of carbon such as the WT and produces a similar amount of biofilm than the WT strain, but it was slightly more transformable. (A) Growth kinetics of *P. stutzeri* WT (grey) and $\Delta endA$ (orange) using DNA as the sole source of carbon in modified M63 did not differ between both strains. Fish DNA (circles) at a concentration of 80 ng/ μ L was only resuspended in water, without further purification. The control (triangles) contained no source of carbon. The dots represent the mean of five replicates, the error bars stand for the standard deviation. (B) Transformability with rifampicin-resistance DNA (orange) relative to transformability without DNA (grey, control) was two times higher for the deletion mutant $\Delta endA$. The dots represent the three replicates and the black lines the means. (C) $endA$ deletion did not impact the production of biofilm. Crystal Violet retention was measured at 595 nm and normalized by culture density in twelve replicates per strain.

We next explored the ability of the $\Delta endA$ mutant to perform natural transformation. The negative mutant of Dns, a homologue of EndA^{Ps}, which is an extracellular nuclease of *V. cholerae*, has an increased transformability to about 300-fold relative to the wild-type (Blokesch and Schoolnik, 2008). To avoid interference with natural transformation, *dns* transcription is downregulated by the quorum sensing regulator HapR, before the onset of natural transformation (Blokesch and Schoolnik, 2008). Hence, we evaluated whether the $\Delta endA$ mutant was also hyper-transformable with a natural transformation experiment. Interestingly, the mutant was about two times more transformable than the wild-type (Figure

2.7B), indicating that the nuclease EndA slightly reduces the transformability of *P. stutzeri* but to a much lower extent than its homologue in *V. cholerae*. Finally, we investigated whether *endA* impacted biofilm formation with crystal violet staining assays as its homologue does in *P. aeruginosa* (Cherny and Sauer, 2019). Regarding *P. stutzeri*, there was no difference in the total biofilm biomass between the wild-type and the $\Delta endA$ mutant (Figure 2.7C). Taken together, our results indicated that *endA* is not an essential gene for nutrition based on DNA or biofilm production but might reduce the transformability of *P. stutzeri*.

2.6 Discussion

This study aimed to investigate whether *P. stutzeri* could use DNA as a nutrient source and whether this ability relied on the competence apparatus or on extracellular degradation by a nuclease. By evaluating bacterial growth in various media and with the help of competence mutants as well as a nuclease mutant, we showed that *P. stutzeri* can use DNA as a sole source of carbon or nitrogen and that neither the competence apparatus nor the nuclease EndA are essential for this phenotype.

We first investigated the ability of *P. stutzeri* to use DNA as a source of nutrients. Interestingly, in our conditions, this species could only use DNA as a source of carbon or nitrogen, unlike *P. aeruginosa*, which can also use it as a source of phosphorus (Mulcahy *et al.*, 2010). In fact, *P. aeruginosa* derives the greatest benefit from using DNA as a source of phosphorus, which is consistent with the transcriptional induction of the extracellular nuclease EddB under phosphate-limiting conditions (Mulcahy *et al.*, 2010). This difference may be due to the distinct environments the two species colonize. Although *P. aeruginosa* and *P. stutzeri* have both been found in environmental and medical samples (Lalucat *et al.*, 2006; Crone *et al.*, 2020; Horcajada *et al.*, 2024), *P. aeruginosa* is the most prevalent *Pseudomonas* species in clinical settings (Ergin and Mutlu, 1999), highlighting its greater ability to colonize humans. An early study reports that when humans are infected with Gram-negative pathogens, the level of inorganic phosphate in plasma decreases to a point that is expected to be suboptimal for bacterial growth (Weinberg, 1974). Consequently, it might be interesting for bacteria to produce a nuclease to acquire phosphorus from DNA under these conditions. Regarding soil nutrient limitations, an experiment with an extract from agricultural soil indicated that soil limitations in both carbon and nitrogen were hindering the proper growth of *P. stutzeri*

(Lorenz and Wackernagel, 1991). It might therefore be interesting to test whether adding DNA under these specific conditions could promote growth. As nutrient limitations vary considerably between soil types (Aldén et al., 2001), it is difficult to conclude whether *P. stutzeri* regularly faces growth difficulties due to a lack of nitrogen or carbon, and copes with this stress by degrading DNA.

Two models were possible to explain nutrition based on DNA: either DNA was first internalized by the competence apparatus and then degraded inside of the cells, or it was degraded outside of the cells and then smaller building blocks were imported. In previous work, *E. coli* was shown to be able to use DNA as a food source and depended for that on homologues of *H. influenzae* competence genes (Finkel and Kolter, 2001; Palchevskiy and Finkel, 2006), whose mutation in *H. influenzae* leads to DNA uptake and transformability impairment (Sinha et al., 2012). When the competence gene homologues were mutated in *E. coli*, the growth yield of mutants was highly reduced in a minimal medium supplemented with DNA, indicating the need of these genes for efficient DNA internalization (Palchevskiy and Finkel, 2006). In the case of *P. stutzeri*, deletion of the natural competence genes did not considerably affect its ability to use DNA as a source of nutrients, but it affected its transformability (Graupner et al., 2000; Graupner and Wackernagel, 2001), showing the role of these genes in extracellular DNA uptake for transformation but not for nutrition. This suggests that the model with extracellular degradation is more likely.

Bioinformatic screening allowed us to identify three potential extracellular nucleases, among which EndA was prioritized for experimental investigation due to its homology with nucleases known to mediate DNA degradation in other species, such as *S. oneidensis* (Heun et al., 2012) and *V. cholerae* (Seper et al., 2011). However, deletion of *endA* did not impair the ability of *P. stutzeri* to grow on DNA. The two other predicted extracellular nucleases therefore remain promising candidates and should be further investigated. Importantly, the lack of DNA-based growth defect in the $\Delta endA$ mutant does not totally exclude a functional role, as redundancy may compensate for its loss (Zhang, 2012; Ghosh and O'Connor, 2017). For this reason, combinatorial deletion of multiple candidate genes should be also considered. Alternatively, a high throughput approach might be used, for example based on a library of mutants with random transposon insertions (Chao et al., 2016). DNase-deficient mutants and hyper DNase-producers might be identified on DNase test agar plates (Waller, Hodel, and Almquist,

1985; Waller, Hodel, and Nuti, 1985), after optimization of the detection method as no clear nuclease activity was visible for *P. stutzeri* (Figure S2.5). While our working model assumes that an extracellular nuclease mediates DNA degradation, a contribution from intracellular nucleases released through spontaneous or stress-induced cell lysis cannot be excluded. Clarifying whether DNA degradation is primarily driven by secreted enzymes, lysis-derived nucleases, or a combination of both remains essential to fully understand the mechanism underlying DNA-based nutrition in *P. stutzeri*.

The observation that *P. stutzeri* exhibits enhanced growth when utilizing its own DNA as a substrate, compared to DNA from two other species, raises intriguing questions. However, this finding is constrained by the limited scope of the experiment, which included only three species. Consequently, it is difficult to conclude whether *P. stutzeri* grows better with its DNA than with any other DNA with a similar base composition. To extrapolate on this observation, a more extensive study involving other species, such as other *Pseudomonas* and *P. stutzeri* strains, would be required. Furthermore, a limitation of this experiment lies in understanding the origin of the observed growth differences. We propose three hypotheses to explain this interesting result. The first hypothesis centers on the composition of DNA, specifically the frequency of each nucleotide. If the internalized DNA is similar to the cell DNA, fewer nucleotides and related building blocks would need to be synthesized *de novo*, if the DNA is not completely degraded during the DNA-based nutrition. The second hypothesis involves the presence of some specific motifs on the DNA that could facilitate its cleavage. For instance, *Serratia* endonuclease NucA preferentially cleaves GC-rich regions over AT-rich regions (Meiss *et al.*, 1995; Friedhoff *et al.*, 1996). The third hypothesis concerns DNA methylation, which can affect DNA stability by making it more resistant to endonucleases, depending on the methylation sites (McClelland, 1983; McClelland *et al.*, 1994). To test these hypotheses, we propose monitoring the growth of *P. stutzeri* fed with various plasmids. For the first and second hypotheses, using different inserts cloned into the same vector would help address questions about DNA composition and motifs. To minimize vector-related influences, it is advisable to use a large insert in combination with a small vector backbone. Additionally, restriction enzyme digestion followed by gel purification can be utilized to isolate the DNA fragment of interest and remove residual vector sequences. Although PCR amplification of the fragment would provide the highest purity, this approach is costly for generating large

quantities of DNA. To examine the effect of methylation, the same plasmid extracted from different species could be used as a substrate. Unraveling the mechanisms that enable *P. stutzeri* to thrive more effectively on self-DNA compared to the tested foreign DNA would help determine whether this observation is due to a specific process, such as targeted degradation of conspecific DNA, or whether it results from intrinsic properties of the DNA itself — independent of its taxonomic origin — such as base composition or methylation status.

To summarize, our work indicates that *P. stutzeri* can use DNA as the sole source of carbon and nitrogen as evidenced in other species such as *P. aeruginosa*, *E. coli*, *S. oneidensis*, and *V. cholerae*. This ability seems not to be related to the competence apparatus in contrast to what was observed in *E. coli*. The deletion of our most promising candidates for predicted extracellular nucleases, EndA, did not lead to a reduction in growth based on DNA consumption. The fact that *P. stutzeri* grows better on its own DNA than on foreign DNA is intriguing and should be further deepened by investigating the role of methylation, DNA composition, and sequence motifs. The investigation presented herein explores the utilization of DNA as a nutrient source by *P. stutzeri*, providing a foundation for future research endeavours aimed at deciphering the ecological and evolutionary significance of this intriguing process.

2.7 References

- Aldén, L., Demoling, F., and Bååth, E. (2001) Rapid Method of Determining Factors Limiting Bacterial Growth in Soil. *Appl Environ Microbiol* **67**: 1830–1838.
- Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., et al. (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**: 420–423.
- Van Alphen, W., Van Boxtel, R., Van Selm, N., and Lugtenberg, B. (1978) Pores in the outer membrane of Escherichia coli K12. Involvement of proteins b and c in the permeation of cephaloridine and ampicillin. *FEMS Microbiol Lett* **3**: 103–106.
- Blokesch, M. and Schoolnik, G.K. (2008) The extracellular nuclease Dns and its role in natural transformation of Vibrio cholerae. *J Bacteriol* **190**: 7232–7240.
- Chao, M.C., Abel, S., Davis, B.M., and Waldor, M.K. (2016) The design and analysis of transposon insertion sequencing experiments. *Nat Rev Microbiol* **14**: 119–128.
- Cherny, K.E. and Sauer, K. (2019) Pseudomonas aeruginosa requires the DNA-specific endonuclease EndA to degrade extracellular genomic DNA to disperse from the biofilm. *J Bacteriol* **201**: 1–18.
- Core R Team (2019) A Language and Environment for Statistical Computing. *R Found Stat Comput* **2**: <https://www.R-project.org>.
- Crone, S., Vives-Flórez, M., Kvich, L., Saunders, A.M., Malone, M., Nicolaisen, M.H., et al. (2020) The environmental occurrence of Pseudomonas aeruginosa. *J Pathol Microbiol Immunol* **128**: 220–231.
- Demoling, F., Figueroa, D., and Bååth, E. (2007) Comparison of factors limiting bacterial growth in different soils. *Soil Biol Biochem* **39**: 2485–2495.
- Doig, P., de Jonge, B.L., Alm, R.A., Brown, E.D., Uria-Nickelsen, M., Noonan, B., et al. (1999) Helicobacter pylori Physiology Predicted from Genomic Comparison of Two Strains . *Microbiol Mol Biol Rev* **63**: 675–707.
- Dubnau, D. (1991) The regulation of genetic competence in Bacillus subtilis. *Mol Microbiol* **5**: 11–18.
- Ely, B. and Scott, L.T.E. (2014) Correction of the Caulobacter crescentus NA1000 genome annotation. *PLoS One* **9**: 1–8.
- Ergin, C. and Mutlu, G. (1999) Clinical distribution and antibiotic resistance of Pseudomonas species. *East J Med* **4**: 65–69.
- Finkel, S.E. and Kolter, R. (2001) DNA as a nutrient: Novel role for bacterial competence gene homologs. *J Bacteriol* **183**: 6288–6293.

- Friedhoff, P., Kolmes, B., Gimadutdinow, O., Wende, W., Krause, K.L., and Pingoud, A. (1996) Analysis of the mechanism of the *Serratia* nuclease using site-directed mutagenesis. *Nucleic Acids Res* **24**: 2632–2639.
- Ghosh, S. and O'Connor, T.J. (2017) Beyond paralogs: The multiple layers of redundancy in bacterial pathogenesis. *Front Cell Infect Microbiol* **7**: 1–14.
- Gödeke, J., Heun, M., Bubendorfer, S., Paul, K., and Thormann, K.M. (2011) Roles of two *Shewanella oneidensis* MR-1 extracellular endonucleases. *Appl Environ Microbiol* **77**: 5342–5351.
- Graupner, S., Frey, V., Hashemi, R., Lorenz, M.G., Brandes, G., and Wackernagel, W. (2000) Type IV pilus genes *pilA* and *pilC* of *Pseudomonas stutzeri* are required for natural genetic transformation, and *pilA* can be replaced by corresponding genes from nontransformable species. *J Bacteriol* **182**: 2184–2190.
- Graupner, S. and Wackernagel, W. (2001) Identification and characterization of novel competence genes *comA* and *exbB* involved in natural genetic transformation of *Pseudomonas stutzeri*. *Res Microbiol* **152**: 451–460.
- Hamilton, H.L., Domínguez, N.M., Schwartz, K.J., Hackett, K.T., and Dillard, J.P. (2005) *Neisseria gonorrhoeae* secretes chromosomal DNA via a novel type IV secretion system. *Mol Microbiol* **55**: 1704–1721.
- Herriott, R.M., Meyer, E.M., and Vogt, M. (1970) Defined nongrowth media for stage II development of competence in *Haemophilus influenzae*. *J Bacteriol* **101**: 517–524.
- Heun, M., Binnenkade, L., Kreienbaum, M., and Thormann, K.M. (2012) Functional specificity of extracellular nucleases of *Shewanella oneidensis* MR-1. *Appl Environ Microbiol* **78**: 4400–4411.
- Horcajada, J.P., Edwards, F., Fonio, S., Montero, M., Harris, P., Paterson, D.L., and Laupland, K.B. (2024) *Pseudomonas stutzeri* bloodstream infection is a prevailing community-onset disease with important mortality rates: results from a retrospective observational study in Australia. *Infect Dis (Auckl)* **56**: 606–615.
- Huang, L., Zhang, Y., Du, X., An, R., and Liang, X. (2022) *Escherichia coli* Can Eat DNA as an Excellent Nitrogen Source to Grow Quickly. *Front Microbiol* **13**: 1–13.
- Hui, C.Y., Guo, Y., Zhang, X., Shao, J.H., Yang, X.Q., and Zhang, W. (2013) Oligodeoxyribonucleotides derived from salmon sperm DNA: An alternative to defibrotide. *Biologicals* **41**: 190–196.
- Lalucat, J., Bennasar, A., Bosch, R., García-Valdés, E., and Palleroni, N.J. (2006) Biology of *Pseudomonas stutzeri*. *Microbiol Mol Biol Rev* **70**: 510–547.
- Liao, S., Klein, M.I., Heim, K.P., Fan, Y., Bitoun, J.P., Ahn, S.J., et al. (2014) *Streptococcus mutans* extracellular DNA is upregulated during growth in biofilms, actively released via membrane vesicles, and influenced by components of the protein secretion machinery. *J*

- Bacteriol* **196**: 2355–2366.
- Liechti, G.W. and Goldberg, J.B. (2013) Helicobacter pylori salvages purines from extracellular host cell DNA utilizing the Outer Membrane-Associated nuclease NucT. *J Bacteriol* **195**: 4387–4398.
- Lorenz, M.G. and Wackernagel, W. (1991) High frequency of natural genetic transformation of Pseudomonas stutzeri in soil extract supplemented with a carbon/energy and phosphorus source. *Appl Environ Microbiol* **57**: 1246–1251.
- Martínez-García, E. and de Lorenzo, V. (2011) Engineering multiple genomic deletions in Gram-negative bacteria: Analysis of the multi-resistant antibiotic profile of Pseudomonas putida KT2440. *Environ Microbiol* **13**: 2702–2716.
- Mcclelland, M. (1983) The effect of site specific methylation on restriction endonuclease cleavage. *Nucleic Acids Res* **11**: 235.
- Mcclelland, M., Nelson, M., and Raschke, E. (1994) Effect of site-specific modification on restriction endonucleases and DNA modification methyltransferases. *Nucleic Acids Res* **22**: 3640–3659.
- Meile, J.C., Wu, L.J., Ehrlich, S.D., Errington, J., and Noirot, P. (2006) Systematic localisation of proteins fused to the green fluorescent protein in Bacillus subtilis: Identification of new proteins at the DNA replication factory. *Proteomics* **6**: 2135–2146.
- Meiss, G., Friedhoff, P., Hahn, M., Gimadutdinow, O., and Pingoud, A. (1995) Sequence Preferences in Cleavage of dsDNA and ssDNA by the Extracellular Serratia marcescens Endonuclease. *Biochemistry* **34**: 11979–11988.
- Michod, R.E. and Wojciechowski, M.F. (1988) DNA repair and the evolution of transformation in the Bacterium Bacillus subtilis. *Genetics* **118**: 31–39.
- Mulcahy, H., Charron-Mazenod, L., and Lewenza, S. (2010) Pseudomonas aeruginosa produces an extracellular deoxyribonuclease that is required for utilization of DNA as a nutrient source. *Environ Microbiol* **12**: 1621–1629.
- Nedenskov, P. (1994) Nutritional requirements for growth of Helicobacter pylori. *Appl Environ Microbiol* **60**: 3450–3453.
- Nielsen, K.M., Johnsen, P.J., Bensasson, D., and Daffonchio, D. (2007) Release and persistence of extracellular DNA in the environment. *Environ Biosafety Res* **6**: 37–53.
- Okonechnikov, K., Golosova, O., Fursov, M., Varlamov, A., Vaskin, Y., Efremov, I., et al. (2012) Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **28**: 1166–1167.
- Palchevskiy, V. and Finkel, S.E. (2006) Escherichia coli competence gene homologs are essential for competitive fitness and the use of DNA as a nutrient. *J Bacteriol* **188**: 3902–3910.

- Pinchuk, G.E., Ammons, C., Culley, D.E., Li, S.M.W., McLean, J.S., Romine, M.F., et al. (2008) Utilization of DNA as a sole source of phosphorus, carbon, and energy by *Shewanella* spp.: Ecological and physiological implications for dissimilatory metal reduction. *Appl Environ Microbiol* **74**: 1198–1208.
- Pisetsky, D.S. (2017) The biological functions of DNA: From the sublime to the slime. *Arthritis Res Ther* **19**: 4–5.
- Redfield, R.J. (1993) Genes for breakfast: The have-your-cake and-eat-it-too of bacterial transformation. *J Hered* **84**: 400–404.
- Sakamoto, J.J., Sasaki, M., and Tsuchido, T. (2001) Purification and Characterization of a *Bacillus subtilis* 168 Nuclease, YokF, Involved in Chromosomal DNA Degradation and Cell Death Caused by Thermal Shock Treatments. *J Biol Chem* **276**: 47046–47051.
- Seper, A., Fengler, V.H.I., Roier, S., Wolinski, H., Kohlwein, S.D., Bishop, A.L., et al. (2011) Extracellular nucleases and extracellular DNA play important roles in *Vibrio cholerae* biofilm formation. *Mol Microbiol* **82**: 1015–1037.
- Shak, S., Capon, D.J., Hellmiss, R., Marsters, S.A., and Baker, C.L. (1990) Recombinant human DNase I reduces the viscosity of cystic fibrosis sputum. *Proc Natl Acad Sci U S A* **87**: 9188–9192.
- Sinha, S., Mell, J., and Redfield, R.J. (2013) The availability of purine nucleotides regulates natural competence by controlling translation of the competence activator Sxy. *Mol Microbiol* **88**: 1106–1119.
- Sinha, S., Mell, J.C., and Redfield, R.J. (2012) Seventeen sxy-dependent cyclic amp receptor protein site-regulated genes are needed for natural transformation in *Haemophilus influenzae*. *J Bacteriol* **194**: 5245–5254.
- Sinha, S. and Redfield, R.J. (2012) Natural DNA uptake by *Escherichia coli*. *PLoS One* **7**: 3–9.
- Sprouffske, K. and Wagner, A. (2016) Growthcurver: An R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics* **17**: 17–20.
- Tanizawa, Y., Fujisawa, T., and Nakamura, Y. (2018) DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* **34**: 1037–1039.
- Waller, J.R., Hodel, S.L., and Almquist, D.B. (1985) An evaluation of five techniques used to detect deoxyribonuclease production by bacteria. *J Microbiol Methods* **3**: 303–309.
- Waller, J.R., Hodel, S.L., and Nuti, R.N. (1985) Improvement of two toluidine blue O-mediated techniques for DNase detection. *J Clin Microbiol* **21**: 195–199.
- Weinberg, E.D. (1974) Iron and susceptibility to infectious disease. *Science (80-)* **184**: 952–9556.
- Whitchurch, C.B., Tolker-Nielsen, T., Ragas, P.C., and Mattick, J.S. (2002) Extracellular DNA required for bacterial biofilm formation. *Science (80-)* **295**: 1487.

- Wickham, H. (2008) *Elegant Graphics for Data Analysis: ggplot2*, Springer (ed).
- Winsor, G.L., Van Rossum, T., Lo, R., Khaira, B., Whiteside, M.D., Hancock, R.E.W., and Brinkman, F.S.L. (2009) *Pseudomonas* Genome Database: Facilitating user-friendly, comprehensive comparisons of microbial genomes. *Nucleic Acids Res* **37**: 483–488.
- Yin, W., Wang, Y., Liu, L., and He, J. (2019) Biofilms: The microbial “protective clothing” in extreme environments. *Int J Mol Sci* **20**: 1–18.
- Zdobnov, E.M. and Apweiler, R. (2001) InterProScan - An integration platform for the signature-recognition methods in InterPro. *Bioinformatics* **17**: 847–848.
- Zhang, J. (2012) Genetic redundancies and their evolutionary maintenance. *Adv Exp Med Biol* **751**: 279–300.

2.8 Supplementary Material

Table S2.1 Bacterial strains, plasmids, and oligonucleotides used in this chapter. Abbreviations: Gm^R, gentamicin resistance; Km^R, kanamycin resistance; MCS, multiple cloning site.

Strains, plasmids or oligonucleotides		Genotype, phenotype or relevant characteristics	Reference
Strains			
<i>P. stutzeri</i>	ATCC 17587	Wild-type phenotype	Laboratory strain
<i>P. stutzeri</i>	$\Delta pilA$	ATCC 17587 $\Delta pilA$	This study
<i>P. stutzeri</i>	$\Delta comA$	ATCC 17587 $\Delta comA$	This study
<i>P. stutzeri</i>	$\Delta endA$	ATCC 17587 $\Delta endA$	This study
<i>E. coli</i>	DH5 α λ pir	Wild-type phenotype	Laboratory strain
<i>C. crescentus</i>	NA1000	Wild-type phenotype	Laboratory strain
<i>S. marcescens</i>	DSM 30121	Wild-type phenotype	Laboratory strain
<i>S. ureilytica</i>	NiVa51	Wild-type phenotype	Laboratory strain
<i>S. ureilytica</i>	Lr 5 / 4	Wild-type phenotype	Laboratory strain
Plasmids			
	pEMG	Km ^R , oriR6K, MCS with two flanking I-SceI sites, used for markerless deletions	(Martínez-García and de Lorenzo, 2011)
	pSW-2	Gm ^R , oriRK2, <i>xyIS</i> , Pm $\rightarrow$ I-sceI (transcriptional fusion of I-sceI to Pm promoter)	(Martínez-García and de Lorenzo, 2011)
	pEMG- $\Delta pilA$	pEMG bearing upstream and downstream flanking regions of <i>pilA</i>	This study
	pEMG- $\Delta comA$	pEMG bearing upstream and downstream flanking regions of <i>comA</i>	This study
	pEMG- $\Delta endA$	pEMG bearing upstream and downstream flanking regions of <i>endA</i>	This study
Oligonucleotides			
	PilA_UFR_F	AAAGAATTCAACTGGTCGGATGCAAGC, with a EcoRI underlined site, located upstream of <i>pilA</i>	This study
	PilA_UFR_R	ACAAACGACAGAGTTTGGATCTCCTTAAGTGTTAA GGC, with overlap to downstream flanking region underlined, located upstream of <i>pilA</i>	This study
	PilA_DFR_F	AAGGAGATCCAACTCTGTCGTTTGAACGCGGTA, with overlap to upstream flanking region underlined, located downstream of <i>pilA</i>	This study
	PilA_DFR_R	AAAGGATCCAGGTTCATTAGCCGAAGGCGTAA, with a BamHI underlined site, located downstream of <i>pilA</i>	This study
	ComA_UFR_F	AAAGAATTCCAGGTGCCGATCTTGTAGTT, with a EcoRI underlined site, located upstream of <i>comA</i>	This study
	ComA_UFR_R	TCATTTTCCCGCTTTCAGTCTCCAGAAATCCTTTCC, with overlap to downstream flanking region underlined, located upstream of <i>comA</i>	This study
	ComA_DFR_F	CTGGAGACTGAAAGCGGGAAAAATGAGAACGG, with overlap to upstream flanking region underlined, located downstream of <i>comA</i>	This study
	ComA_DFR_R	AAAGGATCCATGAAGGCGCTGAAGATCTC, with a BamHI underlined site, located downstream of <i>comA</i>	This study
	EndA_UFR_F	AAAGAATTCCGCCAGGTGGCCTACAAGC, with a EcoRI underlined site, located upstream of <i>endA</i>	This study
	EndA_UFR_R	GCTCTCATATGCAGCACGTCGAGTGTCGCATCCATG CG, with overlap to downstream flanking region underlined, located upstream of <i>endA</i>	This study

EndA_DFR_F	<u>GATGCGACACTC</u> GACGTGCTGCATATGAGAGC, with overlap to upstream flanking region underlined, located downstream of <i>endA</i>	This study
EndA_DFR_R	AAAG <u>GATCC</u> AGTGCAGAAAGCCCAGC, with a BamHI underlined site, located downstream of <i>endA</i>	This study
rpoB_Ps_all_F	AACGTATCCGCAAGGACT, for amplification of <i>rpoB</i> in <i>P. stutzeri</i>	This study
rpoB_Ps_all_R	TATCGATACCGAGCGAGC, for amplification of <i>rpoB</i> in <i>P. stutzeri</i>	This study
rpoB_Cb_all_F	CAATCCTTCACCGCAAG, for amplification of <i>rpoB</i> in <i>C. crescentus</i>	This study
rpoB_Cb_all_R	GAGCGATTTCCTTGACC, for amplification of <i>rpoB</i> in <i>C. crescentus</i>	This study

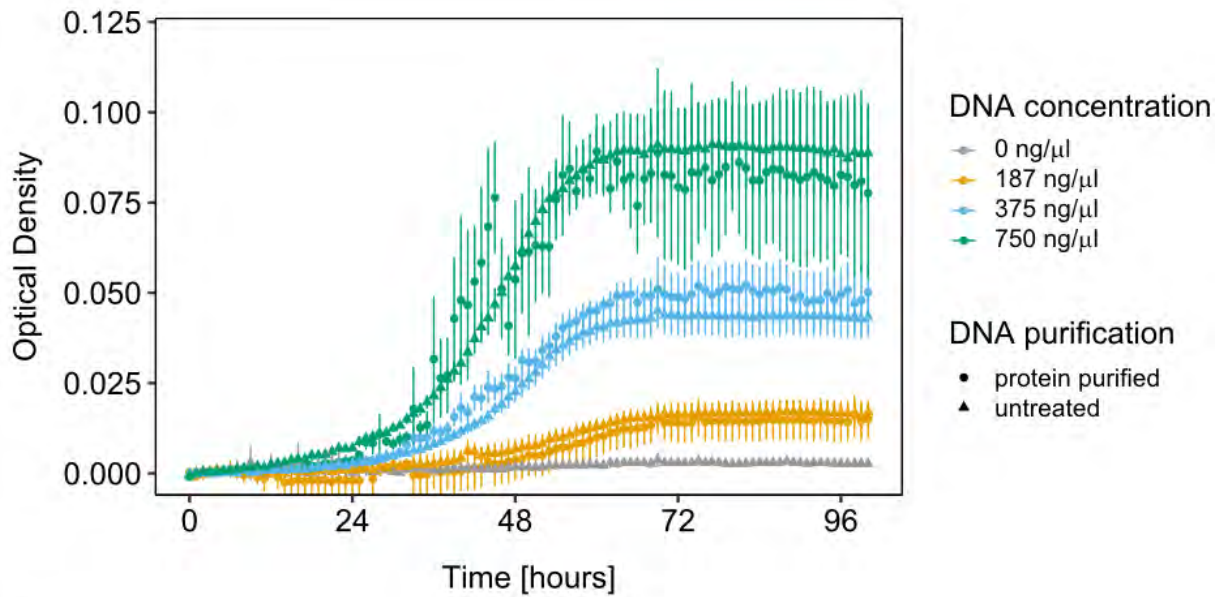
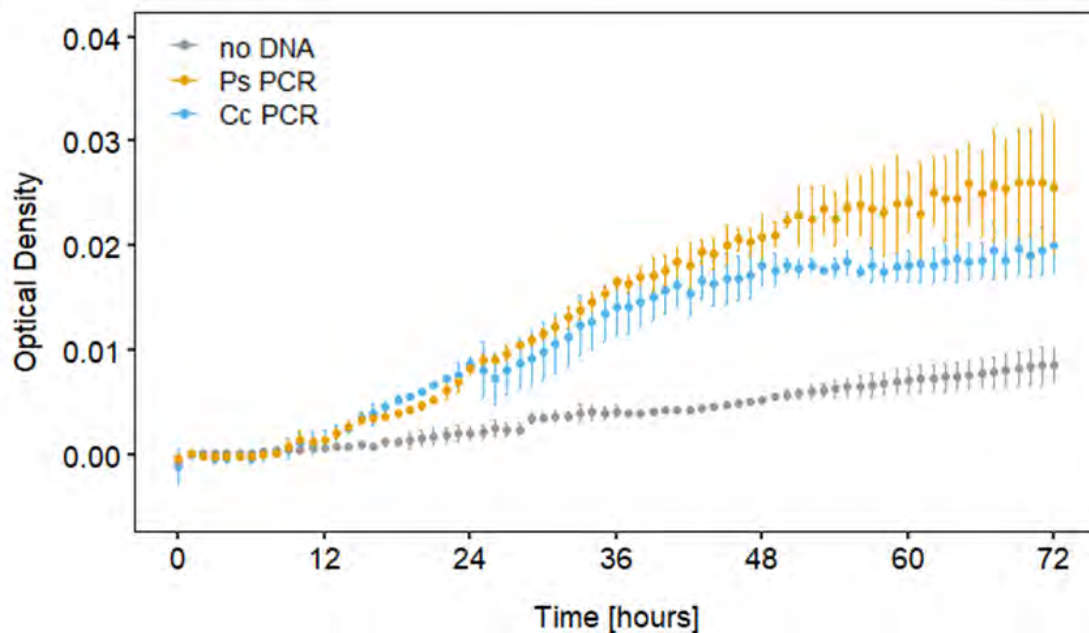
A**B**

Figure S2.1 Growth kinetic of *P. stutzeri* using fish DNA or PCR products as the sole source of carbon indicated that the commercial DNA was free from contaminants and confirmed that growth was due to DNA presence. (A) Concentrations of DNA: 750 ng/μL (purple), 375 ng/μL (blue), 187 ng/μL (green), 0 ng/μL (orange). Higher growth is obtained with higher concentration of DNA. The fish DNA was either purified by ethanol precipitation (triangle) or by proteinase K-digestion followed by phenol/chloroform/isoamyl alcohol purification (circle). The purification does not impact the growth indicating that the DNA was not contaminated by proteins which might have served as a source of nitrogen and carbon. The points represent the mean of four replicates, the error bars stand for the standard deviation. The measures were performed in a 96-well plate reader Synergy HT (BioTek, USA). (B) The observation that PCR products also supported growth strengthens the evidence that nucleic acids, rather than contaminants, enables growth. The PCR products are amplifications of the gene *rpoB* of *P. stutzeri* ATCC 17587 and *C. crescentus* NA1000. The products had a length of 4 kb and a GC content of 59% for *P. stutzeri* and 64% for *C. crescentus*. PCR products were purified with a kit and diluted to a final concentration of 37.5 ng/μL.

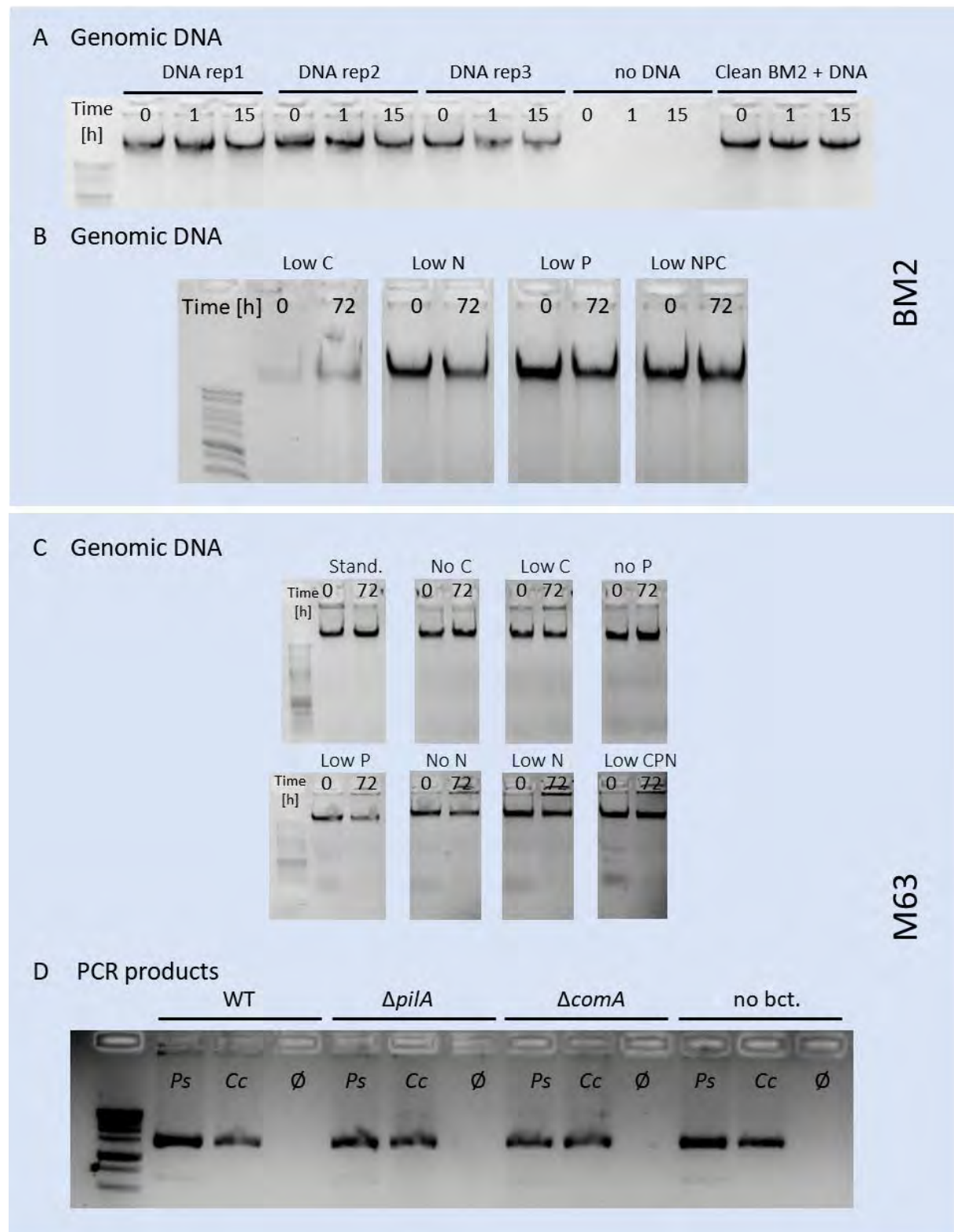


Figure S2.2 No clear degradation of DNA was observed with gel electrophoresis despite using two different media and their variations. Briefly, experiments were performed by mixing either filtered supernatants (A and B) or overnight bacterial cultures (C and D) of *P. stutzeri* wild-type (A, B, C, D) and natural transformation mutants (D) with *P. stutzeri* genomic DNA (A,B,C) or *rpoB* purified PCR products. While the mix was incubated at 30°C, samples were retrieved and stored at -20°C before gel electrophoresis. Incubation time is indicated on the graphs (A, B, C) in hours. For (D), only samples after 7 days of incubation were loaded on a 1% agarose gel. Final DNA concentrations were 250 ng/μL (A), 100 ng/μL (B), 20 ng/μL (C) and 37.5 ng/μL (D).

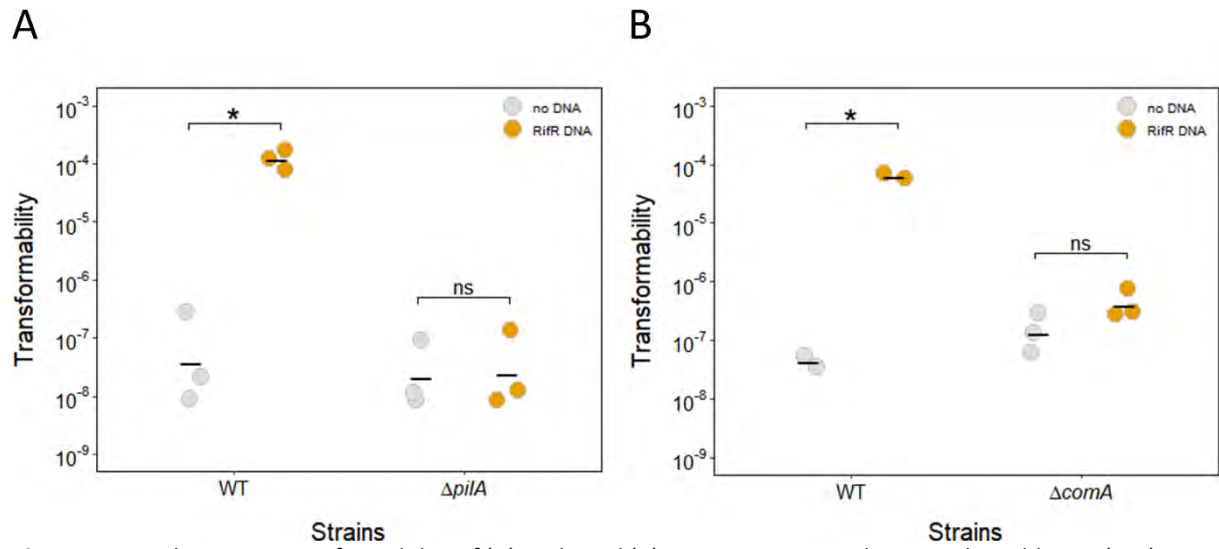


Figure S2.3 Reduction in transformability of (A) $\Delta pilA$ and (B) $\Delta comA$ mutants relative to the wild-type (WT) strain confirms the essentiality of these genes for natural transformation. Each condition had three replicates except WT strain in (B) (only two) indicated with dots, the mean is shown with the black line.

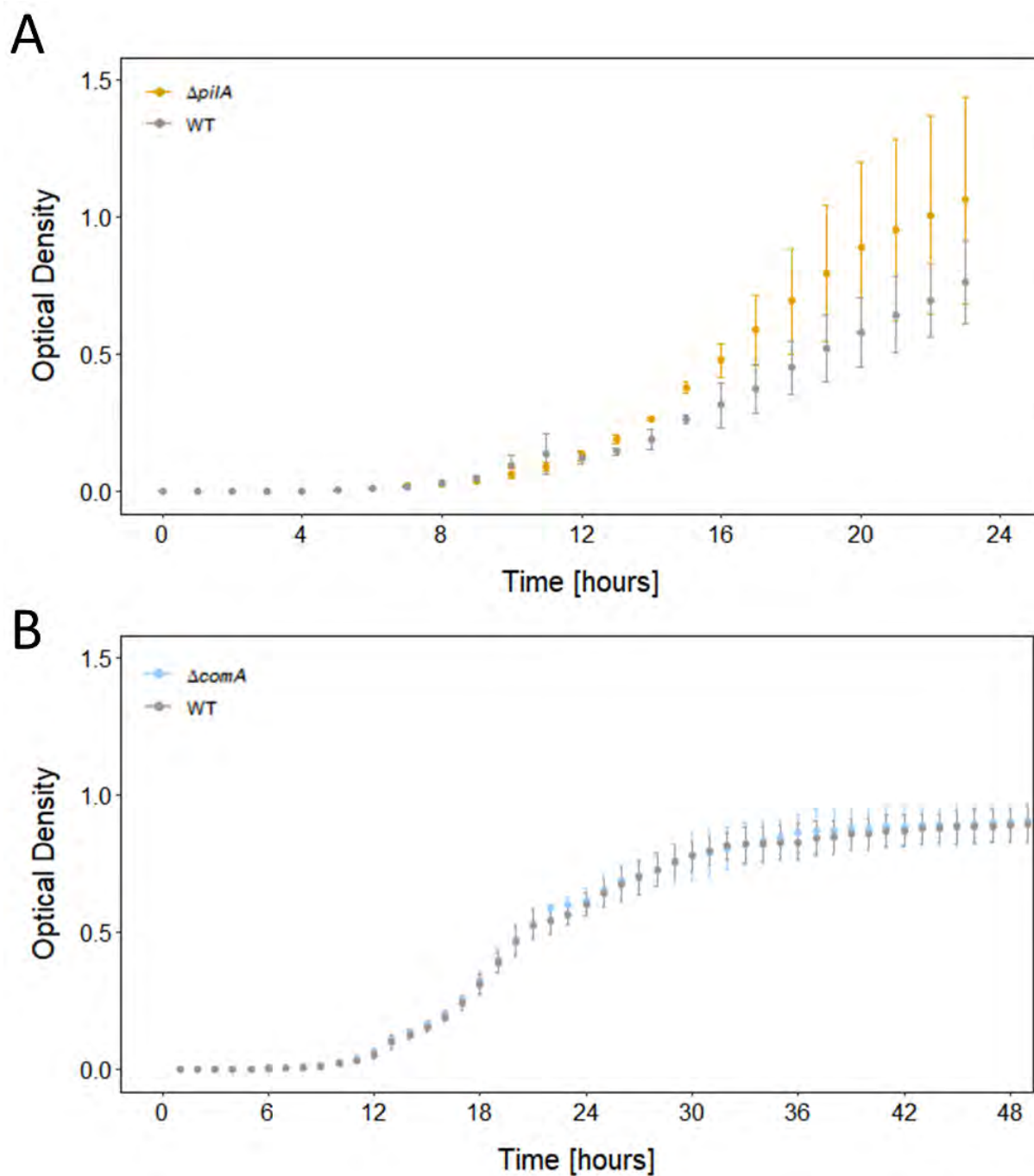


Figure S2.4 Growth kinetics in LB medium of *P. stutzeri*: (A) $\Delta pilA$ mutant (orange) compared to the wild-type strain (grey), and (B) $\Delta comA$ mutant (blue) compared to the wild-type strain (grey). The points represent the mean of two replicates for (A) and three replicates for (B), and the error bars stand for the standard deviation. The measures of each panel were performed at different dates and plates: 384-wells plate for (A) and 96-well plate for (B).

Table S2.2 Genes encoding potential and confirmed nucleases in our genome assembly of *P. stutzeri* ATCC 17587. Amino-acids sequences were submitted to SignalP server to predict whether the proteins had signal peptides. Data can find be found in the 8 columns on the right. The best candidates for extracellular degradation are highlighted in grey.

Locus ID	Gene	Start	End	Description	Prediction	OTHER	SP	LIPO	TAT	TATLIPO	PILIN	CS Position
01350		157201	158514	type I restriction endonuclease subunit S	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
05930	rph	671343	672065	ribonuclease PH	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
06830		763373	764167	nuclease	SP	0.0008	0.9981	0.0003	0.0003	0.0002	0.0002	CS pos: 28-29. Pr: 0.9684
08380	sbcC	924758	928411	nuclease SbcCD subunit C	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
08390	sbcD	928408	929643	nuclease SbcCD subunit D	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
08490		936560	937342	endonuclease	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
11390	sbcB	1234171	1235607	exodeoxyribonuclease I	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
11800	rnt	1275264	1275944	ribonuclease T	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
11860	nth	1280164	1280802	endonuclease III	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
12120	rnc	1305390	1306079	ribonuclease 3	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
12610		1358848	1359255	endoribonuclease	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
13350	ruvC	1425646	1426170	crossover junction endodeoxyribonuclease RuvC	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
13700	rnz	1460332	1461315	ribonuclease Z	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
14200		1507068	1507871	DNA repair exonuclease	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
14420	rnd	1527704	1528828	ribonuclease D	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
15240	recJ	1613888	1615603	single-stranded-DNA-specific exonuclease RecJ	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
16530	rne	1751812	1754937	ribonuclease E	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
24380	fan1	2598814	2600451	Fanconi-associated nuclease	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
24720		2638109	2638912	endonuclease/exonuclease/phosphatase	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
24990		2683626	2684438	exodeoxyribonuclease III	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
25790	rnhA	2773636	2774091	ribonuclease HI	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
26010	wrn	2800591	2801226	ribonuclease D	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
27420		2954833	2955177	HNH endonuclease	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	

27720	rnhB	2985234	2985872	ribonuclease HII	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	CS pos: 19-20. Pr: 0.9738
29570	xseA	3199071	3200447	exodeoxyribonuclease 7 large subunit	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
29660	endA	3210785	3211516	DNA-specific endonuclease I	SP	0.0002	0.9991	0.0002	0.0002	0.0002	0.0002	
30530	nfi	3303294	3304085	nuclease	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
35260		3808759	3809544	TatD-related deoxyribonuclease	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
35300		3813353	3814093	endonuclease V	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
35590		3851528	3854029	ribonuclease R	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
35900		3889430	3890470	endonuclease	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
36170	xseB	3926830	3927072	exodeoxyribonuclease 7 small subunit	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
36610	cas1	3974657	3975526	CRISPR-associated endonuclease Cas1	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
36630	cse3	3976451	3977062	CRISPR-associated endonuclease Cse3	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
36680	cas3	3980940	3983627	CRISPR-associated endonuclease/helicase Cas3	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
36790	ybeY	3995624	3996106	endonuclease YbeY	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
37090	orn	4024300	4024842	oligoribonuclease	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
39260		4287824	4288252	putative pre-16S rRNA nuclease	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
40330		4410392	4411261	endonuclease	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	
40790		4452233	4452709	putative endonuclease	SP	0.0003	0.9989	0.0002	0.0002	0.0002	0.0002	CS pos: 18-19. Pr: 0.9736
42810	rnpA	4667946	4668341	ribonuclease P protein component	OTHER	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	

Table S2.3 Genes encoding potential and confirmed nucleases in the genome assembly of *P. stutzeri* ATCC 17587 (GCF_003410115.1) from the database of <https://www.pseudomonas.com/> (Winsor et al., 2009). Amino-acids sequences from genes whose protein localizations were unknown or extracellular were submitted to SignalP server to predict whether the proteins had signal peptides. A low value indicated that the protein contained a signal peptide. The best candidates for extracellular degradation are highlighted in grey.

Locus Tag	RefSeq Accession	Gene	Start	End	Description	Localization	SignalP likelihood
D0A22_RS05195	WP_013983856.1	sfsA	37779	38486	DNA/RNA nuclease SfsA	cytoplasmic	
D0A22_RS13585	WP_011913281.1		163039	163842	endonuclease/exonuclease/phosphatase family protein	cytoplasmic	
D0A22_RS20220	WP_151321071.1	recC	2192	5611	exodeoxyribonuclease V subunit gamma	cytoplasmic	
D0A22_RS16795	WP_151320855.1	recD	1927	3777	exodeoxyribonuclease V subunit alpha	cytoplasmic	
D0A22_RS03100	WP_023082699.1		608086	609294	GIY-YIG nuclease family protein	cytoplasmic	
D0A22_RS08835	WP_011911698.1		189583	190305	ribonuclease PH	cytoplasmic	
D0A22_RS06710	WP_151320316.1		355469	356194	deoxyribonuclease V	cytoplasmic	
D0A22_RS14055	WP_011913378.1	rnhA	284569	285024	ribonuclease HI	unknown	1.0001
D0A22_RS20565	WP_017244187.1		76373	76855	thermonuclease family protein	unknown	0.0002
D0A22_RS11730	WP_125459841.1	recJ	193441	195156	single-stranded-DNA-specific exonuclease RecJ	cytoplasmic	
D0A22_RS18470			57523	58298	endonuclease (pseudogene)	unknown	1.0001
D0A22_RS03590	WP_151320171.1	nth	698741	699379	endonuclease III	cytoplasmic	
D0A22_RS00495	WP_011913540.1		100778	101119	ribonuclease E inhibitor RraB	unknown	1.0001
D0A22_RS06925	WP_151320321.1	mutL	404371	406248	DNA mismatch repair endonuclease MutL	cytoplasmic	
D0A22_RS16025	WP_151320803.1		72756	74132	exodeoxyribonuclease VII large subunit	cytoplasmic	
D0A22_RS21100	WP_151321103.1		8465	9478	ribonuclease Z	cytoplasmic	
D0A22_RS12215	WP_151320601.1		296231	297247	ligase-associated DNA damage response exonuclease	cytoplasmic	
D0A22_RS18535	WP_151320989.1		67311	70964	exonuclease SbcC	unknown	1
D0A22_RS03815	WP_151320175.1	sbcB	742281	743714	exodeoxyribonuclease I	cytoplasmic	
D0A22_RS01755	WP_011912714.1	rnhB	358489	359127	ribonuclease HII	cytoplasmic	
D0A22_RS07535	WP_011914801.1		534024	534566	oligoribonuclease	cytoplasmic	
D0A22_RS16345	WP_014597419.1		137150	137425	GIY-YIG nuclease family protein	unknown	1
D0A22_RS04660	WP_151320229.1		911281	911709	endonuclease domain-containing protein	unknown	1.0001
D0A22_RS13280			97438	99260	excinuclease ABC subunit UvrC (pseudogene)	unknown	1.0001
D0A22_RS20180	WP_011915212.1		117099	117494	ribonuclease P protein component	cytoplasmic	

D0A22_RS16790			1	1856	exodeoxyribonuclease V subunit beta (pseudogene)	unknown	1
D0A22_RS02720	WP_013982184.1		544066	544869	DNA repair exonuclease	cytoplasmic	
D0A22_RS08825	WP_011911700.1		188344	189123	exodeoxyribonuclease III	cytoplasmic	
D0A22_RS05925	WP_151320283.1		193781	194434	exonuclease domain-containing protein	cytoplasmic	
D0A22_RS13750	WP_013982852.1	xthA	215421	216233	exodeoxyribonuclease III	cytoplasmic	
D0A22_RS20345	WP_011915061.1		35670	36539	endonuclease	cytoplasmic	
D0A22_RS17850	WP_151320922.1		62085	64034	relaxase/mobilization nuclease and DUF3363 domain-containing protein	cytoplasmic	
D0A22_RS03460	WP_013982082.1		674092	674781	ribonuclease III	cytoplasmic	
D0A22_RS11420	WP_011913913.1	ruvC	135568	136092	crossover junction endodeoxyribonuclease RuvC	cytoplasmic	
D0A22_RS00245	WP_023114177.1		44767	46752	relaxase/mobilization nuclease and DUF3363 domain-containing protein	cytoplasmic	
D0A22_RS06875	WP_011914677.1		393519	396020	ribonuclease R	cytoplasmic	
D0A22_RS15965	WP_045158941.1	uvrB	59839	61854	excinuclease ABC subunit B	cytoplasmic	
D0A22_RS20850	WP_151321092.1		44387	45409	UV damage endonuclease UvsE	cytoplasmic	
D0A22_RS12200	WP_013983281.1	pdeM	291131	291820	ligase-associated DNA damage response endonuclease PdeM	unknown	1
D0A22_RS18530	WP_151320988.1	sbcD	66079	67314	exonuclease subunit SbcD	cytoplasmic	
D0A22_RS03620	WP_125870844.1	rnt	703595	704269	ribonuclease T	cytoplasmic	
D0A22_RS01605	WP_011912744.1		327360	327704	HNH nuclease family protein	unknown	1
D0A22_RS07115	WP_003283056.1		450595	450837	exodeoxyribonuclease VII small subunit	cytoplasmic	
D0A22_RS16070	WP_053527841.1		84125	84856	deoxyribonuclease	extracellular	0.0012
D0A22_RS04515			879653	881108	ribonuclease G (pseudogene)	unknown	1
D0A22_RS12365	WP_049324507.1		327821	330979	ribonuclease E	cytoplasmic	1.0001
D0A22_RS18725	WP_151320998.1	uvrA	101205	104042	excinuclease ABC subunit UvrA	membrane	
D0A22_RS02605	WP_013982206.1	rnd	522280	523404	ribonuclease D	cytoplasmic	
D0A22_RS08380	WP_151320409.1		95952	96746	thermonuclease family protein	unknown	0.0005
D0A22_RS16520			179959	180749	exodeoxyribonuclease III (pseudogene)	unknown	1.0001

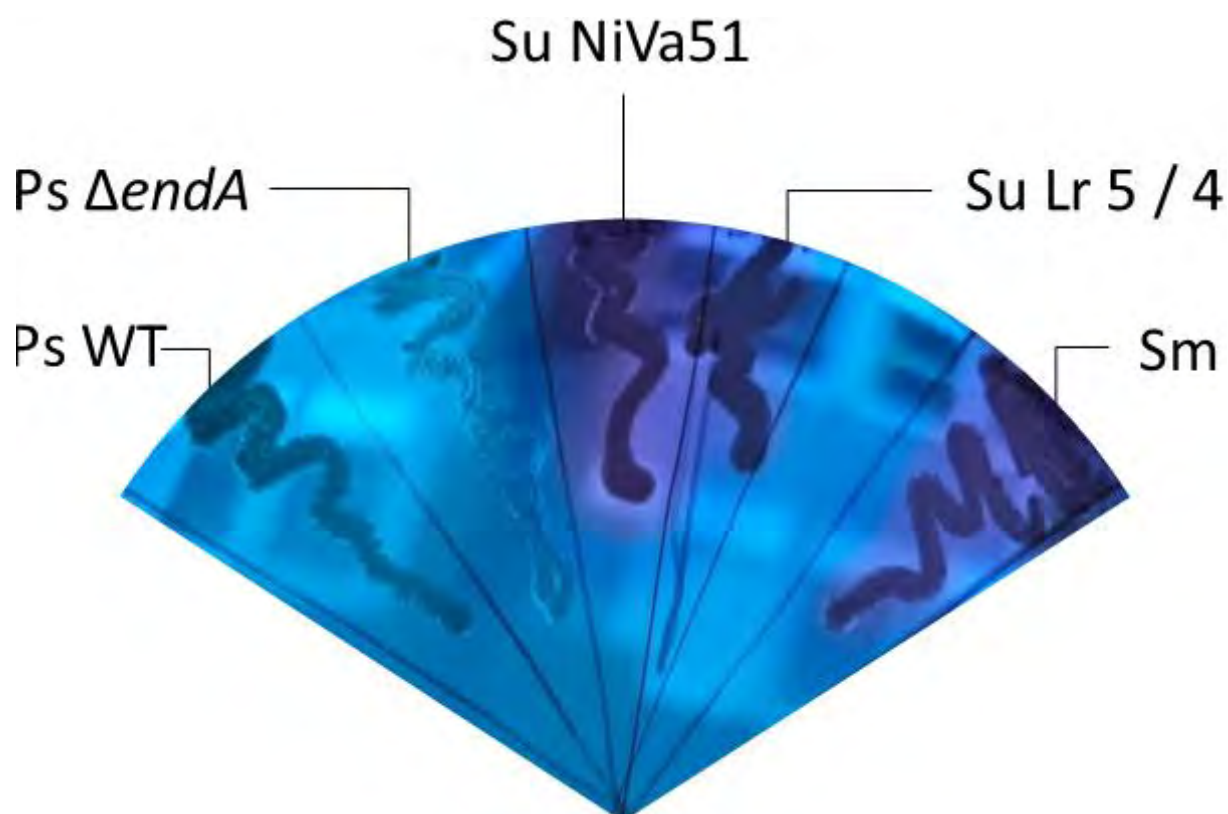


Figure S2.5 No DNA degradation was visible in presence of *P. stutzeri*, unlike certain *Serratia* strains, on agar plates supplemented with DNA and toluidine blue O (TBO) dye. In the presence of unhydrolyzed DNA, TBO exhibits a royal blue color. When TBO is complexed with the oligonucleotides and mononucleotides released during DNA hydrolysis, the color spectrum shifts from blue to pink. The pink halo around *Serratia marcescens* (SM) and *Serratia ureilytica* (Su Niva51) indicates that these strains produced nucleases in contrast to *P. stutzeri* wild-type (Ps WT) and its $\Delta endA$ mutant (Ps $\Delta endA$). The results were less clear for *S. ureilytica* Lr 5 / 4 (Su Lr 5 / 4).

Chapter 3 - Regulation of natural competence in *P. stutzeri*

3.1 Foreword

In this part of the project, we investigated three aspects of the regulation of competence in *P. stutzeri*: the triggers of competence, the expression of competence at the single-cell level, and the role of TfoX-like transcriptional regulators in the regulation of the transformability of *P. stutzeri* and other phenotypes.

The key findings of this chapter are the following:

- Most of the population of *P. stutzeri* expresses competence genes mainly at the onset of the stationary phase, but *pilA* presents significant variation in expression levels across cells.
- Expression of competence genes is reduced when extracellular DNA concentrations are high or under genotoxic stresses caused by antibiotics, but mild ultraviolet radiations significantly increase competence.
- TfoX-like homologues do not regulate natural transformation in *P. stutzeri*, but in their absence, we detected higher biofilm production.

3.2 Abstract

Natural transformation is a key mechanism of horizontal gene transfer, enabling the uptake and integration of extracellular DNA into the host genome. Recent studies have revealed that, although phylogenetically distant species share conserved proteins for DNA uptake and processing, they vary in the triggers and regulatory mechanisms involved. To date, research on the natural transformation of *P. stutzeri* has mainly focused on the machinery rather than on the regulatory aspects. Here we investigate two key aspects of competence regulation in *P. stutzeri*: the expression of competence genes at both the cellular and population levels under various conditions, and the role of TfoX-like homologues in transformability. Using transcriptional reporters for the competence genes *comA* and *pilA*, we found that high concentrations of extracellular DNA decreased the expression of these genes, while mild ultraviolet radiation increased their expression. At the single-cell level, we observed that most cells expressed competence genes during the transition to the stationary phase, although *pilA* showed more heterogeneous expression compared to *comA*. The deletion of two *tfoX* homologues did not significantly affect natural transformation, suggesting that these genes are unlikely to regulate competence in *P. stutzeri*. These findings indicate that *P. stutzeri* regulates competence in a specific way compared to other Gram-negative models and emphasize the importance of studying competence in multiple species to fully understand its variability and ecological significance.

3.3 Introduction

Natural transformation is a horizontal gene transfer mechanism, whereby extracellular DNA is internalized and integrated into the host genome by homologous recombination. While this process potentially confers adaptive advantages through the acquisition of novel genetic material, it is associated with significant metabolic costs and inherent risks for the bacterial host. The integration process can be lethal if unsuccessful (Ambur *et al.*, 2016), and even successful transformation events may not always yield beneficial or even neutral outcomes. For instance, the incorporation of damaged DNA (Overballe-Petersen *et al.*, 2013) or genetic material from less fit individuals is possible, as these are more likely to release their DNA into the environment upon death (Redfield, 1988; Overballe-Petersen *et al.*, 2013). Regardless of the DNA incorporated, natural transformation requires a specific physiological state called competence, which can involve the transcription of more than a hundred genes in some species (Prudhomme *et al.*, 2006; Ambur *et al.*, 2016). Indirect physiological costs have been observed in transforming *Bacillus subtilis* cells, which suffer from a growth delay for several hours after transformation (Haijema *et al.*, 2001). Despite these potential drawbacks, the evolutionary persistence of natural transformation in numerous bacterial lineages suggests that the adaptive benefits conferred by this mechanism generally outweigh the associated costs.

To mitigate the cost of natural competence, transformable bacteria have developed regulatory strategies, often at the transcriptional level. One of these strategies is to develop competence in response to environmental cues, that correlate with times when DNA uptake is favourable (Seitz and Blokesch, 2012). These signals are usually linked to nutrition (Spencer and Herriott, 1965; Sparling, 1966), stresses (Prudhomme *et al.*, 2006; Charpentier *et al.*, 2011; Boutry *et al.*, 2013), and cell-cell signaling (Claverys *et al.*, 2006; Lo Scrudato and Blokesch, 2012). Similarly, for certain species, only a fraction of the population exhibits competence, as seen in *B. subtilis*, where 10 to 20 percent of individuals develop competence (Haijema *et al.*, 2001; Maamar and Dubnau, 2005; Maamar *et al.*, 2007). Phenotypic heterogeneity within a clonal population has been postulated to be advantageous because it mitigates the extinction risks of a population in unpredictably changing environments (Veening *et al.*, 2008; Grimbergen *et al.*, 2015). Another strategy might be to enhance the probability of integrating good-quality species-specific DNA. This mechanism is exemplified by predation of non-kin

neighboring cells performed by *Vibrio cholerae*, which utilizes its type VI secretion system (T6SS) to lyse adjacent cells, thereby facilitating the uptake of freshly released genomic material (Borgeaud *et al.*, 2015; Matthey *et al.*, 2019). Understanding the factors that trigger competence and how it is regulated is essential for determining when natural competence may be beneficial.

Previous studies indicate that the regulation of competence varies significantly between species, whereas DNA uptake and processing mechanisms are relatively conserved (Johnston *et al.*, 2014). Indeed, both competence-inducing cues and regulatory cascades exhibit notable differences across species. Consequently, it is essential to examine a diverse array of models. Much of the existing research on regulatory mechanisms has concentrated on the pathogenic bacteria *V. cholerae*, *Haemophilus influenzae*, *Streptococcus pneumoniae*, and on the model organism *B. subtilis* (Johnsborg *et al.*, 2007; Johnston *et al.*, 2014; Dubnau and Blokesch, 2019). However, there is a need to also investigate environmental transformable bacteria, which face very different conditions than pathogens.

One model of environmental bacteria for natural transformation is *Pseudomonas stutzeri*, which was mainly studied by the group of Wackernagel until 2005. This group demonstrated the role of some genes like *pilA*, *pilC*, *pilT*, *pilU*, *comA*, and *exbB* in the uptake and processing of DNA (Graupner *et al.*, 2000, 2001; Graupner and Wackernagel, 2001). The same group studied conditions affecting transformability such as nutrients, temperature, and pH. As *P. stutzeri* experiences nutrient limitations in soil, Lorenz and Wackernagel tested growth and transformability on soil extract medium (Lorenz and Wackernagel, 1991). Limitations of phosphorus, nitrogen, or carbon increased natural transformation in *P. stutzeri*, but simultaneous restriction of combinations of these nutrients resulted in low to undetectable transformation levels (Lorenz and Wackernagel, 1991, 1992), indicating that an insufficient supply of nutrients prevented competence development. Regarding other physical and chemical parameters, the transformation optimum was measured at around 30°C in neutral to moderately alkaline conditions (Lorenz and Wackernagel, 1990a, 1992). Despite indications about conditions favourable to natural competence, the molecular regulatory network of *P. stutzeri* was never explored.

Molecular regulation of competence in Gram-negative bacteria has been mainly studied in *V. cholerae* and *H. influenzae*. The *tfoX* gene of *V. cholerae* and its ortholog in *H. influenzae*,

sxy, encode functionally equivalent master regulators of natural competence (Williams *et al.*, 1994; Meibom *et al.*, 2005). In both *V. cholerae* and *H. influenzae*, both *tfoX* and *sxy* work in concert with the cAMP receptor protein (CRP) for competence development (Redfield *et al.*, 2005; Lo Scrudato *et al.*, 2014; Wu *et al.*, 2015). The interaction between CRP and Sxy has been demonstrated as well in *E. coli* (Cameron and Redfield, 2006; Søndberg *et al.*, 2019), highlighting the conservation of the mechanism. In the presence of a high concentration of cAMP, the cAMP-CRP complex binds to specific DNA sites, named CRP sites, affecting the transcription of numerous genes. *H. influenzae* has two types of CRP sites, canonical sites like *E. coli*'s and sites requiring both CRP and the coregulator Sxy (Cameron and Redfield, 2006); the latter are called CRP-S sites. Phylogenetic footprinting identified CRP-S sites in the promoter regions of *Enterobacteriaceae*, *Pasteurellaceae*, and *Vibrionaceae* orthologs of *H. influenzae* CRP-S-regulated genes (Cameron and Redfield, 2006). Promotor orthologs of *H. influenzae* CRP-S-regulated genes were also found in *Pseudomonadaceae* and *Xanthomonadaceae*, but no conserved motifs were detected in the promoters from either family (Cameron and Redfield, 2006). Therefore, homolog of TfoX/Sxy might be a good candidate regulator for natural transformation in *P. stutzeri*.

To gain a deeper understanding of competence regulation in *P. stutzeri*, we explored three aspects: the triggers of competence, the expression of competence at the single-cell level, and the role of *tfoX/sxy* (named only *tfoX* hereafter) in the transformability of *P. stutzeri*. Firstly, we evaluated whether the presence of DNA or stressing factors increases the development of competence at a population level by monitoring the activity of the *comA* and *pilA* promoters. We observed interesting effects of extracellular DNA and UV stress on the expression of these genes population-wide. Secondly, we studied the expression of competence at the single-cell level to determine whether the whole population or only a fraction of the population enters the competence state. Our results indicated that most of the cells in the population were expressing competence genes during the transition from the exponential phase to the stationary phase; however, while the expression was homogenous for *comA*, the transcription of *pilA* showed strong heterogeneity within isogenic populations. Finally, we investigated whether *P. stutzeri* homologues of *tfoX*, a key regulator of natural transformation in *V. cholerae* (Meibom *et al.*, 2005) and *H. influenzae* (Williams *et al.*, 1994), also had a role in natural transformation in *P. stutzeri*. The deletion of two *tfoX* homologues in *P. stutzeri* had

no clear effect on the transformability of *P. stutzeri* but increased slightly biofilm production. Collectively, these findings underscore the differences between *P. stutzeri* and well-studied Gram-negative models; they also emphasize the importance of investigating additional models to broaden our knowledge of bacterial competence regulation across diverse species.

3.4 Materials and Methods

3.4.1 Bacterial strains and medium

The bacterial strains, plasmids, and oligonucleotides used in this chapter are listed in Table S 3.1. Strains of *P. stutzeri* and *E. coli* were grown in LB medium (tryptone, 10 g/L; yeast extract 5 g/L; NaCl, 5 g/L) at 30°C overnight (with 180 rpm agitation) or on LB agar plates (solidified with 1.5% agar). When required, antibiotics were added at the following concentration: ampicillin (Amp), 100 µg/mL; gentamicin (Gm), 15 µg/mL for *E. coli* and 30 µg/mL for *P. stutzeri*; kanamycin (Km), 50 µg/mL; rifampicin (Rif), 50 µg/mL. Experiments assessing expression of reporter strain were done in M63 medium (100mM KH₂PO₄, 15mM (NH₄)₂SO₄, 0.5mg/L FeSO₄*7H₂O, 1mM MgSO₄*7H₂O, 0.2% glucose).

3.4.2 Strains and plasmids construction

Two types of mutants were constructed for this project: deletion mutant and transcriptional reporter strains. The steps are briefly described below.

The *P. stutzeri* deletion mutants $\Delta tfoXC$, $\Delta tfoXN$, $\Delta clpV$ were constructed by forced recombination between homologous DNA sequences according to the method of Martínez-García and de Lorenzo (Martínez-García and de Lorenzo, 2011). Briefly, the upstream and downstream flanking regions of the locus of interest were amplified by PCR separately. Then, both fragments were joined together by PCR. The resulting fragment was purified, digested with EcoRI and BamHI restriction enzymes (RE), and ligated to a EcoRI-BamHI-digested pEMG plasmid. The ligation mix was then inserted in *E. coli* DH5 α λ *pir* by heat-shock followed by the selection of complete plasmid and further propagation of the plasmid. Then, *P. stutzeri* was electroporated with the plasmid pEMG- $\Delta tfoXC$ or pEMG- $\Delta tfoXN$ or pEMG- $\Delta clpV$, which inserts itself in the host genome by homologous recombination of the flanking regions insert. To excise the pEMG from the host genome, pSW-2 plasmid encoding I-SceI restriction enzyme was electroporated in a *P. stutzeri* clone obtained from the first electroporation. The loss of the pEMG was first evaluated by restriking Gm-resistant colonies

from the second electroporation on LB-Km. The susceptibility to Km indicates the removal of pEMG from the host genome. The deletion of the gene of interest was finally confirmed by PCR.

P. stutzeri reporter strains were constructed by insertion into the genome of a single copy of a promoter coupled to a *sfGFP* via a Tn7 vector system. For that purpose, the *sfGFP* and the promoters of *pilA* and *comA* were amplified by PCR with the primers listed in Table S 3.1. Then, the PCR fragments were digested with the RE EcoRI, SphI, and HindIII (New England BioLabs) and ligated to the pME6182 plasmid, which was also digested with RE. The ligation mix was then inserted in *E. coli* DH5α λ *pir* by heat-shock followed by the selection of complete plasmid and further propagation of the plasmid. The transposon delivery plasmids pME6182-P*pilA*-*sfGFP* or pME6182-P*comA*-*sfGFP* or pME6182-*sfGFP* and the transposition helper plasmid pUX-BF13 were then introduced by electroporation into *P. stutzeri*. Finally, the transposon insertion was assessed by PCR. Regarding the construct of the unstable *sfGFP*, an RPAANDENYAAAV tag was added to the C-terminal end of the *sfGFP* protein by using an especially long primer (Andersen *et al.*, 1998).

For the bacterial two-hybrid assays, the genes of interest (*tfoXC*, *tfoXN*, CRP) were amplified by PCR from the genomic DNA of *P. stutzeri* ATCC17587 and cloned in both pKT25 and pUT18C plasmids using BamHI and EcoRI RE (New England Biolabs). The integrity of the cloned sequences was verified by Sanger sequencing. The pKT25- and pUT18C-derived plasmids were co-transformed in *E. coli* BTH101; positive clones were selected on both kanamycin and ampicillin.

3.4.3 Isolation and purification of DNA

In transformation assays, we used DNA extracted from a spontaneous rifampicin-resistant mutant of *P. stutzeri* ATCC 17587. To extract the DNA, firstly the bacterial cells from 500 μ L overnight culture were lysed in 500 μ L lysis buffer (10 mM Tris, 100 mM EDTA, 2% SDS, pH7) supplemented with proteinase K (80 μ g/mL) at 55°C for 1 h. Then, the RNA was degraded with the help of RNase (40 μ g/mL) for 30 min incubation at 37°C. The proteins were precipitated by the addition of 1 volume of 7.5 M ammonium acetate. A centrifugation step allowed to separate the precipitated proteins from the supernatant. Sodium acetate at 3 M was added to the supernatant in a ratio of 1:10. Then, the DNA was precipitated with the addition of 2.5 volumes of pure ethanol. A centrifugation step allowed gathering of the precipitated DNA

at the bottom of the tube. Finally, the DNA was washed twice with 70% ethanol before being resuspended in nuclease-free water. The concentrations of DNA were estimated with a Qubit® 2.0 fluorometer (Life Technologies, USA) according to the manufacturer's protocol.

3.4.4 Natural transformation assays

P. stutzeri tetR-kmR (recipient) and $\Delta comA$ (donor) strains were grown in LB medium at 30°C overnight (with 180 rpm agitation). After centrifugation for 1 min at 13'000 rpm, the pellet was washed twice with saline (0.9% w/v NaCl) and resuspended in saline at an OD₆₀₀ of 1 measured with spectrophotometer Genesys 10s UV-VIS (Thermo Fisher Scientific). In a tube, 80 µL of recipient bacteria suspension were mixed either with 20 µL H₂O (control treatment), with 20 µL of *rifR* DNA (final concentration, 20 µg/mL), or with 80 µL of donor bacteria suspension. In the latter case, the final bacterial suspension had a total volume of 100 µL thanks to the concentrating step of recipient's suspension in 20 µL before mixing. Then, 25 µL of the mix was spotted on LB agar plates in triplicates. After the spots had dried, inoculated agar plates were incubated at 30°C for 24 hours. Then, each bacterial spot was collected with an inoculation loop and resuspended in 1 mL saline in a tube. The tubes were vigorously vortexed for 15 s to resuspend the bacteria. Samples were diluted and plated on (1) LB-Rif plates for the determination of the number of transformants and on (2) LB plates for the total colony-forming units (CFU). In case of transformation with donors, plates contained also kanamycin to select only the recipient cells. Transformation frequencies are given as the ratio of rifampicin-resistant CFU/mL divided by the total number of CFU/mL.

3.4.5 Transcriptional reporters with various triggers

Overnight cultures of reporter strains in LB were washed and diluted at an OD₆₀₀ of 1 in M63 medium supplemented or not according to the treatments. The starting OD was high because the high-intensity UV treatment (30 mJ/cm²) leads to high cell death. The control treatment was the M63 medium. Concerning the other treatments, the final concentration of the supplement is the following: 50 µg/mL *P. stutzeri* DNA for *Ps* DNA, 50 µg/mL Fish DNA for Fish DNA, 2 µg/mL of nalidixic acid for Nal, 0.1 µg/mL of ciprofloxacin for Cip. Concerning the UV, 225 µL of a bacterial suspension at an OD₆₀₀ of 1 were irradiated with an intensity of 5 mJ/cm², respectively 30 mJ/cm², in a UV crosslinker (Amersham Life Science). The bacterial survival was evaluated with colonies plating on LB after the radiations. The optical density and fluorescence were measured over time at 30°C in a plate reader (Cytation5 from BioTek, USA;

OD: absorbance at 600 nm; fluorescence: excitation at 485 nm, emission at 520 nm, gain at 100).

3.4.6 Microscopy of transcriptional reporter

Overnight cultures of reporter strains were diluted at an OD₆₀₀ of 0.001 in LB medium and grown aerobically. At specific time points, 200 µL-samples were removed from the liquid culture to measure the optical density in a plate reader (Biotek Synergy HT; OD: absorbance at 600 nm). Ten microlitres of the samples were applied on an agarose pad to assess the fluorescence levels of individual cells with a microscope (Leica DM4 B; PH: 100 ms exposure, a gain of 2, intensity of 200%; fluorescence: 500 ms. exposure, a gain of 2, intensity of 100%). The pictures were analyzed using the *ImageJ* software (version 1.8.0) (Schindelin *et al.*, 2009). After adjusting the threshold of the pictures obtained with the phase-contrast channel, the outlines of cells were detected with ImageJ function “Analyze Particles.” The outlines were then applied to the pictures obtained with the fluorescence channel. The mean and median intensity of each cell was evaluated by the software and reported in a table. To assess whether the non-fluorescent cells were alive or dead, the cells were stained with propidium iodide (30 mM) for 15 min before being observed by microscopy on an agar pad.

3.4.7 Bacterial two-hybrid and β -galactosidase assay

To investigate the possible interactions between *tfoX*-like proteins, we tested them in a heterologous two-hybrid system. The heterologous two-hybrid system is based on the expression in *E. coli* of two proteins of interest that are cloned in-frame with either T18 or T25 split adenyl cyclase domains (Karimova *et al.*, 1998). Domain proximity, due to protein-protein interactions, restores the enzymatic activity and results in an increase in β -galactosidase production which is observed by the degradation of the colorless ortho-nitrophenyl- β -D-galactopyranoside (ONPG) substrate into galactose and the chromophore ortho-nitrophenol, yielding a bright yellow solution during β -galactosidase assay.

Overnight cultures of *E. coli* BTH101 carrying the pKT25 and pUT18C plasmid derivatives of interest were washed with and resuspended in a saline solution. The 5 µL of the cell suspension was spotted in triplicates on LB agar plates, supplemented with 1 mM IPTG, kanamycin, and ampicillin. After 24 h of growth at 30°C, spots were collected and resuspended in 1 mL of a NaCl 0.9% solution. OD₆₀₀ of 200 µL aliquot was measured in a plate-reader

SpectraMax i3x (Molecular Devices, USA). The remaining cell suspension was lysed by the addition of 40 μ L of chloroform, vortexed, and incubated for 15 min at 37°C. The 20 μ L of cell extract was mixed with 180 μ L of Z-buffer (Na_2HPO_4 , 0.06 M; NaH_2PO_4 , 0.04 M; KCl, 0.01 M and MgSO_4 , 0.001 M) supplemented with 3.5 ml/L of β -mercaptoethanol. β -galactosidase assay was initiated by the addition of 40 μ L of ortho-nitrophenyl- β -galactoside (ONPG, 4 mg/ml in Z-buffer) and stopped after an appropriate reaction time with the addition of 100 μ L of Na_2CO_3 , 1 M. OD₄₂₀ was measured in a plate reader SpectraMax i3x (Molecular Devices, USA) and reaction times were recorded for data treatment. Miller units were calculated as follows: $1000 * [\text{OD}_{420}] / ([\text{Time, min.}] * [\text{volume, mL}] * [\text{OD}_{600}])$ (Miller, 1972).

3.4.8 Biofilm assay

As *P. stutzeri* is not specially known to be a producer of biofilm and we did not observe biofilm production in our regular cultures in LB, we first tested the following media prepared according to their references: LB (tryptone, 10 g/L; yeast extract 5 g/L; NaCl, 5 g/L), Y medium (Shimada *et al.*, 2012), Minimal medium K (Shang *et al.*, 2021), Basal Medium (BM) (Ding *et al.*, 2019), BM without carbon (BM-C), Basal Medium 2 (BM2) (Mulcahy *et al.*, 2010). BM2 medium leads to the highest production of biofilm, so we did our experiments with it as described below.

Overnight cultures in LB were washed two times and then normalized to an OD₆₀₀ of 0.4 in 5 mL BM2 medium prepared as previously done (Mulcahy *et al.*, 2010). Then, the new cultures were distributed in flat-bottom transparent 96-well plates and incubated without shaking at 30°C for 5 days. To normalize the production of biofilm with the growth of bacterial culture, the absorbance was first measured at 600nm in a SpectraMax i3X plate-reader (Molecular Devices). Then, for the biofilm staining, liquid cultures were removed by pipetting and wells were washed 3x with 200 μ L of MilliQ water. The biofilm was then stained with 150 μ L crystal violet 1% (w/v) solubilized in a 30% (v/v) acetic acid solution for 15 min at 30°C. Then, crystal violet was removed, and wells were finally washed 3x with 200 μ L MilliQ water. The 150 μ L of 30% acetic acid solution was added to the wells to destain the Crystal violet and incubated at room temperature for 15 min with a 300rpm shaking (Heidolf Vibramax 100). Finally, 100 μ L were transferred to a new flat-bottom 96-well plate and absorbance was measured at 595 nm in a SpectraMax i3X plate-reader (Molecular Devices). Crystal Violet retention was measured at 595 nm and normalized by culture density measured at 600 nm.

3.4.9 Bioinformatics

Homologues of proteins were identified in *P. stutzeri* by blasting the protein sequence of interest on *P. stutzeri* ATCC17587 proteins with BLASTp (Altschul *et al.*, 1990).

For the identification of *P. stutzeri* proteins associated with T6SS, on one side we selected proteins whose annotation by homology suggested a link to T6SS. On the other side, we submitted the protein sequences to the tool SecReT6 (Zhang *et al.*, 2023). SecReT6 predicts two T6SS gene clusters.

For the phylogenetic tree to investigate the annotation of TfoX and TfoY of *V. cholerae*, we downloaded all protein sequences corresponding to the search request “TfoX *Vibrio cholerae*” from the protein database of NCBI and removed entries whose sequences were shorter than 175 amino acids, not identified as TfoX-like or not originating from *V. cholerae*. In addition, homologues of TfoX from *Haemophilus influenzae* (WP_005692773.1) and *Escherichia coli* (WP_000839159.1) were added. All sequences were aligned using MUSCLE on Ugene (Edgar, 2004; Okonechnikov *et al.*, 2012). A phylogenetic tree was constructed using IQ-tree with 1000 ultrafast bootstrap on Ugene (Okonechnikov *et al.*, 2012; Nguyen *et al.*, 2015). The tree was visualized and annotated with FigTree (Rambaut, 2009).

3.4.10 Graphics and data analysis

The statistical tests were performed by using the *RStudio* software (R version 4.0.4 (2021-02-15)) (Core R Team, 2019). Graphics were designed with *ggplot2* and base R functions. Regarding statistical analyses, the normality of the data distribution was tested with Shapiro-Wilk normality test, the equality of variances with F-test, or Levene’s test. Finally, means were compared either with the Student’s t-test/Wilcoxon rank test or ANOVA with post-hoc Tukey or Dunnett tests. Significance levels are indicated on the graphs as follows: *for p-value <0.05; **for p-value <0.01; *** for p-value <0.001; NS for non-significant results (p-value > 0.05). The multiple amino acids alignment was generated by the MUSCLE protein alignment algorithm in Unipro UGENE v45.1 (Okonechnikov *et al.*, 2012).

3.5 Results and Discussion

3.5.1 Does the presence of DNA or physiological stress trigger natural competence?

To elucidate potential competence triggers in *P. stutzeri*, the fluorescence and optical density of strains encoding transcriptional reporters *PcomA-sfGFP* and *PpilA-sfGFP* were monitored in the presence of DNA and genotoxic stress. The reporter plasmids were constructed by fusing the coding sequence of superfolder Green Fluorescent Protein (sfGFP) with the promoters of competence genes *comA* and *pilA* (Pédelacq *et al.*, 2006). The *pilA* gene encodes a prepilin, an essential structural component of the type IV pilus involved in DNA uptake (Graupner *et al.*, 2000). Conversely, *comA* plays a crucial role in the later stages of natural transformation, facilitating DNA translocation from the periplasm to the cytoplasm (Graupner and Wackernagel, 2001).

Firstly, we evaluated the expression of *comA* and *pilA* in a standard M63 medium (0.2% glucose). Both reporter strains with a promoter expressed higher levels of fluorescence than the promoter-less construct, confirming that the promoters were active (Figure 3.1). Interestingly, expression level of the construct with *comA* promoter was higher than with *pilA* promoter (Figure 3.1). In addition, *PcomA-sfGFP* showed an increasing fluorescence over time in the stationary phase in contrast to *PpilA-sfGFP*, which reached a fluorescence plateau at the end of the exponential phase. This suggests that the expression timing of *comA* and *pilA* might be slightly different. As the DNA must be first imported into the periplasm, the proteins building the pilus might be produced before proteins involved in further steps of natural transformation.

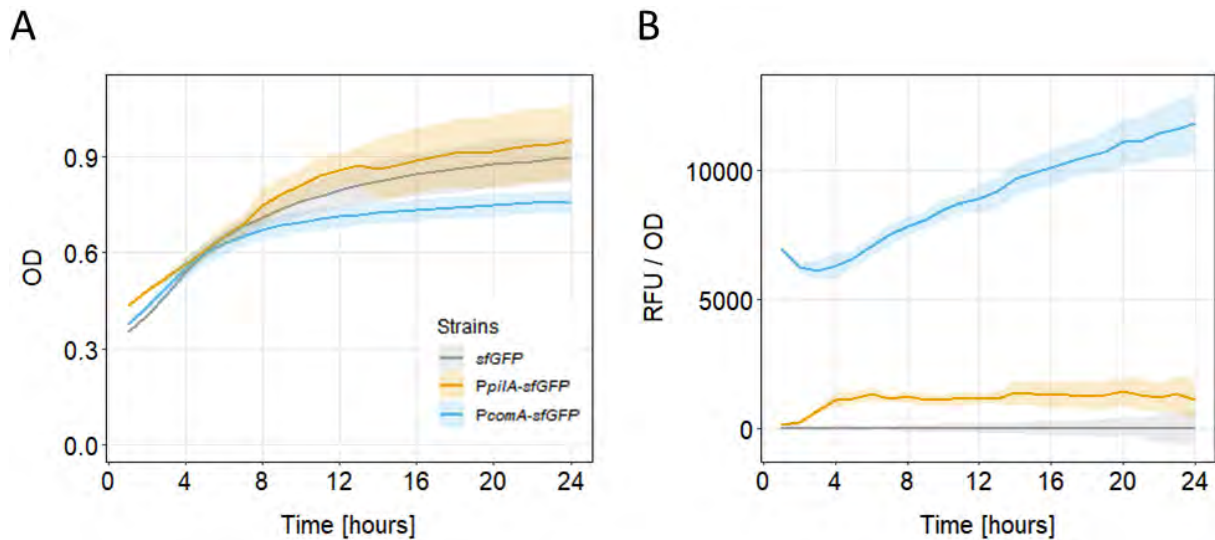


Figure 3.1 (A) Optical density (OD) and (B) relative fluorescence unit (RFU) corrected by OD of the reporter strains *PpilA-sfGFP* (orange), *PcomA-sfGFP* (blue) and the promoter-less strain *sfGFP* (grey) grown in M63 (glucose 0.2%). There are three replicates per strain. The mean values of three replicates are represented with the thick line and the standard deviation with the shade.

We next sought to determine whether *P. stutzeri* induces competence in response to extracellular DNA. Considering the energetic costs associated with entering a competent state (Haijema *et al.*, 2001), we hypothesized that cells might upregulate competence upon detecting extracellular DNA in their surroundings. However, contrary to our initial expectation, exposure to both homologous and heterologous DNA significantly decreased the expression of *pilA* and *comA*, with a more pronounced effect observed for homologous DNA (Figure 3.2). Although seemingly counterintuitive, this result may be attributed to the high concentrations of DNA used in our experiments compared to typical environmental conditions (DeFlaun *et al.*, 1986; Ogram *et al.*, 1987; Karl and Bailiff, 1989; Hurt *et al.*, 2001; Nielsen *et al.*, 2007). These elevated DNA concentrations might have acted as a distinct signal. It is known that high concentrations of extracellular DNA can chelate divalent cations, such as Mg^{2+} , potentially triggering various cellular responses. For instance, in *P. aeruginosa*, cation chelation can activate the cationic antimicrobial peptide resistance operon (Mulcahy *et al.*, 2008), while in *Salmonella enterica*, it induces the *pmr* operon (Johnson *et al.*, 2013). In extreme cases, high extracellular DNA concentrations can cause cell lysis by perturbing the membrane (Mulcahy *et al.*, 2008). Therefore, further experiments are needed to validate the observed decrease in gene expression and to determine whether this reduction is correlated with DNA concentration.

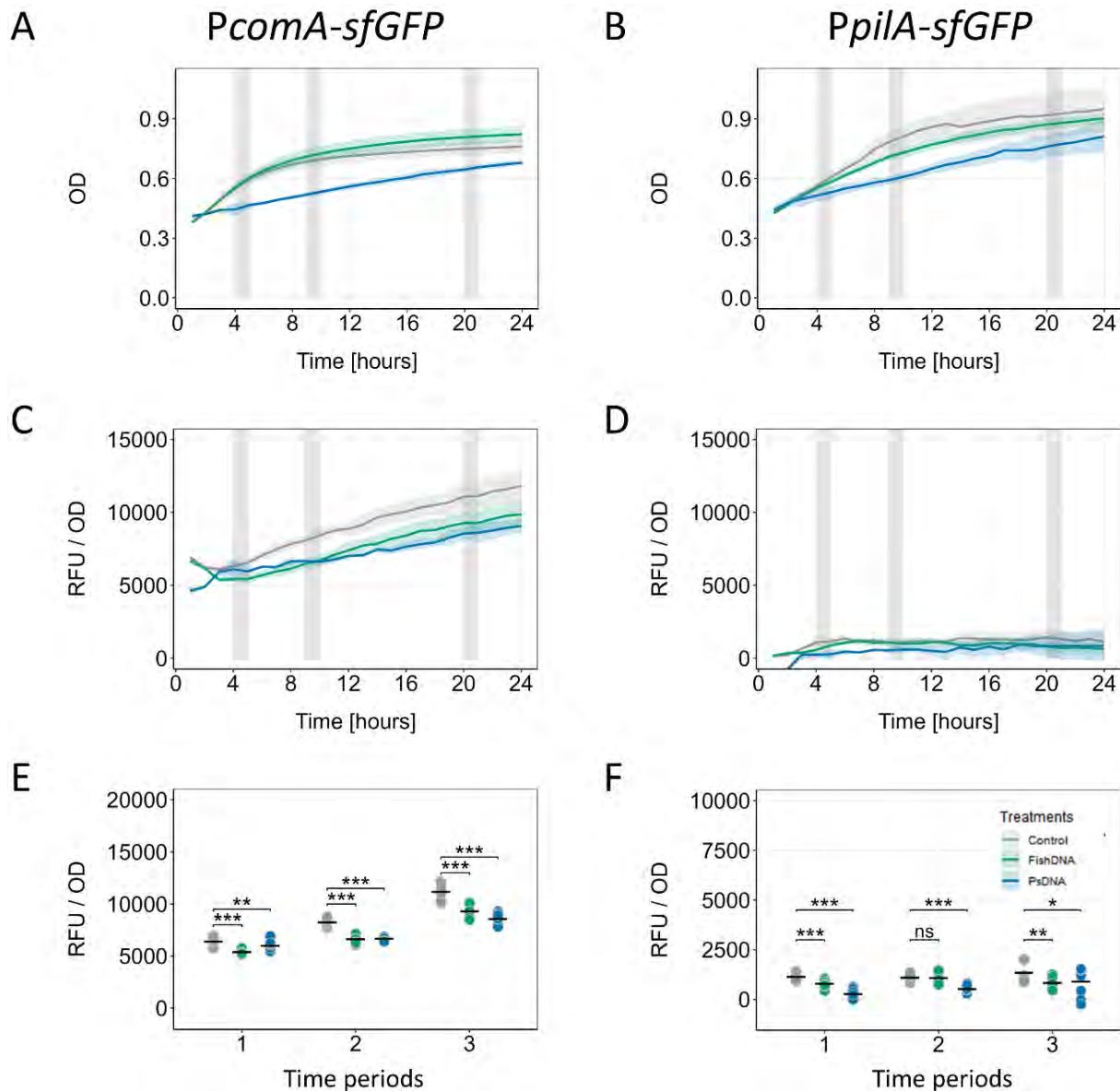


Figure 3.2. Extracellular DNA decreased competence in *P. stutzeri*. Optical density (A, B) and Relative Fluorescence Unit (RFU) corrected by OD (C, D) of the reporter strains *PcomA-sfGFP* (left) and *PpilA-sfGFP* (right) grown in M63 (glucose 0.2%) with no supplement (Control, grey), with fish DNA (green), with *P. stutzeri* DNA (blue). (A, B, C, D) The mean values of three replicates are represented with the thick line and the standard deviation with the shade. (E, F) Results of RFU corrected by OD at three time periods (TP) shaded on graphs upper plots. Data were measured every 15 min. TP 1 corresponds to 4-5 hours of incubation, TP 2 to 9-10 hours, and TP 3 to 20-21 hours. Individual measures are representative with dots, the mean values with the black bar.

One proposed function of natural transformation is facilitating recombinational repair of DNA damage. This hypothesis posits that the acquisition of exogenous DNA provides a template for repairing chromosomal lesions. To investigate this potential function, we examined the effects of genotoxic stress induced by ciprofloxacin and nalidixic acid. These quinolone antibiotics target bacterial gyrase and topoisomerases of type II and IV bacterial type II topoisomerases, leading to DNA strand breaks and activation of the SOS response and other DNA repair pathways (Aldred *et al.*, 2014). Ciprofloxacin was shown to increase natural transformation in

Helicobacter pylori (Dorer *et al.*, 2010), and nalidixic acid to induce competence gene *comEA* expression of *Legionella pneumophila* (Charpentier *et al.*, 2011). Contrary to our expectations, our results indicated that exposure to these antibiotics led to a reduction in competence gene expression in *P. stutzeri*, particularly for *pilA* and to a lesser extent for *comA* (Figure 3.3). Concentrations of antibiotic above the minimum inhibitory concentration tend to decrease transformability in *H. pylori* (Dorer *et al.*, 2010), but the concentrations we used did not affect growth considerably (Figure 3.3). These preliminary results suggest that this type of stress does not trigger competence in *P. stutzeri* but rather has the opposite effect.

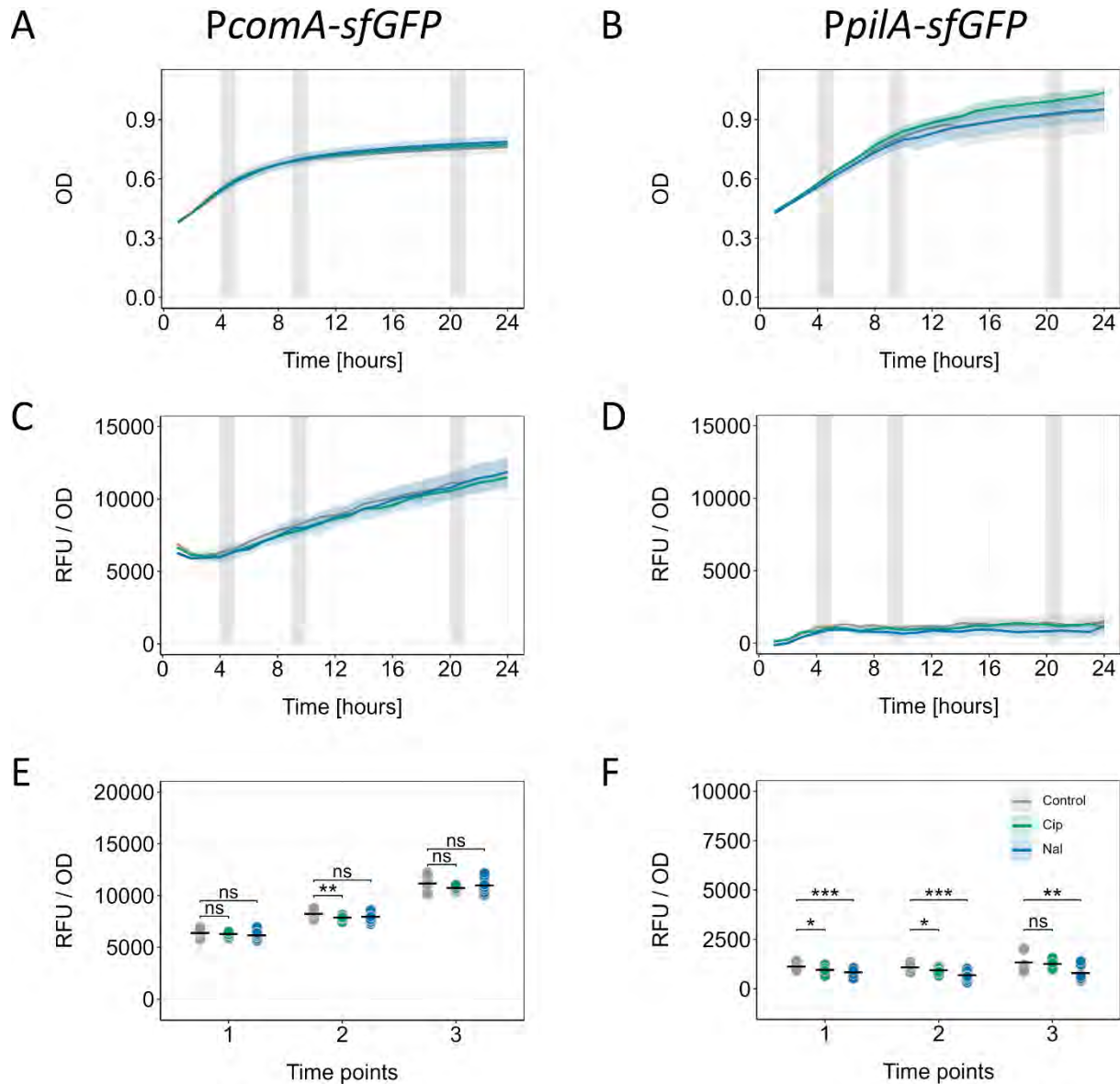


Figure 3.3. Our antibiotic treatments did not affect competence. Optical density (A, B) and Relative Fluorescence Unit (RFU) corrected by OD (C, D) of the reporter strains *PcomA-sfGFP* (left) and *PpilA-sfGFP* (right) grown in M63 (glucose 0.2%) with no supplement (Control, grey), ciprofloxacin (green, conc. = 0.1 $\mu\text{g}/\text{mL}$), with nalidixic acid (blue, conc. = 2 $\mu\text{g}/\text{mL}$). (A, B, C, D) The mean values of three replicates are represented with the thick line and the standard deviation with the shade. (E, F) Results of RFU corrected by OD at three time periods (TP) shaded on graphs upper plots. Data were measured every 15 min. TP 1 corresponds to 4-5 hours of incubation, TP 2 to 9-10 hours, and TP 3 to 20-21 hours. Individual measures are representative with dots, the mean values with the black bar.

High-intensity UV treatment (30 mJ/cm^2) significantly reduced the apparent expression in both reporter strains. However, this treatment resulted in substantial cell death, and the population failed to recover during the experimental period (Figure S 3.1). This high mortality rate complicated the assessment of fluorescence in surviving cells.

Conversely, low-intensity UV radiation (5 mJ/cm^2) led to increased expression in both reporter strains. Further investigation is necessary to determine whether this upregulation of

competence gene expression under mild UV stress correlates with enhanced transformability in *P. stutzeri*. These findings may indicate that *P. stutzeri* benefits from natural transformation, particularly for DNA repair mechanisms.

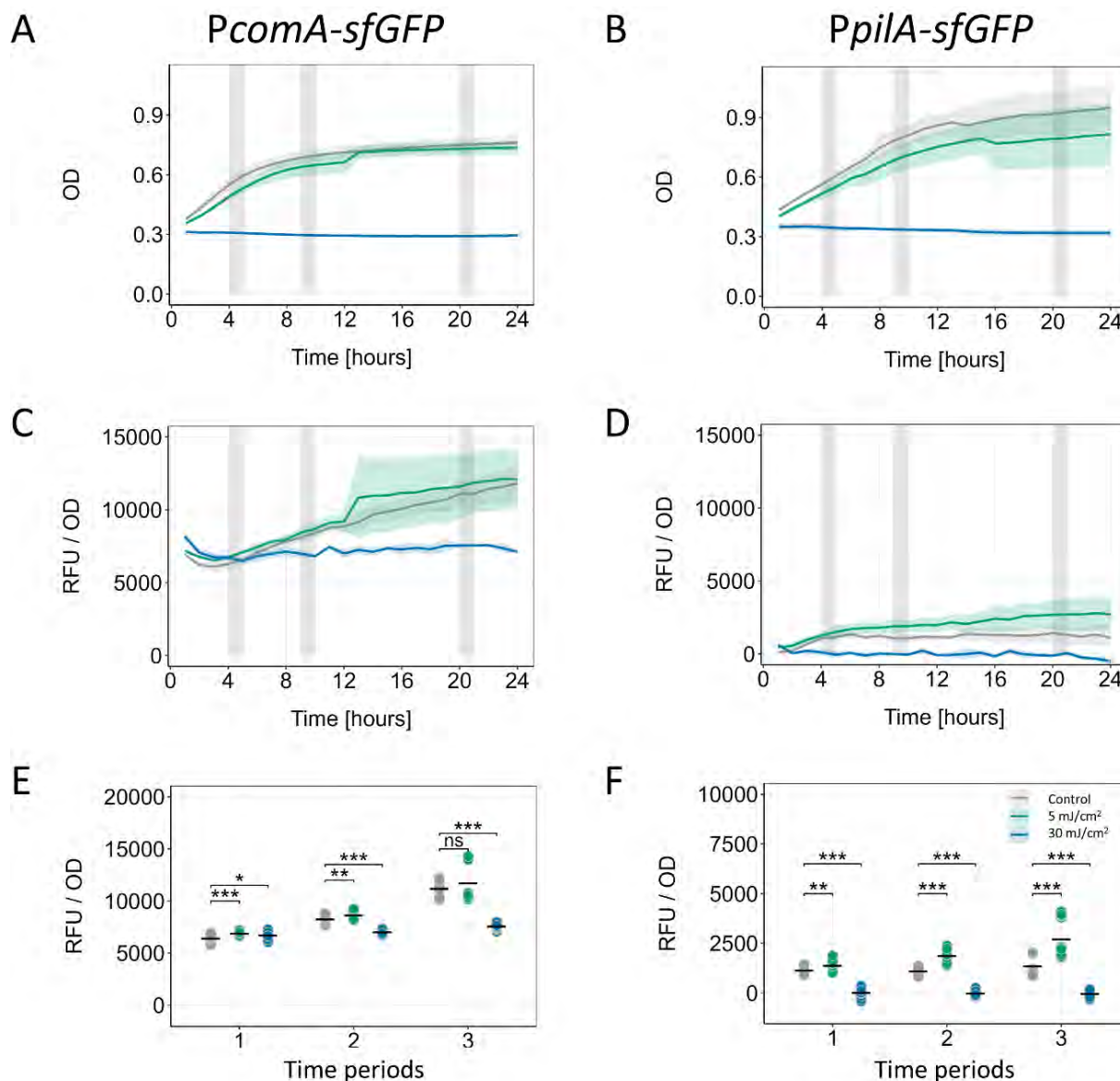


Figure 3.4. UV radiation impacted competence expression differently depending on the intensities of the radiations. Optical density (A, B) and Relative Fluorescence Unit (RFU) corrected by OD (C, D) of the reporter strains *PcomA-sfGFP* (left) and *PpilA-sfGFP* (right) grown in M63 (glucose 0.2%) after UV stresses (5 mJ/cm² in green, 30 mJ/cm² in blue). (A, B, C, D) The mean values of three replicates are represented with the thick line and the standard deviation with the shade. (E, F) Results of RFU corrected by OD at three time periods (TP) shaded on graphs upper plots. Data were measured every 15 min. TP 1 corresponds to 4-5 hours of incubation, TP 2 to 9-10 hours, and TP 3 to 20-21 hours. Individual measures are representative with dots, the mean values with the black bar.

In this preliminary study, we examined the effects of extracellular DNA and genotoxic stresses on the expression of *pilA* and *comA* in *P. stutzeri*. Our study focused on two types of genotoxic stress: antibiotics and UV radiation. Our results showed that high concentrations of

extracellular DNA led to reduced expression of *pilA* and *comA*. Conversely, low-intensity UV radiation increased the expression of both genes. Both antibiotic treatments resulted in decreased expression in our reporter strains.

While these findings are intriguing, several limitations should be addressed in future studies. Most experiments were conducted using a single concentration. Future studies should explore a range of concentrations to establish dose-dependent responses. Moreover, our study lacks direct evidence linking competence gene expression levels to transformability; however, previous research suggests a positive correlation between increased competence gene expression and enhanced transformability (Charpentier *et al.*, 2011). Therefore, future experiments are needed to test if the observed responses are dose-dependent and correlated with transformability. For antibiotic testing, we propose adopting an approach inspired by the antibiogram method. This technique creates a concentration gradient along the agar, allowing for direct evaluation of dose-dependent effects on competence gene expression. Finally, future experiments should include DNA transfer assays to establish the relationship between gene expression and transformability.

3.5.2 Global development of competence in *P. stutzeri* with a maximal expression at the onset of the stationary phase.

A strategy to reduce the costs associated with natural competence is to limit its development to only a subset of the population. This phenotypic heterogeneity is observed in *B. subtilis*, where 10 to 20 percent of individuals develop competence (Haijema *et al.*, 2001; Maamar and Dubnau, 2005; Maamar *et al.*, 2007). In the presence of a saturating concentration of extracellular DNA, *P. stutzeri* ATCC17587 has a transfer frequency of about 10^{-5} , meaning that 1 cell out of 100'000 receives the marker (see Chapter 2). This transformation level could be explained by only a small fraction of the population being highly competent or by an average low rate of competence, which are two distinct strategies with different evolutionary implications.

To determine when the competence genes were expressed and whether the expression of competence genes was homogenous within the cell population, we used the transcriptional reporter strains for *pilA* and *comA* competence genes and assessed their expression over time using microscopy and image analysis. In addition, the optical density of the cultures was measured simultaneously to monitor the growth.

The expression of *pilA* and *comA* varied during the growth phases (Figure 3.5), like the transformability of *P. stutzeri* (Lorenz and Wackernagel, 1990b, 1991). During the exponential phase, all reporter strains expressed similarly low levels of fluorescence, meaning that our genes of interest, *pilA*, and *comA*, are hardly transcribed. During the transition from the exponential to the stationary phase, most of the cells with the promotor construct expressed fluorescence: *PcomA-sfGFP* strain emitted a homogenous signal in contrast to the *PpilA-sfGFP* strain, whose fluorescence intensity was highly heterogeneous within the population. The *PpilA-sfGFP* population included non-fluorescent cells, which were nonetheless viable (Figure S 3.2). During the late stationary phase, the cell distribution shifted to lower fluorescence intensity for *PpilA-sfGFP* and *PcomA-sfGFP* reporters, indicating a decrease in competence gene expression. These results suggest that *P. stutzeri* is mainly competent during the transition from the exponential phase to the stationary phase, which is consistent with transformation assay (Lorenz and Wackernagel, 1990b, 1991). As the GFP protein is quite stable, part of the fluorescence observed in the stationary phase might be due to earlier expression (Andersen *et al.*, 1998).

To assess more precisely the timing of expression, we constructed mutants by adding an RPAANDENYAAAV tag to the C-terminal end of the sfGFP protein. This tag reduces the half-life of the GFPmut3, a mutant approximately 20 times more fluorescent than wild-type GFP, from more than a day to approximately 60 min in *E. coli* and to 190 min in *Pseudomonas putida* due to degradation by indigenous housekeeping proteases (Andersen *et al.*, 1998). Unfortunately, we did not detect any fluorescence in our mutants, even during the exponential phase, by flow cytometry (Figure S 3.3) and microscopy (data not shown). Other tags reducing less the half-life of sfGFP should be considered (Andersen *et al.*, 1998).

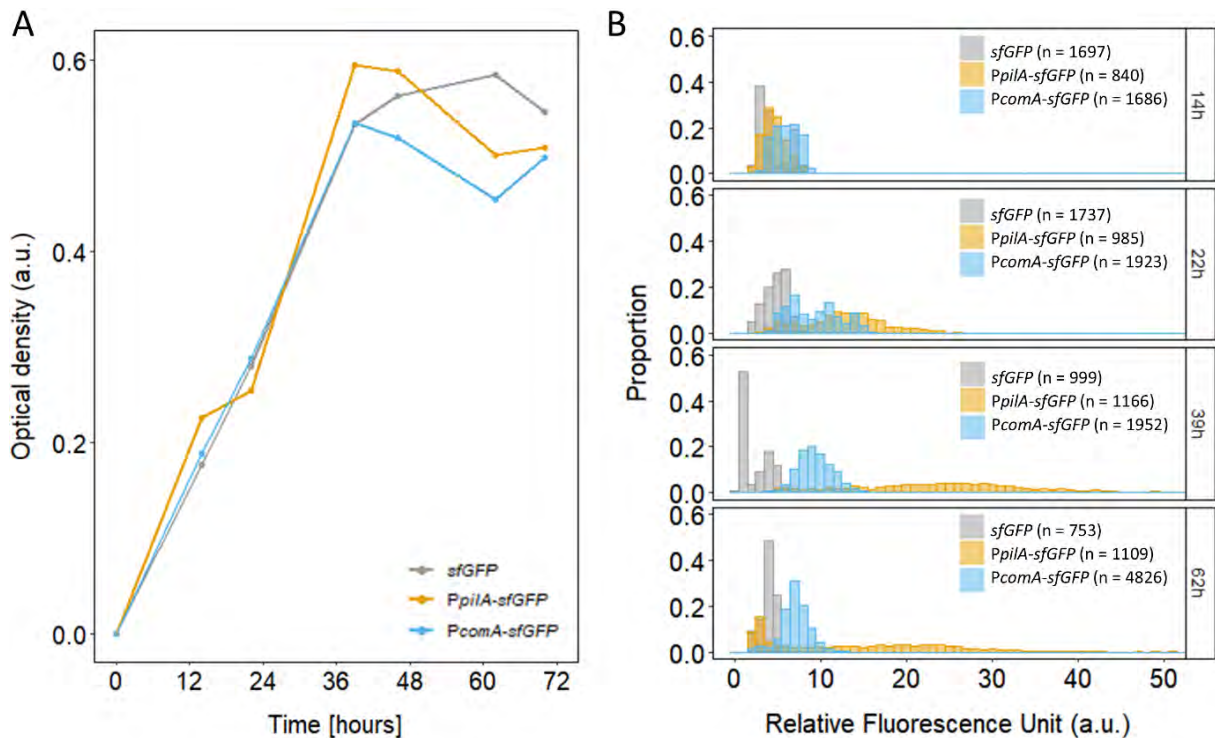


Figure 3.5. Reporter strains develop competence during exponential phase (A) Optical density over time of the transcriptional reporter strains *sfGFP*, *PpilA-sfGFP*, *PcomA-sfGFP*. Each dot represents a single measure by time point. (B) The proportion of mean fluorescence per cell of transcriptional reporter strains *sfGFP*, *PpilA-sfGFP*, *PcomA-sfGFP* during the exponential phase (14h, 22h), the transition from exponential to stationary phase (39h), and the late stationary phase (62h). The bars represent the proportion of cells expressing a certain level of fluorescence (mean). The number of cells whose fluorescence was measured is indicated next to the legend.

Concerning the proportion of cells expressing competence genes, most of the population developed competence (Figure 3.5). Notably, the expression levels differed among the reporters: *comA* was expressed at a low level across most of the cells, whereas *pilA* expression was higher but varied considerably in the population (Figure 3.5, Figure 3.6). Data obtained by flow cytometry also highlighted the same trend (Figure S 3.3). This contrasts with *V. cholerae*, for which similar distribution of signals for *pilA* and *comEA* were reported (Lo Scrudato and Blokesch, 2012). It would be interesting to investigate whether higher *pilA* expression correlates with greater transformation capacity using Fluorescence-Activated Cell Sorting (FACS) coupled with transformation experiments. It might also be relevant to explore whether the expression level of *pilA* influences the number of pili per cell and whether the distribution and number of pili can affect the overall competence of the bacterial cell, as more pili may facilitate greater interaction with extracellular DNA.

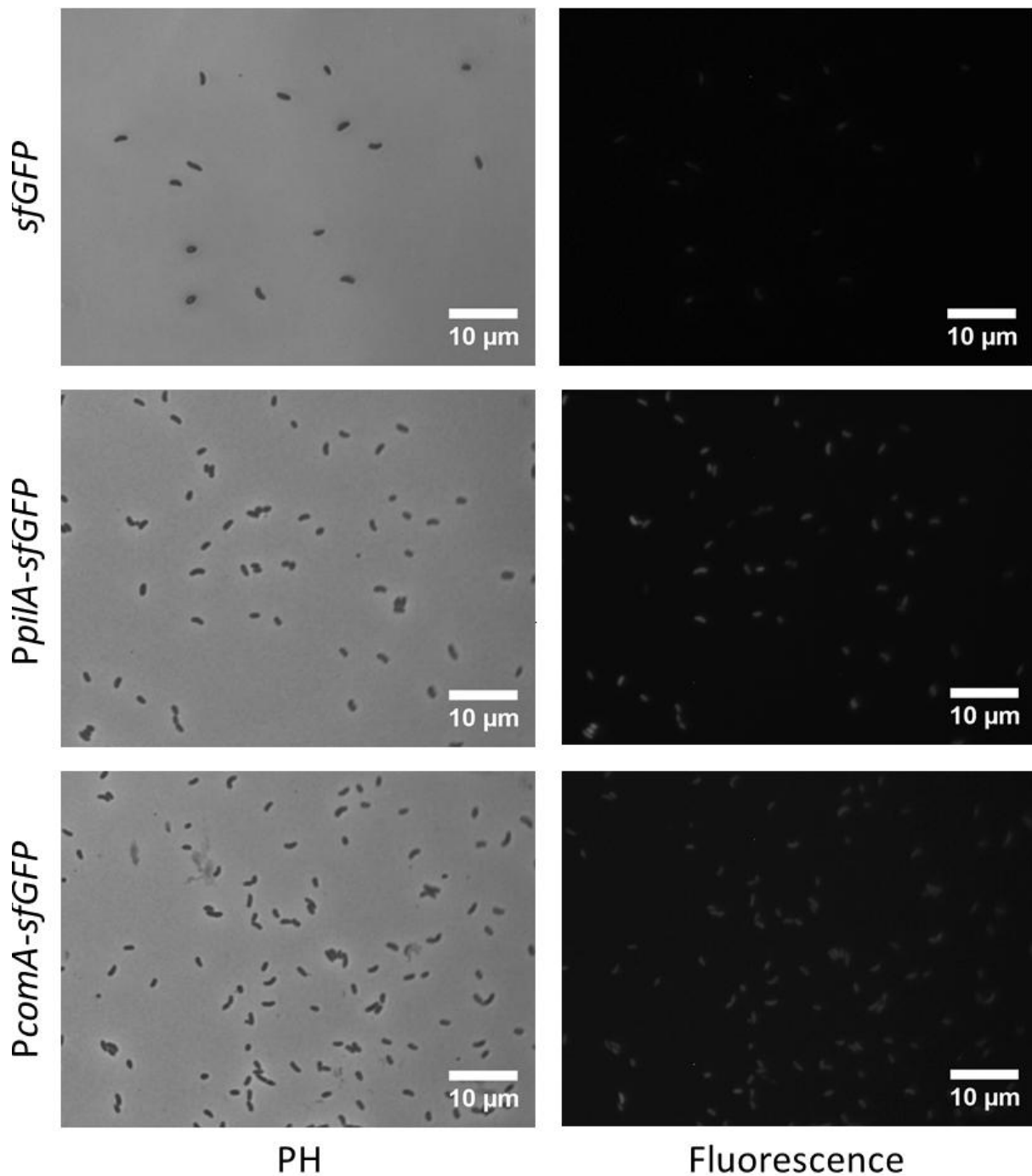


Figure 3.6. Representative micrographs showing cells of the reporter strains *sfGFP*, *PpilA-sfGFP*, and *PcomA-sfGFP* from a liquid culture in LB, after 39h incubation. PH, phase contrast channel; Fluorescence, fluorescence channel. For each time point and strain, the fluorescence intensity of at least 750 cells was assessed.

To summarize, in this part, we investigated at the single-cell level the expression of *pilA* and *comA*, two genes involved in the natural transformation of *P. stutzeri*. We showed that most of the cells expressed these competence genes with a maximum during the transition from exponential to stationary phase. This is consistent with the observations about *P. stutzeri* being highly transformable during this transition (Lorenz and Wackernagel, 1990b, 1991). Interestingly, the competence genes *pilA* and *comA* presented different patterns of expression with high variation in *pilA* expression among cells. FACS coupled with transformation

experiments would allow to investigate whether higher *pilA* expression correlates with greater transformation capacity.

3.5.3 TfoX, a regulator of competence in *P. stutzeri*?

TfoX, also known as Sxy, is a key positive regulator of natural transformation in *V. cholerae* and *H. influenzae* (Williams *et al.*, 1994; Meibom *et al.*, 2005). Homologues of TfoX have been identified in many Gammaproteobacteria, suggesting a possible common competence regulatory network (Cameron and Redfield, 2006).

Interestingly, many *Vibrionaceae* possess two TfoX paralogues: TfoX and TfoY (Pollack-Berti *et al.*, 2010; Metzger *et al.*, 2016; Pursley *et al.*, 2018). While TfoY has been correctly identified in several studies, it is frequently mislabeled as TfoX in databases (Figure S 3.4). Despite their similarity, TfoX and TfoY have diverged in their functions. For example, although both are essential for natural transformation in *V. fischeri* but without being functionally identical (Pollack-Berti *et al.*, 2010), only TfoX is required for natural transformation in *V. cholerae* (Metzger *et al.*, 2016). Their effect was also investigated regarding motility in *V. cholerae* as *tfoY* has a c-di-GMP riboswitch within the 5' UTR and c-di-GMP is often involved in the switch between biofilm and motile lifestyles (Metzger *et al.*, 2016). The *tfoY* mutant showed a slight but significant decrease in motility on soft-agar plates, while TfoY overproduction greatly increased motility. This strong motility effect was not seen with TfoX expression, emphasizing the distinct functions for TfoY and TfoX in *V. cholerae* (Metzger *et al.*, 2016).

A BLASTp search using TfoX from *V. cholerae* revealed two partial homologues of TfoX in *P. stutzeri*, located at genomic loci Ps32070 and Ps26900 (Figure 3.7). The homologues exhibited 27-28% identity to TfoX, with locus Ps32070 showing 37% query coverage in the C-terminal and locus Ps26900 showing 18% query coverage in the N-terminal. Hence, we constructed three markerless in-frame deletion mutants of TfoX: $\Delta tfoXC$ for the locus Ps32070, $\Delta tfoXN$ for the locus Ps26900, and a mutant without both genes ($\Delta tfoXC-N$).

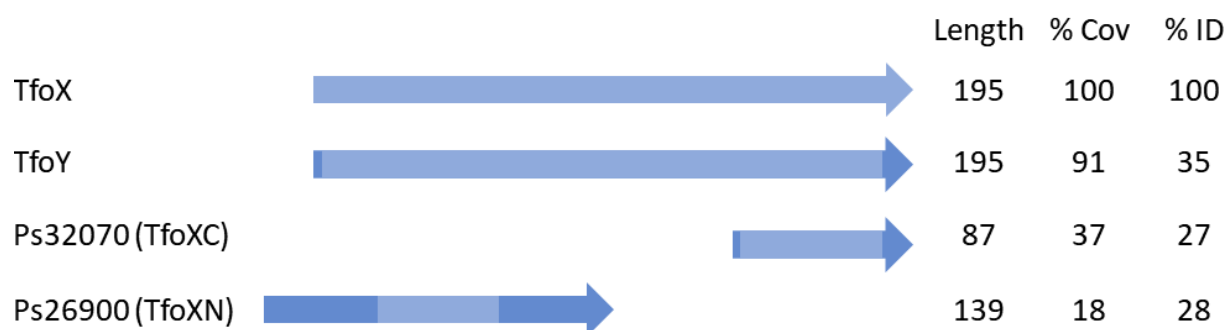


Figure 3.7 Schematic alignment of TfoX-like proteins from *V. cholerae* and *P. stutzeri*. The alignment was done on TfoX (AWB73649.1) from *V. cholerae*. TfoY (AWB74250.1) belongs to *V. cholerae*, Ps32070 and Ps26900 are from *P. stutzeri*. The light blue regions indicate the regions aligning to TfoX, and the dark blue regions show the unaligned sections based on BLASTp (Altschul *et al.*, 1990). The protein length is indicated as the number of amino acids. % Cov: percentage coverage; %ID: percentage identity.

Given that *P. stutzeri* has two partial homologues, potentially resulting from gene fission, we first investigated whether they interact with each other. Gene fissions are relatively common (Karamichali *et al.*, 2014), supported by the observation that these homologues align with different regions of TfoX. To investigate the possible interactions between TfoXC and TfoXN, we tested them in a heterologous two-hybrid system. We also included cAMP receptor protein (CRP) as it seems to interact with *tfoX/sxy* in *V. cholerae*, *H. influenzae*, and *E. coli* (Redfield *et al.*, 2005; Cameron and Redfield, 2006; Lo Scrudato *et al.*, 2014; Wu *et al.*, 2015). The heterologous two-hybrid system is based on the expression in *E. coli* of two proteins of interest that are cloned in-frame with either T18 or T25 split adenylyl cyclase domains (Karimova *et al.*, 1998). Domain proximity, due to protein-protein interactions, restores the enzymatic activity and results in an increase in β -galactosidase production which is observed by the degradation of the colorless ortho-nitrophenyl- β -D-galactopyranoside (ONPG) substrate into galactose and the chromophore ortho-nitrophenol, yielding a bright yellow solution during β -galactosidase assay. We observed that CRP interacts with itself as expected (Figure 3.8, Figure S 3.5), given its known tendency to form homodimers (Popovych *et al.*, 2009). However, CRP did not interact with any of the TfoX-like homologues. Protein TfoXN also showed some interaction with itself in contrast to TfoXC (Figure 3.8, Figure S 3.5). Interestingly, both TfoX-like homologues might interact with each other. However, one control (NP-N) exhibited an unexpectedly high level of ONPG degradation, thereby complicating the interpretation.

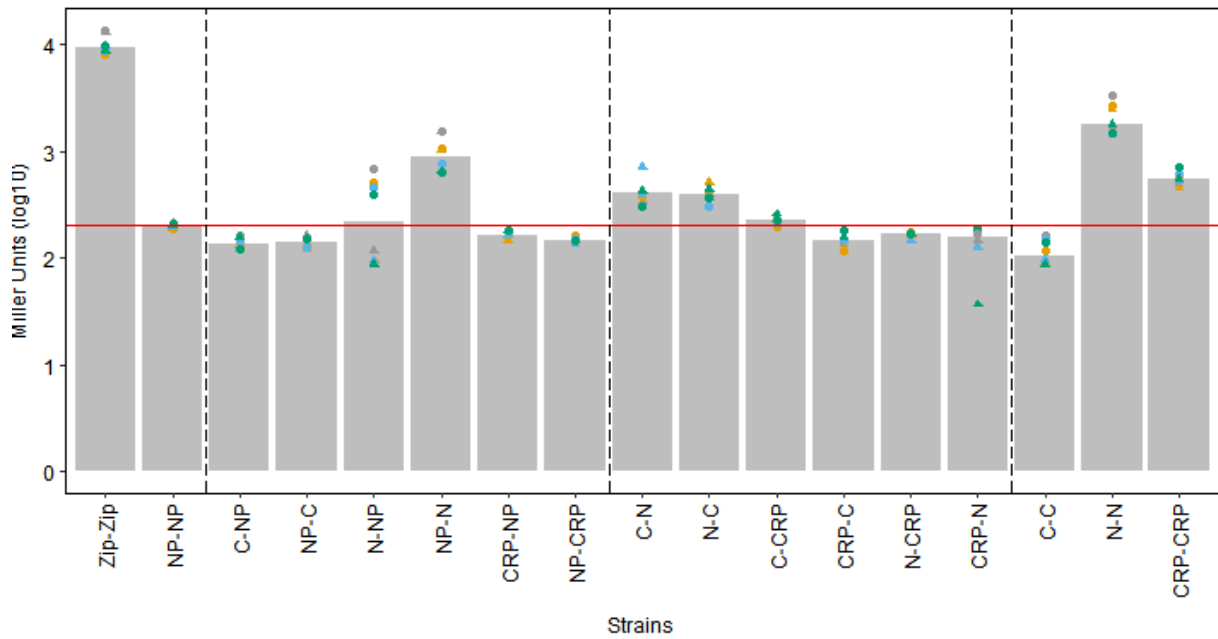


Figure 3.8 TfoX-like homologues of *P. stutzeri* might interact. β -galactosidase assay quantifying the interactions between TfoXC, TfoXN, and CRP in *E. coli* 24 h after spotting. The two technical replicates are shown with different shapes and the three biological replicates with different colors. The bars show the medians, and the horizontal red line indicates the median of the negative control (NP-NP). NP: no protein, negative control; Zip: leucine zipper, positive control; C: TfoXC; N: TfoXN, CRP: cAMP receptor protein.

We investigated next whether the ability of *P. stutzeri* to transform was dependent on TfoX-like homologues with natural transformation assays, given the involvement of the TfoX and TfoY in the natural transformation of *V. fischeri* and *V. cholerae* (Pollack-Berti *et al.*, 2010; Metzger *et al.*, 2016). Although the transformability of the different $\Delta tfoX$ mutants was slightly lower, it was not clearly different from the wild-type strain (Figure 3.9), indicating that these TfoX homologues are not essential for natural transformation in *P. stutzeri*.

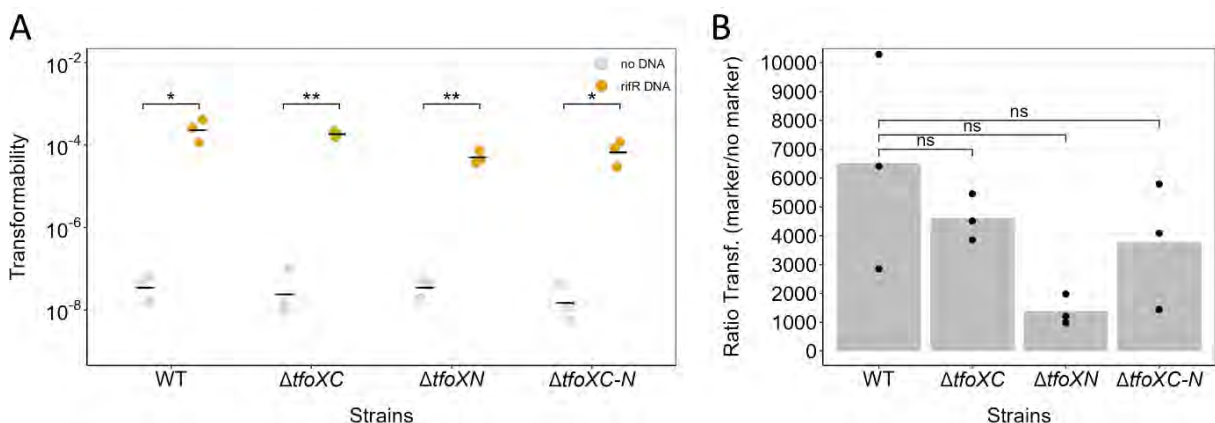


Figure 3.9 Deletion of *tfoX* homologues did not significantly affect the transformability of *P. stutzeri*. (A) Transformability of *P. stutzeri* ATCC 17587 wild-type (WT) and various $\Delta tfoX$ mutants in the presence of DNA from a rifampicin-resistant strain (rif^R DNA, orange) or with water (no DNA, grey) as control treatment. (B) The transformability in presence of the selective marker was normalized with the transformability with no marker. The dots represent the three replicates and the black line/bars the means. The experiments were performed at least two times with similar results.

The ability of *P. stutzeri* to perform natural transformation with donor cells and the role of TfoX and TfoY as regulators of T6SS in *V. cholerae* (Metzger *et al.*, 2016) led us to investigate whether the $\Delta tfoX$ -like mutants would show impaired DNA retrieval from donor cells. T6SS are nanomachines utilized by many Gram-negative bacteria as a weapon against eukaryotic hosts or prokaryotic competitors. The mechanism by which natural transformation occurs in *P. stutzeri* with donor cells is not entirely clear; it is not known whether donor cells are killed like during fratricide performed by *Streptococcus pneumoniae* (Johnsborg *et al.*, 2009) or release DNA such as *Neisseria gonorrhoeae* (Dillard and Seifert, 2001). Also, despite T6SS being common in *Pseudomonas*, not all *P. stutzeri* possess such system (Barret *et al.*, 2011; Navarro-Monserrat and Taylor, 2023). Our genome annotation of *P. stutzeri* ATCC17587 revealed multiple T6SS-linked homologues (Table S 3.2), suggesting the presence of these systems. This was further supported by the T6SS prediction tool SecReT6, which indicated two T6SS in our strain (Zhang *et al.*, 2023). The removal of TfoX-like homologues in *P. stutzeri* did not affect the transformability of the strains with donor cells (Figure 3.10). Similarly, deletion of *clpV* homologue, a T6SS component involved in sheath disassembly in *P. aeruginosa* (LeRoux *et al.*, 2012), had no impact on transformation (Figure S 3.6). Taken together, these results indicate that in *P. stutzeri*, neither TfoX-like homologues nor T6SS play a significant role in natural transformation with donor cells.

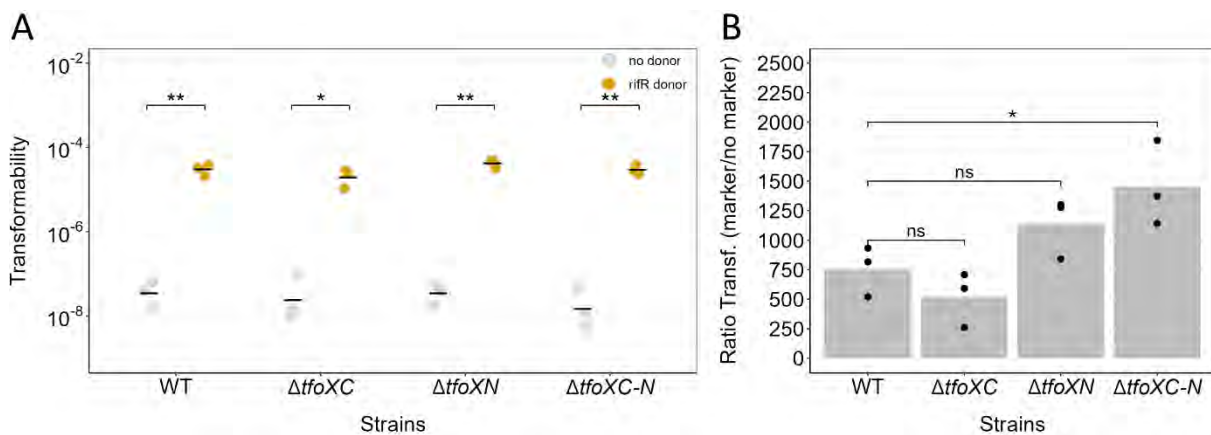


Figure 3.10 Deletion of *tfoX* homologues did not significantly affect the transformability of *P. stutzeri* with donor cells. (A) Transformability of *P. stutzeri* ATCC 17587 wild-type (WT) and various $\Delta tfoX$ mutants in the presence of rifampicin-resistant $\Delta comA$ donor cells (rifR donor, orange) or with water (no donor, grey) as control treatment. To separate the donor from recipient cells, recipients were resistant to kanamycin in contrast to donor cells. To avoid donor cells becoming resistant to kanamycin through natural transformation, a non-transformable $\Delta comA$ mutant was used (Chapter 2). (B) The transformability in presence of the selective marker was normalized with the transformability with wild-type marker. The dots represent the three replicates and the black line/bars the means.

The multiple roles of TfoY in *V. cholerae* prompted us to investigate whether TfoX-like mutants affect biofilm production in *P. stutzeri* (Metzger *et al.*, 2016). Although *P. stutzeri* is not primarily known for biofilm formation, it does produce biofilms in certain media (Figure S 3.7). We compared biofilm accumulation in dense suspensions of TfoX-like mutants and the wild-type strain to assess the involvement of TfoX-like genes in biofilm production. Both $\Delta tfoX$ mutants exhibited significantly increased biofilm production compared to the wild-type strain, indicating that both homologues repress biofilm formation (Figure 3.11).

Interestingly, natural transformation in *P. aeruginosa* has only been observed so far under conditions that promote biofilm production (Nolan *et al.*, 2020). Similarly, biofilm-growth conditions enhance natural transformation efficiency in *Staphylococcus aureus* (Maree *et al.*, 2022). These findings suggest a potential link between biofilm and natural transformation competence. Therefore, evaluating the transformability of *P. stutzeri* under biofilm-promoting conditions could provide valuable insights into this relationship.

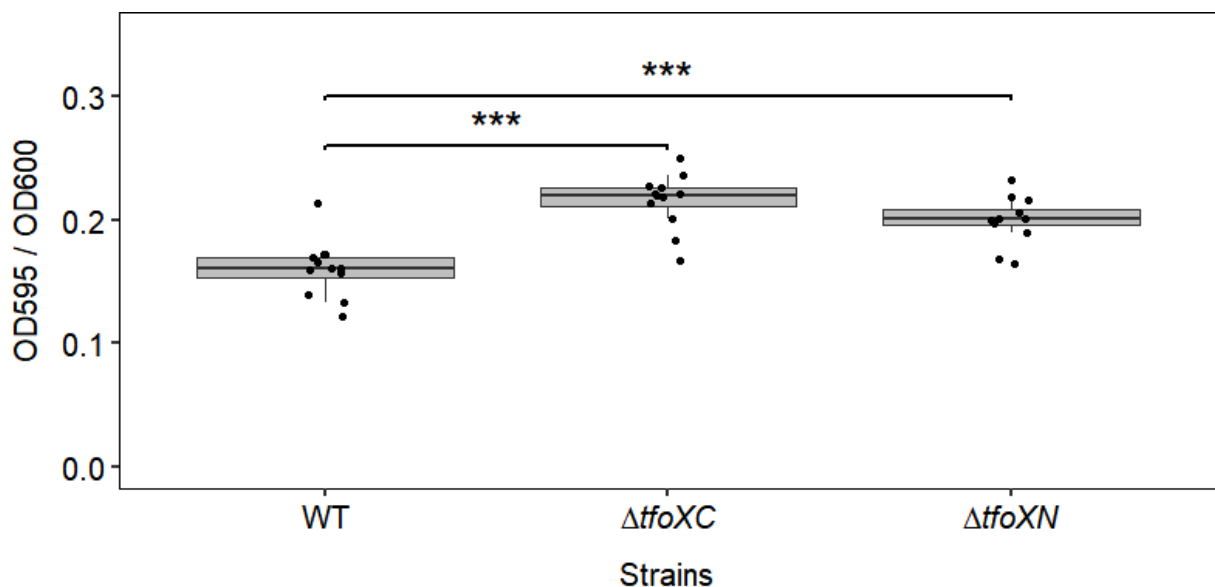


Figure 3.11 TfoX-like homologues deletion increased the production of biofilm. For biofilm quantification, crystal violet retention in a 5-day-old biofilm was measured at 595 nm and normalized by culture density measured at 600 nm in twelve replicates per strain. The dots represent the replicates and the bars the means.

Overall, we identified two short TfoX-like homologues in *P. stutzeri* and investigated their roles in cellular processes linked to natural transformation. Our findings reveal that neither of the TfoX-like homologues plays a significant role in natural transformation, whether with naked DNA or donor cells. Furthermore, the T6SS is unlikely to be involved in natural transformation with donor cells. Intriguingly, both TfoX-like homologues act as repressors of biofilm

production in *P. stutzeri*. These results suggest that while the TfoX-like homologues do not contribute to natural transformation in *P. stutzeri*, they play a role in regulating biofilm formation. This unexpected function highlights the diverse and species-specific roles that TfoX-like proteins can assume in different bacterial species, underscoring the importance of studying homologous proteins across various organisms to fully understand their functional repertoire.

3.6 Conclusion

This work aimed to gain a deeper understanding of competence regulation in *P. stutzeri*. Therefore, we explored three aspects of regulation: the triggers of competence, the expression of competence at the single-cell level, and the role of *tfoX* in the transformability of *P. stutzeri*.

Our results at the population level showed interesting effects, especially with UV-stress for which low doses triggered the expression of competence genes. Regarding the gene expression at the single-cell level, as expected, the general expression was higher at the transition from the exponential phase to the stationary phase. Interestingly, most of the cells were expressing competence genes but the homogenous expression of *comA* contrasts with the expression of *pilA* whose level of expression varies considerably from cell to cell. To investigate if the levels of competence gene expression relate to the transformability in individual cells, it would be interesting to separate the cells into subgroups with the help of FACS before testing their ability to transform.

Finally, we investigated the potential role of *P. stutzeri* homologues of TfoX—a key regulator of natural transformation in *V. cholerae* and *H. influenzae* (Williams *et al.*, 1994; Meibom *et al.*, 2005)—in the natural transformation process. Our results indicated that the deletion of these homologues did not significantly affect the transformability of *P. stutzeri*. However, we observed an increase in biofilm production following the deletion. Given these findings, other potential regulators of natural transformation in *P. stutzeri* might be considered. One candidate is Vfr, a transcription regulator of *P. aeruginosa* that induce Type IV pilus genes, as well as numerous genes involved in motility and adherence (Wolfgang *et al.*, 2003). Additionally, the regulation of certain virulence genes by *vfr* in *P. aeruginosa* is known to require cAMP (Fuchs *et al.*, 2010). This suggests that Vfr, along with cAMP, could potentially

play a role in the regulation of natural transformation in *P. stutzeri*. Taken together, these findings underscore the variations in natural transformation regulation compared to well-known Gram-negative models. They also highlight the importance of studying additional models like *P. stutzeri* to achieve a more comprehensive understanding of competence regulation on a global scale.

3.7 References

- Aldred, K.J., Kerns, R.J., and Osheroﬀ, N. (2014) Mechanism of quinolone action and resistance. *Biochemistry* **53**: 1565–1574.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.
- Ambur, O.H., Engelstädter, J., Johnsen, P.J., Miller, E.L., and Rozen, D.E. (2016) Steady at the wheel: Conservative sex and the benefits of bacterial transformation. *Philos Trans R Soc B Biol Sci* **371**: 1–11.
- Andersen, J.B., Sternberg, C., Poulsen, L.K., Bjørn, S.P., Givskov, M., and Molin, S. (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl Environ Microbiol* **64**: 2240–2246.
- Bao, Y., Lies, D.P., Fu, H., and Roberts, G.P. (1991) An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of Gram-negative bacteria. *Gene* **109**: 167–168.
- Barret, M., Egan, F., Fargier, E., Morrissey, J.P., and O’Gara, F. (2011) Genomic analysis of the type VI secretion systems in *Pseudomonas* spp.: Novel clusters and putative effectors uncovered. *Microbiology* **157**: 1726–1739.
- Borgeaud, S., Metzger, L.C., Scignari, T., and Blokesch, M. (2015) The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. *Science (80-)* **347**: 63–67.
- Boutry, C., Delplace, B., Clippe, A., Fontaine, L., and Hols, P. (2013) SOS Response Activation and Competence Development Are Antagonistic Mechanisms in *Streptococcus thermophilus*. *J Bacteriol* **195**: 696–707.
- Cameron, A.D.S. and Redfield, R.J. (2006) Non-canonical CRP sites control competence regulons in *Escherichia coli* and many other γ -proteobacteria. *Nucleic Acids Res* **34**: 6001–6014.
- Charpentier, X., Kay, E., Schneider, D., and Shuman, H.A. (2011) Antibiotics and UV radiation induce competence for natural transformation in *Legionella pneumophila*. *J Bacteriol* **193**: 1114–1121.
- Claverys, J.P., Prudhomme, M., and Martin, B. (2006) Induction of competence regulons as a general response to stress in gram-positive bacteria. *Annu Rev Microbiol* **60**: 451–475.
- Core R Team (2019) A Language and Environment for Statistical Computing. *R Found Stat Comput* **2**: <https://www.R-project.org>.
- DeFlaun, M.F., Paul, J.H., and Davis, D. (1986) Simplified method for dissolved DNA determination in aquatic environments. *Appl Environ Microbiol* **52**: 654–659.

- Dillard, J.P. and Seifert, H.S. (2001) A variable genetic island specific for *Neisseria gonorrhoeae* is involved in providing DNA for natural transformation and is found more often in disseminated infection isolates. *Mol Microbiol* **41**: 263–277.
- Ding, X.S., Zhao, B., An, Q., Tian, M., and Guo, J.S. (2019) Role of extracellular polymeric substances in biofilm formation by *Pseudomonas stutzeri* strain XL-2. *Appl Microbiol Biotechnol* **103**: 9169–9180.
- Dorer, M.S., Fero, J., and Salama, N.R. (2010) DNA damage triggers genetic exchange in *Helicobacter pylori*. *PLoS Pathog* **6**: 1–10.
- Dubnau, D. and Blokesch, M. (2019) Mechanisms of DNA Uptake by Naturally Competent Bacteria. *Annu Rev Genet* **53**: 217–237.
- Edgar, R.C. (2004) MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 1–19.
- 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.
- Graupner, S., Frey, V., Hashemi, R., Lorenz, M.G., Brandes, G., and Wackernagel, W. (2000) Type IV pilus genes pilA and pilC of *Pseudomonas stutzeri* are required for natural genetic transformation, and pilA can be replaced by corresponding genes from nontransformable species. *J Bacteriol* **182**: 2184–2190.
- Graupner, S. and Wackernagel, W. (2001) Identification and characterization of novel competence genes comA and exbB involved in natural genetic transformation of *Pseudomonas stutzeri*. *Res Microbiol* **152**: 451–460.
- Graupner, S., Weger, N., Sohni, M., and Wackernagel, W. (2001) Requirement of novel competence genes pilT and pilU of *Pseudomonas stutzeri* for natural transformation and suppression of pilT deficiency by a hexahistidine tag on the type IV pilus protein PilA. *J Bacteriol* **183**: 4694–4701.
- Grimbergen, A.J., Siebring, J., Solopova, A., and Kuipers, O.P. (2015) Microbial bet-hedging: The power of being different. *Curr Opin Microbiol* **25**: 67–72.
- Haijema, B.J., Hahn, J., Haynes, J., and Dubnau, D. (2001) A ComGA-dependent checkpoint limits growth during the escape from competence. *Mol Microbiol* **40**: 52–64.
- Humair, B., González, N., Mossialos, D., Reimann, C., and Haas, D. (2009) Temperature-responsive sensing regulates biocontrol factor expression in *Pseudomonas fluorescens* CHA0. *ISME J* **3**: 955–965.
- Hurt, R.A., Qiu, X., Wu, L., Roh, Y., Palumbo, A. V., Tiedje, J.M., and Zhou, J. (2001) Simultaneous Recovery of RNA and DNA from Soils and Sediments. *Appl Environ Microbiol* **67**: 4495–4503.

- Johnsborg, O., Eldholm, V., Håvarstein, L.S., and Sigve Håvarstein, L. (2007) Natural genetic transformation: prevalence, mechanisms and function. *Res Microbiol* **158**: 767–778.
- Johnsborg, O., Håvarstein, L.S., Sigve, L., and Avarstein, H.° (2009) Regulation of natural genetic transformation and acquisition of transforming DNA in *Streptococcus pneumoniae*. *FEMS Microbiol Rev* **33**: 627–642.
- Johnson, L., Horsman, S.R., Charron-Mazenod, L., Turnbull, A.L., Mulcahy, H., Surette, M.G., et al. (2013) Extracellular DNA-induced antimicrobial peptide resistance in *Salmonella enterica* serovar Typhimurium. *BMC Microbiol* **13**: 1–8.
- Johnston, C., Martin, B., Fichant, G., Polard, P., and Claverys, J.P. (2014) Bacterial transformation: Distribution, shared mechanisms and divergent control. *Nat Rev Microbiol* **12**: 181–196.
- Karamichali, I., Koumandou, V.L., Karagouni, A.D., and Kossida, S. (2014) Frequent gene fissions associated with human pathogenic bacteria. *Genomics* **103**: 65–75.
- Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci U S A* **95**: 5752–5756.
- Karl, D.M. and Bailiff, M.D. (1989) The measurement and distribution of dissolved nucleic acids in aquatic environments. *Limnol Oceanogr* **34**: 543–558.
- LeRoux, M., De Leon, J.A., Kuwada, N.J., Russell, A.B., Pinto-Santini, D., Hood, R.D., et al. (2012) Quantitative single-cell characterization of bacterial interactions reveals type VI secretion is a double-edged sword. *Proc Natl Acad Sci U S A* **109**: 19804–19809.
- Lorenz, M.G. and Wackernagel, W. (1991) High frequency of natural genetic transformation of *Pseudomonas stutzeri* in soil extract supplemented with a carbon/energy and phosphorus source. *Appl Environ Microbiol* **57**: 1246–1251.
- Lorenz, M.G. and Wackernagel, W. (1990a) Natural genetic transformation of *Pseudomonas stutzeri* by sand-adsorbed DNA. *Arch Microbiol* **154**: 380–385.
- Lorenz, M.G. and Wackernagel, W. (1990b) Natural genetic transformation of *Pseudomonas stutzeri* by sand-adsorbed DNA. *Arch Microbiol* 380–385.
- Lorenz, M.G. and Wackernagel, W. (1992) Stimulation of natural genetic transformation of *Pseudomonas stutzeri* in extracts of various soils by nitrogen or phosphorus limitation and influence of temperature and pH. *Microb Releases* **1**: 173–176.
- Maamar, H. and Dubnau, D. (2005) Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop. *Mol Microbiol* **56**: 615–624.
- Maamar, H., Raj, A., and Dubnau, D. (2007) Noise in gene expression determines cell fate in *Bacillus subtilis*. *Science (80-)* **317**: 526–529.

- Maree, M., Thi Nguyen, L.T., Ohniwa, R.L., Higashide, M., Msadek, T., and Morikawa, K. (2022) Natural transformation allows transfer of SCCmec-mediated methicillin resistance in *Staphylococcus aureus* biofilms. *Nat Commun* **13**: 1–14.
- Martínez-García, E. and de Lorenzo, V. (2011) Engineering multiple genomic deletions in Gram-negative bacteria: Analysis of the multi-resistant antibiotic profile of *Pseudomonas putida* KT2440. *Environ Microbiol* **13**: 2702–2716.
- Matthey, N., Stutzmann, S., Stoudmann, C., Guex, N., Iseli, C., and Blokesch, M. (2019) Neighbor predation linked to natural competence fosters the transfer of large genomic regions in *Vibrio cholerae*. *Elife* **8**: 1–21.
- Mavridou, D.A.I., Gonzalez, D., Kim, W., West, S.A., and Foster, K.R. (2018) Bacteria Use Collective Behavior to Generate Diverse Combat Strategies. *Curr Biol* **28**: 345–355.
- Meibom, K.L., Blokesch, M., Dolganov, N.A., Wu, C.-Y.Y., and Schoolnik, G.K. (2005) Chitin Induces Natural Competence in *Vibrio cholerae*. *Science (80-)* **310**: 1824–1827.
- Metzger, L.C., Stutzmann, S., Scignari, T., Van der Henst, C., Matthey, N., and Blokesch, M. (2016) Independent Regulation of Type VI Secretion in *Vibrio cholerae* by TfoX and TfoY. *Cell Rep* **15**: 951–958.
- Miller, J.H. (1972) Experiments in molecular genetics, Cold Spring.
- Mulcahy, H., Charron-Mazenod, L., and Lewenza, S. (2008) Extracellular DNA Chelates Cations and Induces Antibiotic Resistance in *Pseudomonas aeruginosa* Biofilms. *PLoS Pathog* **4**: 1000213.
- Mulcahy, H., Charron-Mazenod, L., and Lewenza, S. (2010) *Pseudomonas aeruginosa* produces an extracellular deoxyribonuclease that is required for utilization of DNA as a nutrient source. *Environ Microbiol* **12**: 1621–1629.
- Navarro-Monserrat, E.D. and Taylor, C.G. (2023) T6SS: A Key to *Pseudomonas*’s Success in Biocontrol? *Microorganisms* **11**: 1–17.
- Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., and Minh, B.Q. (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**: 268–274.
- Nielsen, K.M., Johnsen, P.J., Bensasson, D., and Daffonchio, D. (2007) Release and persistence of extracellular DNA in the environment. *Environ Biosafety Res* **6**: 37–53.
- Nolan, L.M., Turnbull, L., Katrib, M., Osvath, S.R., Losa, D., Lazenby, J.J., and Whitchurch, C.B. (2020) *Pseudomonas aeruginosa* is capable of natural transformation in biofilms. *Microbiol (United Kingdom)* **166**: 995–1003.
- Ogram, A., Sayler, G.S., and Barkay, T. (1987) The extraction and purification of microbial DNA from sediments. *J Microbiol Methods* **7**: 57–66.

- Okonechnikov, K., Golosova, O., Fursov, M., Varlamov, A., Vaskin, Y., Efremov, I., et al. (2012) Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **28**: 1166–1167.
- Overballe-Petersen, S., Harms, K., Orlando, L.A.A., Moreno Mayar, J.V., Rasmussen, S., Dahl, T.W., et al. (2013) Bacterial natural transformation by highly fragmented and damaged DNA. *Proc Natl Acad Sci U S A* **110**: 19860–19865.
- Pédelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C., and Waldo, G.S. (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* **24**: 79–88.
- Pollack-Berti, A., Wollenberg, M.S., and Ruby, E.G. (2010) Natural transformation of *Vibrio fischeri* requires *tfoX* and *tfoY*. *Environ Microbiol* **12**: 2302–2311.
- Popovych, N., Tzeng, S.R., Tonelli, M., Ebright, R.H., and Kalodimos, C.G. (2009) Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. *Proc Natl Acad Sci U S A* **106**: 6927–6932.
- Prudhomme, M., Attaiech, L., Sanchez, G., Martin, B., and Claverys, J.-P. (2006) Antibiotic Stress Induces Genetic Transformability in the Human Pathogen *Streptococcus pneumoniae*. *Science (80-)* **313**: 89–92.
- Pursley, B.R., Maiden, M.M., Hsieh, M.L., Fernandez, N.L., Severin, G.B., and Waters, C.M. (2018) Cyclic di-GMP regulates TfoY in *Vibrio cholerae* to control motility by both transcriptional and posttranscriptional mechanisms. *J Bacteriol* **200**: 1–19.
- Rambaut, A. (2009) FigTree. *Comput Progr Distrib by author, website* <http://tree.bio.ed.ac.uk/software/figtree/>.
- Redfield, R.J. (1988) Evolution of Bacterial Transformation: Is Sex With Dead Cells Ever Better Than No Sex at All ? *Genetics* **119**: 213–221.
- Redfield, R.J., Cameron, A.D.S., Qian, Q., Hinds, J., Ali, T.R., Kroll, J.S., and Langford, P.R. (2005) A novel CRP-dependent regulon controls expression of competence genes in *Haemophilus influenzae*. *J Mol Biol* **347**: 735–747.
- Sambrook, J. and Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual*. Third Edition.
- Schindelin, J., Arganda-Carrera, I., Frise, E., Verena, K., Mark, L., Tobias, P., et al. (2009) Fiji - an Open platform for biological image analysis. *Nat Methods* **9**: 1–15.
- Lo Scrudato, M. and Blokesch, M. (2012) The regulatory network of natural competence and transformation of *Vibrio cholerae*. *PLoS Genet* **8**: 1002778.
- Lo Scrudato, M., Borgeaud, S., and Blokesch, M. (2014) Regulatory elements involved in the expression of competence genes in naturally transformable *Vibrio cholerae*. *BMC Microbiol* **14**: 1–13.
- Seitz, P. and Blokesch, M. (2012) Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative

- bacteria. *FEMS Microbiol Rev* **37**: 336–363.
- Shang, L., Yan, Y., Zhan, Y., Ke, X., Shao, Y., Liu, Y., et al. (2021) A regulatory network involving Rpo, Gac and Rsm for nitrogen-fixing biofilm formation by *Pseudomonas stutzeri*. *npj Biofilms Microbiomes* **7**: 1–15.
- Shimada, K., Itoh, Y., Washio, K., and Morikawa, M. (2012) Efficacy of forming biofilms by naphthalene degrading *Pseudomonas stutzeri* T102 toward bioremediation technology and its molecular mechanisms Kohei. *Chemosphere* **87**: 226–233.
- Søndberg, E., Sinha, A.K., Gerdes, K., and Semsey, S. (2019) CRP interacts specifically with *sxy* to activate transcription in *Escherichia coli*. *Front Microbiol* **10**: 1–8.
- Sparling, P.F. (1966) Genetic transformation of *Neisseria gonorrhoeae* to streptomycin resistance. *J Bacteriol* **92**: 1364–1371.
- Spencer, H.T. and Herriott, R.M. (1965) Development of competence of *Haemophilus influenzae*. *J Bacteriol* **90**: 911–920.
- Veening, J.W., Smits, W.K., and Kuipers, O.P. (2008) Bistability, epigenetics, and bet-hedging in bacteria. *Annu Rev Microbiol* **62**: 193–210.
- Williams, P.M., Bannister, L.A., and Redfield, R.J. (1994) The *Haemophilus influenzae* *sxy*-1 mutation is in a newly identified gene essential for competence. *J Bacteriol* **176**: 6789–6794.
- 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.
- Wu, R., Zhao, M., Li, J., Gao, H., Kan, B., and Liang, W. (2015) Direct regulation of the natural competence regulator gene *tfoX* by cyclic AMP (cAMP) and cAMP receptor protein (CRP) in *Vibrios*. *Sci Rep* **5**: 1–15.
- Zhang, J., Guan, J., Wang, M., Li, G., Djordjevic, M., Tai, C., et al. (2023) SecReT6 update: a comprehensive resource of bacterial Type VI Secretion Systems. *Sci China Life Sci* **66**: 626–634.

3.8 Supplementary Material

Table S 3.1 Bacterial strains, plasmids, and oligonucleotides used in this chapter. Abbreviations: AmpR, ampicillin resistance; GmR, gentamicin resistance; KmR, kanamycin resistance; CmR, chloramphenicol resistance; MCS, multiple cloning site.

Strains, plasmids or oligonucleotides		Genotype, phenotype, or relevant characteristics	Reference
Strains			
<i>P. stutzeri</i>	ATCC 17587	Wild type phenotype	Laboratory strain
<i>P. stutzeri</i>	ATCC 17587	Spontaneous rifampicin-resistant phenotype	Chapter 2
<i>P. stutzeri</i>	tetR-kmR	ATCC 17587, miniTn7::Km ^R ::Tet ^R , contains two cassettes of resistance against kanamycin and tetracycline	Chapter 4
<i>P. stutzeri</i>	ΔcomA	ATCC 17587 ΔcomA	Chapter 2
<i>P. stutzeri</i>	ΔcomA rif ^R	ATCC 17587 ΔcomA spontaneous rifampicin-resistant mutant	This study
<i>P. stutzeri</i>	ΔtfoXC	ATCC 17587 ΔtfoXC	This study
<i>P. stutzeri</i>	ΔtfoXN	ATCC 17587 ΔtfoXN	This study
<i>P. stutzeri</i>	ΔtfoXC ΔtfoXN	ATCC 17587 ΔtfoXC ΔtfoXN	This study
<i>P. stutzeri</i>	ΔclpV	ATCC 17587 ΔclpV	This study
<i>P. stutzeri</i>	sfGFP	ATCC 17587, miniTn7::sfGFP, contains a promoterless sfGFP	This study
<i>P. stutzeri</i>	PpilA-sfGFP	ATCC 17587, miniTn7::PpilA-sfGFP, contains the promoter of <i>pilA</i> controlling the sfGFP gene	This study
<i>P. stutzeri</i>	PcomA-sfGFP	ATCC 17587, miniTn7::PcomA-sfGFP, contains the promoter of <i>ComA</i> controlling the sfGFP gene	This study
<i>E. coli</i>	DH5α-λpir	Laboratory cloning strain	(Sambrook and Russell, 2001)
<i>E. coli</i>	BTH101	F ⁻ , <i>cya</i> -99, <i>araD</i> 139, <i>galE</i> 15, <i>galK</i> 16, <i>rpsL</i> 1 (Str ^r), <i>hsdR</i> 2, <i>mcrA</i> 1, <i>mcrB</i> 1, Bacterial two-hybrid assay	BACTH system kit, Euromedex
Plasmids			
	pEMG	Km ^R , oriR6K, MCS with two flanking I-SceI sites, used for markerless deletions	(Martínez-García and de Lorenzo, 2011)
	pEMG-ΔtfoXC	pEMG bearing upstream and downstream flanking regions of <i>tfoXC</i>	This study
	pEMG-ΔtfoXN	pEMG bearing upstream and downstream flanking regions of <i>tfoXN</i>	This study
	pEMG-ΔclpV	pEMG bearing upstream and downstream flanking regions of ΔclpV	This study
	pSW-2	oriRK2, Gm ^R , <i>xylS</i> , Pm→I-sceI (the I-Sce-I gene controlled by Pm promoter)	(Martínez-García and de Lorenzo, 2011)
	pUX-BF13	oriR6K, Amp ^R , helper plasmid providing the Tn7 transposition functions in trans	(Bao <i>et al.</i> , 1991)
	pME6182	Mini-Tn7 gene delivery vector based on pME3280a, ColE1 replicon, Gm ^R , Amp ^R	(Humair <i>et al.</i> , 2009)
	pColE2-sfGFP-Cm ^R	pColE2 plasmid (Mavridou <i>et al.</i> , 2018) with Cm ^R resistance cassette and superfolder GFP gene (<i>sfGFP</i>)	Lab plasmid
	pME6182-sfGFP	pME6182 containing a promoterless sfGFP gene	This study
	pME6182-PpilA-sfGFP	pME6182 containing the <i>pilA</i> promoter region controlling the sfGFP gene	This study
	pME6182-PcomA-sfGFP	pME6182 containing the <i>comA</i> promoter region controlling the sfGFP gene	This study
	pKT25	Km ^R	THE BACTH system kit, Euromedex
	pKT25-Zip	pKT25 derivative; encodes a leucine zipper ; Km ^R	BACTH system kit, Euromedex

pKT25-CRP	pKT25 derivative; encodes the locus Ps07430, coding for cyclic AMP receptor-like protein, from <i>P. stutzeri</i> ATCC17587; Km ^R	This study
pKT25-tfoXC	pKT25 derivative; encodes the locus Ps32070 from <i>P. stutzeri</i> ATCC17587; Km ^R	This study
pKT25-tfoXN	pKT25 derivative; encodes the locus Ps26900 from <i>P. stutzeri</i> ATCC17587; Km ^R	This study
pUT18C	Amp ^R	THE BACTH system kit, Euromedex
pUT18C-Zip	pUT18C derivative; encodes a leucine zipper ; Amp ^R	BACTH system kit, Euromedex
pUT18C-CRP	pUT18C derivative; encodes the locus Ps07430, coding for cyclic AMP receptor-like protein, from <i>P. stutzeri</i> ATCC17587; Amp ^R	This study
pUT18C-tfoXC	pUT18C derivative; encodes the locus Ps32070 from <i>P. stutzeri</i> ATCC17587; Amp ^R	This study
pUT18C-tfoXN	pUT18C derivative; encodes the locus Ps26900 from <i>P. stutzeri</i> ATCC17587; Amp ^R	This study
Oligonucleotides (5' → 3')		
tfoXC_UFR_F	AAAGAATTCGATATTTTCGCCATGTCTTGGC, with a EcoRI underlined site, located upstream of the <i>tfoXC</i> homologue	This study
tfoXC_UFR_R	TTGTGAGCTAGCTCAGTTCGTCGTGCATGGTG, with overlap to downstream flanking region underlined, located upstream of the <i>tfoXC</i> homologue	This study
tfoXC_DFR_F	TGCACGACGAACTGAGCTAGCTCACAACCTCTGC, with overlap to upstream flanking region underlined, located downstream of the <i>tfoXC</i> homologue	This study
tfoXC_DFR_R	AAAGGATCCGAAGTGCTTGAAGCCAGTAGC, with a BamHI underlined site, located downstream of the <i>tfoXC</i> homologue	This study
tfoXN_UFR_F	AAAGAATTCGCTTTACGCCGAAGAACG, with a EcoRI underlined site, located upstream of the <i>tfoXN</i> homologue	This study
tfoXN_UFR_R	GATCCTGCCTGCGTTGCCCGCCTTACACC, with overlap to downstream flanking region underlined, located upstream of the <i>tfoXN</i> homologue	This study
tfoXN_DFR_F	AAGGCGGGCAACGCAGGCAGGATCTCTCG, with overlap to upstream flanking region underlined, located downstream of the <i>tfoXN</i> homologue	This study
tfoXN_DFR_R	AAAGGATCCCTGAAATACGGCACCTTG, with a BamHI underlined site, located downstream of the <i>tfoXN</i> homologue	This study
clpV_UFR_F	AAAGAATTCGGCTTTCCCGGCAGTGATAT, with a EcoRI underlined site, located upstream of the locus Ps41000	This study
clpV_UFR_R	AACATTGCGGACATCTTTTCATCTTCTTAATGGGTCTG G, with overlap to downstream flanking region underlined, located upstream of the locus Ps41000	This study
clpV_DFR_F	ATTAAGGAAGATGAAAGATGTCCGCAATGTTCAATCAG GT, with overlap to upstream flanking region underlined, located downstream of the locus Ps41000	This study
clpV_DFR_R	AAAGGATCCAGCAGATGCAGCTGGGC, with a BamHI underlined site, located downstream of the locus Ps41000	This study

sfGFP_pME_F	AAAGAATT <u>CGATG</u> CAAAATAAGGAGTTTAAGGAG, with a EcoRI underlined site, forward primer to amplify <i>sfGFP</i> from pColE2- <i>sfGFP</i> -Cm plasmid, for construct with promotor	This study
HSKNS-sfGFP	GGGAAGCTTCCCGGGGTACCCCATGGACTAGT GATGCAAAATAAGGAGTTTAAGGAG, with a HindIII underlined site, forward primer to amplify <i>sfGFP</i> from pColE2- <i>sfGFP</i> -Cm plasmid, for construct without promotor	This study
sfGFP_pME_R	AAAGCATG <u>CCCTG</u> TGGATCCTTATTATTTGTACAG, with a SphI underlined site, reverse primer to amplify <i>sfGFP</i> from pColE2- <i>sfGFP</i> -Cm plasmid	This study
sfGFP-AAV_pME_R	AAAGCATG <u>CCCTG</u> TGGATCCTTATTA <u>AACTGCTGCAG</u> CGTAGTTTTCGTCGTTTGCTGCAGGCCTTTGTACAGTT CATCCATA, with a SphI underlined site, reverse primer to amplify <i>sfGFP</i> -AAV from pColE2- <i>sfGFP</i> -Cm plasmid. The amino acids tag added to render the sfGFP less stable is highlighted in grey.	This study
pilA_prom_F	ATAAAGCTTGGCCTCGCTCCAATCTATC, with a HindIII underlined site, forward primer to amplify <i>pilA</i> promotor	This study
pilA_prom_R	ATAGAATTCTTTGGATCTCCTTTAAGTGTTAAGG, with a EcoRI underlined site, reverse primer to amplify <i>pilA</i> promotor	This study
comA_prom_F	ATAAAGCTTAGAAAACGCAGCGACTTGTTAG, with a HindIII underlined site, forward primer to amplify <i>comA</i> promotor	This study
comA_prom_R	ATAGAATTCTTTCAGTCTCCAGAAATCCTTTCC, with a EcoRI underlined site, reverse primer to amplify <i>comA</i> promotor	This study
Ps_tfoXC_F	AAAGGATCCCATGCACGACGAACTGCTTTTCG, with a BamHI underlined site, reverse primer to amplify <i>tfoXC</i> (Ps32070)	This study
Ps_tfoXC_R	AAAGAATTCTAGTTGTGCGGGATTGGTGCC, with a EcoRI underlined site, reverse primer to amplify <i>tfoXC</i> (Ps32070)	This study
Ps_tfoXN_F	AAAGGATCCCGTGGAACGGCCTCGCTTTC, with a BamHI underlined site, forward primer to amplify <i>tfoXN</i> (Ps26900)	This study
Ps_tfoXN_R	AAAGAATTCTAAGTCGCGGCTGCGTC, with a EcoRI underlined site, reverse primer to amplify <i>tfoXN</i> (Ps26900)	This study
PsCRP_F	AAAGGATCCCATGGTAGCTATCACTCTTAC GC, with a BamHI underlined site, forward primer to amplify locus Ps07430, coding for cyclic AMP receptor-like protein.	This study
PsCRP_R	AAAGAATTCTCAGCGAGTACCGAATACCA, with a EcoRI underlined site, reverse primer to amplify locus Ps07430, coding for cyclic AMP receptor-like protein.	This study
pKT25_F	CGATTCGGTGACCGATTACC	Sequencing primer
pKT25_R	CGATTAAGTTGGGTAACGCCAG	Sequencing primer
pUT18C_F	CGCCGGATGTACTGGAAC	Sequencing primer
pUT18C_R	CGGCATCAGAGCAGATTGTA	Sequencing primer
M13F-47	CGCCAGGGTTTTCCAGTCACGAC	Sequencing primer
M13R	TCACACAGGAAACAGCTATGAC	Sequencing primer

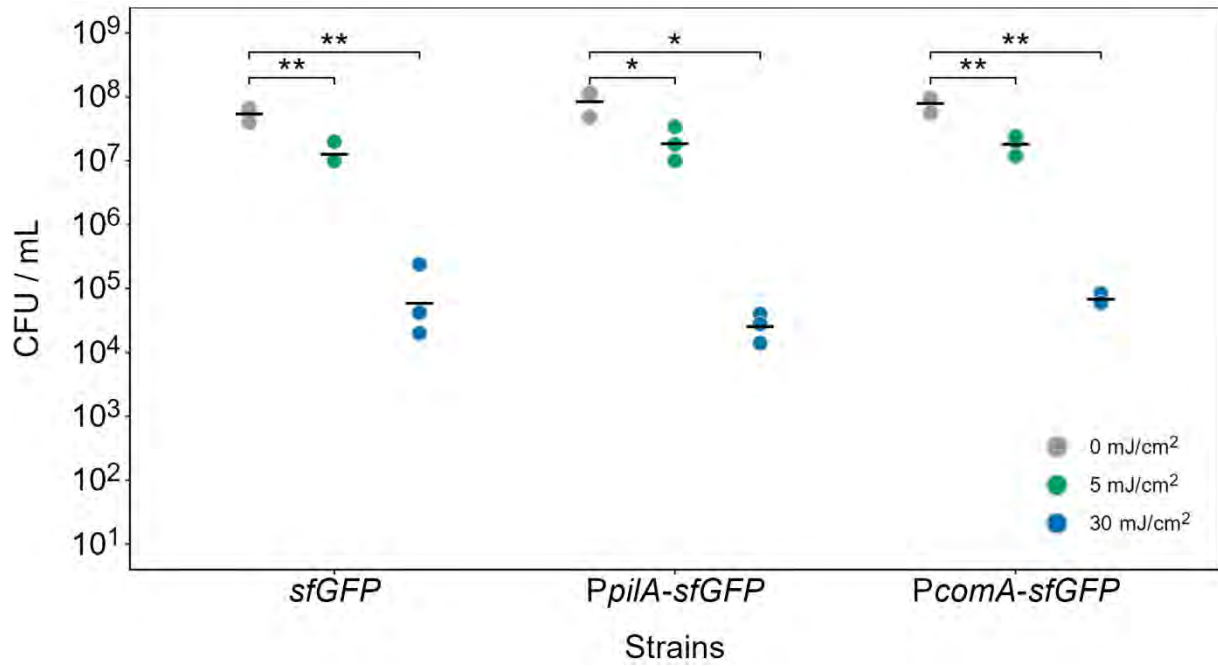


Figure S 3.1. UV treatments reduced cell viability. Directly after being submitted to UV, cells were plated on LB to estimate their survival. The 5 mJ/cm² treatment (green) reduced the cell population by 75% relative to the control (grey), the 30 mJ/cm² treatment reduced the cell population by 1000 times. No clear differences between the strains were noticeable. The three replicates are represented with dots, the mean values with the black bar.

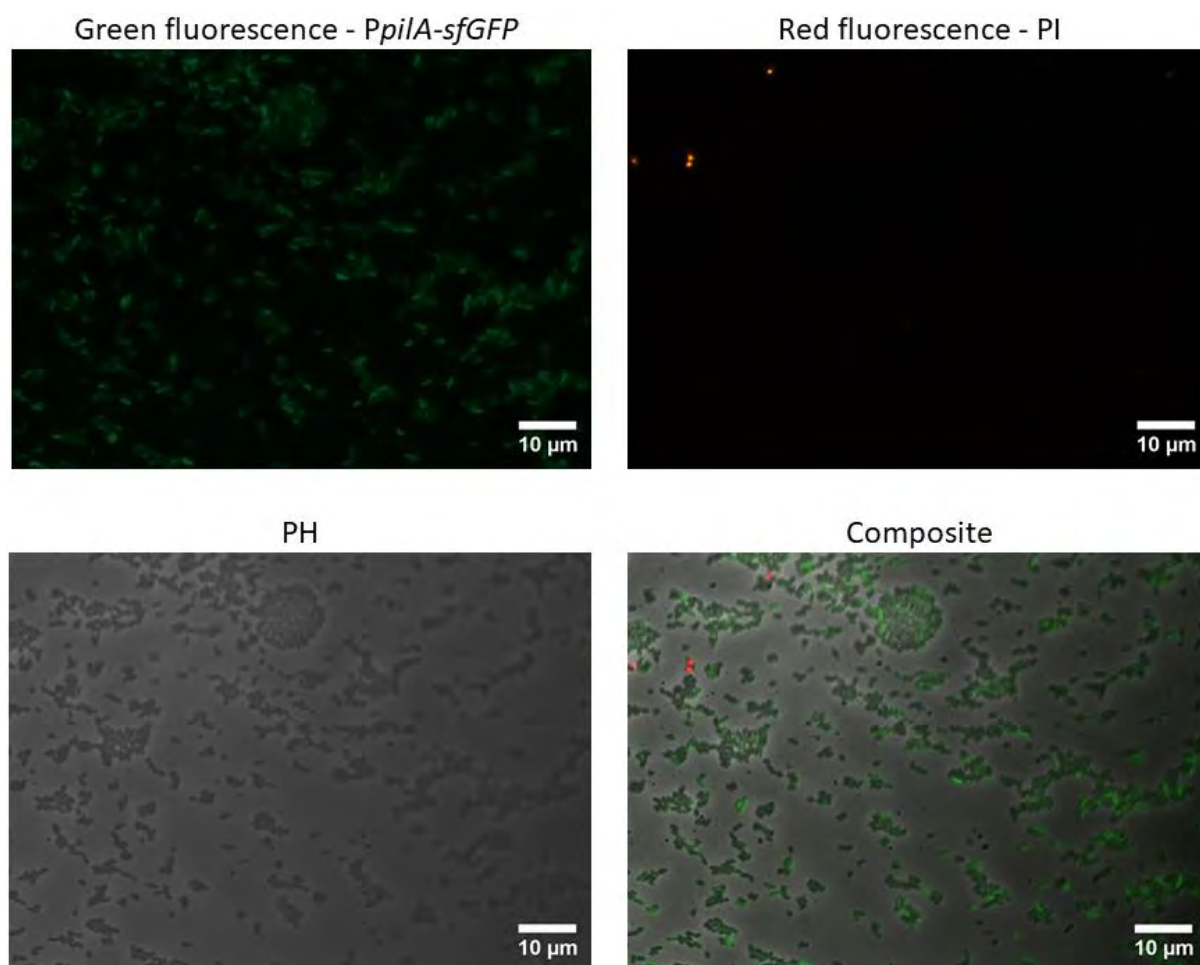


Figure S 3.2 Non-fluorescent cells of reporter strain *PpilA-sfGFP* were mainly alive. Representative micrographs showing cells of the reporter strains *PpilA-sfGFP* stained with propidium iodide (PI) during the exponential phase in a liquid culture with LB. As PI penetrates only bacteria with damaged membranes, dead cells produce a red fluorescence. Composite image results from the superposition of micrographs obtained by phase-contrast (PH) and fluorescence channels.

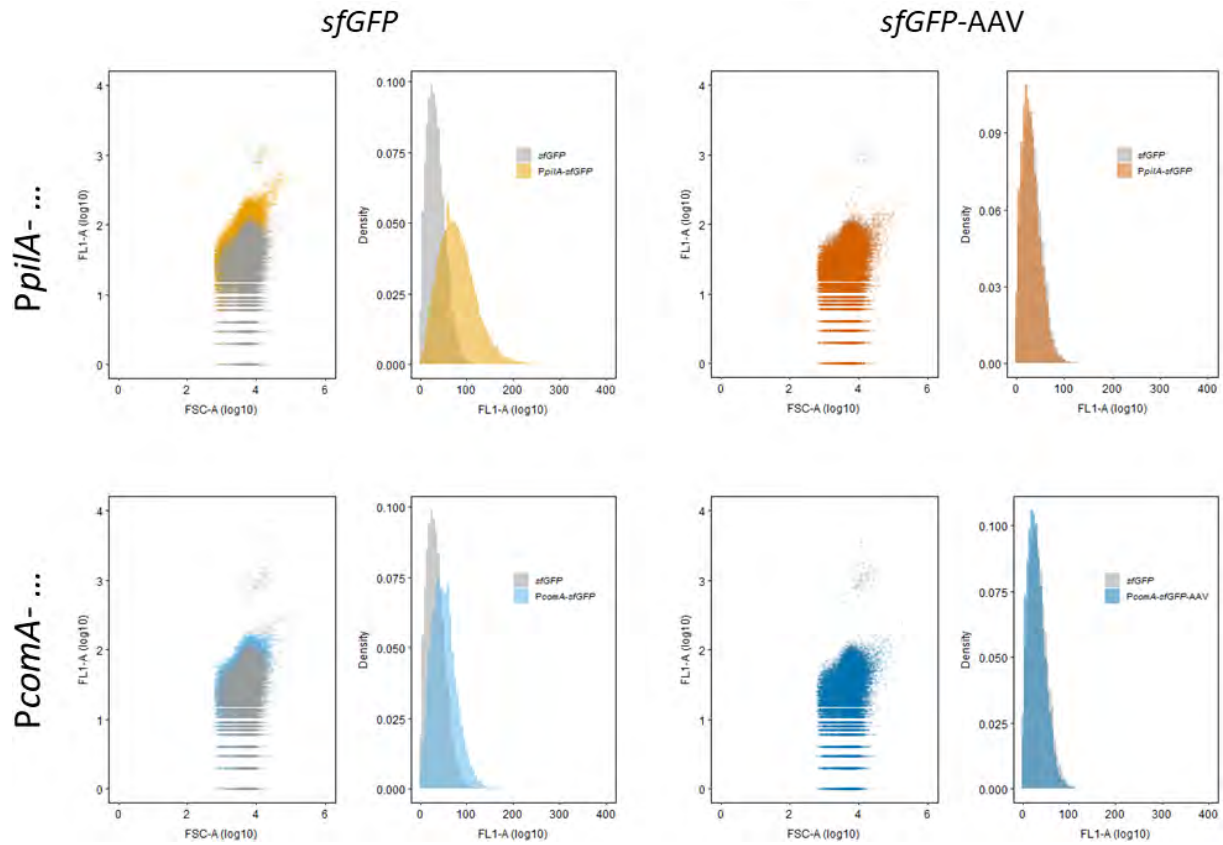
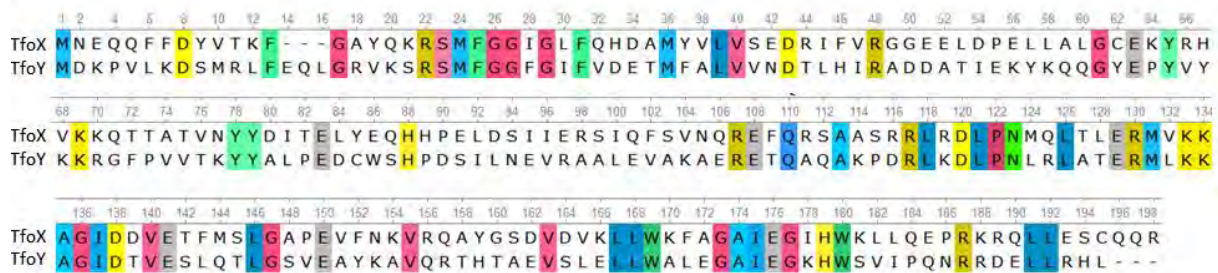


Figure S 3.3 Fluorescence of reporter strains with an unstable *sfGFP* (*sfGFP-AAV*) was not detected by flow cytometry in contrast to reporters with an unmodified *sfGFP*. The strain carrying the promoter-less construct is represented in grey.

A



B

Aligned better to	TfoX	TfoY
Described as		
TfoX or TfoX1	162	99
TfoX2	0	4

C

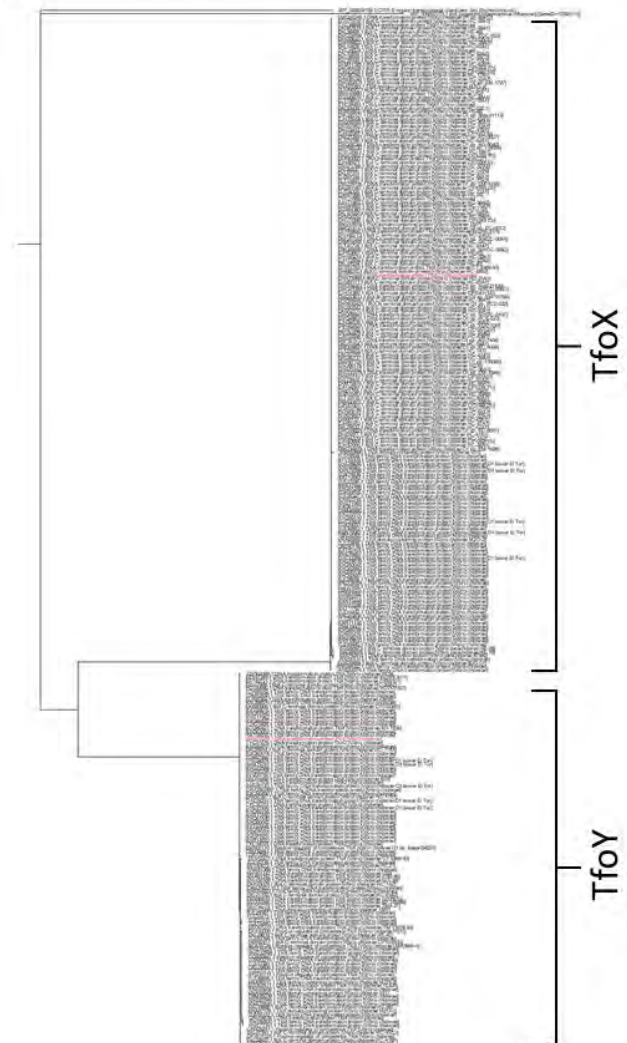


Figure S 3.4 TfoY is often misidentified as TfoX in *V. cholerae*. We analysed protein sequences from the NCBI database corresponding to «TfoX *Vibrio cholerae*» excluding entries shorter than 175 amino acids, not identified as TfoX-like, or not from *V. cholerae*. These sequences were aligned with *V. cholerae* TfoX (AWB73649.1) and TfoY (AWB74250.1) to determine their closest match. The functions of their respective genes have been particularly studied regarding natural transformation and T6SS in *V. cholerae* (Metzger *et al.*, 2016). (A) Alignment of TfoX (AWB73649.1) and TfoY (AWB74250.1) showing the shared amino acids (colored). (B) This table summarizes the accuracy of protein entry naming. (C) Phylogenetic tree based on the entries for “TfoX *Vibrio cholerae*” (all in black) with TfoX (AWB73649.1) and TfoY (AWB74250.1) from *V. cholerae* highlighted in red, rooted with TfoX homologues from *H. influenzae* (WP_005692773.1) and *E. coli* (WP_000839159.1).

Table S 3.2 Homologues of type VI secretion system in *P. stutzeri*. Information regarding the best match to these homologues is displayed in columns 2 to 5. Results of the BLASTp are shown in columns 6 to 8 (Altschul *et al.*, 1990). SecReT6 predicts two T6SS gene clusters which are annotated in the last column (Zhang *et al.*, 2023). %ID: percentage of identity, %QC: percentage query cover *also called *clpV*

<i>P. stutzeri</i>	Reference	Species	Name	Product	%ID	%QC	E-value	SecReT6 prediction
LOCUS_01250	WP_003147991.1	<i>Pseudomonas aeruginosa</i> PAO1	<i>tssI</i>	type VI secretion system tip protein TssI/VgrG	76.7	100.0	5.6e-79	1
LOCUS_01280	WP_003114614.1	<i>Pseudomonas aeruginosa</i> PAO1	<i>tssM</i>	type VI secretion system membrane subunit TssM	70.8	97.1	1.1e-109	1
LOCUS_01300	WP_011105481.1	[<i>Pseudomonas syringae</i>] pv. tomato str. DC3000	<i>tssK</i>	type VI secretion system baseplate subunit TssK	65.3	96.9	5.8e-119	1
LOCUS_01310	WP_003114612.1	<i>Pseudomonas aeruginosa</i> PAO1	<i>tssJ</i>	type VI secretion system lipoprotein TssJ	42.1	32.6	2.5e-11	1
LOCUS_01320	WP_011105483.1	[<i>Pseudomonas syringae</i>] pv. tomato str. DC3000	<i>tagH</i>	type VI secretion system-associated FHA domain protein TagH	45.4	92.9	2.7e-51	1
LOCUS_02010	NP_253954.1	<i>Pseudomonas aeruginosa</i> PAO1	<i>hcpB</i>	secreted protein Hcp	41.8	60.1	1.1e-25	
LOCUS_02020	WP_011105496.1	[<i>Pseudomonas syringae</i>] pv. tomato str. DC3000	<i>tssI</i>	type VI secretion system tip protein TssI/VgrG	49.3	95.2	3.9e-51	
LOCUS_40910	WP_003097814.1	<i>Pseudomonas aeruginosa</i> PAO1	<i>tssA</i>	type VI secretion system protein TssA	71.2	97.6	2.1e-293	2
LOCUS_40920	WP_003114607.1	<i>Pseudomonas aeruginosa</i> PAO1	<i>tssB</i>	type VI secretion system contractile sheath small subunit	80.6	100.0	4.0e-212	2
LOCUS_40930	WP_011105492.1	[<i>Pseudomonas syringae</i>] pv. tomato str. DC3000	<i>tssC</i>	type VI secretion system contractile sheath	63.2	100.0	4.8e-137	2
LOCUS_40940	WP_011105491.1	[<i>Pseudomonas syringae</i>] pv. tomato str. DC3000	<i>tssE</i>	type VI secretion system baseplate subunit TssE	74.9	100.0	0.0e+00	2
LOCUS_40980	WP_003120358.1	<i>Pseudomonas aeruginosa</i> PAO1	<i>tssF</i>	type VI secretion system baseplate subunit TssF	79.9	99.3	1.6e-54	2
LOCUS_40990	WP_003114608.1	<i>Pseudomonas aeruginosa</i> PAO1	<i>tssG</i>	type VI secretion system baseplate subunit TssG	85.8	100.0	4.1e-253	2
LOCUS_41000	WP_011105486.1	[<i>Pseudomonas syringae</i>] pv. tomato str. DC3000	<i>tssH</i> *	type VI secretion system ATPase TssH	70.4	99.0	1.1e-297	2
LOCUS_41020	WP_003114027.1	<i>Pseudomonas aeruginosa</i> PAO1	PA5266	type VI secretion system tip protein VgrG	63.7	82.2	5.7e-56	

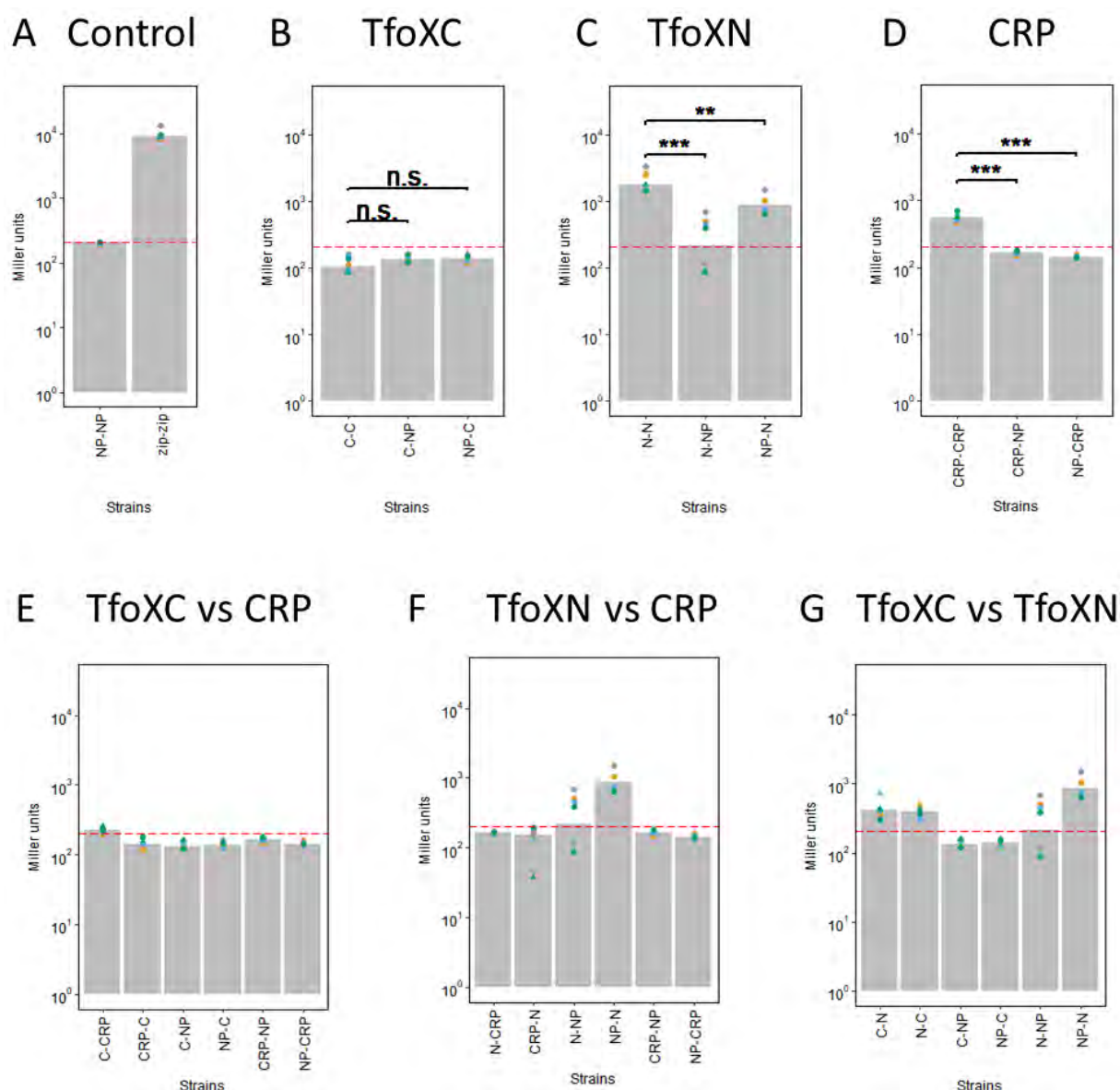


Figure S 3.5 Investigation of TfoX-like homologues interaction. Plots (A) shows results of control strains. Plots (B), (C) and (D) show interaction of the proteins with themselves. Interactions between different proteins are displayed in plots (E), (F) and (G). β -galactosidase assays were used to quantify the interactions between TfoXC, TfoXN, and CRP in *E. coli* 24 h after spotting. The two technical replicates are shown with different shapes and the three biological replicates with different colors. The bars show the medians, and the horizontal red line indicates the median of the negative control (NP-NP). NP: no protein, negative control; Zip: leucine zipper, positive control; C: TfoXC; N: TfoXN, CRP: cAMP receptor protein.

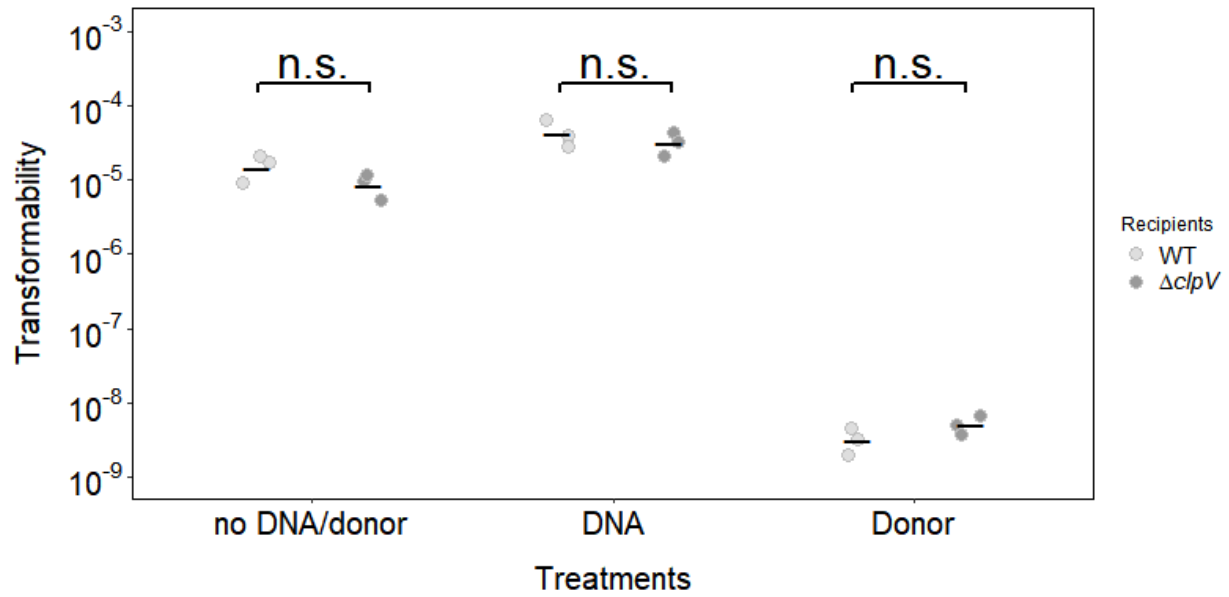


Figure S 3.6 Deletion of *clpV* homologue did not significantly affect the transformability of *P. stutzeri* with donor cells. Transformability of *P. stutzeri* ATCC 17587 wild-type (WT) and $\Delta clpV$ mutant with extracellular DNA (DNA), in the presence of rifampicin-resistant $\Delta comA$ donor cells (Donor) or with water (none) as control treatment. The dots represent the three replicates and the black lines the means.

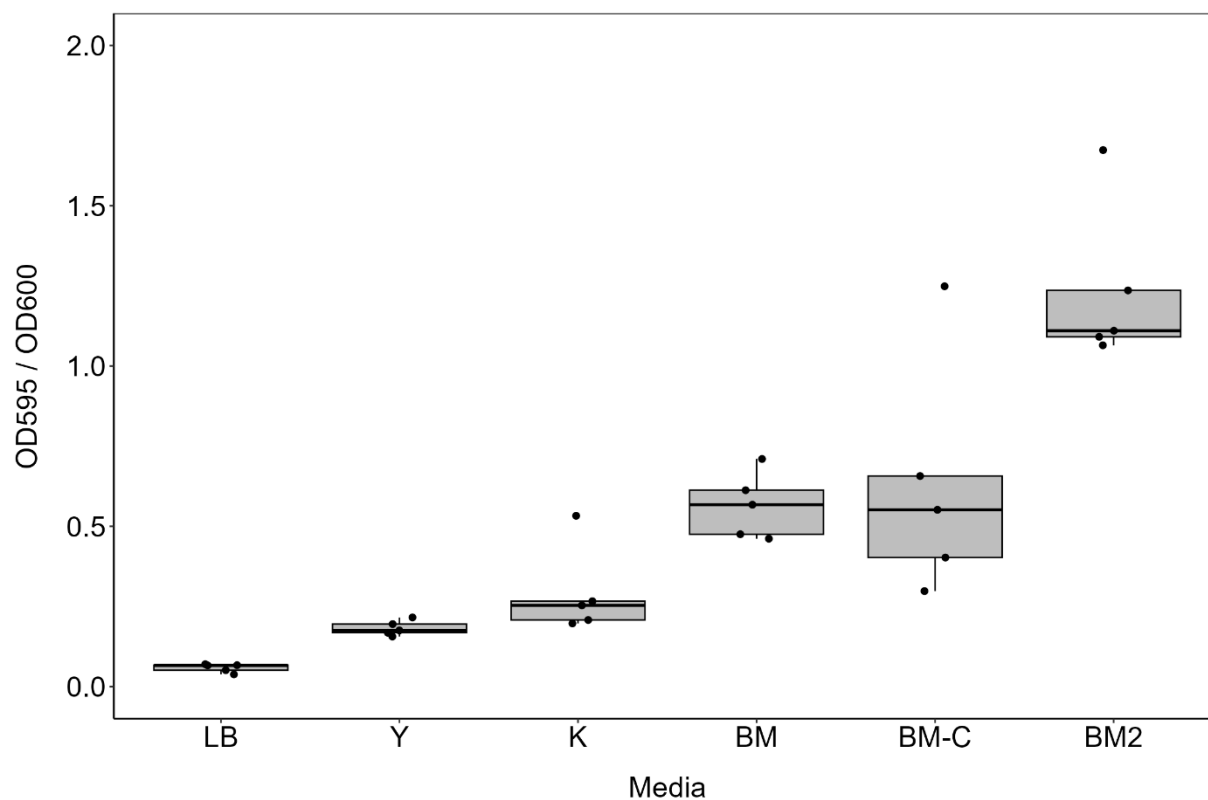


Figure S 3.7 Effect of media on biofilm production of wild-type strain of *P. stutzeri*. *P. stutzeri* was grown statically for 7 days in various media (Y medium (Shimada *et al.*, 2012), Minimal medium K (Shang *et al.*, 2021), Basal Medium (BM) (Ding *et al.*, 2019), BM without carbon (BM-C), Basal Medium 2 (BM2) (Mulcahy *et al.*, 2010)). Then, the biofilm was stained with crystal violet. The retention of crystal violet was quantified at 595 nm and normalized by culture density measured at 600 nm in five replicates for each medium.

Chapter 4 - Broad-scale study of essential genes and natural transformation genetics in *P. stutzeri*

4.1 Foreword

In this part of the project, we tried to set-up a transposon insertion sequencing method to identify the genes impacting natural transformation in *P. stutzeri*.

The key elements of this chapter are the following:

- Transformability of *P. stutzeri* was found to be similar when a genetic marker was provided as extracellular DNA or within intact donor cells.
- Adaptation of transposon insertion sequencing protocol allowed the identification of essential genes for *P. stutzeri*.
- However, low transposon saturation in libraries after natural transformation resulted in limited statistical significance.

4.2 Abstract

Natural transformation, a key mechanism of horizontal gene transfer in bacteria, is typically studied using purified DNA, which may not fully represent natural conditions. In this study, we employed a transposon insertion sequencing (TIS) method to try and identify genes involved in natural transformation in *Pseudomonas stutzeri*, with a particular focus on the role of cell-cell interactions. A transposon mutant library of recipient cells was generated and transformed with either extracellular DNA or donor cells lacking the membrane-channel protein ComA to prevent cross-transformation. Transformants and the entire recipient population were collected for DNA extraction and sequencing. Despite collecting sufficient mutants, our initial TIS methodology resulted in suboptimal transposon density in the initial library. Even after optimizing the preparation of sequencing libraries, we were unable to obtain meaningful results regarding the genetic determinants required for natural transformation. This suggests that further refinement of the TIS methodology is necessary to fully elucidate the genes involved in DNA uptake, translocation, recombination, and cell-cell interaction during natural transformation. Nevertheless, by pooling sequencing libraries, we successfully identified core genome genes in *P. stutzeri*. This work provides valuable insights into this species core genome and underscores the significance of cell-cell interactions during natural transformation.

4.3 Introduction

Natural transformation is a mechanism of horizontal gene transfer (HGT) in bacteria, allowing the uptake and integration of extracellular DNA into the host genome. This process plays a crucial role in bacterial evolution and adaptation by facilitating the acquisition of new traits, repair of damaged genes, and enhancement of genetic diversity (Redfield, 1993; Ambur *et al.*, 2016). In addition, natural transformation distinguishes itself from other HGT mechanisms by being controlled by the recipient cells. Unlike conjugation or transduction, which are driven by plasmids, transposons, or bacteriophages, transformation relies on the recipient's ability to develop competence and actively uptake external DNA.

Typical assays of natural transformation involve mixing recipient cells with purified DNA in specific conditions and then quantifying the number of transformants relative to the total number of cells. However, there is increasing evidence indicating that this conventional methodology does not fully reflect the transformation process happening in the natural environment. In natural settings, extracellular DNA originates from diverse sources and is susceptible to degradation by extracellular nucleases, which can reduce transformation rates (Nielsen *et al.*, 1997, 2007; Sikorski *et al.*, 1998). Given these factors, it is reasonable to hypothesize that bacteria have evolved selective mechanisms to acquire the most beneficial DNA within this heterogeneous pool of DNA from multiple sources and varying quality.

To increase the probability of incorporating genetic material from the same or closely related species, several strategies have evolved (Seitz and Blokesch, 2012). For instance, *Neisseria gonorrhoeae* and *Haemophilus influenzae* uptake preferentially genetic material with specific DNA sequences that are overrepresented in their genome (Sisco and Smith, 1979; Danner *et al.*, 1980; Graves *et al.*, 1982; Smith *et al.*, 1999). In contrast, other species do not show any discrimination between self and foreign DNA regarding the uptake but perform bacterial predation, allowing recipient cells to access fresh DNA. For example, Gram-negative bacteria such as *Vibrio cholerae* and *Acinetobacter baylyi* use type VI secretion systems to lyse genetically different members of their own species (Borgeaud *et al.*, 2015; Ringel *et al.*, 2017). Similarly, Gram-positive *Streptococcus* species produce competence-stimulating peptides (CSPs) that not only trigger competence but also the production of toxins. These toxins lyse related strains and species with different CSPs, thereby enhancing the efficiency of lateral gene transfer (Kreth *et al.*, 2005; Johnsborg *et al.*, 2008). The relationships between donor and

recipient cells do not always involve killing. Indeed, *N. gonorrhoeae* secretes chromosomal DNA for transformation via a type IV secretion system (Dillard and Seifert, 2001; Hamilton *et al.*, 2005). Interestingly, transformation with donor cells fosters the transfer of large continuous genomic regions in *B. subtilis*, *V. cholerae*, and *S. pneumoniae* (Cowley *et al.*, 2018; Matthey *et al.*, 2019; Deng *et al.*, 2023). These examples underscore the complex ecological interactions that may underlie natural transformation in bacterial communities and the importance of considering cell-to-cell interactions in the process of natural transformation.

The importance of donor cells in natural transformation was already mentioned 40 years ago in the environmental model for natural transformation *Pseudomonas stutzeri* (Stewart *et al.*, 1983). Stewart and colleagues observed that transformation efficiency was significantly enhanced when DNA was introduced as a component of intact donor cells rather than as a purified DNA suspension (Stewart *et al.*, 1983). Their findings also emphasized the crucial role of cell-to-cell contact in facilitating this process. In addition, they suspected that the donation happened by secretion rather than cell lysis since no cell death was detectable. However, the mechanism behind was never elucidated. The ecological versatility of *P. stutzeri* makes it relevant for studying natural transformation and understanding how environmental factors affect cell-to-cell interactions. Indeed, *P. stutzeri* is found in various environments, including soil and marine habitats (Lalucat *et al.*, 2006). Moreover, *P. stutzeri* is closely related to other clinically and environmentally important *Pseudomonas* species performing natural transformation (Carlson *et al.*, 1983; Demanèche *et al.*, 2001; Nolan *et al.*, 2020), allowing for comparative studies and potential extrapolation of findings to other organisms. However, our knowledge of the genetics of the natural transformation of *P. stutzeri* is very incomplete. Only a few genes of the competence apparatus have been experimentally identified, especially using deletion mutants (Chapter 1, table 1) (Graupner *et al.*, 2000, 2001; Graupner and Wackernagel, 2001b). A high-throughput method combined with well-designed experiments might allow the investigation of important genes for the recipient as well as for the donor in the process of natural transformation in *P. stutzeri*.

Transposon insertion sequencing (TIS) has revolutionized the field of functional genomics by enabling high-throughput identification of genes essential for various biological processes. This methodology combines high-throughput transposon mutagenesis with next-generation

sequencing technologies. This method focuses on analyzing the effects of gene disruptions caused by transposon insertions.

The TIS workflow typically involves four steps presented in Figure 4.1: (A) the generation of a saturated transposon mutant library by introducing a randomly inserting transposon (commonly a Tn5 or mariner transposon), (B), optional cultivation of the mutant pool under specific conditions of interest, (C) preparation of the sequencing library and massively parallel sequencing of transposon-genome junctions, and (D) computation analysis to quantify mutant frequencies and infer gene essentiality or fitness contributions (Chao *et al.*, 2016; Cain *et al.*, 2020).

TIS methods have diverse applications in microbial genetics research (Kwon *et al.*, 2016). One of the first applications was to identify essential genes, those crucial for growth or survival under optimal conditions. In highly saturated libraries, essential genes can be inferred from genomic regions with few or no transposon insertions, as mutations in these genes would result in non-viable organisms. However, TIS applications extend far beyond essentiality determination. Researchers employ TIS to identify genes required for growth or survival under specific conditions, as well as those involved in various biological processes such as host colonization, adherence, and sporulation (Figure 4.1) (Kwon *et al.*, 2016). Furthermore, TIS can elucidate genetic interaction networks by comparing wild-type strains with mutants (van Opijnen *et al.*, 2009). The power of TIS lies in its ability to provide nuanced results beyond simple essential/non-essential classifications. By analysing significant changes in the relative frequency of mutants before and after selection steps or between different conditions, researchers can infer the fitness contribution of specific genes. This quantitative approach allows for a more comprehensive understanding of gene function and importance in various contexts. In conclusion, TIS has revolutionized our ability to conduct high-throughput functional genomics studies in microorganisms, offering unprecedented insights among others into gene essentiality, conditional fitness, and complex genetic interactions. As the technique continues to evolve, it promises to remain a cornerstone in unravelling the intricacies of microbial genomes and their functions.

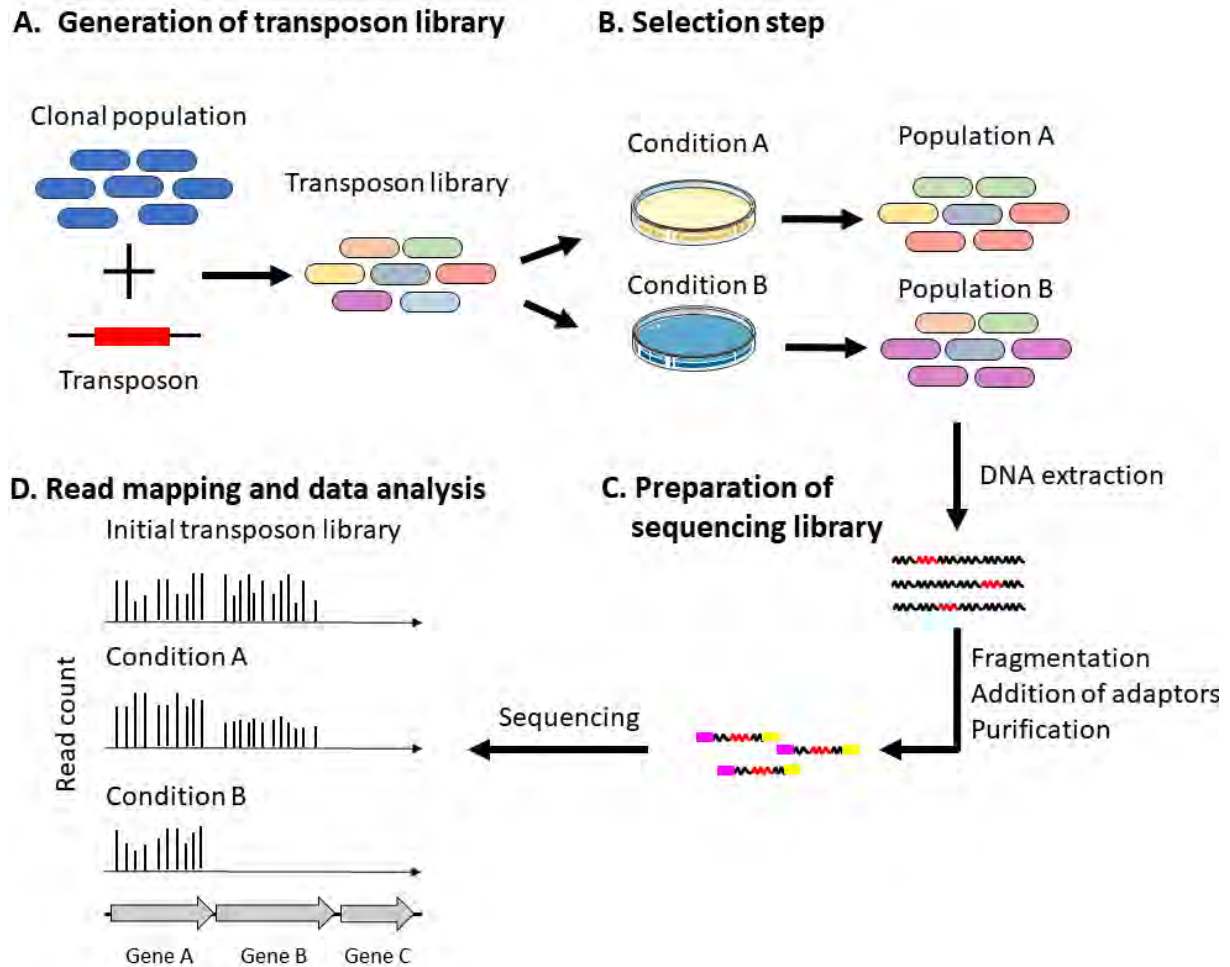


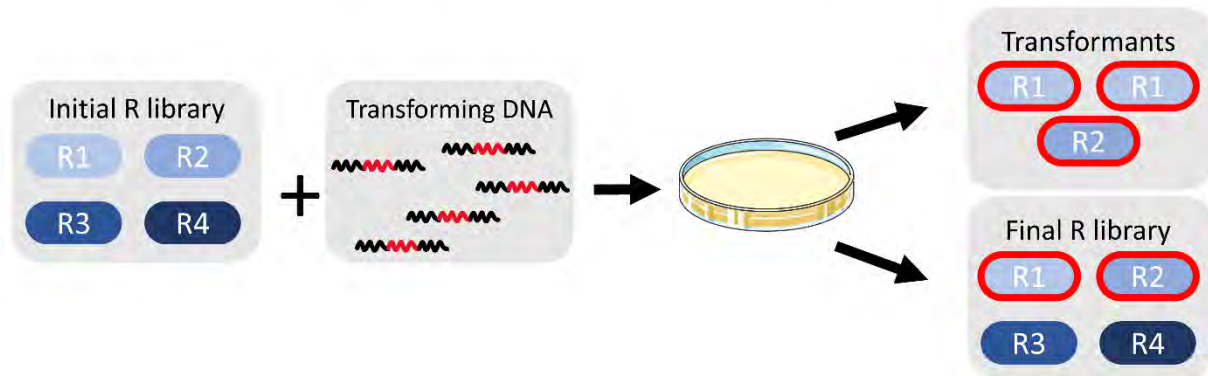
Figure 4.1 Schematic of a typical workflow of TIS experiments to determine conditionally essential genes. (A) Creation of mutants library due to random insertion of transposon in a clonal population. (B) The transposon library is for example grown under conditions A and B which leads to the selection of different mutants. Mutants with transposon insertion in genes essential for growth in condition B will be absent in the population after the growth. (C) DNA from both selected populations and the initial transposon library is extracted, and usually fragmented by enzyme or sheared (depending on the version of transposon insertion sequencing (Kwon *et al.*, 2016)). Sequencing adaptors (pink and yellow rectangles) are then added to allow high-throughput sequencing. More details can be found in Figure S 4.1. (D) The last step consists of mapping the reads to the genome of interest and analyzing the data. The vertical lines denote the number of reads inserted at each insertion site. The absence of reads for a specific gene indicates that no mutant with this specific transposon insertion was present in the library and so the essentiality of this gene under a given condition. Specifically, gene A is non-essential for growth in all conditions, as well as for the initial conditions of the preparation of the transposon library. Gene B is essential for growth in condition B, as evidenced by the absence of reads, i. e. mutants bearing transposon inserted in this gene, in that condition. Gene C is essential for growth in general, as it is already absent from the initial transposon library. The relative change in read count for gene B between the initial library and the library grown in condition A indicates that transposon mutants of gene B exhibit reduced fitness in condition A.

In this part of the project, we aimed to identify new genes involved in the natural transformation of *P. stutzeri* and to investigate the role of cell-cell interactions with the help of TIS. For that purpose, we first prepared a transposon mutants library of recipient cells. The recipient library was transformed either with extracellular DNA (eDNA) or in the presence of donor cells bearing an antibiotic resistance marker (Figure 4.2). To avoid cross-transformation, the donor cells were non-transformable due to the absence of the membrane-channel protein

ComA. Then, the transformants—the cells that have acquired the resistance—and the entire population of recipients were collected so that the DNA could be extracted and prepared for transposon sequencing (Figure S 4.1). Comparison of the relative abundance of each mutant between the transformant pool and the whole population should have allowed us to identify genes involved in natural transformation such as the ones coding for the DNA uptake, translocation, and recombination (Chapter 1, table 1). By comparing the two pools of transformants—those resulting from donor cell transformation and those from eDNA transformation—we should have been able to identify genes required for example for predation, such as the type VI secretion system in *V. cholerae* (Matthey *et al.*, 2019). In scenarios where DNA release from donor cells occurs through mechanisms other than recipient-induced lysis, such as autolysis or active secretion (e.g., via the type IV secretion system like in *N. gonorrhoeae*), an analogous experimental design utilizing a transposon mutagenesis library in donor strain could elucidate the genetic determinants of DNA donation. This approach would involve tracking the transfer of transposon-tagged DNA sequences, enabling the identification of transposon insertions that disrupt the DNA donation process. However, this methodology necessitates a substantial number of transformants to achieve comprehensive genomic coverage and statistical significance.

The initial TIS methodology used for sequencing library preparation proved suboptimal for *P. stutzeri*. Adaptations in the protocol allowed to obtain an initial transposon library with about 40% transposon density but the libraries resulting from the transformation experiments were less dense, preventing a proper identification of genes essential for natural transformation. However, pooling of all sequencing libraries allowed us to identify genes belonging to the core genome of *P. stutzeri*. The subsequent results section presents our methodology, including optimization attempts, presents the obtained data, and offers potential interpretations within the constraints of the experimental limitations encountered.

A. Transformation of recipient library with DNA



B. Transformation of recipient library with donor cells

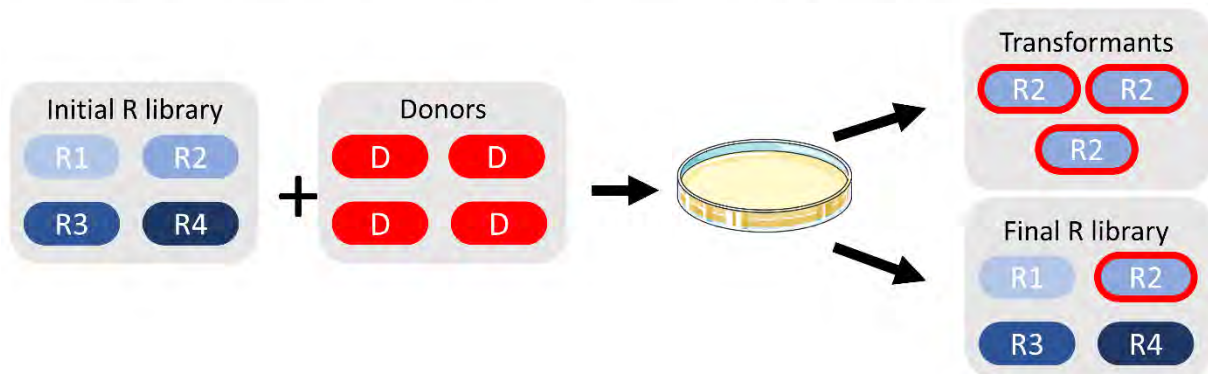


Figure 4.2 Schematic representation of the transformation assays with the recipient TIS library (blue, mutants R1-R4) mixed with (A) extracellular DNA or (B) donor cells (red). After transformation, cultures were plated on non-selective media (all recipients) and selective media (transformants, red outline). Here, comparisons of transformant versus recipient sequencing libraries reveal genes essential for transformation: R3/R4 failed to transform, indicating essential transformation genes. R1 transformed efficiently with DNA but not donors, while only R2 transformed with donors, revealing donor-specific requirements. Mutants absent from the initial library would reflect essential growth/survival genes in *P. stutzeri*.

4.4 Materials and Methods

4.4.1 Bacterial strains and medium

The bacterial strains, plasmids, and oligonucleotides used in this chapter are listed in Table S 4.1. Strains of *P. stutzeri* and *Escherichia coli* were grown overnight in LB medium at 30°C (with 180 rpm agitation) or on LB agar plates. When required, antibiotics were added at the following concentration: ampicillin (Amp), 100 µg/mL; gentamicin (Gm), 15 µg/mL for *E. coli* and 30 µg/mL for *P. stutzeri*; kanamycin (Km), 50 µg/mL; tetracycline (Tet), 12 µg/mL; rifampicin (Rif), 50 µg/mL.

4.4.2 Strains and plasmid construction

Two types of mutants were used for this project: $\Delta comA$ deletion mutant, and a wild-type and $\Delta comA$ strains bearing a tetracycline and kanamycin resistance cassette inserted via a Tn7 vector system. Construction of the *P. stutzeri* deletion mutant $\Delta comA$ is presented in chapter 2.

P. stutzeri tetR-kmR was constructed by insertion of a single copy of a cassette of tetracycline and kanamycin resistance via a Tn7 vector system into the genome of wild-type or $\Delta comA$ strains. For that purpose, the *tetR* cassette was amplified by PCR with the primers listed in Table S 4.1. Then, the PCR fragments were digested with NotI restriction enzymes and ligated to the pMP7607 plasmid, which was also digested with NotI. The ligation mix was then inserted in *E. coli* DH5 α λ *pir* by heat-shock followed by selection of complete plasmid and further propagation of the plasmid. The transposon delivery plasmids pMP7607-tetR-kmR and the helper plasmid pUX-BF13 were then introduced by electroporation into *P. stutzeri*. The presence of the transposon insertion was assessed by PCR.

4.4.3 Isolation and purification of DNA

In the transformation assays, DNA extraction was done on cells from 500 μ L of overnight culture harvested by centrifugation at 13'000 rpm for 3 min and lysed in 500 μ L lysis buffer (10 mM Tris, 100 mM EDTA, 2% SDS, pH 7) supplemented with proteinase K (80 μ g/mL) at 55°C. Then, the RNA was degraded with the help of RNase (40 μ g/mL) for 30 min incubation at 37°C. The proteins were precipitated by the addition of 1 volume of 7.5 M ammonium acetate. A centrifugation step allowed to separate the precipitated proteins from the supernatant. Sodium acetate at 3 M was added to the supernatant in a ratio of 1:10. Then, the DNA was precipitated with the addition of 2.5 volumes of pure ethanol. A centrifugation step allowed to gather the precipitated DNA at the bottom of the tube. Finally, the DNA was washed twice with 70% ethanol before being resuspended in nuclease-free water. The concentrations of DNA were estimated with a Qubit® 2.0 fluorometer (Life Technologies, USA) according to the manufacturer's protocol.

4.4.4 Natural transformation assays

P. stutzeri tetR-kmR recipient and *rifR* $\Delta comA$ (donor) strains were grown overnight in LB medium at 30°C (with 180 rpm agitation). After centrifugation for 1 min at 13'000 rpm, the pellet was washed twice with saline (0.9% w/v NaCl) and resuspended in saline at optical

density measured at 600nm (OD₆₀₀) of 1. In a tube, 80 µL of recipient bacteria suspension were mixed either with 20 µL H₂O (control treatment), with 20 µL of rif^R DNA (stock DNA concentration: 100, 10, 1, 0,1 ng/µL), or with 80 µL of rif^R- or km^R-resistant donor bacteria suspension. In the latter case, the final bacterial suspension had a total volume of 100 µL thanks to a concentrating step of recipient's suspension in 20 µL before mixing. Regarding the assay performed with DNase treatment (Figure 4.3), 10 units of RQ1 DNase (Promega) were added to the tube containing 100 µL of donor and recipient bacteria. Then, three drops of 25 µL of the suspension were spotted on a LB agar plate. After the spots had dried, inoculated agar plates were incubated at 30°C for 24 hours. Then, each bacterial spot was collected with an inoculation loop and resuspended in 1 mL saline. The tubes were vigorously vortexed for 15 s to resuspend the bacteria. Samples were diluted and plated. In mix with only recipient bacteria, the total number of viable cells was estimated on LB plates and the number of transformants was determined on LB-Rif plates. In mix of recipient and donor bacteria, the total number of viable recipient cells was estimated on LB-Km plates and the number of transformants was determined on LB plates containing rifampicin and kanamycin. All cell numbers are expressed in colony-forming units (CFU) per ml. The antibiotics concentrations were 25 µg/mL for kanamycin and 50 µg/mL for rifampicin in plates. Transformation frequencies are given as number of rif^R colonies per recipient colonies.

The natural transformation assay for the transposon insertion sequencing follows a similar protocol than described above except that the recipient cells was a library of transposon mutants bearing Gm^R and the donor cells were Tet^R-Km^R due to the insertion of resistances cassette by a Tn7 vector system (see sections "Transposon mutagenesis" and "Strains and plasmid construction"). Hence, the transfer of the kanamycin resistance was monitored instead of rifampicin-resistance. This change was made to avoid collecting spontaneous resistant mutants given that the spontaneous mutation rate to kanamycin was much lower than to rifampicin. Moreover, to limit bacterial multiplication, the spots were incubated for 8 hours instead of 24 hours. To obtain enough transformants, transformation assay was performed with much high volumes.

4.4.5 Transposon mutagenesis

Transposon mutagenesis was carried out by conjugation of the *P. stutzeri* receptor strain with donor *E. coli* JKE201 harboring pBT20, which encodes the mariner Himar1 transposon,

followed by selection for *P. stutzeri* cells harboring the Transposon-containing gentamicin-resistance on selective agar plates. *E. coli* JKE201 has a diaminopimelic acid (DAP) auxotrophic phenotype, which facilitates post-conjugational counterselection. In brief, donor *E. coli* JKE201 [pBT20] was cultivated overnight at 37°C in liquid LB with 1 mM DAP and 100 µg/mL ampicillin and the recipient *P. stutzeri* at 30°C in liquid LB. For conjugation, OD₆₀₀ of overnight cultures of the donor strain and recipient strain were measured with the spectrophotometer Genesys 10S UV-VIS (Thermo Scientific) and adjusted to an OD₆₀₀ of 0.5. Then, bacteria contained in 5ml were washed one time for *P. stutzeri* and three-time for *E. coli* before being pooled together in 50ul saline (NaCl 0.9% w/v) and spotted on dry LB 1 mM DAP plate. The plate was incubated for 3 hours at 30°C. The mating spot was scraped and resuspended in 4 mL saline. Then, 10 µL of the bacteria suspension were diluted and plated to estimate the number of conjugants, the rest was plated on LB-Gm agar with 100 µL suspension per plate. All the plates were incubated 24 hours at 30°C. The colonies were collected by flooding each plate with 2 mL LB, scratching them with an inoculation loop, and collecting the bacterial suspension. All the bacteria suspensions were pooled. The OD₆₀₀ was adjusted to 2 with glycerol at a final concentration of 20% before aliquoting and freezing at -80°C. The number of conjugants was estimated by CFU counting on the plates containing only 10 µL of the conjugation mix and CFU estimation based on pictures of the plates containing 100 µL of the conjugation mix. If we estimate that one colony corresponds to one mutant, the library should contain about 200'000 mutants (only non-essential genes).

4.4.6 Preparation of sequencing libraries

The preparation of the sequencing libraries followed mainly the protocol Hardy *et al.* (Hardy *et al.*, 2021) used to identify essential genes and determinants of natural transformation in *Legionella pneumophila*. The main steps are described in Figure S 4.1.

The first protocol we performed for the initial library was the following. Pellets of mutant libraries from the -80°C were thawed. Genomic DNA (gDNA) extraction was carried out directly on the pellet cells with the QIAamp DNA Kit (Qiagen) according to the manufacturer's instructions. Approximately 10 mg of DNA were mechanically sheared by sonication using a Branson sonifier for 5 min (10 s. on and 50 s. off; 20% intensity) in 1.5-ml tubes kept on ice water. Small gDNA fragments were removed by mixing sonicated gDNA with 1.2x Agencourt Ampure XL magnetic beads (Beckman Coulter) according to the manufacturer's instructions.

This treatment selected gDNA fragments being between 250 and 750 bp. Homopolymeric cytosine tails (C-tails) were then added to the 3' ends of all fragments by incubation of 3 mg of size-selected DNA fragments with the recombinant terminal deoxynucleotidyl transferase (rTdT; 30 U/mL; Promega) at 37°C for 1 h, followed by heat inactivation at 75°C for 20 min. The rTdT reagents were then removed by purifying the rTdT reaction mixture firstly with 0.5x of Ampure XL beads. The beads containing high-molecular-weight fragments were discarded, and the supernatant was mixed again with 0.5x of Ampure XL beads. To amplify transposon junctions, a first PCR (PCR1) was performed in a final volume of 50 µL by mixing 500 ng C-tailed DNA, 1 µL biotinylated pBT20-PCR1 primer (30 µM), 3 µL olj376 primer (30 µM), 2.5 µL deoxynucleoside triphosphates (dNTPs) (10 mM), 10 µL Q5 reaction buffer, and 0.75 µL Q5 high-fidelity DNA polymerase (New England Biolabs). The program of PCR1 was the following: 5 min at 98°C, 20x (30 s at 98°C, 30 s at 65°C, 30 s at 72°C), 5 min at 72°C. PCR1 products were purified using 1.2x Ampure beads. Biotinylated and purified PCR1 products were then selectively captured using streptavidin-containing Dynabeads (M-280; Invitrogen) according to the manufacturer's instructions. A second PCR was performed in a final volume of 50 µL by resuspending Dynabeads (which had the PCR1 products bound to them) in the preprepared PCR2 reaction mix consisting of 1 µL pBT20-PCR2_mod primer (30mM), 1 µL TdT_index_X primer (30mM), 2.5 µL dNTPs (10 mM), 10 µL Q5 reaction buffer, and 0.75µL Q5 high-fidelity DNA polymerase. The program of PCR2 was the following: 5 min at 98°C, 20x (30 s at 98°C, 30 s at 55°C, 30 s at 72°C), 5 min at 72°C. PCR2 products were purified with 1.2 Ampure XL beads. The obtained library was sent to Novogen (England) to be sequenced on an Illumina Novaseq with 150 bp pair-end reads using the standard sequencing primer.

Due to unsaturated library with the first protocol, we did some modifications. To increase the number of fragments with transposon, an additional linear PCR (linPCR) was performed in triplicates after DNA shearing step (Dawoud *et al.*, 2014). Each tube contained a PCR mix having a final volume of 50 µL by mixing 500 ng sheared DNA, 1 µL biotinylated pBT20-PCR1 primer (30 µM), 1 µL deoxynucleoside triphosphates (dNTPs) (10 mM), 10 µL Colorless GoTaq Reaction Buffer, and 0.5 µL GoTaq DNA Polymerase (Promega). The program of linPCR was the following: 2 min at 95°C, 50x (30 s at 95°C, 45 s at 60°C, 10 s at 72°C), 5 min at 72°C. Then, the PCR products were pooled and purified with 1.2 Ampure XL beads before C-tailing as in the regular protocol. The PCR2 was also modified relative to the initial protocol by reducing the

number of cycles from 20 to 15 and reducing the volume of primers from 1 μL to 0.66 μL because we regularly encountered issue of primers dimers formation in our samples.

4.4.7 qPCR

The quantitative PCR was performed with the Rotor-Gene SYBR Green PCR kit (Qiagen) in Rotor-Gene Q real-time PCR cycler (Qiagen). For each sample, we mixed 10 μL Rotor-Gene SYBR Green PCR Master Mix, 1 μL primer A (4 μM), 1 μL primer B (4 μM), 3 μL H_2O , 5 μL DNA (conc. 0.5 or 5 $\text{ng}/\mu\text{L}$). Two pairs of primers were designed (Table S 4.1). The *plagen* pair amplified a fragment in the region with reads enrichment. The *gen* pair amplified a specific region of the chromosome where the read coverage of this region is representative of the mean read coverage (Figure S 4.2).

4.4.8 Sequencing data analysis

For each condition, 58 to 85 million reads were obtained (see Table 4.2). The reads were trimmed and mapped on the genome using a series of command-line tools. Firstly, *cutadapt* (Martin, 2011) removed reads without a match to transposon and reads smaller than 50 bp. Then, *fastp* (Chen *et al.*, 2018) removed the adaptors and trimmed the 3'end to 72 bp. Trimmed reads were mapped to the genome of *P. stutzeri* with the help of *bowtie2* (Langmead and Salzberg, 2012). Samtools (Li *et al.*, 2009) was used to gather the following information in a file: contig name, transposon insertion position, insertion code, and read length. Three files were prepared: 1. the count of all reads (including reads mapping multiple times on the genome), 2. the count of all reads mapping once on the genome, 3. the count of single insertion sites (from reads mapping only once on the genome). Output wig files were analyzed using the TRANSIT pipeline (DeJesus *et al.*, 2015). Insertion counts at each TA site were normalized using TTR and averaged across libraries. Single-condition essentiality analysis was performed with the HMM, Gumbel and Tn5gap methods (DeJesus and Ioerger, 2013; DeJesus *et al.*, 2013). Conditional essentiality analysis was performed with the “resampling” method according to the Transit software documentation. Summary of analysis can be found in Table S 4.2. The distribution of TA sites on the genome of *P. stutzeri* ATCC 17587 (unpublished) was realized with the help of the *RStudio* package *Biostrings* (Pagès *et al.*, 2022).

4.4.9 Graphics and data analysis

The statistical tests were performed by using the *RStudio* software (Core R Team, 2019). Graphics were designed with *ggplot2* and base R functions (Wickham, 2008). Regarding

statistical analyses, the normality of the data distribution was tested with Shapiro-Wilk normality test, the equality of variances with F-test, or Levene's test. Finally, the means were compared either with the Student's t-test/Wilcoxon rank test or ANOVA with post-hoc Tukey or Dunnett tests. The results of the statistical tests are indicated in the figures accordingly: *p-value <0.05; **p-value <0.01; ***p-value <0.001.

4.5 Results and Discussion

4.5.1 *P. stutzeri* ATCC 17587 can be transformed efficiently with donor cells.

P. stutzeri was shown to perform natural transformation efficiently with donor cells about four decades ago (Stewart *et al.*, 1983). Given that we were working with a different strain than the one used in the original study, we conducted transformation assays with donor cells to confirm that the *P. stutzeri* ATCC 17587 strain is also capable of transformation with donors. The inhibition of natural transformation with donor cells by DNase confirmed that transferred DNA was free and accessible to DNase (Figure 4.3).

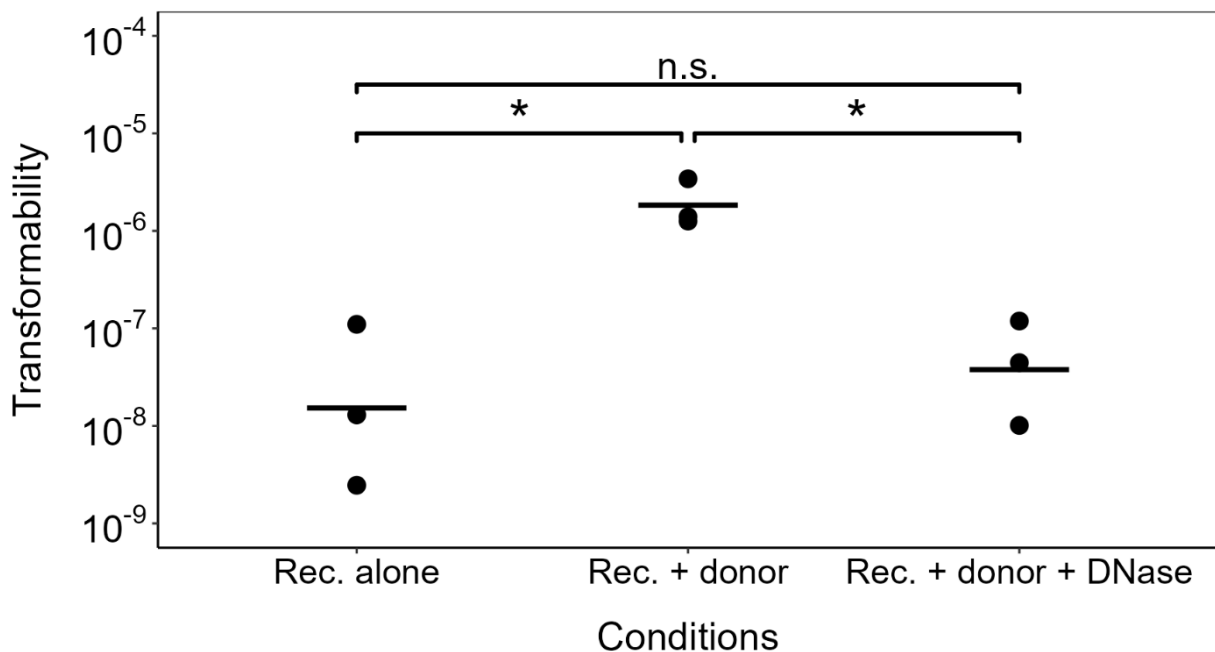


Figure 4.3. Transformation of recipient with donor cells is inhibited in presence of DNase. The recipient cells of *P. stutzeri* ATCC 17587 *tetR kmR* (Rec.) was incubated with water (alone) as a control treatment, with donor bacteria from the strain $\Delta comA$ in absence or presence of DNase. Each condition had three replicates indicated with dots, the mean is shown with the black line.

Our study aimed to employ TIS method concurrently on transformation assays with both extracellular DNA and donor cells to elucidate genes specific to each transformation modality. To establish comparable transformation conditions, we exposed recipient cells to equivalent

ratios of donors or varying concentrations of exogenous DNA. The transformation efficiency increased proportionally with DNA concentration until reaching a saturation plateau (Figure 4.4). Maximal transformation occurred at approximately 2 ng/ μ L DNA, consistent with previously reported saturating concentrations ranging from 1 ng/ μ L (Carlson *et al.*, 1983) to 3 ng/ μ L (Berndt *et al.*, 2003). Notably, recipient cells demonstrated comparable transformability when exposed to saturating DNA concentrations or donor cells. This observation suggests either substantial DNA release by donor cells or the presence of factors enhancing DNA transfer efficiency.

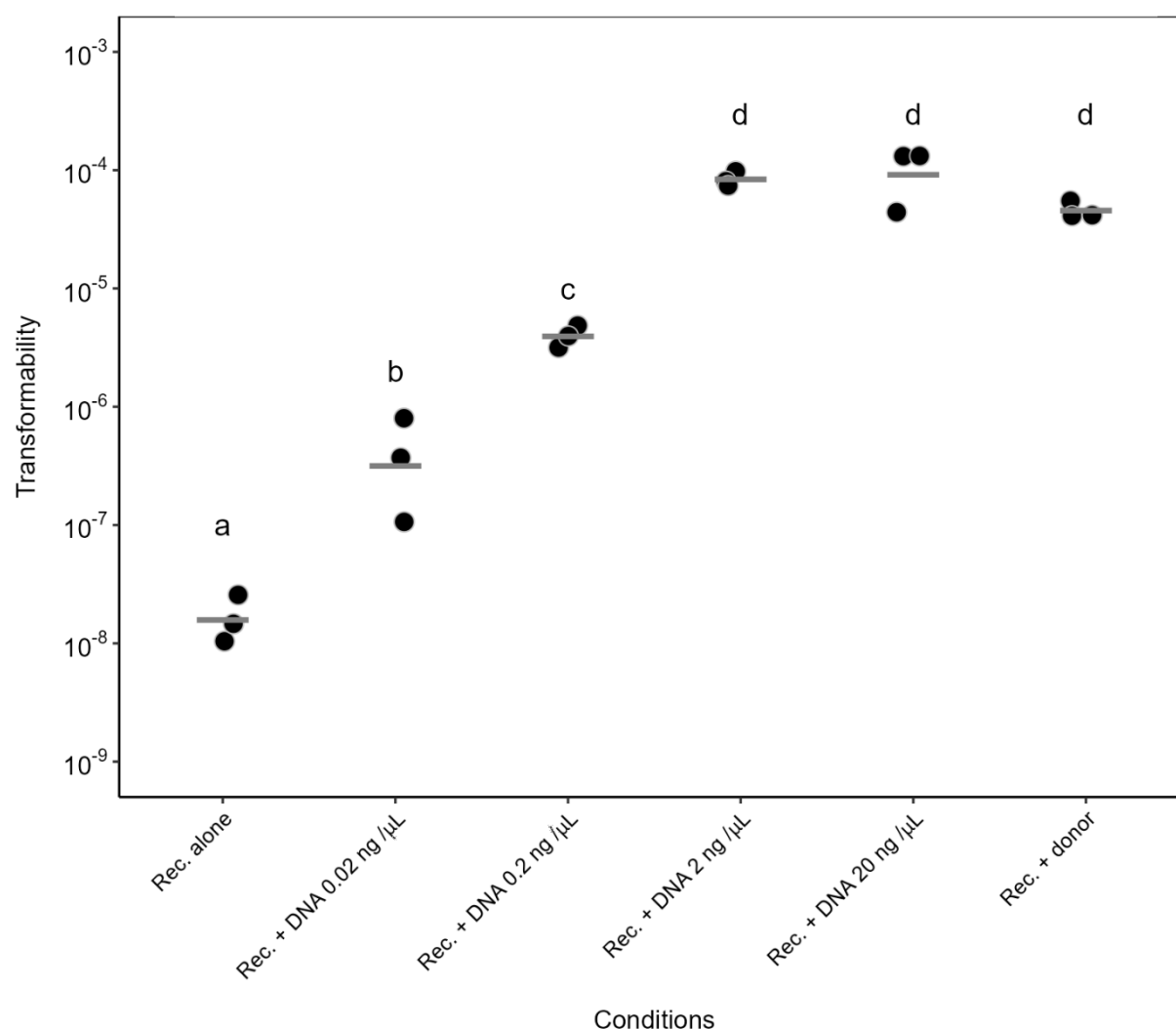


Figure 4.4. Transformability of recipient with donor cells is similar to transformability with saturating concentration of extracellular DNA. The recipient cells of *P. stutzeri* ATCC 17587 *tetR kmR* (Rec.) was incubated with water (alone) as a control treatment, with rifampicin-resistant DNA at various concentration (0.02-20 ng/ μ L), or with donor bacteria of the strain $\Delta comA$. Each condition had three replicates indicated with dots, the mean is shown with the black line. Different letters indicate that the conditions significantly differ based on Tukey Honest Significant Difference (HSD) post hoc test.

4.5.2 Evaluation of the recipient transposon library regarding its mutant richness

The first step of TIS methodology involves generating a highly saturated transposon library. High saturation yields more data per gene, thereby enhancing the statistical power to detect significant differences in insertion frequencies between conditions or to accurately identify essential genes (Chao *et al.*, 2016). To achieve higher saturation, mariner-based transposons are frequently employed due to their specific insertion at TA sites (Lampe *et al.*, 1996). As example, Hardy and colleagues developed a method using mariner-based transposons that improved saturation from 7% to 43% in *L. pneumophila* when studying genes required for natural transformation (Hardy *et al.*, 2021).

We adapted this protocol for our library preparation, initiating with the construction of a transposon mutant library via conjugation of the *P. stutzeri* recipient strain with donor *E. coli* JKE201 harboring pBT20, a plasmid containing the mariner Himar1 transposon (Kulasekara *et al.*, 2005). Approximately 200'000 mutant colonies were recovered.

The sequencing of the transposon insertion sites revealed a maximum of 12'286 unique insertions out of about 200'000 possible insertions, indicating a 6% saturation. This unexpectedly low saturation, compared to the 43% achieved by Hardy and colleagues (Hardy *et al.*, 2021), prompted further investigation. Analysis of the 39 million reads post-trimming and sorting revealed that 5% of reads did not map, 55% mapped to multiple positions and were consequently discarded, and 40% mapped to unique positions.

When reads matching multiple regions were retained during the trimming process and mapped against the *P. stutzeri* genome, we observed distinct regions with higher read densities, particularly evident in a peak on the chromosome and on the second plasmid (Figure 4.5). Both regions contained genes belonging to the *mer* gene cluster, which provide mercuric resistance to *P. stutzeri* (Zheng *et al.*, 2018). The unequal distribution of reads between the regions dramatically reduced the global sequencing depth and indicated a strong bias in the representation of different insertions.

Given the homogeneous distribution of TA sites along the genome (Figure S 4.3), we hypothesized that the observed bias could originate from either amplification bias during library preparation (PCRs) or differential copy numbers between plasmids and the chromosome.

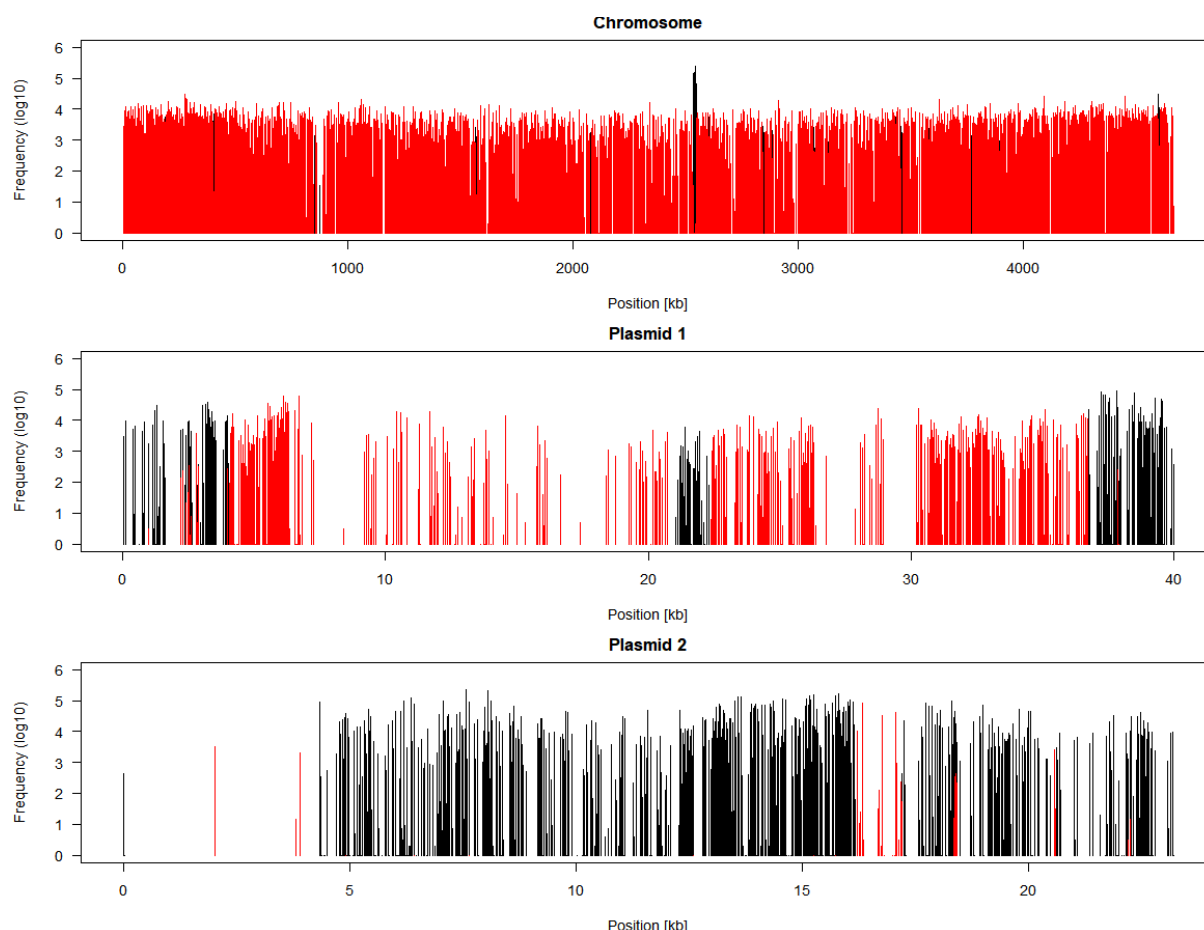


Figure 4.5 Log10 of all trimmed transposon insertions read counts of initial recipient library in the chromosome (top) and both plasmids (middle and bottom). The reads matching to one position are represented in red, and the reads matching to multiple positions are in black.

To examine the hypothesis of preferential enrichment during sequencing library preparation, we employed quantitative PCR (qPCR) analysis on the library both pre- and post-processing for sequencing. Two primer pairs were utilized: one specific to the putatively enriched region and another annealing outside this region, serving as a representative of the average genomic content (Figure S 4.2). The cycle threshold (Ct) value inversely correlates with the copy number of the targeted region. Prior to sequencing library preparation, Ct values for both regions were in the same range, indicating no enrichment of a specific region (Table 4.1). However, after preparation, Ct values differed by up to 2^{20} -fold between targeted regions. This substantial divergence suggests disproportionate amplification of the enriched region during the sequencing library preparation process, potentially due to PCR bias or other amplification-related factors.

Table 4.1 Mean Ct values ($\pm$ standard deviation) obtained by qPCR with two primer pairs on two replicates. The difference in Ct values was calculated with the means. The average region refers to the region having a read coverage within the average (Figure S 4.2). The enriched region refers to the region having a particularly high read coverage. Ct: Cycle threshold.

Samples	Ct of the average region	Ct of the enriched region	Δ Ct
Library before preparation	15.83 $\pm$ 0.32	14.33 $\pm$ 0.08	1.5
Library after preparation	27.60 $\pm$ 0.75	7.36 $\pm$ 0.14	20.24

To elucidate the origin of the amplification bias, we systematically modified critical parameters in the initial sequencing library preparation protocol and quantitatively assessed the outcomes via qPCR (Table S 4.3). The library preparation workflow includes several sequential steps: genomic DNA extraction, DNA fragmentation, PCR amplification of transposon-genome junctions, selective capture of the fragments of interest, PCR amplification for Illumina adaptor incorporation, and high-throughput sequencing (Figure S 4.1). Our analyses revealed that the bias predominantly manifested during the second PCR amplification step. While reducing the number of thermal cycles attenuated this bias, it concomitantly resulted in a substantial decrease in DNA yield, rendering the final concentration insufficient for downstream applications. Another cause of bias that we explored was the effect of dNTP imbalance during the second PCR. Given the GC-rich genome of *P. stutzeri*, extensive PCR amplification could lead to preferential depletion of G/C nucleotides relative to A/T nucleotides. Such nucleotide limitation may introduce amplification bias, favouring sequences with lower requirements for the depleted nucleotides, thereby distorting the representation of diverse sequences in the final library (Benjamini and Speed, 2012).

The optimized protocol for sequencing library preparation diverged from the initial methodology in several key aspects: addition of a linear PCR before C-tailing to increase the proportion of fragments with transposon, decreasing the numbers of cycles and lowering concentration of primers of the second PCR. Our optimization efforts successfully reduced the disparity between targeted regions from Δ Ct of 20.24 to 8.1 for the initial library (Table 4.1, Table S 4.3). As further improvements to the protocol proved to be challenging, we opted to increase the raw data amount per sample during sequencing to enhance overall sequencing depth and mitigate the remaining bias.

4.5.3 Final results analysis

This study aimed to utilize TIS to identify genetic determinants of natural transformation in *P. stutzeri*. The primary objectives were to determine genes essential for natural transformation using both purified DNA and donor cells, and, in a second step, to investigate if specific genes were required for DNA acquisition from donors through for example predation (Figure 4.2). Additionally, the study sought to elucidate genes essential for growth and survival, providing a comprehensive understanding of the genetic factors influencing natural transformation and overall fitness of *P. stutzeri*. The initial sequencing of the transposon mutant library revealed inadequate recovery of mutant diversity. Consequently, the protocol for preparing sequencing libraries was optimized. Five new libraries were then sequenced, representing different stages of the transformation experiment: the initial library, recipient pools from transformations with DNA and donor cells, and transformants obtained from DNA and donor cell transformations. The analysis and results from these optimized libraries are presented below and summarized in Table S 4.2.

The adaptation of the sequencing library protocol and increased total sequencing reads led to the recovery of more unique insertions for the initial library with a saturation of 40% (Table 4.2). Nevertheless, the saturation of the other libraries was quite low, indicating a considerable loss in mutant diversity during the natural transformation assay, which happened probably during the sampling, although the number of collected colonies was considerable.

Table 4.2 Key numbers regarding the different sequencing libraries. *During the trimming process, reads without adapters, too short, matching to none or multiple sites were removed.

	Initial library 1st protocol	Initial library 2nd protocol	Transf. with DNA		Transf. with donors	
			All cells	Transf.	All cells	Transf.
Total reads	57'846'729	64'469'577	70'975'559	63'154'702	58'822'132	84'465'208
Total reads after trimming*	15'466'375	9'105'060	9'824'131	8'554'713	13'203'387	14'093'755
Total nr. insertion	12'286	81'854	16'058	15'354	16'154	12'605
% TA site saturation	6.0	40.3	7.9	7.6	7.9	6.2

4.5.3.1 Essential genes investigation

Despite the saturation of the libraries after transformation assay was much lower than the initial library (Table 4.2), we evaluated the gene essentiality with the TRANSIT software (Ioerger, 2022). This software performs single-condition essentiality analysis with mainly two methods: the Gumbel method and the Hidden Markov Model (HMM) method (DeJesus and

Ioerger, 2013; DeJesus *et al.*, 2013). The Gumbel method focuses on genes with statistically significant gaps, or consecutive sequences of TA sites lacking insertions. In case of sub-saturated libraries, empty sites will appear randomly in non-essential regions, but long stretches of empty sites are statistically improbable, and genes containing such gaps are classified as essential. This method is not really powerful with very low-saturated libraries as it will classify many genes as “uncertain”, especially for genes containing less than 10 TA sites (Ioerger, 2022). The HMM method works on a different principle, using a probabilistic model to assess the state at each TA based on the counts and consistency with neighbouring sites (Ioerger, 2022). This allows the HMM to smooth individual outlier values. The key distinction from Gumbel analysis is that the HMM considers insertion count magnitudes, which can provide insights into a gene fitness effect. HMM and Gumbel methods have been developed for Mariner transposons which have the specificity to insert only in TA sites (Lampe *et al.*, 1996; Ioerger, 2022). Another method available in the TRANSIT software is the Tn5Gaps method. It is based on a Gumbel method but adapted to Tn5 transposons, which can insert everywhere in the genome, thus creating less-saturated libraries. Despite the recipient mutant library was generated using a Mariner transposon, we applied this method to our sequencing libraries as well, due to the low saturation levels of our libraries.

To identify essential genes, which are required for growth, we used the three statistical methods (Gumbel, HMM and Tn5Gaps) on all our sequencing libraries pooled together (Table S 4.2). Out of the 4368 genes predicted in the genome of *P. stutzeri*, the Gumbel method classified 500 genes as essential, the HMM method 4324 genes, the Tn5Gap only 109 genes (Figure 4.6, Table S 4.5). The overestimation of essential genes by the HMM is probably due to the skewed read count distribution (DeJesus and Ioerger, 2013), as indicated by Transit's quality control analysis of the sequencing libraries (Table S 4.6). Typically, between 5 to 15% of genes are essential for growth *in vitro* in bacterial genomes (Luo *et al.*, 2014; Turner *et al.*, 2015; Ioerger, 2022). The Gumbel method and its derivative approach offer a credible means of classifying gene essentiality. This observation is further reinforced by the fact numerous genes coding for example for 30S and 50S ribosomal proteins and peptidoglycan biosynthesis were identified as essential by the Gumbel method, yet classified as non-essential by the Tn5Gap method, despite being part of the core essential genome of *Pseudomonas aeruginosa* (Table S 4.7) (Poulsen *et al.*, 2019). When examining the genes considered essential by all

three methods, there was limited consensus in their results, with only 10 genes receiving consistent classification across both approaches (**Figure 4.6**). Most of the identified genes have been shown to be essential by means of transposon mutagenesis in other research. Three open reading frame (ORFs) particularly caught our attention as they belong to different classes of mobile genetic elements. The LOCUS_00110 and LOCUS_001210 blasted to IS5 family transposase and were also classified as essential in *E. coli*, *Burkholderia pseudomallei*, and *Ralstonia solanacearum* (Gerdes *et al.*, 2003; Moule *et al.*, 2014; Su *et al.*, 2020) in contrast to LOCUS_01220 for which no significant match was found. Further investigations would be required to confirm the essentiality of these unexpected results.

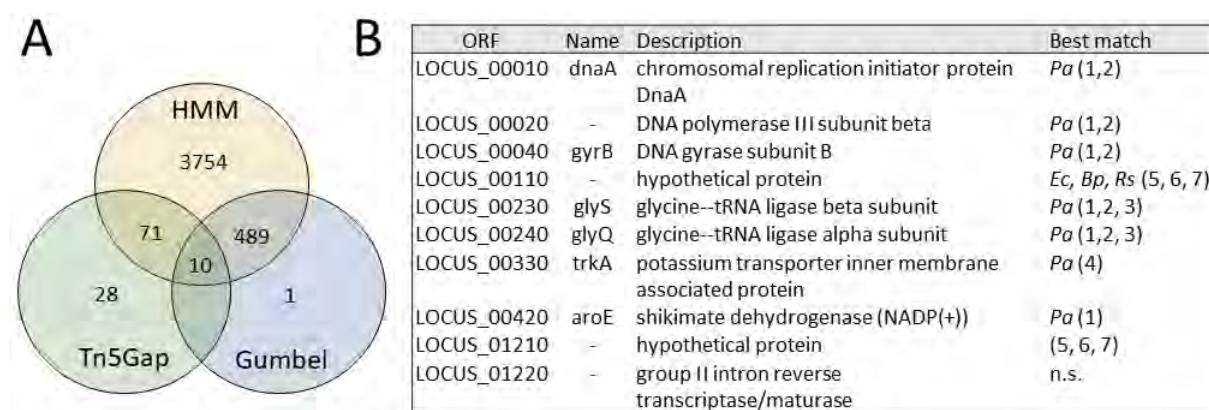


Figure 4.6 (A) Venn diagram showing the overlap of essential genes predicted by three different statistical methods (Gumbel, HMM, and Tn5Gap). Numbers represent the count of genes identified as essential by each method or combination of methods. (B) The ten genes consistently identified as essential across all three statistical methods (Gumbel, HMM, and Tn5gap) in our TIS analysis. The amino acids sequences of these genes were blasted in database of essential genes (Luo *et al.*, 2014) and best matches are shown. *Pa*: *Pseudomonas aeruginosa*; *Ec*: *Escherichia coli*; *Bp*: *Burkholderia pseudomallei*; *Rs*: *Ralstonia solanacearum*; n.s.: no significant match. In brackets, references: 1 (Lee *et al.*, 2015), 2 (Turner *et al.*, 2015), 3 (Liberati *et al.*, 2006), 4 (Gallagher *et al.*, 2011), 5 (Gerdes *et al.*, 2003), 6 (Moule *et al.*, 2014), 7 (Su *et al.*, 2020).

4.5.3.2 Essential genes for transformation

To identify genes specific to natural transformation, we compared mutant pools across different conditions. This approach distinguishes between genes essential for transformation and those generally required for bacterial survival. Mutants defective in competence, DNA uptake, processing, or recombination should be absent from the transformed population but present in the general recipient pool. We performed conditional essentiality analysis using the “resampling” method in Transit software (Table S 4.2) (Ioerger, 2022). Comparing transformant pools with DNA or donor cells against potential recipients provided inconclusive results (Table S 4.8, Table S 4.9). The identified loci were mainly clustered together and some already considered essential for growth by the Gumbel method. More importantly, known

essential genes for natural transformation in *P. stutzeri* were not detected (Vosman *et al.*, 1991; Graupner *et al.*, 2000, 2001; Graupner and Wackernagel, 2001a), further indicating unreliable results. Pooling transformant and recipient libraries did not improve outcomes (Table S 4.2). The data quality and loss of mutant richness during transformation assays may have prevented definitive identification of loci essential for natural transformation (Table S 4.6).

4.6 Conclusion

Adapting a TIS method from a *L. pneumophila* protocol proved to be more challenging than anticipated. Despite numerous tests, the final results did not meet our expectations: identifying genes essential for the recipients in natural transformation assay with DNA or donor cells. First, the challenge concerned the sequencing libraries preparation which suffered from a bias in reads distribution, considerably reducing our ability to evaluate mutant richness in the initial recipient mutant library. Our investigations revealed that the problematic step in sequencing preparation was primarily the second PCR.

To mitigate potential PCR bias, alternative sequencing library preparation methods should be explored, including those that employ uniform amplicon lengths or utilize PCR-free techniques (Kwon *et al.*, 2016; Alkam *et al.*, 2021). Then, the low saturation of the libraries after transformation assay might have reduced our ability to identify correctly loci involved in natural transformation. This might have been corrected by collecting more colonies after transformation assays but also having replicates for the transformation assay and sequencing libraries. Nevertheless, this study successfully identified genes essential for *P. stutzeri* growth, contributing to the general identification of core genome genes of *Pseudomonas*.

4.7 References

- Alkam, D., Wongsurawat, T., Nookaew, I., Richardson, A.R., Ussery, D., Smeltzer, M.S., and Jenjaroenpun, P. (2021) Is amplification bias consequential in transposon sequencing (Tnseq) assays? a case study with a staphylococcus aureus tnseq library subjected to pcr-based and amplification-free enrichment methods. *Microb Genomics* **7**: 1–13.
- Ambur, O.H., Engelstädter, J., Johnsen, P.J., Miller, E.L., and Rozen, D.E. (2016) Steady at the wheel: Conservative sex and the benefits of bacterial transformation. *Philos Trans R Soc B Biol Sci* **371**: 1–11.
- Bao, Y., Lies, D.P., Fu, H., and Roberts, G.P. (1991) An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of Gram-negative bacteria. *Gene* **109**: 167–168.
- Benjamini, Y. and Speed, T.P. (2012) Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids Res* **40**: 1–14.
- Berndt, C., Meier, P., and Wackernagel, W. (2003) DNA restriction is a barrier to natural transformation in *Pseudomonas stutzeri* JM300. *Microbiology* **149**: 895–901.
- Borgeaud, S., Metzger, L.C., Scignari, T., and Blokesch, M. (2015) The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. *Science (80-)* **347**: 63–67.
- Cain, A.K., Barquist, L., Goodman, A.L., Paulsen, I.T., Parkhill, J., and van Opijnen, T. (2020) A decade of advances in transposon-insertion sequencing. *Nat Rev Genet* **21**: 526–540.
- Carlson, C.A., Pierson, L.S., Rosen, J.J., and Ingraham, J.L. (1983) *Pseudomonas stutzeri* and related species undergo natural transformation. *J Bacteriol* **153**: 93–99.
- Chao, M.C., Abel, S., Davis, B.M., and Waldor, M.K. (2016) The Design and Analysis of Transposon-Insertion Sequencing Experiments. *Nat Rev Microbiol* **14**: 119–128.
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018) Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: i884–i890.
- Core R Team (2019) A Language and Environment for Statistical Computing. *R Found Stat Comput* **2**: <https://www.R-project.org>.
- Cowley, L.A., Petersen, F.C., Junges, R., Jimson D. Jimenez, M., Morrison, D.A., and Hanage, W.P. (2018) Evolution via recombination: Cell-to-cell contact facilitates larger recombination events in *Streptococcus pneumoniae*. *PLoS Genet* **14**: 1–23.
- Danner, D.B., Deich, R.A., Sisco, K.L., and Smith, H.O. (1980) An eleven-base-pair sequence determines the specificity of DNA uptake in *Haemophilus* transformation. *Gene* **11**: 311–318.
- Dawoud, T.M., Jiang, T., Mandal, R.K., Ricke, S.C., and Kwon, Y.M. (2014) Improving the Efficiency of Transposon Mutagenesis in *Salmonella Enteritidis* by Overcoming Host-Restriction Barriers. *Mol Biotechnol* **56**: 1004–1010.

- DeJesus, M.A., Ambadipudi, C., Baker, R.E., Sassetti, C., and Ioerger, T.R. (2015) TRANSIT - A Software Tool for Himar1 TnSeq Analysis. *PLoS Comput Biol* **11**: 1–17.
- DeJesus, M.A. and Ioerger, T.R. (2013) A Hidden Markov Model for identifying essential and growth-defect regions in bacterial genomes from transposon insertion sequencing data. *BMC Bioinformatics* **14**: 1–12.
- DeJesus, M.A., Zhang, Y.J., Sassetti, C.M., Rubin, E.J., Sacchettini, J.C., and Ioerger, T.R. (2013) Bayesian analysis of gene essentiality based on sequencing of transposon insertion libraries. *Bioinformatics* **29**: 695–703.
- Demanèche, S., Kay, E., Gourbière, F., and Simonet, P. (2001) Natural Transformation of *Pseudomonas fluorescens* and *Agrobacterium tumefaciens* in Soil. *Appl Environ Microbiol* **67**: 2617–2621.
- Deng, L., Wang, C., Zhang, X., Yang, W., Tang, H., Chen, Xinru, et al. (2023) Cell-to-cell natural transformation in *Bacillus subtilis* facilitates large scale of genomic exchanges and the transfer of long continuous DNA regions. *Nucleic Acids Res* **51**: 3820–3835.
- Dillard, J.P. and Seifert, H.S. (2001) A variable genetic island specific for *Neisseria gonorrhoeae* is involved in providing DNA for natural transformation and is found more often in disseminated infection isolates. *Mol Microbiol* **41**: 263–277.
- Gallagher, L.A., Shendure, J., and Manoil, C. (2011) Genome-scale identification of resistance functions in *Pseudomonas aeruginosa* using Tn-seq. *MBio* **2**: 1–18.
- Gerdes, S.Y., Scholle, M.D., Campbell, J.W., Balázs, G., Ravasz, E., Daugherty, M.D., et al. (2003) Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655. *J Bacteriol* **185**: 5673–5684.
- Graupner, S., Frey, V., Hashemi, R., Lorenz, M.G., Brandes, G., and Wackernagel, W. (2000) Type IV pilus genes pilA and pilC of *Pseudomonas stutzeri* are required for natural genetic transformation, and pilA can be replaced by corresponding genes from nontransformable species. *J Bacteriol* **182**: 2184–2190.
- Graupner, S. and Wackernagel, W. (2001a) Identification and characterization of novel competence genes comA and exbB involved in natural genetic transformation of *Pseudomonas stutzeri*. *Res Microbiol* **152**: 451–460.
- Graupner, S. and Wackernagel, W. (2001b) *Pseudomonas stutzeri* has two closely related pilA genes (Type IV pilus structural protein) with opposite influences on natural genetic transformation. *J Bacteriol* **183**: 2359–2366.
- Graupner, S., Weger, N., Sohni, M., and Wackernagel, W. (2001) Requirement of novel competence genes pilT and pilU of *Pseudomonas stutzeri* for natural transformation and suppression of pilT deficiency by a hexahistidine tag on the type IV pilus protein PilA. *J Bacteriol* **183**: 4694–4701.

- Graves, J.F., Biswas, G.D., and Sparling, P.F. (1982) Sequence-specific DNA uptake in transformation of *Neisseria gonorrhoeae*. *J Bacteriol* **152**: 1071–1077.
- Hamilton, H.L., Domínguez, N.M., Schwartz, K.J., Hackett, K.T., and Dillard, J.P. (2005) *Neisseria gonorrhoeae* secretes chromosomal DNA via a novel type IV secretion system. *Mol Microbiol* **55**: 1704–1721.
- Hardy, L. and Charpentier, X. (2019) Querying *Legionella* genomes using transposition-sequencing. *Methods Mol Biol* **1921**: 107–122.
- Hardy, L., Juan, P.A., Coupat-Goutaland, B., and Charpentier, X. (2021) Transposon insertion sequencing in a clinical isolate of *legionella pneumophila* identifies essential genes and determinants of natural transformation. *J Bacteriol* **203**: 1–16.
- Harms, A., Liesch, M., Körner, J., Québatte, M., Engel, P., and Dehio, C. (2017) A bacterial toxin-antitoxin module is the origin of inter-bacterial and inter-kingdom effectors of *Bartonella*. *PLoS Genet* **13**: 1–22.
- Hernández-Plaza, A., Szklarczyk, D., Botas, J., Cantalapiedra, C.P., Giner-Lamia, J., Mende, D.R., et al. (2023) eggNOG 6.0: enabling comparative genomics across 12 535 organisms. *Nucleic Acids Res* **51**: 389–394.
- Ioerger, T.R. (2022) Analysis of Gene Essentiality from TnSeq Data Using Transit. *Methods Mol Biol* **2377**: 391–421.
- Johnsborg, O., Eldholm, V., Bjørnstad, M.L., and Håvarstein, L.S. (2008) A predatory mechanism dramatically increases the efficiency of lateral gene transfer in *Streptococcus pneumoniae* and related commensal species. *Mol Microbiol* **69**: 245–253.
- Kreth, J., Merritt, J., Shi, W., and Qi, F. (2005) Co-ordinated bacteriocin production and competence development: A possible mechanism for taking up DNA from neighbouring species. *Mol Microbiol* **57**: 392–404.
- 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.
- Kwon, Y.M., Ricke, S.C., and Mandal, R.K. (2016) Transposon sequencing: methods and expanding applications. *Appl Microbiol Biotechnol* **100**: 31–43.
- Lagendijk, E.L., Validov, S., Lamers, G.E.M., De Weert, S., and Bloemberg, G. V (2010) Genetic tools for tagging Gram-negative bacteria with mCherry for visualization in vitro and in natural habitats, biofilm and pathogenicity studies. *FEMS Microbiol Lett* **305**: 81–90.
- Lalucat, J., Bennasar, A., Bosch, R., García-Valdés, E., and Palleroni, N.J. (2006) Biology of *Pseudomonas stutzeri*. *Microbiol Mol Biol Rev* **70**: 510–547.
- Lampe, D.J., Churchill, M.E.A., and Robertson, H.M. (1996) A purified mariner transposase is sufficient to mediate transposition in vitro. *EMBO J* **15**: 5470–5479.

- Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.
- Lee, S.A., Gallagher, L.A., Thongdee, M., Staudinger, B.J., Lippman, S., Singhc, P.K., and Manoila, C. (2015) General and condition-specific essential functions of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* **112**: 5189–5194.
- Lefebvre, M.D. and Valvano, M.A. (2002) Construction and evaluation of plasmid vectors optimized for constitutive and regulated gene expression in *Burkholderia cepacia* complex isolates. *Appl Environ Microbiol* **68**: 5956–5964.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
- Liberati, N.T., Urbach, J.M., Miyata, S., Lee, D.G., Drenkard, E., Wu, G., et al. (2006) An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proc Natl Acad Sci U S A* **103**: 2833–2838.
- Luo, H., Lin, Y., Gao, F., Zhang, C.T., and Zhang, R. (2014) DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. *Nucleic Acids Res* **42**: 574–580.
- Martin, M. (2011) Cutadapt removed adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**: 10–12.
- Martínez-García, E. and de Lorenzo, V. (2011) Engineering multiple genomic deletions in Gram-negative bacteria: Analysis of the multi-resistant antibiotic profile of *Pseudomonas putida* KT2440. *Environ Microbiol* **13**: 2702–2716.
- Matthey, N., Stutzmann, S., Stoudmann, C., Guex, N., Iseli, C., and Blokesch, M. (2019) Neighbor predation linked to natural competence fosters the transfer of large genomic regions in *vibrio cholerae*. *Elife* **8**: 1–21.
- Moule, M.G., Hemsley, C.M., Seet, Q., Guerra-Assunção, J.A., Lim, J., Sarkar-Tyson, M., et al. (2014) Genome-wide saturation mutagenesis of *Burkholderia pseudomallei* K96243 predicts essential genes and novel targets for antimicrobial development. *MBio* **5**: 1–9.
- Nielsen, K.M., Johnsen, P.J., Bensasson, D., and Daffonchio, D. (2007) Release and persistence of extracellular DNA in the environment. *Environ Biosafety Res* **6**: 37–53.
- Nielsen, K.M., Van Weerelt, M.D.M., Berg, T.N., Bones, A.M., Hagler, A.N., and Van Elsas, J.D. (1997) Natural Transformation and Availability of Transforming DNA to *Acinetobacter calcoaceticus* in Soil Microcosms. *Appl Environ Microbiol* **63**: 1945–1952.
- Nolan, L.M., Turnbull, L., Katrib, M., Osvath, S.R., Losa, D., Lazenby, J.J., and Whitchurch, C.B. (2020) *Pseudomonas aeruginosa* is capable of natural transformation in biofilms. *Microbiol (United Kingdom)* **166**: 995–1003.
- van Opijnen, T., Bodi, K.L., and Camilli, A. (2009) Tn-seq: High-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat Methods* **6**: 767–772.

- Pagès, H., Aboyoun, P., Gentleman, R., and DebRoy, S. (2022) Biostrings: Efficient manipulation of biological strings. <https://bioconductor.org/packages/Biostrings>.
- Poulsen, B.E., Yang, R., Clatworthy, A.E., White, T., Osmulski, S.J., Li, L., et al. (2019) Defining the core essential genome of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* **116**: 10072–10080.
- Redfield, R.J. (1993) Genes for breakfast: The have-your-cake and-eat-it-too of bacterial transformation. *J Hered* **84**: 400–404.
- Ringel, P.D., Hu, D., and Basler, M. (2017) The Role of Type VI Secretion System Effectors in Target Cell Lysis and Subsequent Horizontal Gene Transfer. *Cell Rep* **21**: 3927–3940.
- Seitz, P. and Blokesch, M. (2012) Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. *FEMS Microbiol Rev* **37**: 336–363.
- Sikorski, J., Graupner, S., Lorenz, M.G., and Wackernagel, W. (1998) Natural genetic transformation of *Pseudomonas stutzeri* in a non-sterile soil. *Microbiology* **144**: 569–576.
- Sisco, K.L. and Smith, H.O. (1979) Sequence-specific DNA uptake in *Haemophilus* transformation *Microbiology* : *Proc Natl Acad Sci* **76**: 972–976.
- Smith, H.O., Gwinn, M.L., and Salzberg, S.L. (1999) DNA uptake signal sequences in naturally transformable bacteria. *Res Microbiol* **150**: 603–616.
- Stewart, G.J., Carlson, C.A., and Ingraham, J.L. (1983) Evidence for an active role of donor cells in natural transformation of *Pseudomonas stutzeri*. *J Bacteriol* **156**: 30–35.
- Su, Y., Xu, Y., Li, Q., Yuan, G., and Zheng, D. (2020) The essential genome of *Ralstonia solanacearum*. *Microbiol Res* **238**: 1–6.
- Turner, K.H., Wessel, A.K., Palmer, G.C., Murray, J.L., and Whiteley, M. (2015) Essential genome of *Pseudomonas aeruginosa* in cystic fibrosis sputum. *Proc Natl Acad Sci U S A* **112**: 4110–4115.
- Vosman, B., Hellingwerf, K.J., Heuingwerf, K.J., and Hellingwerf, K.J. (1991) Molecular cloning and functional characterization of a *recA* analog from *Pseudomonas stutzeri* and construction of a *P. stutzeri recA* mutant. *Antonie Van Leeuwenhoek* **59**: 115–123.
- Wickham, H. (2008) *Elegant Graphics for Data Analysis: ggplot2*, Springer (ed).
- Zheng, R., Wu, S., Ma, N., and Sun, C. (2018) Genetic and physiological adaptations of marine bacterium *Pseudomonas stutzeri* 273 to mercury stress. *Front Microbiol* **9**: 1–14.

4.8 Supplementary Material

Table S 4.1 Bacterial strains, plasmids and oligonucleotides used in this chapter. Abbreviations: Amp^R, ampicillin resistance; Gm^R, gentamicin resistance; Km^R, kanamycin resistance; Cm^R, chloramphenicol resistance; Tet^R, tetracycline resistance; MCS, multiple cloning site.

Strains, plasmids or oligonucleotides		Genotype, phenotype or relevant characteristics	References
Strains			
<i>P. stutzeri</i>	ATCC 17587	Wild type phenotype	Laboratory collection
<i>P. stutzeri</i>	ATCC 17587 rif ^R	Spontaneous rifampicin-resistant phenotype	Chapter 2
<i>P. stutzeri</i>	tet ^R -km ^R	ATCC 17587, miniTn7::Km ^R ::Tet ^R , contains two cassettes of resistance against kanamycin and tetracycline	This study
<i>P. stutzeri</i>	ΔcomA	ATCC 17587 ΔcomA	Chapter 2
<i>P. stutzeri</i>	ΔcomA rif ^R	ATCC 17587 ΔcomA spontaneous rifampicin-resistant mutant	Chapter 3
<i>P. stutzeri</i>	ΔcomA tet ^R -km ^R	ATCC 17587 ΔcomA, miniTn7::Km ^R ::Tet ^R , contains two cassettes of resistance against kanamycin and tetracycline	This study
<i>E. coli</i>	JKE201-pBT20	JKE201 strain containing the pBT20 plasmid	(Harms <i>et al.</i> , 2017)
Plasmids			
pBT20	Gm ^R , Amp ^R ; mariner transposon		(Kulasekara <i>et al.</i> , 2005)
pEMG	Km ^R , oriR6K, MCS with two flanking I-SceI sites, used for markerless deletions		(Martínez-García and de Lorenzo, 2011)
pSW-2	oriRK2, Gm ^R , xylS, Pm→I-sceI (the I-Sce-I gene controlled by Pm promoter)		(Martínez-García and de Lorenzo, 2011)
pUX-BF13	oriR6K, Amp ^R , helper plasmid providing the Tn7 transposition functions in trans		(Bao <i>et al.</i> , 1991)
pME6000	pBBR1 ori, Tet ^R <i>lacZ mob</i> ⁺ , provider of tetracycline resistance		(Lefebvre and Valvano, 2002)
pMP7607	pBK-miniTn7-KmΩSm1 derivative harboring the <i>mCherry</i> gene under the control of the <i>Ptac</i> promoter		(Lagendijk <i>et al.</i> , 2010)
pME7607-tet ^R	pME7607 in which the <i>mCherry</i> gene and the <i>tac</i> promoter was replaced by a <i>tetR</i> gene.		This study
Oligonucleotides (5'→3')			
pME6000_tetR_F	ATAAGAATGCGGCCGCGATGTTGTCTACATGGCTCTGC, with a NotI underlined site, to amplify <i>tetR</i>		This study
pME6000_tetR_R	ATAAGAATGCGGCCGCGCTATCGTTTCCACGATCAGC, with a NotI underlined site, to amplify <i>tetR</i>		This study
olj376	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGGGGGGGGGGGGGGG		(Hardy <i>et al.</i> , 2021)
pBT20-PCR1	Biotin-TCGTATAATGTGTGGAATTGTGAGCGG		(Hardy <i>et al.</i> , 2021)
pBT20-PCR2_mod	AATGATACGGCGACACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCTGGACTCTAGAGGATCACCCAGCTTCTTGCAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT		This study
TdT_index_1	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT		(Hardy <i>et al.</i> , 2021)
gen_F	TCGAAGCATCCGATCTGACG		This study
gen_R	ACCTCATCACTGAAGGCTGC		This study
plagen_F	CATGCTTGGTGATCGCCTTG		This study
plagen_R	CTTCCGCTTCGATTGGCATG		This study

Genomic DNA fragmentation by shearing

Poly C tail addition to the 3' end

PCR 1 : amplification of junctions and addition of biotin

Capture of junctions with biotin – streptavidin beads interaction

PCR 2: addition of Illumina adaptors

Sequencing

Figure S 4.1 Transposon sequencing library preparation, adapted from Hardy and Charpentier (Hardy and Charpentier, 2019). After DNA extraction and fragmentation, the rTdT enzyme adds a poly C-tail to DNA fragments. Transposon junctions (red) are amplified by PCR using a 5'-biotinylated transposon-specific primer (pBT20-PCR1) and the poly-G primer olj376, which also adds a region for Illumina sequencing (blue). Fragments are then purified with streptavidin beads (grey). A second PCR adds Illumina adapters (pink, yellow) and index for multiplexing (orange). Sequencing uses Read1TnLp for transposon junction reads and standard Illumina primers for index reading. AMPure XP bead purification is performed between steps.

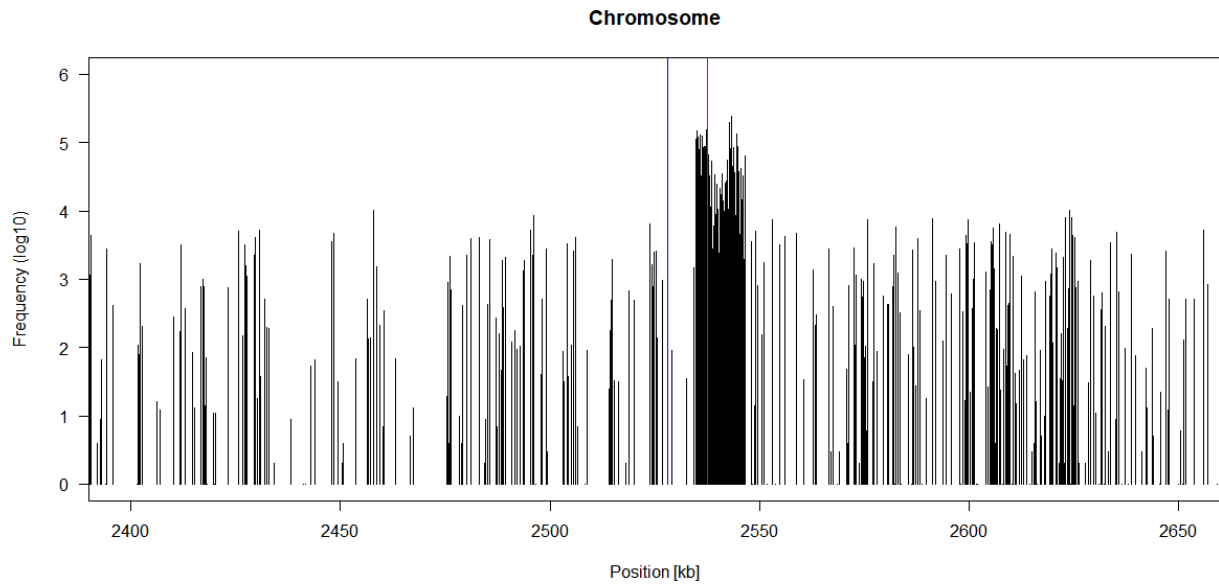


Figure S 4.2 Log₁₀ of transposon insertions reads of initial recipient library in the chromosome. Only a subset of the chromosome is shown. The blue line shows the sequence amplified by qPCR to represent regions with an average read coverage. The red line indicates the sequence amplified by qPCR to represent regions with enriched read coverage.

Table S 4.2. Summary of the analysis realized and their results regarding the number of essential genes (adjusted p-value <0.05). Libraries names: all recipients from transformations with DNA (DNA_all) and donor cells (cells_all), transformants obtained from DNA (DNA_trans) and donor cell (cells_trans) transformations. *These results are presented in Figure 4.6.

Goal: Identify loci essential for...	Method	Contrast		Nr. essential genes
		Control	Treatment	
Growth	HMM*	All 6 libraries	-	4324
	Gumbel*	all 6 libraries	-	500
	Tn5gap*	all 6 libraries	-	109
Transformation	Resampling	DNA_all, cells_all	DNA_trans, cells_trans	0
Transformation with DNA	Resampling	DNA_all	DNA_trans	16
Transformation with donors	Resampling	cells_all	cells_trans	0

Table S 4.3 Mean Ct values obtained by qPCR with two primer pairs on two replicates to evaluate the overrepresentation of a region relative to another regarding conditions and step in the sequencing libraries preparation (Figure S 4.2). The qPCR was performed directly after the step of interest. Ct: Cycle threshold.* indicates the condition used in the initial protocol.

Goal	Tested step	Rational	Conditions	Δ mean Cycle threshold
Effect of kit for gDNA extraction	DNA extraction	Minimization of plasmid extraction	Genomic-tip 20 kit (Qiagen)	-0.2
			Wizard® HMW DNA Extraction Kit (Promega)	2.4
			QIAamp DNA kit (Qiagen) *	1.3
Effect of kit for gDNA extraction	DNA extraction	Minimization of plasmid extraction	Genomic-tip 20 kit (Qiagen)	1.1
			Wizard® HMW DNA Extraction Kit (Promega)	1.2
			QIAamp DNA kit (Qiagen) *	2.3
Effect of nr. of cycles and pooling of multiples tubes	PCR1	Minimization of bias due to overamplification and single tube	20 cycles	2.3
			25 cycles *	2.2
			20 cycles with 5 tubes pooled	2.7
Effect of nr. of cycles and pooling of multiples tubes	PCR1	Minimization of bias due to overamplification and single tube	20 cycles	1.1
			20 cycles with pooling	2.5
Effect of the nr. of cycles	PCR2	Minimization of bias due to overamplification	10 cycles	9.7
			15 cycles	14.5
Effect of nr. of cycles and presence of beads	PCR2	Minimization of bias due to overamplification, due the lower amount of PCR1 prod., due to presence of beads	25 cycles*	18.9
			20 cycles	19.5
			20 cycles with 2x more beads	19.5
			20 cycles with beads removed	17.5
Effect of the ratio of dNTP	PCR2	Minimization of bias due to dNTP ratio imbalance	standard dNTP mix*	15.7
			biased dNTP mix	13.7

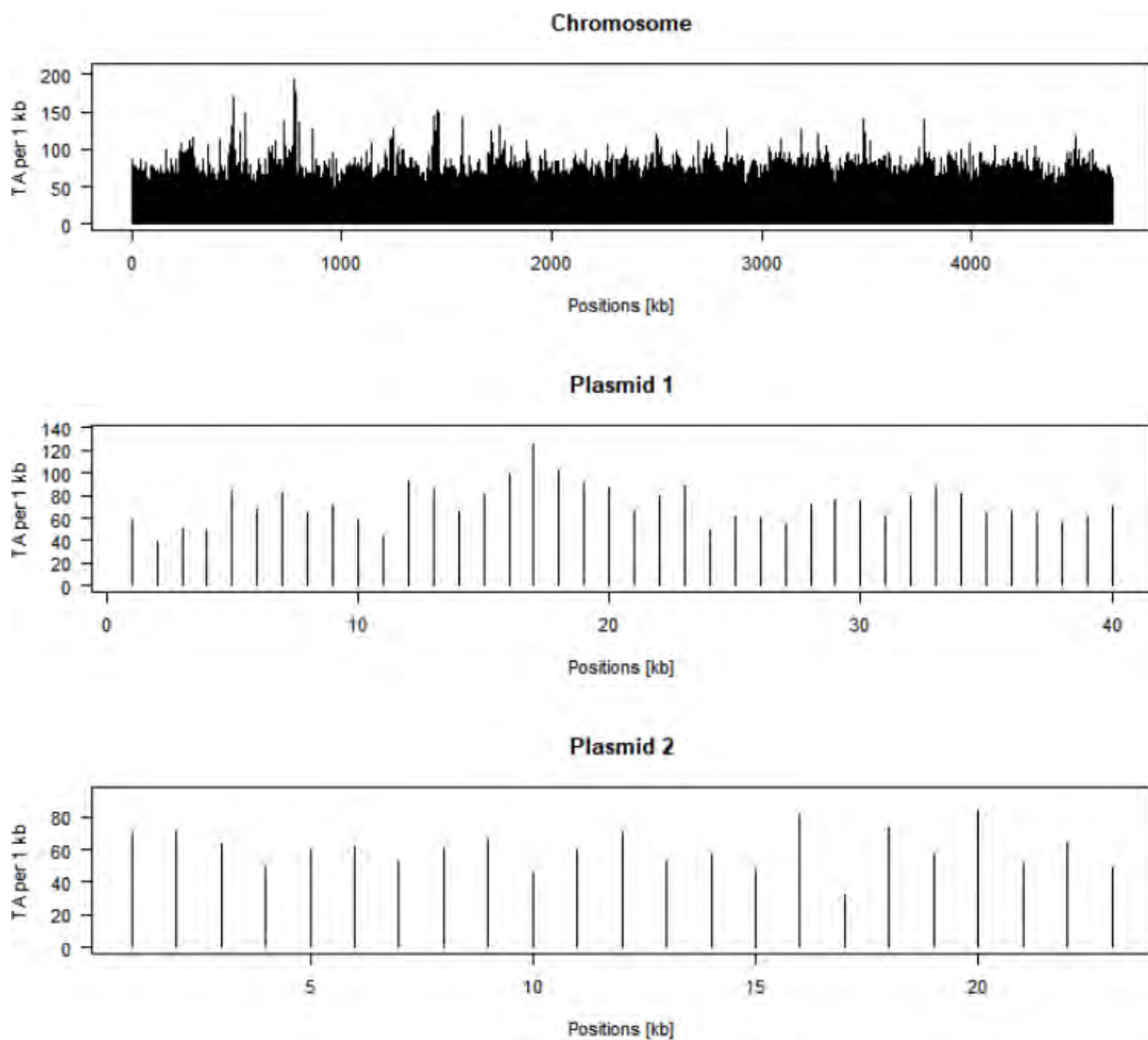


Figure S 4.3 Number of possible transposon insertion sites (TA sites) per kb in the chromosome (top) and both plasmids (middle and bottom). Each bar corresponds to the number of TA sites per kb.

Table S 4.4 Ct values obtained by qPCR with two primer pairs on the final sequencing libraries. The average region refers to the region having a read coverage within the average during the first sequencing of the initial library. The enriched region refers to the region having a particularly high read coverage average during the first sequencing of the initial library. Ct: Cycle threshold.

Samples	Ct enriched region	Ct average region	Δ Ct	Fold change
Initial library	25.7	17.6	8.1	274.4
Transformants with DNA	23.6	17.5	6.1	68.6
All recipients from transf. with DNA	24.1	17.6	6.4	86.8
Library of transformants with donors	23.1	17.9	5.2	37.8
All recipients from transf. with donors	26.0	15.2	10.8	1782.9

Table S 4.5 Summary of genes essentiality for growth based on three methods. *Genes with 0 TA sites were not analyzed by HMM method.

Methods	Categories				
Gumbel	Essential	Non-essential		Uncertain	Too short
	500	3297		177	394
HMM	Essential	Growth-Defect	Growth-advantage	Non-essential	Not analyzed*
	4324	2	23	13	6
Tn5	Essential			Non-essential	
	109			4259	

Table S 4.6 Summary of the quality control performed by TRANSIT software (DeJesus et al., 2015) on the sequencing libraries. Data out of the recommendation are highlighted in bold.

Parameters	Libraries					Definition	Comment
	Initial	Transf. With DNA		Transf. with donors			
		All cells	Transf.	All cells	Transf.		
Density	0.403	0.079	0.076	0.079	0.062		“Well saturated” Himar1 datasets have >30% saturation. Beneath this, statistical methods may have trouble.
Mean Read	44.8	48.3	42.1	65	69.3	Average read-count, including empty sites.	
NZmean	111.2	611.8	557.2	817.3	1118.1	Average read-count, excluding empty sites.	A value between 30-200 is usually good for Himar1 datasets. Too high or too low can indicate problems.
NZmedian	11	87	48	128	56	Median read-count, excluding empty sites	As read-counts can often have spikes, median serves as a good robust estimate.
Max ct	665’080	471’223	474’879	312’385	427’321	Largest read-count in the dataset.	Useful to determine whether there are outliers/spikes, which may indicate sequencing issues.
Total cts	9’105’060	9’824’131	8’554’713	13’203’387	14’093’755	Sum of total read-counts in the dataset.	Indicates how much sequencing material was obtained. Typically >1M reads is desired for Himar1 datasets.
Skewness	180.3	50.4	89.5	42.3	48.2	Skew of read-counts in the dataset.	Large skew may indicate issues with a dataset. Typically a skew < 50 is desired. May be higher when library is under strong selection
Kurtosis	34442.9	3475.3	9663.2	2564.6	3601.1	Kurtosis of the read-counts in the dataset.	

Table S 4.7 Essential open reading frames (ORFs) in *P. stutzeri* identified by Gumbel analysis but classified as non-essential by Tn5 method, based on six sequencing libraries. COG annotations added using EggNOG tool (Hernández-Plaza *et al.*, 2023). C: Energy production and conversion; D: Cell cycle control, cell division, chromosome partitioning; E: Amino Acid transport and metabolism; F: Nucleotide transport and metabolism; G: Carbohydrate transport and metabolism; H: Coenzyme transport and metabolism; I: Lipid transport and metabolism; J: Translation, ribosomal structure and biogenesis; K: Transcription; L: Replication, recombination and repair; M: Cell wall/membrane/envelop biogenesis; N: Cell motility; O: Post-translational modification, protein turnover, chaperones; P: Inorganic ion transport and metabolism; Q: Secondary metabolites biosynthesis, transport and catabolism; S: Function Unknown; T: Signal transduction mechanisms; U: Intracellular trafficking, secretion, and vesicular transport; V: Defense mechanisms.

ORF	Name	Description	COG
LOCUS_00090	-	hypothetical protein	L
LOCUS_00100	-	group II intron reverse transcriptase/maturase	L
LOCUS_00210	<i>lptA_1</i>	lysophosphatidic acid acyltransferase	I
LOCUS_00220	-	D%2CD-heptose 1%2C7-bisphosphate phosphatase	G
LOCUS_00390	<i>tsaC</i>	threonylcarbamoyl-AMP synthase	J
LOCUS_00810	<i>np20</i>	transcriptional repressor	K
LOCUS_01230	-	hypothetical protein	L
LOCUS_01710	-	hypothetical protein	K
LOCUS_01900	<i>rpmB</i>	50S ribosomal protein L28	J
LOCUS_02000	<i>polA</i>	DNA polymerase I	L
LOCUS_02560	-	hypothetical protein	-
LOCUS_03300	-	hypothetical protein	L
LOCUS_03310	-	group II intron reverse transcriptase/maturase	L
LOCUS_03320	-	hypothetical protein	L
LOCUS_03410	-	sodium:proton antiporter	P
LOCUS_03420	-	hypothetical protein	P
LOCUS_03430	<i>Int_1</i>	apolipoprotein N-acyltransferase	M
LOCUS_03440	<i>lspA_1</i>	lipoprotein signal peptidase	MU
LOCUS_03530	-	transposase	L
LOCUS_03540	-	transposase	L
LOCUS_03550	-	hypothetical protein	L
LOCUS_03770	-	IS5 family transposase	L
LOCUS_03790	-	hypothetical protein	L
LOCUS_03800	-	transposase	L
LOCUS_04500	<i>fbp</i>	fructose-1%2C6-bisphosphatase class 1	G
LOCUS_04640	<i>ubiB</i>	putative protein kinase UbiB	H
LOCUS_04650	-	sterol-binding protein	S
LOCUS_04660	<i>ubiE</i>	ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE	H
LOCUS_04680	-	hypothetical protein	S
LOCUS_04790	<i>glnA</i>	glutamine synthetase	E
LOCUS_05300	<i>serA</i>	D-3-phosphoglycerate dehydrogenase	EH
LOCUS_05370	<i>ilvA</i>	L-threonine dehydratase	E
LOCUS_05450	<i>thyA</i>	thymidylate synthase	F
LOCUS_05560	-	hypothetical protein	L
LOCUS_05570	-	transposase	L
LOCUS_05890	<i>spoT</i>	guanosine-3'-bis(diphosphate) 3'-pyrophosphohydrolase	KT
LOCUS_05910	<i>gmk</i>	guanylate kinase	F
LOCUS_06010	<i>dut</i>	deoxyuridine 5'-triphosphate nucleotidohydrolase	F
LOCUS_06020	<i>coaC</i>	phosphopantothienoylcysteine decarboxylase	H
LOCUS_06280	<i>rep</i>	ATP-dependent DNA helicase Rep	L
LOCUS_06330	<i>glnK</i>	nitrogen regulatory protein P-II 2	K
LOCUS_06410	<i>dapF</i>	diaminopimelate epimerase	E
LOCUS_06420	<i>lysA</i>	diaminopimelate decarboxylase	E
LOCUS_06530	<i>argH</i>	argininosuccinate lyase	E
LOCUS_06560	<i>hemC</i>	prophobilinogen deaminase	H
LOCUS_06570	<i>hemD</i>	uroporphyrinogen-III synthase	H

LOCUS_06720	<i>hemB</i>	delta-aminolevulinic acid dehydratase	H
LOCUS_06760	<i>rho</i>	transcription termination factor Rho	K
LOCUS_06770	<i>ubiD</i>	3-octaprenyl-4-hydroxybenzoate carboxy-l yase	H
LOCUS_06800	<i>argS</i>	arginine--tRNA ligase	J
LOCUS_06810	<i>priA</i>	primosomal protein N'	L
LOCUS_06960	<i>hemE</i>	uroporphyrinogen decarboxylase	H
LOCUS_07030	<i>trx</i>	thiol reductase thioredoxin	O
LOCUS_07080	<i>rpoD</i>	RNA polymerase sigma factor RpoD	K
LOCUS_07090	<i>dnaG</i>	DNA primase	L
LOCUS_07110	<i>tsaD</i>	tRNA N6-adenosine threonylcarbamoyltransferase	O
LOCUS_07130	-	7%2C8-dihydroneopterin aldolase	H
LOCUS_07230	<i>pdxA1</i>	4-hydroxythreonine-4-phosphate dehydrogenase 1	H
LOCUS_07240	<i>surA</i>	chaperone SurA	O
LOCUS_07460	-	IS5 family transposase	L
LOCUS_07470	<i>coq7</i>	2-nonaprenyl-3-methyl-6-methoxy-1%2C4-benzoquinol hydroxylase	H
LOCUS_07510	-	membrane protein	S
LOCUS_07520	<i>argC</i>	N-acetyl-gamma-glutamyl-phosphate reductase	E
LOCUS_07530	<i>erpA</i>	iron-sulfur cluster insertion protein ErpA	C
LOCUS_07550	-	peptidase M23	M
LOCUS_07560	<i>tyrS</i>	tyrosine--tRNA ligase	J
LOCUS_07570	<i>birA</i>	bifunctional ligase/repressor BirA	H
LOCUS_07600	<i>tuf_1</i>	elongation factor Tu	J
LOCUS_07610	<i>secE</i>	protein translocase subunit SecE	U
LOCUS_07620	<i>nusG</i>	transcription termination/antitermination protein NusG	K
LOCUS_07640	<i>rplA</i>	50S ribosomal protein L1	J
LOCUS_07660	<i>rplL</i>	50S ribosomal protein L7/L12	J
LOCUS_07680	<i>rpoC</i>	DNA-directed RNA polymerase subunit beta'	K
LOCUS_07690	<i>rpsL</i>	30S ribosomal protein S12	J
LOCUS_07700	<i>rpsG</i>	30S ribosomal protein S7	J
LOCUS_07720	<i>tuf_2</i>	elongation factor Tu	J
LOCUS_07730	<i>rpsJ</i>	30S ribosomal protein S10	J
LOCUS_07740	<i>rplC</i>	50S ribosomal protein L3	J
LOCUS_07760	<i>rplW</i>	50S ribosomal protein L23	J
LOCUS_07770	<i>rplB</i>	50S ribosomal protein L2	J
LOCUS_07780	<i>rpsS</i>	30S ribosomal protein S19	J
LOCUS_07790	<i>rplV</i>	50S ribosomal protein L22	J
LOCUS_07800	<i>rpsC</i>	30S ribosomal protein S3	J
LOCUS_07830	<i>rpsQ</i>	30S ribosomal protein S17	J
LOCUS_07850	<i>rplX</i>	50S ribosomal protein L24	J
LOCUS_07870	<i>rpsN</i>	30S ribosomal protein S14	J
LOCUS_07880	<i>rpsH</i>	30S ribosomal protein S8	J
LOCUS_07890	<i>rplF</i>	50S ribosomal protein L6	J
LOCUS_07900	<i>rplR</i>	50S ribosomal protein L18	J
LOCUS_07920	<i>rpmD</i>	50S ribosomal protein L30	J
LOCUS_07940	<i>secY</i>	protein translocase subunit SecY	U
LOCUS_07960	<i>rpsK</i>	30S ribosomal protein S11	J
LOCUS_07970	<i>rpsD</i>	30S ribosomal protein S4	J
LOCUS_07980	<i>rpoA</i>	DNA-directed RNA polymerase subunit alpha	K
LOCUS_07990	<i>rplQ</i>	50S ribosomal protein L17	J
LOCUS_08080	-	DUF3450 domain-containing protein	D
LOCUS_08090	-	biopolymer transporter ExbB	U
LOCUS_08100	-	hypothetical protein	U
LOCUS_08120	-	hypothetical protein	U
LOCUS_08260	<i>ribD</i>	riboflavin biosynthesis protein RibD	H
LOCUS_08290	<i>ribH</i>	6%2C7-dimethyl-8-ribityllumazine synthase	H
LOCUS_08300	<i>nusB</i>	N utilization substance protein B	K

LOCUS_08310	<i>thiL</i>	thiamine-monophosphate kinase	H
LOCUS_08340	<i>ribA</i>	GTP cyclohydrolase-2	F
LOCUS_08410	-	hypothetical protein	-
LOCUS_09200	<i>glyA2</i>	serine hydroxymethyltransferase 3	E
LOCUS_09280	-	hypothetical protein	L
LOCUS_09290	-	transposase	L
LOCUS_09380	<i>rplU</i>	50S ribosomal protein L21	J
LOCUS_09400	<i>obg</i>	GTPase Obg	S
LOCUS_09430	<i>rpsT</i>	30S ribosomal protein S20	J
LOCUS_09440	<i>mviN_1</i>	putative lipid II flippase MurJ	S
LOCUS_09460	<i>ileS</i>	isoleucine--tRNA ligase	J
LOCUS_09490	<i>ispH</i>	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	IM
LOCUS_09710	<i>fruK</i>	phosphofructokinase	H
LOCUS_09930	<i>gatA</i>	glutamyl-tRNA(Gln) amidotransferase subunit A	H
LOCUS_09940	<i>gatC</i>	glutamyl-tRNA(Gln) amidotransferase subunit C	H
LOCUS_09950	<i>mreB</i>	rod shape-determining protein	D
LOCUS_10110	-	ABC transporter ATP-binding protein	S
LOCUS_10120	<i>lptA_2</i>	lipopolysaccharide export system protein LptA	S
LOCUS_10140	-	hypothetical protein	S
LOCUS_10230	<i>murA</i>	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	M
LOCUS_10370	<i>coaE</i>	dephospho-CoA kinase	F
LOCUS_10410	<i>rplM</i>	50S ribosomal protein L13	J
LOCUS_10420	<i>rpsI</i>	30S ribosomal protein S9	J
LOCUS_10490	<i>gmhA</i>	phosphoheptose isomerase	G
LOCUS_10540	<i>rsmH</i>	ribosomal RNA small subunit methyltransferase H	J
LOCUS_10550	<i>ftsL</i>	cell division protein FtsL	D
LOCUS_10560	<i>ftsI</i>	penicillin-binding protein 3	M
LOCUS_10570	<i>murE</i>	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2%2C6-di aminopimelate ligase	M
LOCUS_10580	<i>murF</i>	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase	M
LOCUS_10590	<i>mraY</i>	phospho-N-acetylmuramoyl-pentapeptide-transferase	M
LOCUS_10600	<i>murD</i>	UDP-N-acetylmuramoylalanine--D-glutamate ligase	M
LOCUS_10610	<i>ftsW</i>	putative lipid II flippase FtsW	M
LOCUS_10620	<i>murG</i>	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	M
LOCUS_10630	<i>murC</i>	UDP-N-acetylmuramate--L-alanine ligase	M
LOCUS_10640	<i>ddlB</i>	D-alanine--D-alanine ligase B	F
LOCUS_10660	<i>ftsA</i>	cell division protein FtsA	D
LOCUS_10670	<i>ftsZ</i>	cell division protein FtsZ	D
LOCUS_10680	<i>lpxC</i>	UDP-3-O-acyl-N-acetylglucosamine deacetylase	M
LOCUS_11400	-	IS5 family transposase	L
LOCUS_11430	-	hypothetical protein	-
LOCUS_11700	<i>ffh</i>	signal recognition particle protein	U
LOCUS_11710	<i>rpsP</i>	30S ribosomal protein S16	J
LOCUS_11720	<i>rimM</i>	ribosome maturation factor RimM	J
LOCUS_11780	<i>hom</i>	homoserine dehydrogenase	E
LOCUS_11790	<i>thrC</i>	threonine synthase	E
LOCUS_11830	<i>argG</i>	argininosuccinate synthase	E
LOCUS_12000	-	tRNA-modifying protein YgfZ	S
LOCUS_12100	<i>lepB</i>	signal peptidase I	U
LOCUS_12130	<i>era</i>	GTPase Era	S
LOCUS_12150	<i>pdxJ</i>	pyridoxine 5'-phosphate synthase	H
LOCUS_12360	<i>metG</i>	methionine--tRNA ligase	J
LOCUS_12370	-	iron-sulfur cluster carrier protein	D
LOCUS_12380	<i>capB</i>	cold shock protein CapB	K
LOCUS_12390	<i>dcd</i>	dCTP deaminase	F
LOCUS_12620	-	IS5 family transposase	L

LOCUS_13010	-	hypothetical protein	S
LOCUS_13030	<i>dapD</i>	2%2C3%2C4%2C5-tetrahydropyridine-2%2C6-dicarboxylate N-succinyltransferase	E
LOCUS_13160	-	DnaA regulatory inactivator Hda	L
LOCUS_13260	<i>proS</i>	proline--tRNA ligase	J
LOCUS_13330	<i>aspS</i>	aspartate--tRNA(Asp/Asn) ligase	J
LOCUS_13390	<i>tolQ</i>	protein TolQ	U
LOCUS_13400	<i>tolR</i>	protein TolR	U
LOCUS_13410	-	hypothetical protein	M
LOCUS_13420	<i>tolB</i>	protein TolB	U
LOCUS_13430	-	peptidoglycan-associated lipoprotein	M
LOCUS_13480	<i>nadA</i>	quinolinate synthase A	H
LOCUS_13540	<i>dapA</i>	4-hydroxy-tetrahydrodipicolinate synthase	E
LOCUS_13580	-	transposase	L
LOCUS_13590	-	transposase	L
LOCUS_13600	-	hypothetical protein	L
LOCUS_13610	-	IS5 family transposase	L
LOCUS_13650	<i>rlmE</i>	ribosomal RNA large subunit methyltransferase E	J
LOCUS_14480	<i>alaS</i>	alanine--tRNA ligase	J
LOCUS_14500	<i>csrA</i>	carbon storage regulator	J
LOCUS_14620	-	flagellar biosynthesis anti-sigma factor FlgM	N
LOCUS_14830	-	group II intron reverse transcriptase/maturase	L
LOCUS_14850	-	hypothetical protein	L
LOCUS_14860	-	transposase	L
LOCUS_14870	-	transposase	L
LOCUS_15120	-	IS5 family transposase	L
LOCUS_15280	-	hypothetical protein	-
LOCUS_15290	<i>prfB</i>	peptide chain release factor 2	J
LOCUS_15300	<i>lysS</i>	lysine--tRNA ligase	J
LOCUS_15390	<i>adk</i>	adenylate kinase	F
LOCUS_15400	-	tRNA (adenosine(37)-N6)-threonylcarbamoyltransferase complex dimerization subunit type 1 TsaB	O
LOCUS_15710	-	hypothetical protein	P
LOCUS_15940	<i>dapE</i>	succinyl-diaminopimelate desuccinylase	E
LOCUS_16020	-	hypothetical protein	S
LOCUS_16140	-	cysteine synthase	E
LOCUS_16280	-	glyceraldehyde-3-phosphate dehydrogenase	G
LOCUS_16410	-	multidrug ABC transporter substrate-binding protein	M
LOCUS_16420	<i>lolD_1</i>	lipoprotein-releasing system ATP-binding protein LolD	V
LOCUS_16430	-	multidrug ABC transporter substrate-binding protein	M
LOCUS_16440	-	hypothetical protein	S
LOCUS_16460	-	translocation protein TolQ	U
LOCUS_16470	-	biopolymer transporter ExbD	U
LOCUS_16480	<i>lpxK</i>	tetraacyldisaccharide 4'-kinase	F
LOCUS_16490	-	UPF0434 protein	S
LOCUS_16500	<i>kdsB</i>	3-deoxy-manno-octulosonate cytidyltransferase	M
LOCUS_16520	<i>murB</i>	UDP-N-acetylenolpyruvoylglucosamine reductase	M
LOCUS_16530	<i>rne</i>	ribonuclease E	J
LOCUS_16580	-	hypothetical protein	S
LOCUS_16610	<i>fabD</i>	malonyl CoA-acyl carrier protein transacylase	I
LOCUS_16620	<i>fabG_1</i>	3-oxoacyl-[acyl-carrier-protein] reductase FabG	IQ
LOCUS_16630	<i>acpP1</i>	acyl carrier protein 1	IQ
LOCUS_16640	-	3-oxoacyl-[acyl-carrier-protein] synthase 2	I
LOCUS_16670	<i>tmk</i>	thymidylate kinase	F
LOCUS_16680	-	DNA polymerase III subunit delta'	L
LOCUS_16800	<i>etfA</i>	electron transfer flavoprotein subunit alpha	C
LOCUS_17050	<i>fliN</i>	flagellar motor switch protein FliN	N

LOCUS_17070	<i>fliP</i>	flagellar biosynthetic protein FliP	N
LOCUS_17120	<i>fliH</i>	flagellar biosynthesis protein FliH	N
LOCUS_17290	<i>ccmA</i>	cytochrome c biogenesis ATP-binding export protein CcmA	O
LOCUS_17310	<i>ccmC</i>	heme exporter protein C	U
LOCUS_17320	<i>ccmD</i>	heme exporter protein D	U
LOCUS_17340	<i>ccmF</i>	cytochrome c-type biogenesis protein CcmF	O
LOCUS_17700	-	hypothetical protein	S
LOCUS_17960	-	hypothetical protein	S
LOCUS_17990	<i>efp</i>	elongation factor P	J
LOCUS_18060	-	4'-phosphopantetheinyl transferase	Q
LOCUS_18500	<i>trkH</i>	Trk system potassium uptake protein	P
LOCUS_18880	<i>nadK</i>	NAD kinase	H
LOCUS_18980	<i>nirD_1</i>	assimilatory nitrite reductase small subunit	P
LOCUS_19370	<i>thrS</i>	threonine--tRNA ligase	J
LOCUS_19380	<i>infC</i>	translation initiation factor IF-3	J
LOCUS_19390	<i>rpmI</i>	50S ribosomal protein L35	J
LOCUS_19400	<i>rplT</i>	50S ribosomal protein L20	J
LOCUS_19410	<i>pheS</i>	phenylalanine--tRNA ligase alpha subunit	J
LOCUS_19420	<i>pheT</i>	phenylalanine--tRNA ligase beta subunit	J
LOCUS_19460	<i>gyrA</i>	DNA gyrase subunit A	L
LOCUS_19470	<i>serC_1</i>	phosphoserine aminotransferase	E
LOCUS_19480	<i>pheA</i>	P-protein	E
LOCUS_19490	<i>hisC2</i>	histidinol-phosphate aminotransferase 2	E
LOCUS_19500	<i>aroA</i>	3-phosphoshikimate 1-carboxyvinyltransferase	E
LOCUS_19520	<i>rpsA</i>	30S ribosomal protein S1	J
LOCUS_19660	<i>ubiG</i>	ubiquinone biosynthesis O-methyltransferase	H
LOCUS_19960	<i>fadB</i>	fatty acid oxidation complex subunit alpha	I
LOCUS_19970	<i>fadA</i>	3-ketoacyl-CoA thiolase	I
LOCUS_19990	<i>topA</i>	DNA topoisomerase 1	L
LOCUS_20350	-	hypothetical protein	-
LOCUS_20450	<i>asd</i>	aspartate-semialdehyde dehydrogenase	E
LOCUS_20500	<i>accD</i>	acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	I
LOCUS_20770	<i>pdxB</i>	erythronate-4-phosphate dehydrogenase	H
LOCUS_20780	<i>dnaX</i>	DNA polymerase III subunit gamma/tau	H
LOCUS_20890	<i>nadE</i>	NH(3)-dependent NAD(+) synthetase	F
LOCUS_21060	<i>tusA</i>	sulfurtransferase TusD	O
LOCUS_21170	<i>ligA</i>	DNA ligase	L
LOCUS_21190	-	transcriptional regulator	K
LOCUS_21370	<i>sucB</i>	dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex	C
LOCUS_21380	<i>lpdG</i>	dihydrolipoyl dehydrogenase	C
LOCUS_21540	<i>fabB</i>	beta-ketoacyl-[acyl-carrier-protein] synthase I	IQ
LOCUS_21550	<i>fabA</i>	3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase	I
LOCUS_21560	<i>gpsA</i>	glycerol-3-phosphate dehydrogenase [NAD(P)+]	I
LOCUS_22320	<i>hemH</i>	ferrochelataase	H
LOCUS_22330	-	hypothetical protein	T
LOCUS_22500	-	hypothetical protein	P
LOCUS_22520	<i>tusD</i>	sulfurtransferase TusD	P
LOCUS_23020	-	putative methylaconitate Delta-isomerase PrpF	S
LOCUS_23070	<i>crfX</i>	crfX protein	-
LOCUS_23110	<i>acnB</i>	aconitate hydratase B	C
LOCUS_23150	<i>lpxH</i>	UDP-2%2C3-diacylglucosamine hydrolase	S
LOCUS_23160	<i>ppiB</i>	peptidyl-prolyl cis-trans isomerase	G
LOCUS_23170	<i>glnS</i>	glutamine--tRNA ligase	J
LOCUS_23180	<i>cysS</i>	cysteine--tRNA ligase	J
LOCUS_23190	<i>fold</i>	bifunctional protein FOLD	F
LOCUS_23230	<i>exaC_1</i>	aldehyde dehydrogenase	C

LOCUS_23270	-	hypothetical protein	-
LOCUS_23520	-	hypothetical protein	-
LOCUS_23600	<i>hupB</i>	DNA-binding protein HU-beta	L
LOCUS_23610	<i>ppiD</i>	peptidylprolyl isomerase	O
LOCUS_23670	<i>exaC_2</i>	aldehyde dehydrogenase	C
LOCUS_23700	-	MerR family transcriptional regulator	K
LOCUS_23710	-	mercury transporter	P
LOCUS_23720	<i>merP_1</i>	mercuric transport protein periplasmic component MerP	P
LOCUS_23730	-	membrane protein	S
LOCUS_23800	<i>tniA_1</i>	transposase	L
LOCUS_24090	-	IS5 family transposase	L
LOCUS_24150	-	protein-tyrosine-phosphatase	T
LOCUS_24420	-	hypothetical protein	-
LOCUS_24430	-	hypothetical protein	-
LOCUS_24440	-	hypothetical protein	-
LOCUS_25490	<i>nfuA</i>	Fe/S biogenesis protein NfuA	C
LOCUS_25780	<i>dnaQ</i>	DNA polymerase III subunit epsilon	L
LOCUS_25890	-	enoyl-[acyl-carrier-protein] reductase [NADH]	I
LOCUS_26150	<i>serS</i>	serine--tRNA ligase	J
LOCUS_26200	<i>trxB1</i>	thioredoxin reductase	O
LOCUS_26300	<i>mnmA</i>	tRNA-specific 2-thiouridylase MnmA	J
LOCUS_26320	<i>purB</i>	adenylosuccinate lyase	F
LOCUS_26410	-	hypothetical protein	H
LOCUS_26600	-	anti-anti-sigma factor	T
LOCUS_26610	-	fused response regulator/phosphatase	KT
LOCUS_26640	-	hypothetical protein	S
LOCUS_26810	<i>nrdA</i>	ribonucleoside-diphosphate reductase	F
LOCUS_26880	-	LysR family transcriptional regulator	K
LOCUS_27220	-	glycerol-3-phosphate dehydrogenase	C
LOCUS_27660	<i>eno</i>	enolase	G
LOCUS_27680	<i>pyrG</i>	CTP synthase	F
LOCUS_27690	<i>tilS</i>	tRNA(Ile)-lysidine synthase	D
LOCUS_27700	<i>accA</i>	acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha	I
LOCUS_27710	<i>dnaE</i>	DNA polymerase III subunit alpha	L
LOCUS_27730	<i>lpxB</i>	lipid-A-disaccharide synthase	M
LOCUS_27740	<i>lpxA</i>	acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	I
LOCUS_27750	<i>fabZ</i>	3-hydroxyacyl-[acyl-carrier-protein] dehydratase FabZ	I
LOCUS_27760	<i>lpxD</i>	UDP-3-O-acylglucosamine N-acyltransferase	I
LOCUS_27770	-	hypothetical protein	M
LOCUS_27800	<i>dxr</i>	1-deoxy-D-xylulose 5-phosphate reductoisomerase	I
LOCUS_27810	<i>cdsA-1</i>	phosphatidate cytidyltransferase	I
LOCUS_27820	<i>uppS</i>	ditrans%2Cpolycis-undecaprenyl-diphosphate synthase	H
LOCUS_27830	<i>frr</i>	ribosome-recycling factor	J
LOCUS_27840	<i>pyrH</i>	uridylate kinase	F
LOCUS_27850	<i>tsf</i>	elongation factor Ts	J
LOCUS_27860	<i>rpsB</i>	30S ribosomal protein S2	J
LOCUS_28430	-	hypothetical protein	L
LOCUS_28440	-	hypothetical protein	L
LOCUS_28450	-	IS5 family transposase	L
LOCUS_28460	-	hypothetical protein	L
LOCUS_28470	-	hypothetical protein	L
LOCUS_28920	<i>lpxL</i>	lipid A biosynthesis lauroyltransferase	M
LOCUS_29020	-	IS5 family transposase	L
LOCUS_29030	-	transposase	L
LOCUS_29040	-	transposase	L
LOCUS_29280	-	hypothetical protein	-
LOCUS_29450	<i>gltX</i>	glutamate--tRNA ligase	J

LOCUS_29520	<i>tadA</i>	tRNA-specific adenosine deaminase	FJ
LOCUS_29540	<i>guaA</i>	GMP synthase [glutamine-hydrolyzing]	F
LOCUS_29550	<i>guaB</i>	inosine-5'-monophosphate dehydrogenase	F
LOCUS_29730	<i>cysE_1</i>	serine O-acetyltransferase	E
LOCUS_29740	<i>der</i>	GTPase Der	S
LOCUS_29750	<i>bamB</i>	outer membrane protein assembly factor BamB	M
LOCUS_29770	<i>hisS</i>	histidine--tRNA ligase	J
LOCUS_29780	<i>ispG</i>	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (flavodoxin)	I
LOCUS_29850	<i>hscA</i>	chaperone protein HscA	O
LOCUS_29870	<i>iscA</i>	iron-binding protein IscA	S
LOCUS_29880	-	iron-sulfur cluster assembly scaffold protein IscU	C
LOCUS_29890	<i>iscS</i>	cysteine desulfurase IscS	E
LOCUS_29930	<i>suhB</i>	inositol-1-monophosphatase	G
LOCUS_29960	<i>secD</i>	protein translocase subunit SecD	U
LOCUS_29970	-	preprotein translocase subunit YajC	U
LOCUS_30020	-	LPS export ABC transporter permease LptG	S
LOCUS_30030	-	LPS export ABC transporter permease LptF	S
LOCUS_30070	<i>valS</i>	valine--tRNA ligase	J
LOCUS_30670	-	hypothetical protein	S
LOCUS_30900	<i>groL</i>	60 kDa chaperonin	O
LOCUS_30910	<i>groS</i>	10 kDa chaperonin	O
LOCUS_31130	<i>secA</i>	protein translocase subunit SecA	U
LOCUS_31150	-	hypothetical protein	S
LOCUS_31300	<i>prfA</i>	peptide chain release factor 1	J
LOCUS_31310	<i>hemaA</i>	glutamyl-tRNA reductase	H
LOCUS_31330	<i>lolB</i>	outer-membrane lipoprotein LolB	M
LOCUS_31340	<i>ispE</i>	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	F
LOCUS_31350	<i>prs</i>	ribose-phosphate pyrophosphokinase	F
LOCUS_31370	<i>pth</i>	peptidyl-tRNA hydrolase	J
LOCUS_31420	-	IS5 family transposase	L
LOCUS_31800	-	hypothetical protein	S
LOCUS_31920	-	hypothetical protein	L
LOCUS_31930	-	transposase	L
LOCUS_31940	-	transposase	L
LOCUS_31980	<i>ilvC</i>	ketol-acid reductoisomerase (NADP(+))	EH
LOCUS_32000	<i>ilvI</i>	acetolactate synthase	E
LOCUS_32140	<i>aroQ1</i>	3-dehydroquinate dehydratase 1	E
LOCUS_32150	<i>accB</i>	biotin carboxyl carrier protein of acetyl-CoA carboxylase	I
LOCUS_32160	<i>accC</i>	biotin carboxylase	I
LOCUS_32270	-	hypothetical protein	-
LOCUS_32380	<i>folK-2</i>	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine diphosphokinase	H
LOCUS_32400	<i>panC</i>	pantothenate synthetase	H
LOCUS_32490	<i>pnp</i>	polyribonucleotide nucleotidyltransferase	J
LOCUS_32500	<i>rpsO</i>	30S ribosomal protein S15	J
LOCUS_32530	<i>infB</i>	translation initiation factor IF-2	J
LOCUS_32540	<i>nusA</i>	transcription termination/antitermination protein NusA	K
LOCUS_32590	-	dihydropteroate synthase	H
LOCUS_32600	<i>ftsH_1</i>	ATP-dependent zinc metalloprotease FtsH	O
LOCUS_32640	<i>carB</i>	carbamoyl-phosphate synthase large chain	F
LOCUS_32650	<i>carA</i>	carbamoyl-phosphate synthase small chain	F
LOCUS_32680	<i>dnaK</i>	chaperone protein DnaK	O
LOCUS_32690	<i>grpE</i>	protein GrpE	O
LOCUS_32720	<i>bamE</i>	outer membrane protein assembly factor BamE	M
LOCUS_32840	-	hypothetical protein	L
LOCUS_32850	-	transposase	L
LOCUS_33010	-	group II intron reverse transcriptase/maturase	L
LOCUS_33500	-	hypothetical protein	C

LOCUS_34800	-	hypothetical protein	-
LOCUS_34820	-	hypothetical protein	-
LOCUS_34840	-	IS5 family transposase	L
LOCUS_34850	-	hypothetical protein	L
LOCUS_34860	-	transposase	L
LOCUS_35330	<i>trpS_2</i>	tryptophan--tRNA ligase	J
LOCUS_35360	-	nicotinate-nucleotide diphosphorylase (carboxylating)	H
LOCUS_35510	<i>alr</i>	alanine racemase%2C biosynthetic	E
LOCUS_35520	<i>dnaB</i>	replicative DNA helicase	L
LOCUS_35550	<i>rpsR</i>	30S ribosomal protein S18	J
LOCUS_35560	<i>rpsF</i>	30S ribosomal protein S6	J
LOCUS_35600	<i>purA</i>	adenylosuccinate synthetase	F
LOCUS_35620	-	protein HflC	O
LOCUS_35630	-	protease modulator HflK	O
LOCUS_35650	<i>hfq</i>	RNA-binding protein Hfq	J
LOCUS_35680	<i>amiB</i>	N-acetylmuramoyl-L-alanine amidase	M
LOCUS_35830	<i>ppa</i>	inorganic pyrophosphatase	C
LOCUS_35940	-	transposase	L
LOCUS_35950	-	transposase	L
LOCUS_36180	<i>dxs</i>	1-deoxy-D-xylulose-5-phosphate synthase	H
LOCUS_36790	<i>ybeY</i>	endoribonuclease YbeY	J
LOCUS_36840	<i>leuS</i>	leucine--tRNA ligase	J
LOCUS_36850	<i>lptE</i>	LPS-assembly lipoprotein LptE	M
LOCUS_36860	<i>holA</i>	DNA polymerase III subunit delta	L
LOCUS_36900	<i>lipA</i>	lipoyl synthase	H
LOCUS_36910	<i>lipB</i>	octanoyltransferase	H
LOCUS_36960	-	rod shape-determining protein RodA	M
LOCUS_36970	<i>pbpA</i>	penicillin-binding protein 2	M
LOCUS_37000	<i>nadD</i>	putative nicotinate-nucleotide adenyltransferase	H
LOCUS_37170	-	hypothetical protein	I
LOCUS_37220	-	hypothetical protein	T
LOCUS_37240	<i>parC</i>	DNA topoisomerase 4 subunit A	L
LOCUS_37280	<i>parE</i>	DNA topoisomerase 4 subunit B	L
LOCUS_37560	<i>waaA</i>	3-deoxy-D-manno-octulosonic acid transferase	M
LOCUS_37580	-	FAD-dependent oxidoreductase	E
LOCUS_37590	-	aldo/keto reductase	C
LOCUS_37610	<i>hldE</i>	bifunctional protein HldE	H
LOCUS_37640	-	IS5 family transposase	L
LOCUS_37680	-	hypothetical protein	KLMT
LOCUS_37700	-	IS5 family transposase	L
LOCUS_37710	-	hypothetical protein	L
LOCUS_37720	-	transposase	L
LOCUS_37730	-	transposase	L
LOCUS_37810	-	hypothetical protein	KLT
LOCUS_37830	-	hypothetical protein	KLT
LOCUS_37840	<i>rfaP</i>	lipopolysaccharide core heptose(I) kinase RfaP	H
LOCUS_37850	<i>waaG</i>	UDP-glucose--(heptosyl) LPS alpha 1%2C3-glucosyltransferase WaaG	M
LOCUS_37870	<i>rfaF</i>	lipopolysaccharide heptosyltransferase II	M
LOCUS_38470	-	transcriptional regulator	-
LOCUS_38730	-	IS5 family transposase	L
LOCUS_38900	<i>epd</i>	D-erythrose-4-phosphate dehydrogenase	H
LOCUS_38910	<i>tktA</i>	transketolase	G
LOCUS_38960	<i>metK</i>	S-adenosylmethionine synthase	H
LOCUS_39260	-	putative pre-16S rRNA nuclease	L
LOCUS_39380	-	membrane protein	S
LOCUS_39590	<i>rpoH</i>	RNA polymerase sigma factor RpoH	K
LOCUS_39620	<i>ftsY</i>	signal recognition particle receptor FtsY	U

LOCUS_39760	<i>coaD</i>	phosphopantetheine adenylyltransferase	H
LOCUS_39770	<i>fdx1</i>	4Fe-4S ferredoxin	C
LOCUS_39860	<i>ilvD</i>	dihydroxy-acid dehydratase	E
LOCUS_39970	<i>pstB</i>	phosphate import ATP-binding protein PstB	P
LOCUS_39990	-	phosphate ABC transporter permease	P
LOCUS_40420	<i>ubiF</i>	2-octaprenyl-3-methyl-6-methoxy-1%2C4-benzoquinol hydroxylase	CH
LOCUS_40580	<i>gpmI</i>	2%2C3-bisphosphoglycerate-independent phosphoglycerate mutase	G
LOCUS_41050	-	IS5 family transposase	L
LOCUS_41570	-	hypothetical protein	K
LOCUS_41930	-	hypothetical protein	L
LOCUS_42210	-	Cd(II)/Pb(II)-responsive transcriptional regulator	K
LOCUS_42310	<i>lspA_3</i>	lipoprotein signal peptidase	MU
LOCUS_42320	<i>Int_4</i>	apolipoprotein N-acyltransferase	M
LOCUS_42330	-	hypothetical protein	P
LOCUS_42340	-	sodium:proton antiporter	P
LOCUS_42350	-	Cd(II)/Pb(II)-responsive transcriptional regulator	K
LOCUS_42420	-	hypothetical protein	-
LOCUS_42560	<i>glmS</i>	glutamine--fructose-6-phosphate aminotransferase [isomerizing]	M
LOCUS_42580	<i>glmU</i>	bifunctional protein GlmU	M
LOCUS_42590	<i>atpC</i>	ATP synthase epsilon chain	C
LOCUS_42600	<i>atpD</i>	ATP synthase subunit beta	C
LOCUS_42620	<i>atpA</i>	ATP synthase subunit alpha	C
LOCUS_42630	<i>atpH</i>	ATP synthase subunit delta	C
LOCUS_42650	<i>atpE</i>	ATP synthase subunit c	C
LOCUS_42660	<i>atpB</i>	ATP synthase subunit a	C
LOCUS_42800	<i>yidC</i>	membrane protein insertase YidC	U
LOCUS_42810	<i>rnpA</i>	ribonuclease P protein component	J
LOCUS_42820	-	invertase	L
LOCUS_42840	-	DDE transposase	L
LOCUS_43080	-	hypothetical protein	-
LOCUS_43140	-	hypothetical protein	-
LOCUS_43150	-	hypothetical protein	-
LOCUS_43160	-	hypothetical protein	-
LOCUS_43170	-	hypothetical protein	-
LOCUS_43220	-	hypothetical protein	-
LOCUS_43310	-	hypothetical protein	-
LOCUS_43350	-	cupin	S
LOCUS_43370	-	DUF86 domain-containing protein	S
LOCUS_43400	-	hypothetical protein	-
LOCUS_43410	-	peptide transporter	D
LOCUS_43420	-	hypothetical protein	-
LOCUS_43480	<i>merE_2</i>	mercury resistance protein	S
LOCUS_43490	<i>merD_3</i>	transcriptional regulator	K
LOCUS_43500	<i>merA_3</i>	putative mercuric reductase	C
LOCUS_43510	-	membrane protein	S
LOCUS_43520	<i>merP_3</i>	mercuric transport protein periplasmic component MerP	P
LOCUS_43590	-	hypothetical protein	-
LOCUS_43620	-	cupin	S
LOCUS_43640	-	DUF86 domain-containing protein	S
LOCUS_43660	-	invertase	L
LOCUS_43670	-	hypothetical protein	S

Table S 4.8 Significant and natural transformation associated (highlighted in grey) loci identified by comparing DNA-transformed and recipient libraries with Resampling method. The loci with adjusted p-value lower than 0.05 are highlighted in bold. Results from the survival essentiality analysis with Gumbel method is also presented. E: essential; NE: non-essential; U: uncertain.

ORF	Name	Description	P-value	P-value adjusted	Gumbel analysis
LOCUS_00010	dnaA	chromosomal replication initiator protein DnaA	0.01	0.87	E
LOCUS_00020	-	DNA polymerase III subunit beta	0.01	0.87	E
LOCUS_00030	recF	DNA replication and repair protein RecF	0.01	0.87	NE
LOCUS_00040	gyrB	DNA gyrase subunit B	0.01	0.87	E
LOCUS_00050	-	hypothetical protein	0.01	0.87	NE
LOCUS_00060	-	hypothetical protein	0.01	0.87	NE
LOCUS_00070	-	hypothetical protein	0.01	0.87	NE
LOCUS_00080	-	hypothetical protein	0.01	0.87	NE
LOCUS_00090	-	hypothetical protein	0.01	0.87	E
LOCUS_00100	-	group II intron reverse transcriptase/maturase	0.01	0.87	E
LOCUS_00110	-	hypothetical protein	0.01	0.87	E
LOCUS_00120	-	hypothetical protein	0.01	0.87	NE
LOCUS_00130	-	hypothetical protein	0.01	0.87	NE
LOCUS_00140	-	hypothetical protein	0.01	0.87	NE
LOCUS_00150	-	hypothetical protein	0.01	0.87	NE
LOCUS_00160	-	hypothetical protein	0.01	0.87	NE
LOCUS_00170	-	transcriptional regulator	0.01	0.87	U
LOCUS_00180	-	RelE toxin	0.00	0.03	NE
LOCUS_00190	-	DEAD/DEAH box helicase	0.01	0.87	NE
LOCUS_00200	ahpF_1	alkyl hydroperoxide reductase subunit F	0.00	0.00	NE
LOCUS_00210	lptA_1	lysophosphatidic acid acyltransferase	0.00	0.00	E
LOCUS_00220	-	D%2CD-heptose 1%2C7-bisphosphate phosphatase	0.00	0.03	E
LOCUS_00230	glyS	glycine--tRNA ligase beta subunit	0.00	0.00	E
LOCUS_00240	glyQ	glycine--tRNA ligase alpha subunit	0.00	0.00	E
LOCUS_00250	tag	DNA-3-methyladenine glycosidase I	0.00	0.03	NE
LOCUS_00260	-	lipid A biosynthesis lauroyl acyltransferase	0.00	0.03	NE
LOCUS_00270	-	PilZ domain-containing protein	0.00	0.03	NE
LOCUS_00280	-	TPR repeat-containing protein	0.00	0.03	NE
LOCUS_00290	-	putative amino-acid ABC transporter ATP-binding protein y4tH	0.00	0.03	NE
LOCUS_00300	-	ectoine/hydroxyectoine ABC transporter permease subunit EhuD	0.00	0.00	NE
LOCUS_00310	-	ectoine/hydroxyectoine ABC transporter permease subunit EhuC	0.00	0.00	NE
LOCUS_00320	-	ectoine/hydroxyectoine ABC transporter substrate-binding protein EhuB	0.00	0.03	NE
LOCUS_00330	trkA	potassium transporter inner membrane associated protein	0.00	0.03	E
LOCUS_00340	-	ribosomal RNA small subunit methyltransferase B	0.00	0.00	U
LOCUS_00350	fmt	methionyl-tRNA formyltransferase	0.00	0.05	U
LOCUS_10320	pilA	fimbrial protein	0.76	0.87	NE
LOCUS_10350	pilC	type II secretion system protein F	0.75	0.87	NE
LOCUS_16450	-	DNA internalization-related competence protein ComEC/Rec2	0.78	0.87	NE
LOCUS_28150	recA	protein RecA	0.74	0.87	U
LOCUS_39340	pilU	twitching motility protein PilU	0.76	0.87	NE
LOCUS_39350	pilT	twitching mobility protein	0.60	0.87	NE

Table S 4.9 Significant and natural transformation associated (highlighted in grey) loci identified by comparing donor-transformed and recipient libraries with Resampling method. No locus has an adjusted p-value is lower than 0.05. Results from the survival essentiality analysis with Gumbel method is also presented. E: essential; NE: non-essential; U: uncertain.

ORF	Name	Description	P-value	P-value adjusted	Gumbel analysis
LOCUS_00200	ahpF_1	alkyl hydroperoxide reductase subunit F	0.04	0.31	NE
LOCUS_00210	lptA_1	lysophosphatidic acid acyltransferase	0.04	0.31	E
LOCUS_00230	glyS	glycine--tRNA ligase beta subunit	0.04	0.31	E
LOCUS_00240	glyQ	glycine--tRNA ligase alpha subunit	0.04	0.31	E
LOCUS_00250	tag	DNA-3-methyladenine glycosidase I	0.04	0.31	NE
LOCUS_00260	-	lipid A biosynthesis lauroyl acyltransferase	0.04	0.31	NE
LOCUS_00270	-	PilZ domain-containing protein	0.03	0.31	NE
LOCUS_00280	-	TPR repeat-containing protein	0.03	0.31	NE
LOCUS_00290	-	putative amino-acid ABC transporter ATP-binding protein y4tH	0.03	0.31	NE
LOCUS_00300	-	ectoine/hydroxyectoine ABC transporter permease subunit EhuD	0.05	0.31	NE
LOCUS_00310	-	ectoine/hydroxyectoine ABC transporter permease subunit EhuC	0.04	0.31	NE
LOCUS_00320	-	ectoine/hydroxyectoine ABC transporter substrate-binding protein EhuB	0.04	0.31	NE
LOCUS_00330	trkA	potassium transporter inner membrane associated protein	0.04	0.31	E
LOCUS_00340	-	ribosomal RNA small subunit methyltransferase B	0.04	0.31	U
LOCUS_00350	fmt	methionyl-tRNA formyltransferase	0.03	0.31	U
LOCUS_10320	pilA	fimbrial protein	0.21	0.31	NE
LOCUS_10350	pilC	type II secretion system protein F	0.23	0.31	NE
LOCUS_39350	pilT	twitching mobility protein	0.21	0.31	NE
LOCUS_39340	pilU	twitching motility protein PilU	0.22	0.31	NE
LOCUS_16450	-	DNA internalization-related competence protein ComEC/Rec2	0.23	0.31	NE
LOCUS_28150	recA	protein RecA	0.20	0.31	U

Part II:

**Phage and Phage-like elements
of *Paracoccus versutus***

Chapter 5 - Characterization of a novel bacteriophage, PvEP4A, with lytic activity against *P. versutus*

5.1 Foreword

In this chapter, I describe PvEP4A, a new-isolated phage infecting *Paracoccus versutus*, and investigate a mechanism conferring resistance to its host.

The key findings of this chapter are the following:

- PvEP4A is a double-stranded DNA tailed bacteriophage of the *Caudoviricetes* class.
- PvEP4A is more closely related to *Rhodobacter* phages than to the *Paracoccus* phages described to date.
- Resistance to PvEP4A emerged quickly in the host *P. versutus* through a short deletion in the *cpdB* gene, coding for a cyclic-dinucleotide diesterase.
- Deletion in the *cpdB* gene leads to an increased intracellular concentration of c-di-GMP, whose effects are pleiotropic, based on comparative transcriptomics of the wild-type and *cpdB* strains.
- The expression of flagella, capsular polysaccharides, and outer membrane receptors is decreased in the *cpdB* mutant, which could explain the phage-resistance phenotype.

5.2 Abstract

Paracoccus species, ubiquitous in terrestrial and aquatic environments, face various challenges including predation by bacteriophages. This study reports the isolation and characterization of PvEP4A, a novel double-stranded DNA tailed bacteriophage belonging to the *Caudoviricetes* class, from a Swiss wastewater treatment plant. This phage infects only its initial host *Paracoccus versutus* DSM582 out of the 14 *Paracoccus* strains. While PvEP4A shares greater genetic similarity with *Rhodobacter* phages than with characterized *Paracoccus* phages, it showed no infectivity toward the tested *Rhodobacter capsulatus* strain. Host resistance to PvEP4A emerged rapidly in its host through a small deletion in the *cpdB* gene, encoding a cyclic-dinucleotide diesterase. This deletion resulted in elevated intracellular concentration of c-di-GMP, a bacterial second messenger that regulates various phenotypes, including motility, biofilm formation, virulence, and cell cycle progression. A transcriptomic analysis showed that the expression of flagella, capsular polysaccharides, and outer membrane receptors is decreased in the *cpdB* mutant, which could explain the phage-resistance phenotype. This research contributes to our understanding of *Paracoccus* phages and provides insights into the complex interactions between these bacteria and their viral parasites.

5.3 Introduction

The genus *Paracoccus*, a member of the alphaproteobacteria class, encompasses a number of metabolically adaptable species. These organisms have demonstrated an extraordinary capacity to thrive in a wide range of environmental niches such as soil, rhizosphere, groundwater, sea water, marine sediments, and sewage treatment plants; some species are even known to cause infections in humans (Daneshvar *et al.*, 2003; Lee *et al.*, 2004; Kelly *et al.*, 2006; Liu *et al.*, 2008; Lasek *et al.*, 2018). In these various environments, bacteria encounter predators and parasites, particularly bacteriophages, which are the most abundant biological entities on Earth (Stone *et al.*, 2019).

Phages play a critical role in shaping bacterial populations. These viral predators require live bacterial hosts for efficient replication, thus exerting a significant influence on host abundance, diversity, physiology, and metabolism (Koskella *et al.*, 2022). Furthermore, through horizontal gene transfer via transduction, phages contribute to microbial diversity, making them essential drivers of bacterial evolution (Chiang *et al.*, 2019). Despite significant advances in our understanding, the complex dynamics of phage ecology and the intricacies of their infection mechanisms remain active areas of research, offering rich opportunities for further scientific exploration.

Research on phages specific to *Paracoccus* has only recently begun (Table 5.1). The first documented *Paracoccus* phage, ϕ Pam-6, a temperate phage infecting *Paracoccus aminophilus*, was described in 2014 (Dziewit *et al.*, 2014). Subsequently, other researchers identified additional phages, including the lytic phage Shpa, which infects *Paracoccus marinus*, and the lytic phage, vB_PmaS-R3, associated with *Paracoccus aminophilus* (Xu *et al.*, 2015; Van Zyl *et al.*, 2016). A recent comprehensive study has significantly expanded our understanding of the *Paracoccus* virome, revealing five novel active temperate phages and identifying 53 prophages through genomic analysis of various *Paracoccus* species (Decewicz *et al.*, 2019). This growing body of research is beginning to illuminate the diverse and complex phage-host interactions in *Paracoccus* genus with for instance two prophages sharing 99% sequence identity, present in two separate species, from different continents and isolated 10 years apart (Decewicz *et al.*, 2019).

Table 5.1 General characteristics and features of known *Paracoccus* phage genomes.

Phage	Host	Type	Morph.	Size (bp)	GC (%)	Nr. Genes	Reference
φPam-6	<i>P. aminophilus</i>	temp.	S	38'771*	61.4*	46	(Dziewit <i>et al.</i> , 2014)
vB_PmaS-R3	<i>P. marcusii</i>	lytic	S	42'093	56.4	52	(Xu <i>et al.</i> , 2015)
Shpa	<i>P. marinus</i>	lytic	S	38'261	64.5	56	(Van Zyl <i>et al.</i> , 2016)
vB_PbeS_Pben1	<i>P. bengalensis</i>	temp.	S	39'879	64.3	71	(Decewicz <i>et al.</i> , 2019)
vB_PkoS_Pkon1	<i>P. kondratievae</i>	temp.	S	49'723	60.6	79	(Decewicz <i>et al.</i> , 2019)
vB_PsuS_Psul1	<i>P. sulfuroxidans</i>	temp.	S	37'901	60.9	57	(Decewicz <i>et al.</i> , 2019)
vB_PthS_Pthi1	<i>P. thiocyanatus</i>	temp.	S	39'547	63.8	52	(Decewicz <i>et al.</i> , 2019)
vB_PyeM_Pyei1	<i>P. yeei</i>	temp.	M	50'161	65.5	75	(Decewicz <i>et al.</i> , 2019)

* These data were estimated based on the phage region prediction of PHASTEST (Wishart *et al.*, 2023) as no sequence was provided in the reference. Abbreviations: Morph: morphotype; S: Siphovirus; M : Myovirus.

In this study, we report the isolation and characterization of a new bacteriophage, PvEP4A, infecting *P. versutus* DSM582. This phage has a narrow host range, targeting only the host used for its isolation among 14 *Paracoccus* strains tested. Interestingly, *P. versutus* DSM582 quickly developed phage resistance through a deletion of nucleotides in the *cpdB* gene, coding for a cyclic-dinucleotide diesterase. This mutation led to an increase in intracellular c-di-GMP concentration affecting multiple aspects of cell physiology, notably motility, biogenesis of cell wall and membrane, and metabolism. Comparative transcriptomics of the wild-type and *cpdB* strains showed that the expression of flagella, capsular polysaccharides, and outer membrane receptors is decreased in the *cpdB* mutant, which could explain the phage-resistance phenotype. This research reports the description and characterization of a new phage, expanding the limited collection of phages that infect members of the *Paracoccus* genus and providing insights into the interactions between *Paracoccus* species and this phage.

5.4 Materials & Methods

5.4.1 Bacterial strains and culture conditions

The bacterial strains, plasmids, and primers used in this chapter are listed in Table S 5.1. The bacteria were routinely grown in LB broth (tryptone, 10 g/L; yeast extract 5 g/L; NaCl, 5 g/L; tryptone and yeast extract from Oxoid) or on LB supplemented with 1.5% agar. When required, the media were supplemented with kanamycin (Km, 50 µg/mL), gentamicin (Gm,

30 µg/mL), and isopropyl-β-d-thiogalactopyranoside (IPTG, 1mM). Saline magnesium buffer (SM buffer) (50 mM Tris at pH 7.5, 100 mM NaCl, and 8 mM MgSO₄) was used to collect phages from double-layer agar assays and for dilution (Carroll-Portillo *et al.*, 2021). In the case of experiments with phages, the medium was supplemented with 10 mM MgSO₄ and 1 mM CaCl₂ to facilitate phage infection (Tucker, 1961; Fortier and Moineau, 2009). *E. coli* strains were grown at 37°C; *Paracoccus* strains were grown at 30°C; for liquid cultures, an agitation of 180 rpm was applied.

5.4.2 Strains and plasmid construction

Plasmid construction and amplification were done in *E. coli* DH5α-λpir. *E. coli* chemocompetent cells were prepared and transformed according to standard CaCl₂-based protocols (Sambrook and Russell, 2001). Electroporation into *P. versutus* was done at 25 µF, 400 Ω, and 2.5 kV after two washes of an overnight culture with a 15% glycerol, 1 mM MOPS buffer. Stocks of all strains were conserved at –80° C in 20% glycerol.

The plasmid pIND4-cpdB was constructed by Gibson assembly with the NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs)(Gibson *et al.*, 2009). Briefly, plasmid pIND4 sequence and *cpdB* were amplified separately with overlapping regions by PCR with the Q5® High-Fidelity DNA Polymerase (New England Biolabs). After purification of PCR products with NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel), they were assembly by Gibson assembly, and the obtained plasmids were transformed into chemocompetent *E. coli*. The transformants were selected on kanamycin. The constructs were confirmed by Sanger sequencing (Microsynth) before being transformed into *P. versutus* strains. The plasmid pIND4 was also transformed into *P. versutus* strains as a control.

5.4.3 Isolation of environmental strains of *Paracoccus versutus*

Paracoccus versutus strains were isolated from the wastewater treatment plant of Vuippens, in Switzerland, specifically from the sludge digestion tank. Samples from different basins were collected and spread on a minimal medium for thiosulfate-oxidizing *Paracoccus* species with thiosulfate and gentamicin to favor the growth of *P. versutus* (Kelly *et al.*, 2006). We isolated dominant colonies and performed PCR first with primers specific to *P. versutus*, and in case of a positive result, we then amplified the 16S rDNA and *rpoB* genes, that were sequenced by Sanger sequencing (Microsynth). Sequence identification was performed using the basic local

alignment search tool (BLAST) (Altschul *et al.*, 1990). With this procedure, we confirmed the isolation of three new strains of *P. versutus*: P2, P45, and P46.

5.4.4 Phage isolation and propagation

To isolate new phages, samples were collected from various tanks (aeration tank, anoxic tank, tank for sludges) of the wastewater treatment plant of la Chaux-de-Fonds, in Switzerland. They were passed through a filter with a 0.22 µm pore size and mixed (10%, v/v) with an overnight culture of *P. versutus* in liquid LB in multiple replicates to promote enrichment of diverse bacteriophages. After 24h of incubation at 30°C, the cultures were mixed, filtered and used in a conventional double-layer agar plate assay for phage plaque detection. Briefly, 10 µL of serially diluted phage suspension was mixed with 50 µL of log-phase host bacteria and 5 mL of molten LB soft agar (0.5% w/v) supplemented with 10 mM MgSO₄ and 1 mM CaCl₂, then overlaid onto 1.5% w/v LB agar plates. Individual plaques were collected, resuspended in SM buffer, and streaked out for subculture. This was repeated three times to obtain pure phage isolates. The phages were propagated using double-agar overlays with multiple Petri dishes to achieve the confluent lysis of the bacterial lawn. Top soft agar layers containing lysed bacterial lawns and phage progeny were collected, vortexed for ~30 s after the addition of 5 mL SM buffer per soft agar plate, incubated at RT for ~ 15 min, and then centrifuged at 6000 g for 10 min before carefully decanting supernatant which was subsequently filtered through a 0.22 µm pore size syringe filter. This method allows the obtention of concentrated phage lysate.

5.4.5 Transmission Electron Microscopy (TEM) of PvEP4A

A concentrated phages lysate of PvEP4A in SM buffer was attached to the grid and negatively stained with 1% uranyl acetate for 1 min. After air-drying for 10 min, TEM (Philips CM100 TEM) was performed at 80 kV. The average size of the head and contractile tail was measured using seven phages from the TEM scans with Fiji (Schindelin *et al.*, 2012).

5.4.6 Phage genome extraction, sequencing and annotations

Phage DNA extraction was performed as described hereafter. First, 1 mL of concentrated phage lysate was first treated with 2 µL DNase RQ1 (Promega)(1 U/µL) and 2 µL RNase A (10 mg/mL) for 1.5 h at 37°C to remove any residual bacterial DNA and RNA. After inactivation of the enzymes with EDTA (20 mM) and heating at 65°C for 10 min, the capsid was digested with 1.25 µL Proteinase K (20 mg/mL) for 1.5 h at 56°C. Then, the phage DNA was extracted

with the Genomic-tip 20/G (Qiagen) according to the manufacturer's guidelines. The complete genome was sequenced by Novogene (UK), using the Illumina Hiseq2500 platform. The genome was assembled using SPAdes (3.15.5) (Prjibelski *et al.*, 2020) and circularity was controlled with Bandage (Wick *et al.*, 2015).

To predict and annotate the open reading frames (ORFs), the complete genome was submitted to various annotation tools: Phastest (Wishart *et al.*, 2023), PhaBOX (Shang, Peng, *et al.*, 2023), PhageScope (Wang, Yang, *et al.*, 2024), GeneMark (Besemer and Borodovsky, 2005), NCBI BLASTp (Altschul *et al.*, 1997), HHPred (Söding *et al.*, 2005), Glimmer (Delcher *et al.*, 2007), and Conserved Domain Database (Marchler-Bauer *et al.*, 2011). Transfer RNA was detected using tRNAscan-SE v.2.0 (Chan and Lowe, 2019) and ARAGORN (Laslett and Canback, 2004). The genomic data were manually annotated using DNA master (Lawrence, 2007).

5.4.7 Comparative genome analysis and phylogenetic tree

Genome sequences of the phages were obtained from the GenBank database: RcCronus (GCF_002756615.1), RcRhea (GCF_001501915.1), RcSaxon (GCA_002756895.1), RcSpartan (GCF_002623345.1), RcTitan (GCF_001551025.1), RcZahn (GCA_020492535.1), vB_PbeS_Pben1 (GCA_004015625.1), vB_PkoS_Pkon1 (GCA_004015645.1), vB_PsuS_Psul1 (GCA_004015665.1), vB_PthS_Pthi1 (GCA_004015685.1), vB_PyeM_Pyei1 (GCA_004015805.1), vB_PmaS-R3 (GCF_000954255.1), Shpa (GCF_002605745.1). Phylogenetic analysis was conducted using the Virus Classification and Tree Building Online Resource (VICTOR) (Meier-Kolthoff and Göker, 2017), which employs the Genome-BLAST Distance Phylogeny method. VICTOR calculated intergenomic distances between the input sequences, performing 100 pseudo-bootstrap replicates to ensure statistical robustness. These distances were then used to infer a balanced minimum evolution tree using FASTME, with branch support values included. The tree construction process incorporated subtree pruning and regrafting postprocessing for the D0 formula. The resulting phylogenetic tree was subsequently visualized and formatted using RStudio (RStudio Team, 2020). For conserved gene phylogeny construction, we utilized terminase large subunit genes and did a CLUSTAL Multiple sequence alignment using the MUSCLE alignment software (Edgar, 2004). The phylogenetic tree was built according to maximum likelihood algorithm on IQ Tree (Nguyen *et al.*, 2015) and visualized using RStudio (RStudio Team, 2020) with the help of package *ggtree* (Yu, 2020).

5.4.8 Phage host range determination

To determine bacterial susceptibility to phage PvEP4A, we used double-layer agar plates. For that purpose, 14 *Paracoccus* strains (*Paracoccus denitrificans* DSM1407; *Paracoccus pantotrophus* DSM1404 and DSM65; *Paracoccus aminophilus* DSM8538; *Paracoccus aestuarii* DSM19484; *Paracoccus alkenifer* DSM11593M; *Paracoccus siganidrum* DSM26381; *Paracoccus yeei* KB2, KB3, and KB4; *P. versutus* DSM582 and environmental strains P2, P45, P46) and *R. capsulatus* DSM1710 were grown in liquid LB medium, diluted 100-fold in soft LB agar (0.5% agar) supplemented with 10 mM MgSO₄ and 1mM CaCl₂, and finally plated onto LB agar plates. After solidification of the top agar, a drop of phage suspension (concentration of about 10⁹ PFU/mL) was spotted on the bacterial lawns, and the plates were incubated at 30°C. The plates were examined for evidence of bacterial lysis after 18-24h.

5.4.9 Isolation of a phage-resistant mutant, its genetic and phenotypic characterization

To evaluate the phage infectivity, we mixed diluted overnight culture of *P. versutus* DSM582 at an optical density (OD) measured at 600nm of 0.6 with multiple concentrations of phage and monitored the OD₆₀₀ at 30°C in a plate reader (Biotek Cytation5) for two days. As we observed regrowth in the treatment with high phage concentration (corresponding to approximately 10⁶ PFU/mL), we isolated individual colonies and tested again their susceptibility to phage with double-agar overlays. They were all resistant. We extracted the DNA from one phage-resistant clone and from the phage-susceptible strain as control with QiaAMP DNA mini kit (Qiagen) and sent it to Illumina whole genome sequencing to evaluate the variations between both strains. Sequence alignment was performed using Bowtie2 (Langmead and Salzberg, 2012) to map reads from the mutant and the wild-type to the wild-type reference genome. Variant calling and annotation were conducted using Freebayes (Garrison and Marth, 2012) through the Bacterial and Viral Bioinformatics Resource Center platform (Olson *et al.*, 2023). To validate the results, a comparative mapping analysis was carried out using BWA-MEM (Md *et al.*, 2019), which yielded consistent outcomes. The analysis identified a single significant variation between the wild-type and phage-resistant strains: a deletion in LOCUS_38800 resulting in a frameshift mutation and an earlier stop codon. To confirm the effect of the deletion in LOCUS _38800, we constructed strains with

complementation of the deletion and tested their phage adsorption, motility and biofilm production abilities.

To evaluate the adsorption of phages, overnight cultures of bacterial strains were diluted to an OD₆₀₀ of 1. Phage lysate was added to a final multiplicity of infection (MOI) of 0.00001, followed by a 10 min adsorption period. After 5 min of centrifugation at 13'000 rpm, the supernatant was removed and used in double-layer plaque assay to quantify the residual plaque forming units (PFU).

To evaluate the production of biofilm, overnight cultures in LB were washed two times and then normalized to an OD₆₀₀ of 0.1 in 5 mL LB medium. Then, the new cultures were distributed in flat-bottom transparent 96-well plates and incubated without shaking at 30°C for 2 days. To normalize the production of biofilm with the growth of bacterial culture, the absorbance was first measured at 600nm in a SpectraMax i3X plate-reader (Molecular Devices). For biofilm staining, liquid cultures were removed by pipetting and wells were washed 3x with 200 µL of MilliQ water. The biofilm was then stained with 150 µL crystal violet 1% (w/v) solubilized in a 30% (v/v) acetic acid solution for 15 min at 30°C. Then, crystal violet was removed, and wells were washed 3x with 200 µL MilliQ water. The 150 µL of 30% acetic acid solution was added to the wells to destain the Crystal violet and incubated at room temperature for 15 min with shaking at 300 rpm (Heidolf Vibramax 100). Finally, 100 µL were transferred to a new flat-bottom 96-well plate and absorbance was measured at 595 nm in a SpectraMax i3X plate-reader (Molecular Devices). Crystal Violet retention was measured at 595 nm and normalized by culture density measured at 600 nm.

Bacterial motility was assessed using a double layer agar plate method. The plates were prepared with a base layer consisting of 1.5% (w/v) agar in LB diluted 5-fold, overlaid with a thin layer (8mL) of 0.3% (w/v) soft agar in LB diluted 5-fold. Overnight bacterial cultures were normalized to an OD₆₀₀ of 0.3 and inoculated into the agar by stabbing vertically through the top layer using a sterile straight pin. The inoculated plates were incubated for 48 hours at 20°C to favor motility over growth. Following incubation, motility was quantified by measuring the diameter of the bacterial spread zone surrounding the inoculation site. To minimize variation due to plate humidity, all strains were inoculated on the same plate. Additionally, all assays were performed in triplicate to ensure reproducibility.

5.4.10 Quantification of c-di-NMP and their metabolites

To quantify concentration of c-di-NMP and their metabolites, overnight cultures of *P. versutus* wild-type and its *cpdB* mutant were used to inoculate three new cultures at an $OD_{600} = 0.001$ in LB and incubated overnight at 30°C, until cultures reached an OD_{600} of 0.5. The extraction of c-di-NMP and their metabolites was based on the protocol provided by Biolog Life Science Institute GmbH & Co. KG. Briefly, bacterial cells from 5 mL culture were washed twice with LB medium before being pelleted. The pellets were resuspended in 300 µL ice-cold extraction solvent (acetonitrile/methanol/water; 2/2/1, v/v/v (HPLC grade)) and incubated for 15 min on ice. Then, the suspension was incubated for 10 min at 95°C before being cooled again on ice. After 10 min centrifugation at 4°C and 20'800 g, the supernatant was transferred into a 2 mL tube. The extraction was repeated twice with 200 µL extraction solvent without heating at 95°C. The residual pellet was kept aside for protein quantification. Combined supernatants were stored at -20°C overnight to precipitate residual proteins which were removed the following day by 10 min centrifugation at 4°C and 20'800 g. The supernatant was transferred to a new tube, evaporated at 45°C and 1 mbar in CentriVap Benchtop Vacuum Concentrator (Labconco) to dryness before sending to Biolog Life Science Institute GmbH & Co. KG for quantification of 3'3' c-di-AMP, 3'3' c-di-GMP, 3'3' cGAMP, 2'3'-cGAMP, 3',5'-cGMP, 3',5'-cAMP, pGpG, and pApA by LC-MS/MS. Protein quantification of cell cultures was performed with the pellets from the metabolites' extraction. Pellets were resuspended in 800 µL 0.1 M NaOH, vortexed and heated for 15 min at 95°C. Then, 10 µL supernatant were used for BCA assay (triplicates), following instructions of BCA Protein Assay Kit (Thermoscientific).

5.4.11 RNA extraction, sequencing and annotation

Three cultures (A, B, C) of *P. versutus* wild-type and its *cpdB* mutant, inoculated in LB, were grown overnight. Each culture was used to inoculate three new cultures at an $OD_{600} = 0.01$ in LB and incubated for 20h at 30°C. The cultures were pooled three by three by mixing initial cultures (for example A1 with B1 and C1), sampled ($OD_{600} = 0.8$, 1 ml) and condensed in a pellet immediately frozen at -80°C. For RNA extraction, samples were lysed with 1 mL TRIzol Reagent (Thermo Fisher Scientific) at 65°C for 10 min; 200 µL of chloroform was added and, after centrifugation, 600 µL of the upper phase were transferred to a new tube. To precipitate the RNA, the same volume of LiCl (7.5 M) was added to the upper phase, mixed and incubated

at -20°C for at least 30 minutes. After centrifugation at maximum speed for 10 minutes at 4°C, the precipitated RNA was washed with 70% EtOH and air dried before being resuspended in RNase-free H₂O. RNA was treated with DNase I (RNase-Free DNase Set, Qiagen) and RNasin® Ribonuclease Inhibitor (Promega) before clean-up with RNeasy Mini Kit (Qiagen) and a DNase I on-column treatment. RNA was quantified using Nanodrop2000 (Thermo Fisher Scientific) and Qubit 2.0 fluorometer (Qubit RNA Broad Range Assay kit, Thermo Fisher Scientific), checked for degradation on an agarose gel, and checked for possible DNA contamination by qPCR using SybrGreen (QuantiNova SYBR Green PCR kit, Qiagen) on a Rotor-Gene Q thermocycler (Qiagen). Total RNA-sequencing was carried out by Novogene after ribosomal RNA depletion on Illumina NovaSeq systems after library construction. An average of 6-10 million 150 bps pair-end reads (Novogene) were obtained per sample. Reads were filtered, trimmed of the first 12 bps low-quality sequences, and searched for residual adapter sequences using *fastp* (v. 0.20.0) (Chen *et al.*, 2018). Remaining reads were mapped onto the *P. versutus* DSM581 genomes using *bowtie2* (v. 2.3.4.3) (Langmead and Salzberg, 2012). Reads were counted using *featureCounts* (v. 2.0.0) (Liao *et al.*, 2014) with default count parameters. Analysis was performed in RStudio (RStudio Team, 2020) using the *DESeq2* (v. 1.38.3) package (Love *et al.*, 2014) with *lfcShrink* function (type=«apeglm»). Enrichment analysis was performed in R using the *clusterProfiler* package (v. 4.6.2) (Wu *et al.*, 2021) with *pAdjustMethod* = 'BH' on genes annotated with EGGnog V6.0 (Hernández-Plaza *et al.*, 2023).

The following analysis was performed to identify several genes potentially associated with phage reception. The Type IV pilus homologs were identified by blasting proteins sequences of PilA (NP_253215.1), PilB (NP_253216.1), PilC (WP_216625754.1), PilD (NP_253218.1), PilE (NP_253246.1) and PilT (NP_249086.1) from *Pseudomonas aeruginosa* and pilA (YP_002518416.1), CpaA (YP_002518415.1), CpaB (YP_002518414.1), CpaC (YP_002518413.1), CpaD (YP_002518412.1), CpaE (YP_002518411.1), CpaF (YP_002518410.1), CpaG (YP_002518409.1), CpaH (YP_002518408.1), CpaI (YP_002518417.1), CpaJ (YP_002518418.1), and CpaK (YP_002518419.1) from *Caulobacter vibriodes* NA1000 on proteins sequences of *P. versutus* DSM582. Genes involved in lipopolysaccharide biogenesis and transport were identified based on EGGnog annotation. Capsule-associated genes were detected with the tool CapsuleFinder (Rendueles *et al.*, 2017). This tool detects various capsule systems (Table S 5.2). To identify other potential phage

receptor candidates, we blasted the protein sequences of known phage receptors from the phageReceptor database of the dsDNA viruses (Table S 5.3) on the protein sequences of *P. versutus* DSM582 (Zhang *et al.*, 2020). Transcription analysis combined with literature review allows us to determine the potential of some loci as playing a role in the reception of PvEP4 phage.

5.4.12 Visual representation and statistical analysis

The statistical analyses and graphics were carried out using *RStudio* software (RStudio Team, 2020). Graphics were designed with *ggplot2* and base R functions (Wickham, 2008). Regarding statistical analyses, the normality of the data distribution was tested with Shapiro-Wilk normality test, the equality of variances with F-test (when two groups) or Levene's test (more than two groups). Finally, means were compared either with Student's t-test/Wilcoxon rank test or ANOVA with post-hoc Tukey or Dunnett tests. The significance is indicated as following: ns: p-value >0.05; *: p-value <0.05; **: p-value <0.01; ***: p-value <0.001. Protein three-dimensional structures were predicted using AlphaFold2 via ColabFold (Mirdita *et al.*, 2022). Molecular graphics and analyses of these protein structures were performed with the UCSF ChimeraX (Meng *et al.*, 2023).

5.5 Results and Discussion

5.5.1 Characterization of the bacteriophage PvEP4A: genomics, morphology, and interaction with hosts

A novel bacteriophage, designated PvEP4A, was isolated from the wastewater treatment plant of La Chaux-de-Fonds (Switzerland), using *Paracoccus versutus* DSM582 as the host organism. To characterize this new phage, we conducted a comprehensive phenotypic and genomic analysis.

Transmission electron microscopy (TEM) examination of seven virions revealed that PvEP4A exhibits a tailed morphology. The phage possesses an icosahedral capsid with a diameter of 43 ± 5 nm (mean $\pm$ standard deviation) and a tail measuring 232 ± 25 nm in length and 21 ± 3 nm in width (Figure 5.1A). These morphological characteristics are consistent with the myovirus morphotype, which is characterized by a contractile, rigid tail ranging from 80 to 455 nm in length and 16 to 20 nm in width (King *et al.*, 2011). All seven observed virions displayed a straight tail, indicative of its rigidity, and one micrograph revealed what appeared to be

contracted tail structures (Figure 5.1B). Overall, all these characteristics are consistent with the morphotype of myoviruses.

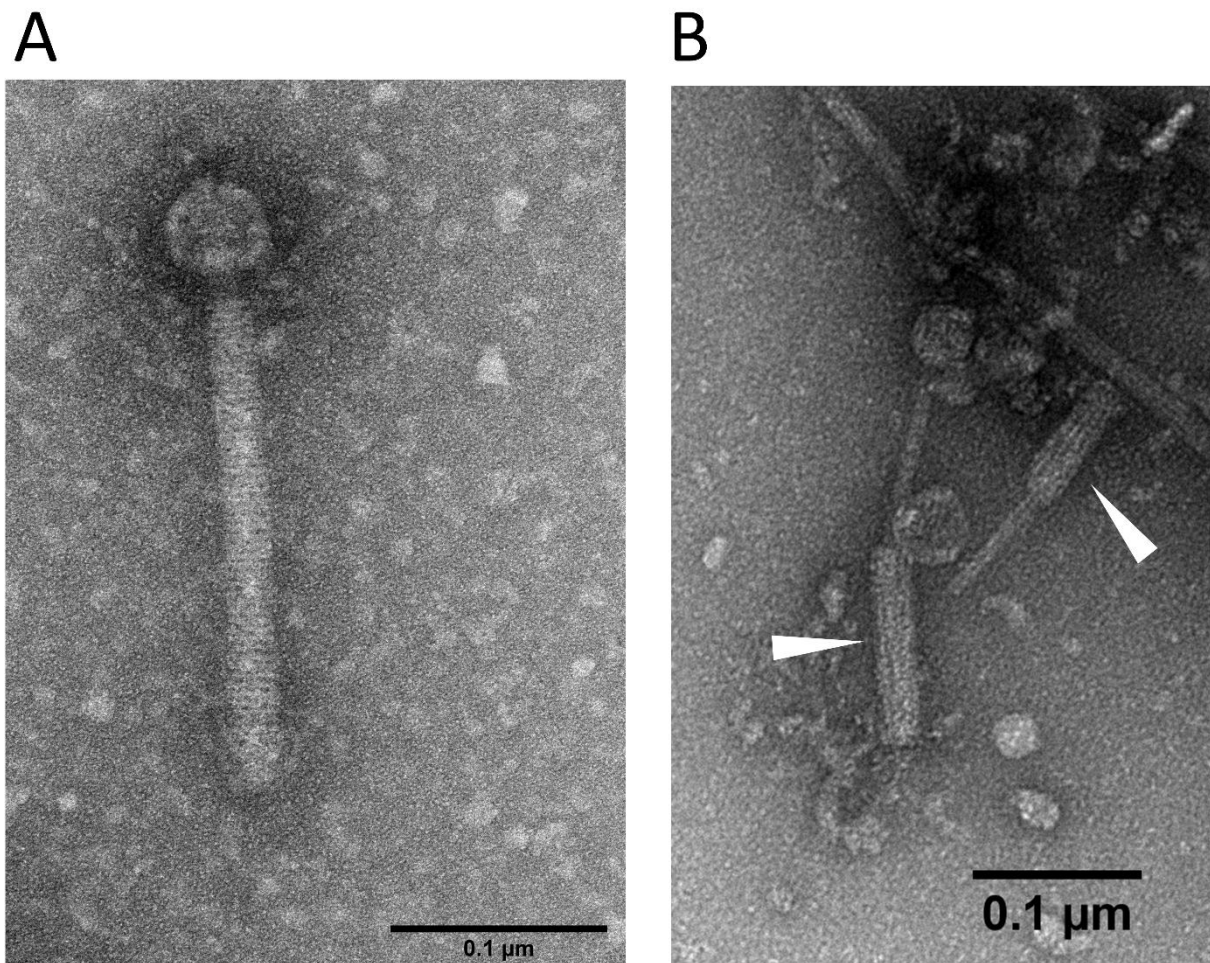


Figure 5.1 TEM observation of the phage PvEP4A with (A) complete virion and (B) contracted tail (indicated by white arrows). The phage virions were stained with 1% uranyl acetate and observed using transmission electron microscopy. The black bar represents a scale of 100nm.

To complement the morphologic characterization, we performed whole-genome sequencing of the PvEP4A phage. Genomic analysis revealed that PvEP4A possesses a genome of 35'870 nucleotides with a GC content of 65%, which closely aligns with the GC content of its host (67%) (Table S 5.4). In silico analysis identified 45 putative open reading frames (ORFs), with 35 (78%) located on the sense strand and 10 (22%) on the antisense strand (Figure 5.2, Table S 5.5). Protein sequence homology searches enabled the annotation of 27 ORFs (60%). All annotated proteins exhibited homology to proteins in other phages (Table S 5.5), except for ORF 28, for which no phage homolog was identified despite extensive database searches. The genome map of PvEP4A, including predicted functions and functional categories, is presented in Figure 5.2.

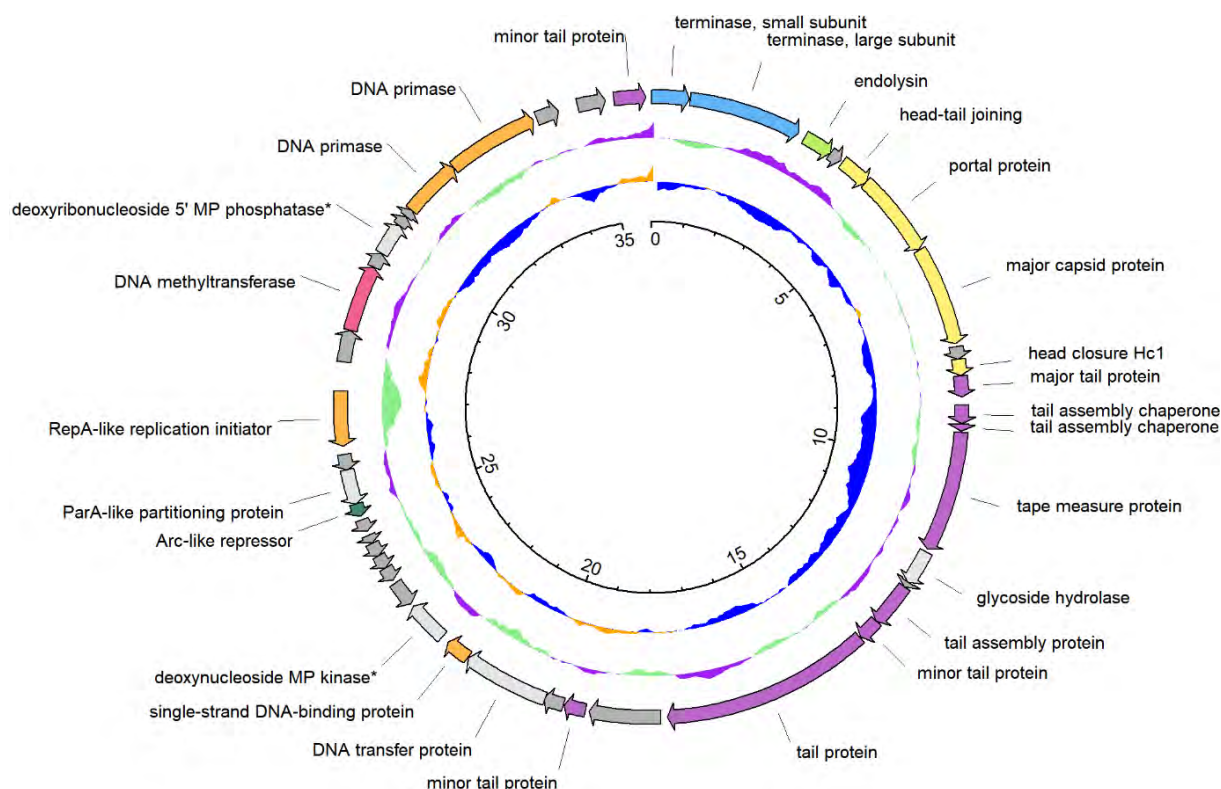


Figure 5.2 Genome map of phage PvEP4A. The outermost ring presented the ORFs of PvEP4A with the predicted functions of the coding DNA sequences indicated on the outside of the ring. The asterisks indicate the two genes whose names have been shortened: “MP” refers to “monophosphate”. Colors refer to specific functional categories attributed by PhageScope (Wang, Yang, *et al.*, 2024) : Yellow: assembly- related ORFs; Pink: immunity-related ORFs; Violet: infection-related ORFs; Light green: lysis-related ORFs; Blue: packaging-related ORFs; Dark green: regulation-related ORFs; Orange: replication-related ORFs; Light grey: unsorted ORFs; Dark grey: hypothetical ORFs. The “unsorted” category refers to genes or proteins that could not be confidently assigned to a specific functional category based on the tool's annotation methods. The middle ring indicated the GC content with, in purple, values above the median (65.6%) and green below the median. The innermost ring represented the positive (orange) and negative (blue) GC skew of the PvEP4A genome. The scale is in kbp.

Computational tools can predict bacteriophage lifestyles. Analysis of the PvEP4A phage genome using multiple tools suggests a lysogenic lifestyle. PhageBOX and PhageScope both indicate lysogeny based on protein composition and the presence of lysogeny-associated proteins, respectively (Shang, Peng, *et al.*, 2023; Shang, Tang, *et al.*, 2023; Wang, Ng, *et al.*, 2024; Wang, Yang, *et al.*, 2024). Additionally, the absence of tRNA or tmRNA genes in the PvEP4A genome, as determined by Aragorn and tRNAscan-SE (Laslett and Canback, 2004; Chan and Lowe, 2019), aligns with this prediction. While the lack of these genes is not uncommon due to the variability in phage tRNA gene content, ranging from zero to many tRNA genes, it is more frequently associated with lysogenic phages, while lytic phages tend to possess tRNA genes more often (Bailly-Bechet *et al.*, 2007)(Bailly-Bechet *et al.*, 2007). These computational predictions and genomic characteristics collectively provide initial insights into PvEP4A's potential lifestyle and host interactions, suggesting a lysogenic nature.

Further genomic analysis revealed that PvEP4A shows closer relatedness to *Rhodobacter* phages than to known *Paracoccus* phages (Figure 5.3). Specifically, PvEP4A exhibits 99% sequence identity with RcRhea, RcCronus, and RcSaxon, which are double-stranded DNA tailed phages infecting *R. capsulatus* (Rapala *et al.*, 2021). Likewise, the best BLASTp matches of PvEP4A proteins belong mostly to RcRhea, RcSaxon, and RcCronus (Table S 5.5).

These phages, including PvEP4A and other phages infecting *Paracoccus*, belong to the *Caudoviricetes* class according to PhaBOX tool (Shang, Peng, *et al.*, 2023). More precisely, PvEP4A was classified within the genus *Cronusvirus*, alongside its closely-related *R. capsulatus* infecting phages. However, the classification of some phages was limited to the class level due to the novelty of the classification system, which is no longer based on the morphotype of tailed phages but on genetic information, and the ongoing creation of new families and genera (Zhu *et al.*, 2022; Turner *et al.*, 2023).

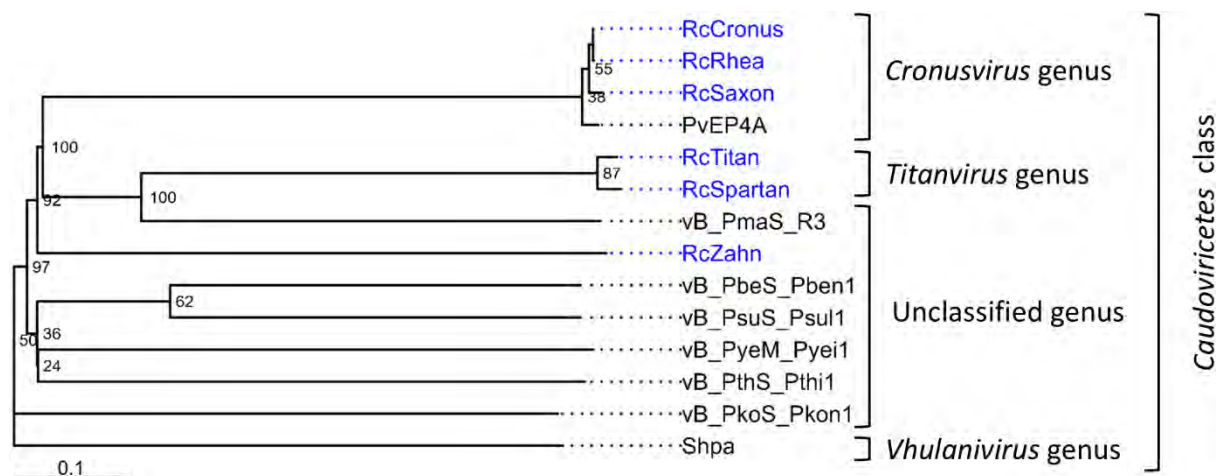


Figure 5.3 Phylogenetic tree constructed using VICTOR (Meier-Kolthoff and Göker, 2017) with whole genome sequences of PvEP4A, some *Rhodobacter* phages (blue), and *Paracoccus* phages (black). A tree based on the terminase large subunit, also used to study phage phylogeny (Van Zyl *et al.*, 2016; Benler *et al.*, 2021), gives similar results (Figure S 5.1). Scale bar: average number of substitutions per site. Numbers at the nodes are support values (percent of generated trees supporting the branch).

We next conducted host range studies using a panel of 14 *Paracoccus* strains and one *R. capsulatus* strain, given the phylogenetic relationship between PvEP4A and phages known to infect *R. capsulatus*. The results revealed a narrow host specificity with consistent lysis observed exclusively in the original host, *P. versutus* DSM582 (Table S 5.6).

Notably, two strains, *P. aestuarii* DSM19484 and *P. versutus* P46, exhibited partial susceptibility. *P. aestuarii* DSM19484 showed inconsistent lysis across multiple assays, while *P. versutus* P46 produced turbid plaques (Figure S 5.2). The inconsistent plaques formation

may be attributed to variations in bacterial growth phases that can influence susceptibility to phage infection (Attrill *et al.*, 2023). The turbid plaques suggest partial lysis of the bacterial population, which could be due to rapid development of phenotypic resistance, abortive infection, or lysogeny (Gordillo Altamirano and Barr, 2021; Makky *et al.*, 2021). Given the temperate lifestyle prediction by PhageScope and PhageBOX, lysogeny is the most plausible explanation for the observed turbidity. The mechanisms underlying both partial and inconsistent complete resistance were not further investigated due to the potential complexity of factors involved. Transient phage resistance has been documented in various Gram-negative bacteria-phage systems without elucidation of the underlying biological mechanisms (Kim and Ryu, 2012).

Taken together, these findings indicate that PvEP4A is a dsDNA tailed phage belonging to the *Cronusvirus* genus and characterized by a highly specific host range. Host range experiments and genomic predictions suggest that PvEP4A may be a temperate phage. To confirm its lysogenic potential, further experiments are necessary, particularly focusing on the two strains that exhibited inconsistent lysis or produced turbid plaques.

5.5.2 High intracellular cyclic di-GMP concentration renders *P. versutus* resistant to the phage PvEP4A.

During phage infection assay with PvEP4A and its host, bacterial regrowth was observed following initial population collapse (Figure S 5.3), suggesting potential phage-resistant mutants. To distinguish between transient (phenotypic) and constitutive (genetic) resistance, we isolated colonies from the infection assay and re-challenged them with PvEP4A using double-layer agar assays. We selected one colony that was persistently resistant to PvEP4A and sequenced it. Mapping of the reads onto the *P. versutus* DSM582 genome revealed a single consistent variation present in the mutant and absent in the wild type: a 22bp deletion in LOCUS_38800, which showed homology (95% query cover, 37% identity) to the *cpdB* gene of *E. coli* (Trülsch *et al.*, 2001; Ribeiro *et al.*, 2023). CpdB is a 3'-nucleotidase/2'3'-cyclic nucleotide phosphodiesterase that has a hydrolytic activity on a variety of substrates such as cyclic and linear dinucleotides and 2',3'-cyclic mononucleotides in *E. coli* (López-Villamizar *et al.*, 2016). Complementation of the mutant with *cpdB* restored its susceptibility to the phage PvEP4A (Figure 5.4A), confirming the role of this mutation in phage resistance.

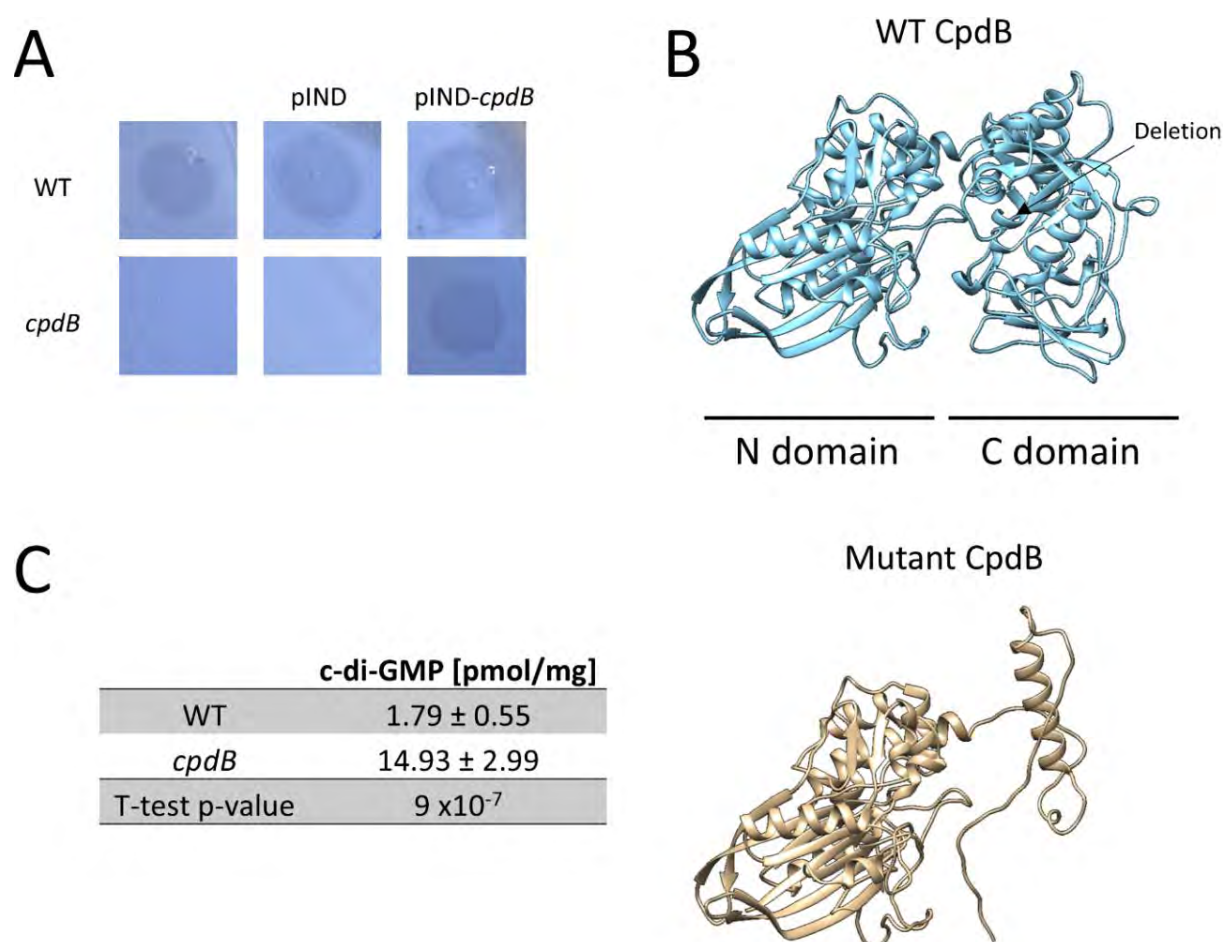


Figure 5.4 CpdB acts as a phosphodiesterase in *P. versutus*, degrading c-di-GMP. Its inactivation caused an increase in c-di-GMP concentration leading to phage resistance. (A) Dark spots of PvEP4A phage on bacteria lawn indicate zones with no bacterial growth. (B) The mutation in *cpdB* should lead to a truncation of the protein as shown on the predicted protein structure by AlphaFold2 with the N-terminus domains on the left (Jumper *et al.*, 2021). (C) Mutation in *cpdB* impaired the ability of the protein to degrade c-di-GMP and so, led to its accumulation in the *cpdB* mutant.

The deletion resulted in a frameshift, significantly truncating the protein due to a premature stop codon (Figure 5.4B, Figure S 5.4). CpdB is composed of two domains, with the N-terminal domain containing the catalytic site while the C-terminal domain containing a substrate-binding site (López-Villamizar *et al.*, 2021). This truncation affected the C-terminal domain of CpdB, hence potentially compromising its substrate-binding function.

To determine the natural substrate of CpdB in *P. versutus*, we used liquid chromatography coupled to mass spectrometry to quantify the concentrations of several cyclic dinucleotides and their degradation products in both the wild-type and the *cpdB* mutant strains. This included 3'3' c-di-AMP, 3'3' c-di-GMP, 3'3' cGAMP, 2'3'-cGAMP, 3',5'-cGMP, 3',5'-cAMP, pGpG, and pApA. Our analysis detected only c-di-GMP, cAMP, and pApA (Table S 5.7). Notably, the concentration of c-di-GMP was significantly altered between the two strains with the *cpdB*

mutant exhibiting an approximately 8-fold increase in c-di-GMP concentration compared to the wild-type strain (Figure 5.4C).

C-di-GMP is known to regulate various functions in bacteria, including motility, biofilm formation, virulence, and cell cycle progression (Hengge, 2009). High concentration of c-di-GMP is generally associated with the biosynthesis of biofilms while low c-di-GMP concentration usually correlates with motility in several bacteria species, such as *R. capsulatus*, *E. coli*, *Pseudomonas aeruginosa*, and *Salmonella enterica* (Simm *et al.*, 2004; Hickman *et al.*, 2005; Pallegar *et al.*, 2020; Shimizu *et al.*, 2022). However, among *Paracoccus* species, the effect of c-di-GMP has only been studied in *P. denitrificans*, and specifically in connection with biofilm formation (Kumar and Spiro, 2017; Morinaga *et al.*, 2020).

To elucidate the impact of elevated c-di-GMP concentrations on *P. versutus* and its resistance to PvEPA4 phage, we performed RNA sequencing on both wild-type and *cpdB* mutant strains. The analysis revealed significant differential expression in 668 out of 5622 transcripts, with 154 genes showing more than two-fold downregulation and 110 genes exhibiting more than two-fold upregulation in the *cpdB* mutant.

Interestingly, many loci identified as diguanylate cyclases (DGC), the enzymes responsible for the production of c-di-GMP, were significantly downregulated whereas phosphodiesterases, which break down c-di-GMP, were not (Table S 5.8) (Römling *et al.*, 2013). This pattern suggests a potential regulatory mechanism where c-di-GMP concentration influences DGC expression. Such regulation has been reported in *Pseudomonas putida* where the expression of the DGC GcbA is regulated by the c-di-GMP effector FleQ: lower c-di-GMP levels promotes *gcbA* transcription in a FleQ-dependent manner, while higher c-di-GMP levels lead to the downregulation of *gcbA* (Xiao *et al.*, 2016). This regulatory mechanism may serve to maintain balanced intracellular c-di-GMP concentrations.

Our analysis also revealed a highly significant downregulation of genes associated with cell motility (Table S 5.9), biogenesis of cell wall/membrane/envelope (Table S 5.10), and signal transduction (Table S 5.11). Conversely, genes involved in energy production and conversion, carbohydrate transport and metabolism, and inorganic ion transport and metabolism exhibited upregulated expression patterns in the *cpdB* strain (Figure 5.5). It is noteworthy that, aside from motility-related genes, the proportion of dysregulated genes in these categories was relatively low, with most fold changes being less than two (Figure S 5.5). The

downregulated motility genes primarily encompassed flagellar components, motors, and export genes (Kalir *et al.*, 2001). These findings underscore the global influence of c-di-GMP on bacterial physiology, affecting multiple cellular processes and potentially contributing to the observed phage resistance phenotype.

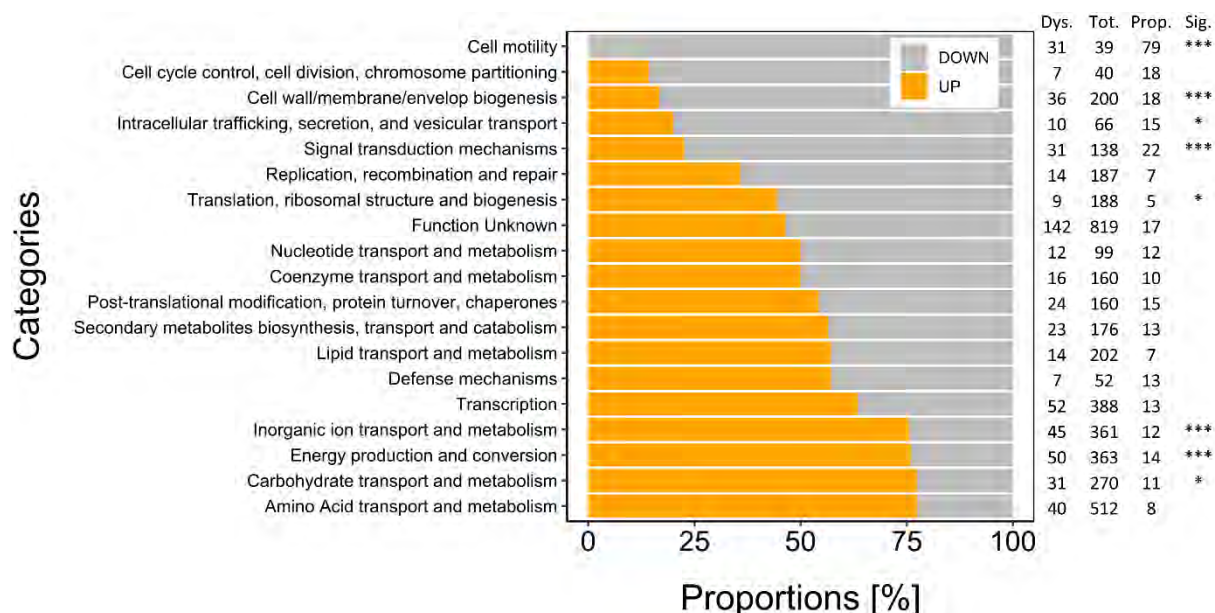


Figure 5.5 Differences in gene expression in *P. versutus*. The proportion of genes exhibiting significantly higher (up-regulated, orange) or lower (down-regulated, grey) transcript levels in the *cpdB* strain are categorized based on Clusters of Orthologous Groups (COG) annotations. Prop.: The proportion of dysregulated genes within each COG category, expressed as a percentage. This is calculated by dividing the number of dysregulated genes (“dys.”) by the total number of genes (“Tot.”) in each COG category. Sig.: Categories with significant gene enrichments between the *cpdB* strain and the wild type, determined through gene set enrichment analysis. Significance is indicated by p-values, with thresholds set at p-value < 0.05 (*), p-value < 0.01 (**), and p-value < 0.001 (***).

Given the global impact of altered c-di-GMP levels in *P. versutus*, we sought to identify which stage of infection was disrupted in the *cpdB* mutant, thereby narrowing down the number of candidates leading to resistance to the PvEP4A phage. The bacteriophage infection cycle comprises several critical steps (Azam and Tanji, 2019). First, the phage attaches to the bacterial cell surface via its ligand protein binding to a receptor, usually a sugar or protein. Next, the phage injects its genome into the bacterial cytoplasm, where phage proteins are then synthesized and assembled. Finally, new phage particles are released, generally causing the bacterial cell to lyse. Phage infection can be impeded at any of these stages, with adsorption being a common point of interference.

To elucidate the stage at which infection was disrupted, we first examined the impact of the *cpdB* mutation on phage adsorption. We quantified the number of unattached phages following incubation with wild-type and *cpdB* mutant cells. Our results revealed a significantly

higher recovery of free phages in samples containing the mutant strain, indicating that the bacterial surface of the *cpdB* mutant either adsorbed or retained bacteriophages less efficiently compared to the wild-type (Figure 5.6).

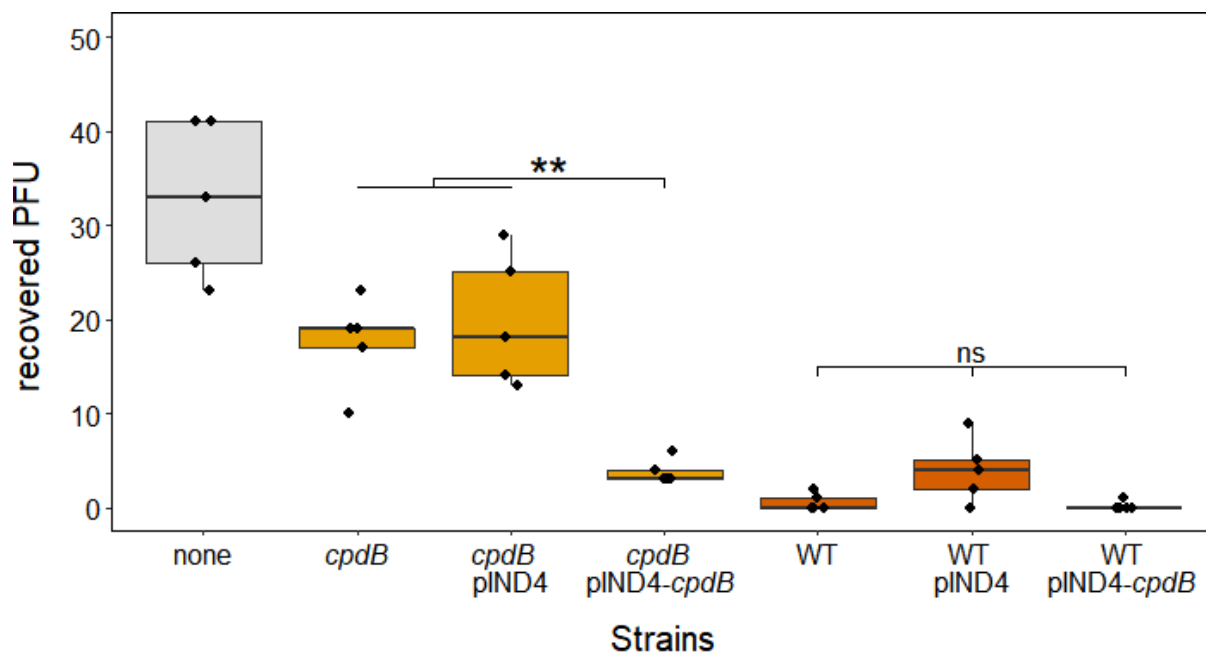


Figure 5.6 Phage resistance of *cpdB* mutant was caused by a decreased phage absorption as shown by the number of residual phages present in the medium after contact with bacteria. The following assays were performed with wild-type strain (WT), the phage resistant mutant (*cpdB*) and complemented strains with the inducible-expression plasmid pIND4 empty or carrying *cpdB*. The assay was performed with 5 replicates per strain.

Impaired phage adsorption can arise from various mechanisms, including alteration, removal, or masking of phage receptors (Azam and Tanji, 2019). Given that the *cpdB* mutation is the sole genomic difference between the mutant and wild-type strains, we hypothesized that phage receptors were more likely either downregulated or masked in the mutant (Scholl *et al.*, 2005).

Receptor masking can occur through various means, such as protein-protein interactions (e.g., TraT binding to OmpA in *E. coli*) or the production of extracellular matrices that act as physical barriers (Riede and Eschbach, 1986; Forde and Fitzgerald, 2003). Biofilm matrices, composed of extracellular polysaccharides, proteins, and DNA, can provide protection against phages (Guttenplan and Kearns, 2013; Visnapuu *et al.*, 2022). Given the association between elevated c-di-GMP levels and biofilm formation, coupled with the upregulation of certain genes involved in carbohydrate transport and metabolism (Table S 5.12) (Guttenplan and Kearns, 2013), we evaluated the production of biofilm in the mutant strain.

Contrary to expectations, the *cpdB* mutant produced less biofilm than the wild-type (Figure 5.7). This reduction in biofilm formation, despite elevated c-di-GMP concentrations, aligns with previous findings in *Paracoccus denitrificans*, where c-di-GMP was shown to downregulate biofilm formation by inhibiting BapA, a key protein in the initial stages of biofilm development (Kumar and Spiro, 2017; Yoshida *et al.*, 2017). To the best of our knowledge, none of the upregulated genes in the carbohydrate transport and metabolism category has been directly associated with phage resistance. Consequently, we postulated that phage resistance in the *cpdB* mutant was more likely attributable to the downregulation of phage receptors rather than receptor masking or biofilm-mediated protection.

Phage receptors in Gram-negative bacteria are primarily located on the outer membrane and consist of outer membrane proteins (OMPs) and lipopolysaccharides (LPS), but flagella, pili, and capsular proteins can also serve as phage receptors (Bertozzi Silva *et al.*, 2016). As previously mentioned, our transcriptomic analysis revealed significant downregulation of motility-related genes, particularly those encoding flagellar components (Table S 5.9). Flagella is frequently targeted by bacteriophages as an initial adsorption site. This strategy enhances infection probability, allowing phages to move along the flagella and access host receptors localized at cell poles (Guerrero-Ferreira *et al.*, 2011; Choi *et al.*, 2013; Bertozzi Silva *et al.*, 2016).

To assess the functional impact of flagellar gene downregulation, we examined swimming motility, a phenotype closely associated with flagellar function. Surprisingly, the *cpdB* mutant strain exhibited enhanced motility compared to the wild-type (Figure S 5.6A). This unexpected result, coupled with the downregulation of flagellar genes, suggests that the observed movement is likely mediated by a flagella-independent mechanism. Intriguingly, the swimming pattern resembled that typically observed in twitching motility assays, despite experimental conditions designed to test swimming motility (Figure S 5.6B) (Huang *et al.*, 2016).

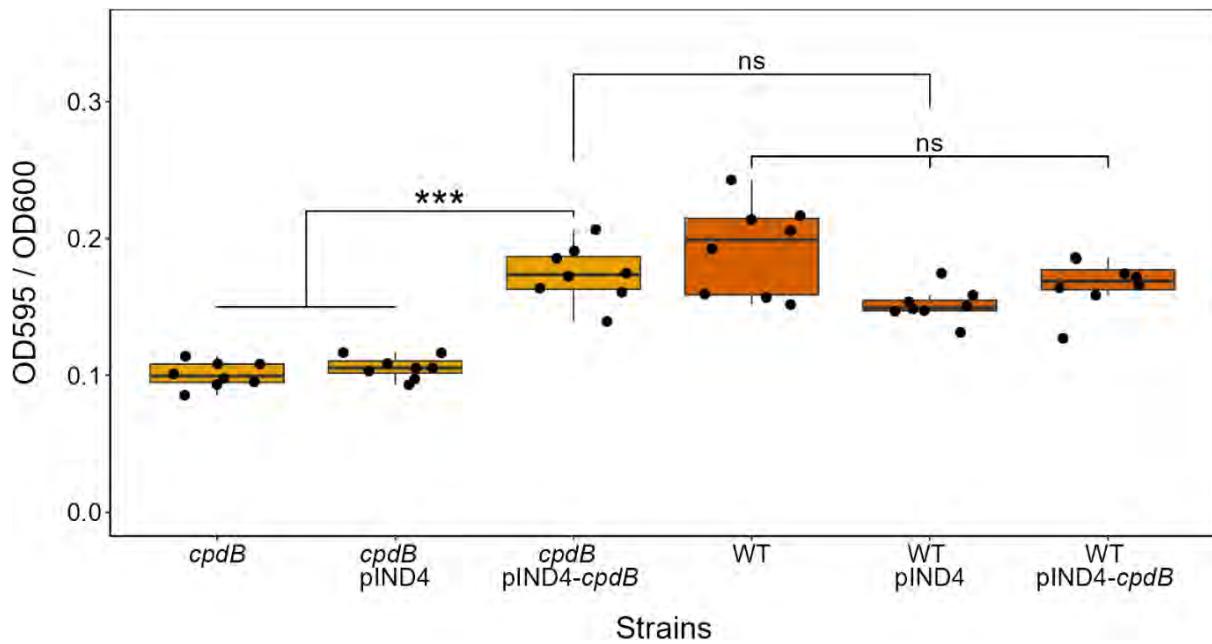


Figure 5.7 Increased concentration of c-di-GMP due to nucleotides deletion in *cpdB* in *P. versutus* led to a decrease in biofilm production. The following assay was performed with wild-type strain (WT), the phage resistant mutant (*cpdB*) and complemented strains with the inducible-expression plasmid pIND4 empty or carrying *cpdB*. Crystal Violet retention was measured at 595 nm and normalized by culture density in 8 replicates (dots) per strain.

This observation prompted us to investigate the presence of type IV pili (T4P) in the *P. versutus* genome, given that T4P are known to facilitate twitching motility (Denise *et al.*, 2019; Cai *et al.*, 2021; Whitfield and Brun, 2024). Protein sequence analysis revealed only weak homology to T4P components (Table S 5.13), consistent with the lack of T4P detection in some alphaproteobacteria (Nudleman and Kaiser, 2004; Denise *et al.*, 2019). These findings suggest that the enhanced motility observed in the *cpdB* mutant is likely mediated by an unknown mechanism, independent of both flagella and T4P. While a comprehensive investigation of this novel motility mechanism is beyond the scope of our current study, the downregulation of flagellar genes remains a plausible contributor to the observed phage resistance phenotype.

We next explore other potential attachment targets for phages. The investigation of lipopolysaccharides (LPS) yielded no significant findings, as expression levels of identified loci remained largely consistent between wild-type and *cpdB* mutant strains (Table S 5.14). In contrast, the search for capsule-related genes revealed promising candidates. Bacterial capsules have a dual role in phage interactions, acting both as potential receptors and as barriers to phage access (Scholl *et al.*, 2005; Bertozzi Silva *et al.*, 2016). While several genes showed approximately 2-fold up- or down-regulation, three loci were particularly noteworthy

due to their significant downregulation (Table 5.2, Table S 5.15). The LOCUS_15920, annotated as glycosyl transferase, showed the most striking result with a 16-fold reduction in expression. Notably, mutations in glycosyl transferases have been linked to phage resistance in various bacterial species, including *Pseudomonas syringae*, *Klebsiella pneumoniae*, and *Xanthomonas oryzae* (Debray *et al.*, 2022; Fang *et al.*, 2022; Zhang *et al.*, 2022). Furthermore, a study on *X. oryzae* demonstrated reduced phage adsorption for the glycosyl transferase mutant (Zhang *et al.*, 2022). The LOCUS_02710, annotated as GalE epimerase, is notable as mutations in GalE epimerase of *Enterococcus faecalis* have been shown to prevent phage adsorption and confer resistance to the siphovirus vB_EfaS_efap05-1 (Huang *et al.*, 2022). The third locus, LOCUS_01740, annotated as sugar transferase, shares a domain with WbaP, which is involved in the first step of O-antigen synthesis (Wang *et al.*, 1996; Kaszowska *et al.*, 2021). Interestingly, mutation in WbaP from *K. pneumoniae* results in resistance to a siphovirus and podovirus while increasing sensitivity to two myoviruses (Kaszowska *et al.*, 2021).

Table 5.2 Most promising candidates in term of differential expression between wild-type and *cpdB* strains, predicted to belong to capsule synthetic pathways by CapsuleFinder (Rendueles *et al.*, 2017). All dysregulated genes predicted by CapsuleFinder are presented in Table S 5.15.

HMM profile	%cov		Locus	L2FC	Padj	Gene	Product
	Prof.	Seq.					
Epimerase	100	69	LOCUS_02710	-1.93	8E-11	<i>galE1</i>	dTDP-4-dehydrorhamnose 3,5-epimerase
WchA	99	82	LOCUS_01740	-1.04	8E-04		sugar transferase
Cps3S	88	63	LOCUS_15920	-4.02	4E-79		glycosyl transferase

Abbreviations: Coverage percentage (%cov) of profile and sequence (sequence) refer to the percentage of the HMM profile or the input sequence that is covered by the match. padj: DESeq2 p-value adjusted for multiple testing; L2FC, logarithm in base 2 of the expression ratio of the wild-type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant).

To identify other potential phage receptor candidates, we combined BLASTp searches for homologs of known dsDNA phage receptors from the phageReceptor database with transcriptomics data from wild-type and *cpdB* mutant strains (Table S 5.3) (Zhang *et al.*, 2020). Our analysis revealed several noteworthy candidates (Table 5.3). The most significantly downregulated gene, LOCUS_27950, encoding flagellin, a subunit protein that polymerizes to form bacterial flagella filaments, which we have already discussed. The second, LOCUS_37280, codes for OmpA, a major outer membrane protein in *E. coli* and other enterobacteria, typically present in approximately 100'000 copies per cell (Ortiz-Suarez *et al.*, 2016). OmpA exhibits multifunctional properties, serving as an adhesin and invasin, biofilm formation contributor, immune system target, and bacteriophage receptor (Smith *et al.*, 2007). Two other

downregulated loci, LOCUS_00270 and LOCUS_29480, matched to OmpA as well. The third most downregulated gene, LOCUS_24780, aligns with *btuB*, which encodes an outer membrane transporter for vitamin B12 and acts as a receptor for phages infecting *E. coli* and *Salmonella typhimurium* (Kim and Ryu, 2012; Shin et al., 2012; Bertozzi Silva et al., 2016). The LOCUS_01810 aligned with the outer membrane protein Ail, which is required for the infection of *Yersinia pestis* by the phage Yep-phi (Zhao et al., 2013). The outer membrane protein NfrA, required for the infection of *E. coli* by the well-studied phage N4, blasted to LOCUS_25650. (Kiino et al., 1993; Sellner et al., 2021). However, both NfrA and Ail are targets of podoviruses so far. Interestingly, the phageReceptor database indicates that OmpA predominantly serves as a receptor for myovirus, while flagellin and BtuB are more commonly associated with siphovirus (Table S 5.3) (Zhang et al., 2020). Flagellar structures are also rather the target of siphovirus. Given that PvEP4A likely has a myovirus morphotype, OmpA homologs emerge as the most likely receptor candidates. Intriguingly, some *P. versutus* homologs to receptors were upregulated, suggesting a potential increase in *cpdB* mutant sensitivity to certain phages (Table S 5.16).

Table 5.3 Significantly downregulated homologs of known phage receptors in *cpdB* mutant.

Gene	Refseq	Locus	%ID	E-val	L2FC	Padj	Product	COG	Morph.
<i>ompA</i>	NP_706879.2	LOCUS_37280	44	4E-04	-3.23	2E-65	hypothetical protein	M	M
		LOCUS_00270	31	3E-08	-1.01	7E-06	peptidoglycan-assoc. lipoprotein	M	M
	NP_415477.1	LOCUS_29480	36	1E-11	-0.55	2E-02	membrane protein	M	M
<i>fliC</i>	NP_460912.1	LOCUS_27950	25	2E-02	-5.32	2E-71	flagellin	N	S
<i>btuB</i>	NP_463009.1	LOCUS_24780	33	4E-02	-1.96	6E-08	hypothetical protein	Q	S
<i>ail</i>	WP_002227864.1	LOCUS_01810	33	6E-01	-1.24	7E-06	membrane protein	M	P
<i>nfrA</i>	NP_415100.1	LOCUS_25650	29	1E-01	-0.55	2E-02	hypothetical protein	S	P

Abbreviation: BLASTp of reference sequence (Refseq) on *P. versutus* proteins with percentage identity (%ID) et e-value (E-val) indicated. Transcripts analysis with DESeq2 p-value adjusted for multiple testing (padj) and logarithm in base 2 of the expression ratio of the wild-type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant) (L2FC). The Clusters of Orthologous Genes (COG) categories are listed as follows: M : Cell wall/membrane/envelop biogenesis; N : Cell motility; Q : Secondary metabolites biosynthesis, transport and catabolism; S : Function Unknown. Typical morphotypes (Morph.) of phages using the receptors.

Unsurprisingly, several candidates for phage reception belong to the COG category cell wall/membrane/envelope (Table S 5.10). In addition to the previously mentioned loci within this COG category, the downregulated LOCUS_20740, identified as an outer-membrane

lipoprotein carrier protein, could be a promising candidate for phage receptors (Bertozzi Silva *et al.*, 2016).

Taken together, our analysis identified several potential phage receptor candidates for PvEP4A in *P. versutus*, including flagella, outer membrane proteins, and capsule biogenesis participants. This diversity reflects the widespread impact of elevated c-di-GMP levels on the cell. Although phage receptor specificity can vary significantly between different phages, those within the same morphotype often share structural similarities that influence their receptor binding mechanisms, leading to trends in receptor utilization among morphotypes (Table S 5.3). Given that PvEP4A has rather a myovirus morphotype and considering the downregulation of certain loci, we believe that the downregulation of glycosyl transferases (LOCUS_15920) and the hypothetical protein matching OmpA (LOCUS_37280) are the most likely explanations for the observed phage-resistance phenotype. While these findings provide valuable insights, the precise mechanisms of phage resistance remain to be fully elucidated. The mutation in *cpdB* appears to influence multiple phenotypic traits, potentially contributing to phage resistance through various pathways. To further refine our prediction, the nature of the phage receptor might be determined by treating bacteria with either proteinase K (destroying surface proteins) or periodate (disrupting surface carbohydrates), prior to phage adsorption assay (Kiljunen *et al.*, 2011; Zhao *et al.*, 2013). Additionally, confirmatory tests for the presence of flagella and capsule in the wild-type and mutant strains are necessary to definitively link c-di-GMP levels to specific phenotypic changes. These future investigations will be essential in unravelling the complex interplay between c-di-GMP signalling, cellular structures, and phage susceptibility in *P. versutus*.

5.6 Conclusion

The study of the bacteriophage PvEP4A infecting *P. versutus* contributes to expand our knowledge of phage-host interactions. This research not only adds a new phage to the limited collection of known *Paracoccus* phages but also sheds light on the complex evolutionary strategies bacteria employ to counteract viral invasion. The rapid emergence of phage resistance through a small genetic deletion affecting not only phage susceptibility but also the cell physiology, highlights the dynamic nature of these interactions and their potential impact on bacterial adaptation. The deletion in the *cpdB* gene, leading to increased intracellular

c-di-GMP levels, demonstrates how a mutation can have pleiotropic effects on bacterial phenotypes. By exploring the molecular mechanisms underlying phage resistance, including alterations in cyclic-di-GMP signalling, this work opens new avenues for understanding bacterial adaptation and survival strategies in microbial communities.

5.7 References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.
- Attrill, E.L., Łapińska, U., Westra, E.R., Harding, S. V, and Pagliara, S. (2023) Slow growing bacteria survive bacteriophage in isolation. *ISME Commun* **3**: 1–9.
- Azam, A.H. and Tanji, Y. (2019) Bacteriophage-host arm race: an update on the mechanism of phage resistance in bacteria and revenge of the phage with the perspective for phage therapy. *Appl Microbiol Biotechnol* **103**: 2121–2131.
- Bailly-Bechet, M., Vergassola, M., and Rocha, E. (2007) Causes for the intriguing presence of tRNAs in phages. *Genome Res* **17**: 1486–1495.
- Benler, S., Yutin, N., Antipov, D., Rayko, M., Shmakov, S., Gussow, A.B., et al. (2021) Thousands of previously unknown phages discovered in whole-community human gut metagenomes. *Microbiome* **9**: 1–17.
- Bertozzi Silva, J., Storms, Z., and Sauvageau, D. (2016) Host receptors for bacteriophage adsorption. *FEMS Microbiol Lett* **363**: 1–11.
- Besemer, J. and Borodovsky, M. (2005) GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res* **33**: 451–454.
- Bradley, E.L., Ökmen, B., Doehlemann, G., Henrissat, B., Bradshaw, R.E., and Mesarich, C.H. (2022) Secreted Glycoside Hydrolase Proteins as Effectors and Invasion Patterns of Plant-Associated Fungi and Oomycetes. *Front Plant Sci* **13**: 1–16.
- Cai, L., Jain, M., Sena-Vélez, M., Jones, K.M., Fleites, L.A., Heck, M., and Gabriel, D.W. (2021) Tad pilus-mediated twitching motility is essential for DNA uptake and survival of *Liberibacters*. *PLoS One* **16**: 1–23.
- Carroll-Portillo, A., Coffman, C.N., Varga, M.G., Alcock, J., Singh, S.B., and Lin, H.C. (2021) Standard bacteriophage purification procedures cause loss in numbers and activity. *Viruses* **13**: 1–16.
- Chan, P.P. and Lowe, T.M. (2019) tRNAscan-SE: Searching for tRNA genes in genomic sequences. *Methods Mol Biol* **1962**: 1–14.
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018) Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: 884–890.
- Chiang, Y.N., Penadés, J.R., and Chen, J. (2019) Genetic transduction by phages and chromosomal islands: The new and noncanonical. *PLoS Pathog* **15**: 1–7.

- Choi, Y., Shin, H., Lee, J.H., and Ryua, S. (2013) Identification and characterization of a novel flagellum-dependent Salmonella-infecting bacteriophage, iEPS5. *Appl Environ Microbiol* **79**: 4829–4837.
- Daneshvar, M.I., Hollis, D.G., Weyant, R.S., Steigerwalt, A.G., Whitney, A.M., Douglas, M.P., et al. (2003) *Paracoccus yeeii* sp. nov. (Formerly CDC group EO-2), a novel bacterial species associated with human infection. *J Clin Microbiol* **41**: 1289–1294.
- Debray, R., De Luna, N., and Koskella, B. (2022) Historical Contingency Drives Compensatory Evolution and Rare Reversal of Phage Resistance. *Mol Biol Evol* **39**: 1–12.
- Decewicz, P., Dziewit, L., Golec, P., Kozłowska, P., Bartosik, D., and Radlinska, M. (2019) Characterization of the virome of *Paracoccus* spp. (Alphaproteobacteria) by combined in silico and in vivo approaches. *Sci Rep* **9**: 1–17.
- Delcher, A.L., Bratke, K.A., Powers, E.C., and Salzberg, S.L. (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **23**: 673–679.
- Denise, R., Abby, S.S., and Rocha, E.P.C. (2019) Diversification of the type IV filament superfamily into machines for adhesion, protein secretion, DNA uptake, and motility. *PLoS Biol* **17**: 1–34.
- Dziewit, L., Czarnecki, J., Wibberg, D., Radlinska, M., Mrozek, P., Szymczak, M., et al. (2014) Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids. *BMC Genomics* **15**: 1–16.
- Edgar, R.C. (2004) MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 1–19.
- Fang, Q., Feng, Y., McNally, A., and Zong, Z. (2022) Characterization of phage resistance and phages capable of intestinal decolonization of carbapenem-resistant *Klebsiella pneumoniae* in mice. *Commun Biol* **5**: 1–14.
- Forde, A. and Fitzgerald, G.F. (2003) Molecular organization of exopolysaccharide (EPS) encoding genes on the lactococcal bacteriophage adsorption blocking plasmid, pCI658. *Plasmid* **49**: 130–142.
- Fortier, L.C. and Moineau, S. (2009) Phage production and maintenance of stocks, including expected stock lifetimes. *Methods Mol Biol* **501**: 203–219.
- Garrison, E. and Marth, G. (2012) Haplotype-based variant detection from short-read sequencing. *arXiv* 1–9.
- Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**: 343–345.
- Gordillo Altamirano, F.L. and Barr, J.J. (2021) Screening for lysogen activity in therapeutically relevant bacteriophages. *Bio-protocol* **11**: 1–9.

- Guerrero-Ferreira, R.C., Viollier, P.H., Ely, B., Poindexter, J.S., Georgieva, M., Jensen, G.J., and Wright, E.R. (2011) Alternative mechanism for bacteriophage adsorption to the motile bacterium *Caulobacter crescentus*. *Proc Natl Acad Sci U S A* **108**: 9963–9968.
- Guttenplan, S.B. and Kearns, D.B. (2013) Regulation of flagellar motility during biofilm formation. *FEMS Microbiol Rev* **37**: 849–871.
- Hengge, R. (2009) Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol* **7**: 263–273.
- Hernández-Plaza, A., Szklarczyk, D., Botas, J., Cantalapiedra, C.P., Giner-Lamia, J., Mende, D.R., et al. (2023) eggNOG 6.0: enabling comparative genomics across 12 535 organisms. *Nucleic Acids Res* **51**: 389–394.
- Hickman, J.W., Tifrea, D.F., and Harwood, C.S. (2005) A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels. *Proc Natl Acad Sci U S A* **102**: 14422–14427.
- Huang, L., Guo, W., Lu, J., Pan, W., Song, F., and Wang, P. (2022) Enterococcus faecalis Bacteriophage vB_EfaS_efap05-1 Targets the Surface Polysaccharide and ComEA Protein as the Receptors. *Front Microbiol* **13**: 1–7.
- Huang, Y.H., Lin, J.S., Ma, J.C., and Wang, H.H. (2016) Functional characterization of triclosan-resistant enoyl-acyl-carrier protein reductase (fabV) in *Pseudomonas aeruginosa*. *Front Microbiol* **7**: 1–13.
- Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., et al. (2019) EggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**: D309–D314.
- Ind, A.C., Porter, S.L., Brown, M.T., Byles, E.D., De Beyer, J.A., Godfrey, S.A., and Armitage, J.P. (2009) Inducible-expression plasmid for *Rhodobacter sphaeroides* and *Paracoccus denitrificans*. *Appl Environ Microbiol* **75**: 6613–6615.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**: 583–589.
- Kalir, S., McClure, J., Pabbaraju, K., Southward, C., Ronen, M., Leibler, S., et al. (2001) Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. *Science (80-)* **292**: 2080–2083.
- Kaszowska, M., Majkowska-skrobek, G., Markwitz, P., Lood, C., Jachymek, W., Maciejewska, A., et al. (2021) The mutation in wbp cps gene cluster selected by phage-borne depolymerase abolishes capsule production and diminishes the virulence of *Klebsiella pneumoniae*. *Int J Mol Sci* **22**: 1–15.
- Kelly, D.P., Rainey, F.A., and Wood, A.N.N.P. (2006) The Genus *Paracoccus*. *The Prokaryotes* **5**: 232–249.

- Kiino, D.R., Singer, M.S., and Rothman-Denes, L.B. (1993) Two overlapping genes encoding membrane proteins required for bacteriophage N4 adsorption. *J Bacteriol* **175**: 7081–7085.
- Kiljunen, S., Datta, N., Dentovskaya, S. V., Anisimov, A.P., Knirel, Y.A., Bengoechea, J.A., et al. (2011) Identification of the lipopolysaccharide core of *Yersinia pestis* and *Yersinia pseudotuberculosis* as the receptor for bacteriophage ϕ A1122. *J Bacteriol* **193**: 4963–4972.
- Kim, M. and Ryu, S. (2012) Spontaneous and transient defence against bacteriophage by phase-variable glucosylation of O-antigen in *Salmonella enterica* serovar Typhimurium. *Mol Microbiol* **86**: 411–425.
- King, A.M.Q., Adams, M.J., Carstens, E.B., and Lefkowitz, E.J. (2011) Virus taxonomy: Ninth report of the international committee on taxonomy of viruses. *Virus Taxon Ninth Rep Int Comm Taxon Viruses* 1221–1234.
- Koskella, B., Hernandez, C.A., and Wheatley, R.M. (2022) Understanding the Impacts of Bacteriophage Viruses: From Laboratory Evolution to Natural Ecosystems. *Annu Rev Virol* **9**: 57–78.
- Kumar, S. and Spiro, S. (2017) Environmental and Genetic Determinants of Biofilm Formation in *Paracoccus denitrificans*. *mSphere* **2**: 1–11.
- Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.
- Lasek, R., Szuplewska, M., Mitura, M., Decewicz, P., Chmielowska, C., Pawlot, A., et al. (2018) Genome structure of the opportunistic pathogen *Paracoccus yeei* (Alphaproteobacteria) and identification of putative virulence factors. *Front Microbiol* **9**: 1–17.
- Laslett, D. and Canback, B. (2004) ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* **32**: 11–16.
- Lawrence, J. (2007) DNA Master. <http://cobamide2.bio.pitt.edu/computer.html>.
- Lee, J.H., Kim, Y.S., Choi, T.J., Lee, W.J., and Kim, Y.T. (2004) *Paracoccus haeundaensis* sp. nov., a Gram-negative, halophilic, astaxanthin-producing bacterium. *Int J Syst Evol Microbiol* **54**: 1699–1702.
- Liao, Y., Smyth, G.K., and Shi, W. (2014) FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.
- Liu, Z.P., Wang, B.J., Liu, X.Y., Dai, X., Liu, Y.H., and Liu, S.J. (2008) *Paracoccus halophilus* sp. nov., isolated from marine sediment of the South China Sea, China and emended description of genus *Paracoccus* Davis 1969. *Int J Syst Evol Microbiol* **58**: 257–261.
- López-Villamizar, I., Cabezas, A., Pinto, R.M., Canales, J., Ribeiro, J.M., Cameselle, J.C., and Costas, M.J. (2016) The Characterization of *Escherichia coli* CpdB as a Recombinant Protein Reveals that, besides Having the Expected 3-Nucleotidase and 2,3-Cyclic

- Mononucleotide Phosphodiesterase Activities, It Is Also Active as Cyclic Dinucleotide Phosphodiesterase. *PLoS One* **11**: 1–22.
- López-Villamizar, I., Cabezas, A., Pinto, R.M., Canales, J., Ribeiro, J.M., Rodrigues, J.R., et al. (2021) Molecular dissection of Escherichia coli Cpdb: Roles of the N domain in catalysis and phosphate inhibition, and of the C domain in substrate specificity and adenosine inhibition. *Int J Mol Sci* **22**: 1–18.
- 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**: 1–21.
- Makky, S., Dawoud, A., Safwat, A., Abdelsattar, A.S., Rezk, N., and El-Shibiny, A. (2021) The bacteriophage decides own tracks: When they are with or against the bacteria. *Curr Res Microb Sci* **2**: 1–9.
- Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., et al. (2011) CDD: A Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* **39**: 225–229.
- Md, V., Misra, S., Li, H., and Aluru, S. (2019) Efficient architecture-aware acceleration of BWA-MEM for multicore systems. *2019 IEEE Int Parallel Distrib Process Symp* 314–324.
- Meier-Kolthoff, J.P. and Göker, M. (2017) VICTOR: genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics* **33**: 3396–3404.
- Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H., and Ferrin, T.E. (2023) UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci* **32**: 1–13.
- Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022) ColabFold: making protein folding accessible to all. *Nat Methods* **19**: 679–682.
- Morinaga, K., Yoshida, K., Takahashi, K., Nomura, N., and Toyofuku, M. (2020) Peculiarities of biofilm formation by Paracoccus denitrificans. *Appl Microbiol Biotechnol* **104**: 2427–2433.
- Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., and Minh, B.Q. (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**: 268–274.
- Nikaido, H. (2018) RND transporters in the living world. *Res Microbiol* **169**: 363–371.
- Nudleman, E. and Kaiser, D. (2004) Pulling together with type IV pili. *J Mol Microbiol Biotechnol* **7**: 52–62.
- Olson, R.D., Assaf, R., Brettin, T., Conrad, N., Cucinell, C., Davis, J.J., et al. (2023) Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC, IRD and VIPR. *Nucleic Acids Res* **51**: D678–D689.
- Ortiz-Suarez, M.L., Samsudin, F., Piggot, T.J., Bond, P.J., and Khalid, S. (2016) Full-Length OmpA: Structure, Function, and Membrane Interactions Predicted by Molecular

- Dynamics Simulations. *Biophys J* **111**: 1692–1702.
- Pallegar, P., Langille, E., Gomelsky, M., and Lang, A.S. (2020) Cyclic di-GMP-Mediated Regulation of Gene Transfer and Motility in *Rhodobacter capsulatus*. *J Bacteriol* **202**: 1–17.
- Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. (2020) Using SPAdes De Novo Assembler. *Curr Protoc Bioinforma* **70**: 1–29.
- Rapala, J., Miller, B., Garcia, M., Dolan, M., Bockman, M., Hansson, M., et al. (2021) Genomic diversity of bacteriophages infecting *Rhodobacter capsulatus* and their relatedness to its gene transfer agent RcGTA. *PLoS One* **16**: 1–24.
- Rendueles, O., Garcia-Garcerà, M., Néron, B., Touchon, M., and Rocha, E.P. (2017) Abundance and co-occurrence of extracellular capsules increase environmental breadth: Implications for the emergence of pathogens. *PLoS Pathog* **13**: 1–30.
- Ribeiro, J.M., Canales, J., Costas, M.J., Cabezas, A., Pinto, R.M., García-Díaz, M., et al. (2023) Genomic Distribution of Pro-Virulent cpdB-like Genes in Eubacteria and Comparison of the Enzyme Specificity of CpdB-like Proteins from *Salmonella enterica*, *Escherichia coli* and *Streptococcus suis*. *Int J Mol Sci* **24**: 1–26.
- Riede, I. and Eschbach, M.-L. (1986) Evidence that TraT interacts with OmpA of *Escherichia coli*. *Febs* **205**: 241–246.
- Römling, U., Galperin, M.Y., and Gomelsky, M. (2013) Cyclic di-GMP: the First 25 Years of a Universal Bacterial Second Messenger. *Microbiol Mol Biol Rev* **77**: 1–52.
- RStudio Team (2020) RStudio: Integrated Development for R. RStudio. <http://www.rstudio.com/>.
- Sambrook, J. and Russell, D.W. (2001) Molecular Cloning: A Laboratory Manual. Third Edition.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012) Fiji: An open-source platform for biological-image analysis. *Nat Methods* **9**: 676–682.
- Scholl, D., Adhya, S., and Merril, C. (2005) *Escherichia coli* K1's capsule is a barrier to bacteriophage T7. *Appl Environ Microbiol* **71**: 4872–4874.
- Sellner, B., Prakapaite, R., Van Berkum, M., Heinemann, M., Harms, A., and Jenal, U. (2021) A New Sugar for an Old Phage: A c-di-GMP-Dependent Polysaccharide Pathway Sensitizes *Escherichia coli* for Bacteriophage Infection. *MBio* **12**: 1–17.
- Shang, J., Peng, C., Liao, H., Tang, X., and Sun, Y. (2023) PhaBOX: A web server for identifying and characterizing phage contigs in metagenomic data. *Bioinforma Adv* **3**: 1–5.
- Shang, J., Tang, X., and Sun, Y. (2023) PhaTYP: predicting the lifestyle for bacteriophages using BERT. *Brief Bioinform* **24**: 1–11.

- Shimizu, T., Aritoshi, T., Beatty, J.T., and Masuda, T. (2022) Persulfide-Responsive Transcription Factor SqrR Regulates Gene Transfer and Biofilm Formation via the Metabolic Modulation of Cyclic di-GMP in *Rhodobacter capsulatus*. *Microorganisms* **10**: 1–13.
- Shin, H., Lee, J.H., Kim, H., Choi, Y., Heu, S., and Ryu, S. (2012) Receptor diversity and host interaction of bacteriophages infecting *Salmonella enterica* Serovar Typhimurium. *PLoS One* **7**: 1–11.
- Simm, R., Morr, M., Kader, A., Nimtz, M., and Römling, U. (2004) GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* **53**: 1123–1134.
- Smith, S.G.J., Mahon, V., Lambert, M.A., and Fagan, R.P. (2007) A molecular Swiss army knife: OmpA structure, function and expression. *FEMS Microbiol Lett* **273**: 1–11.
- Söding, J., Biegert, A., and Lupas, A.N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**: 244–248.
- Stone, E., Campbell, K., Grant, I., and McAuliffe, O. (2019) Understanding and exploiting phage–host interactions. *Viruses* **11**: 1–26.
- Trülsch, K., Roggenkamp, A., Pelludat, C., Rakin, A., Jacobi, C.A., and Heesemann, J. (2001) Cloning and characterization of the gene encoding periplasmic 2',3'-cyclic phosphodiesterase of *Yersinia enterocolitica* O:8. *Microbiology* **147**: 203–213.
- Tucker, R.G. (1961) The Role of Magnesium Ions in the Growth of *Salmonella* Phage Anti -R. *J Gen Microbiol* **3390**: 313–323.
- Turner, D., Shkoporov, A.N., Lood, C., Millard, A.D., Dutilh, B.E., Alfenas-Zerbini, P., et al. (2023) Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. *Arch Virol* **168**: 1–9.
- Visnapuu, A., Der Gucht, M., Wagemans, J., and Lavigne, R. (2022) Deconstructing the Phage–Bacterial Biofilm Interaction as a Basis to Establish New Antibiofilm Strategies. *Viruses* **14**: 1–19.
- Wang, L., Liu, D., and Reeves, P.R. (1996) C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. *J Bacteriol* **178**: 2598–2604.
- Wang, R.H., Ng, Y.K., Zhang, X., Wang, J., and Li, S.C. (2024) Coding genomes with gapped pattern graph convolutional network. *Bioinformatics* **40**: 1–8.
- Wang, R.H., Yang, S., Liu, Z., Zhang, Y., Wang, X., Xu, Z., et al. (2024) PhageScope: a well-annotated bacteriophage database with automatic analyses and visualizations. *Nucleic Acids Res* **52**: D756–D761.
- Whitfield, G.B. and Brun, Y. V. (2024) The type IVc pilus: just a Tad different. *Curr Opin Microbiol* **79**: 1–11.

- Wick, R.R., Schultz, M.B., Zobel, J., and Holt, K.E. (2015) Bandage: Interactive visualization of de novo genome assemblies. *Bioinformatics* **31**: 3350–3352.
- Wickham, H. (2008) *Elegant Graphics for Data Analysis: ggplot2*, Springer (ed).
- Wishart, D.S., Han, S., Saha, S., Oler, E., Peters, H., Grant, J.R., et al. (2023) PHASTEST: Faster than PHASTER, better than PHAST. *Nucleic Acids Res* **51**: 443–450.
- Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021) clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2**: 1–11.
- Xiao, Y., Nie, H., Liu, H., Chen, W., and Huang, Q. (2016) Expression of the diguanylate cyclase GcbA is regulated by FleQ in response to cyclic di-GMP in *Pseudomonas putida* KT2440. *Environ Microbiol Rep* **8**: 993–1002.
- Xu, Y., Zhang, R., and Jiao, N. (2015) Complete genome sequence of *Paracoccus marcusii* phage vB_PmaS-R3 isolated from the South China Sea. *Stand Genomic Sci* **10**: 2–7.
- Yoshida, K., Toyofuku, M., Obana, N., and Nomura, N. (2017) Biofilm formation by *Paracoccus denitrificans* requires a type I secretion system-dependent adhesin BapA. *FEMS Microbiol Lett* **364**: 1–7.
- Yu, G. (2020) Using ggtree to Visualize Data on Tree-Like Structures. *Curr Protoc Bioinforma* **69**: 1–18.
- Zhang, M., Qian, J., Xu, X., Ahmed, T., Yang, Y., Yan, C., et al. (2022) Resistance of *Xanthomonas oryzae* pv. *oryzae* to Lytic Phage X2 by Spontaneous Mutation of Lipopolysaccharide Synthesis-Related Glycosyltransferase. *Viruses* **14**: 1–16.
- Zhang, Z., Yu, F., Zou, Y., Qiu, Y., Wu, A., Jiang, T., and Peng, Y. (2020) Phage protein receptors have multiple interaction partners and high expressions. *Bioinformatics* **36**: 2975–2979.
- Zhao, X., Cui, Y., Yan, Y., Du, Z., Tan, Y., Yang, H., et al. (2013) Outer Membrane Proteins Ail and OmpF of *Yersinia pestis* Are Involved in the Adsorption of T7-Related Bacteriophage Yep-phi. *J Virol* **87**: 12260–12269.
- Zhu, Y., Shang, J., Peng, C., and Sun, Y. (2022) Phage family classification under Caudoviricetes: A review of current tools using the latest ICTV classification framework. *Front Microbiol* **13**: 1–14.
- Van Zyl, L.J., Nemavhulani, S., Cass, J., Cowan, D.A., and Trindade, M. (2016) Three novel bacteriophages isolated from the East African Rift Valley soda lakes. *Virol J* **13**: 1–14.

5.8 Supplementary Material

Table S 5.1 Bacterial strains, plasmids, and oligonucleotides used in this chapter.

Strains, plasmids or oligonucleotides		Relevant characteristics	Reference
Strains			
<i>Paracoccus versutus</i>	DSM582	Wild-type phenotype	DSMZ collection
<i>Paracoccus versutus</i>	cpdB	spontaneous PvEP4A phage resistant strain DSM582 due to a mutation in <i>cpdB</i>	This study
<i>Paracoccus versutus</i>	<i>cpdB</i> pIND4	<i>cpdB</i> mutant strain complemented with pIND4 plasmid; control for presence of pIND	This study
<i>Paracoccus versutus</i>	cpdB pIND4-cpdB	<i>cpdB</i> mutant strain complemented with pIND4- <i>cpdB</i> plasmid	This study
<i>Paracoccus versutus</i>	P2	newly isolated environmental strain	This study
<i>Paracoccus versutus</i>	P45	newly isolated environmental strain	This study
<i>Paracoccus versutus</i>	P46	newly isolated environmental strain	This study
<i>Paracoccus denitrificans</i>	DSM1407	Wild-type phenotype	DSMZ collection
<i>Paracoccus pantotrophus</i>	DSM1404	Wild-type phenotype	DSMZ collection
<i>Paracoccus pantotrophus</i>	DSM 65	Wild-type phenotype	DSMZ collection
<i>Paracoccus aminophilus</i>	DSM8538	Wild-type phenotype	DSMZ collection
<i>Paracoccus aestuarii</i>	DSM19484	Wild-type phenotype	DSMZ collection
<i>Paracoccus alkenifer</i>	DSM11593	Wild-type phenotype	DSMZ collection
<i>Paracoccus siganidrum</i>	DSM26381	Wild-type phenotype	DSMZ collection
<i>Rhodobacter capsulatus</i>	DSM1710	Wild-type phenotype	DSMZ collection
<i>Paracoccus yeei</i>	KB 2	Wild-type phenotype	P. Junier
<i>Paracoccus yeei</i>	KB 3	Wild-type phenotype	P. Junier
<i>Paracoccus yeei</i>	KB 4	Wild-type phenotype	P. Junier
<i>Escherichia coli</i>	DH5α-λpir	Laboratory cloning strain	(Sambrook and Russell, 2001)
Plasmids			
pIND4	IPTG-inducible expression vector designed for use in alphaproteobacteria, carrying a kanamycin resistance marker		(Ind <i>et al.</i> , 2009)
pIND4- <i>cpdB</i>	pIND4 containing <i>cpdB</i> from DSM582		This study
Oligonucleotides		Sequence	Purpose
16s_Paracoccus_F	5'- GTT GTA AAG CTC TTT CAG CTG GGA AG - 3'		<i>P. versutus</i> specific primer
16s_Paracoccus_R	5'- GCT TTT GGG GAT TAA CCC ACT GT - 3'		
GM3F	5'- AGA GTT TGA TC(AC) TGG C -3'		Amplification of 16S rDNA
GM4R	5'- TAC CTT GTT ACG ACT T -3'		
Pv_rpoB_F	5'- ATG GCG CAA GCT TAT GTC G -3'		Amplification of <i>rpoB</i>
Pv_rpoB_R	5'- TTA CTC GTC CTC CTC CGC ATC -3'		
gib_cpdB_with_pIND4_F	5'- GAA ATT AAC CAT GCA ATC TGC GGT TGC ATT ATG G -3'		Amplification of <i>cpdB</i> in DSM582 for assembly
gib_cpdB_with_pIND4_R	5'- GAT GGT GAT GCA GAT GCA GGC GCA GGC G -3'		
gib_pIND4_with_cpdB_F	5'- CCT GCA TCT GCA TCA CCA TCA CCA TCA C -3'		Amplification of <i>pIND4</i> for gibson assembly
gib_pIND4_with_cpdB_R	5'- CAG ATT GCA TGG TTA ATT TCT CCT CTT TAA TTC -3'		
Pv_qPCR_cysE-PyrD_F	5'-CGT CGA ACA CGA TGA ACC AGC-3'		Control for DNA contamination in RNA samples
Pv_qPCR_cysE-PyrD_FR	5'-CTC GGC CGT CCA GAT CTA TTC C-3'		

Table S 5.2 Distribution of proteins classified by CapsuleFinder across capsule systems with mandatory (biologically essential components for a putative functional system, a majority of which, if not all, are required to identify and classify the systems), accessory (non-essential components used to improve the annotation of the system), and forbidden components (e.g., those found in other capsule groups and not in the focal one) (Rendueles et al., 2017).

Capsule systems	Classification		
	mandatory	accessory	forbidden
ABC	92	84	12
GroupIV_e_stricte	12	1	178
GroupIV_f	0	2	178
GroupIV_s_stricte	2	22	178
PGA	5	2	0
Syn_cps3	3	19	180
Syn_has	0	22	180

Table S 5.3 Summary table of dsDNA phage and their host from the phageReceptor database (access on November 5, 2024) (Zhang et al., 2020).

Gene	Receptor	Accession Nb	Host	Virus	Fam	Reference
ail	Attachment invasion locus protein	WP_002227864.1	Yersinia pestis	Yersinia phage Yep-phi	P	PMID: 27722870 PMID: 24006436 PMID: 26965493
BA_3367	LPXTG-motif cell wall anchor domain protein	NP_845653.1	Bacillus anthracis	Bacillus phage Gamma	S	PMID: 26755501 PMID: 16166537
bamA	Outer membrane protein assembly factor BamA	NP_308206.1	Escherichia coli	Escherichia virus 933W	P	PMID: 29896160 PMID: 29973923 PMID: 17693515 PMID: 12761125
btuB	Vitamin B12 transporter BtuB	NP_418401.1	Escherichia coli	Escherichia virus BF23	S	PMID: 29523063 PMID: 1254550
				Escherichia virus EPS7	S	PMID: 29523063 PMID: 19025561
				Escherichia virus DT57C	S	PMID: 26805872
				Enterobacteria phage DT571/2	S	PMID: 26805872
		NP_463009.1	Salmonella enterica	Escherichia phage C1	S	PMID: 12558182
				Escherichia virus H8	S	PMID: 29523063
				Escherichia virus EPS7	S	PMID: 29523063 PMID: 19025561
				Salmonella virus SPC35	S	PMID: 29523063
				Salmonella phage BSP22A	S	PMID: 30297056
				Salmonella phage S113	S	PMID: 30888719
				Salmonella phage S116	S	PMID: 30888719
				Salmonella phage S126	S	PMID: 30888719
				Salmonella phage S132	S	PMID: 30888719
				Salmonella phage S147	S	PMID: 30888719
		NP_709770.2	Shigella flexneri	Salmonella phage BSP22A	S	PMID: 30297056
ECs0179	Outer membrane protein assembly factor BamA	NP_308206.1	Escherichia coli	Escherichia phage vB_EcoP_24B	P	PMID: 29896160 PMID: 29973923 PMID: 17693515 PMID: 12761125

ECs3104	Outer membrane porin C	NP_311131.1	Escherichia coli	Escherichia phage AR1	M	PMID: 11029414
				Escherichia phage PP01	M	PMID: 12435509
fadL	Long-chain fatty acid transport protein	NP_311254.2	Escherichia coli	Escherichia virus 933W	P	PMID: 9712754
		NP_416846.2	Escherichia coli	Enterobacteria phage T2	M	PMID: 26755501
				Enterobacteria phage vB_EcoM_IME341	M	PMID: 3286621
fepA	Ferrienterobactin receptor	NP_415116.1	Escherichia coli	Escherichia virus H8	S	PMID: 31105661
				Escherichia phage vB_EcoS-IME253	S	PMID: 17526714
		NP_459577.1	Salmonella enterica	Escherichia virus H8	S	PMID: 31105661
fhuA	Ferrichrome outer membrane transporter/phage receptor	NP_414692.1	Escherichia coli	Enterobacteria phage phi80	S	PMID: 18516490
						PMID: 26755501
						PMID: 1089064
						PMID: 11344156
				Escherichia virus T5	S	PMID: 29523063
						PMID: 18516490
				Escherichia virus HK022	S	PMID: 18516490
						PMID: 994348
				Escherichia virus N15	S	PMID: 18516490
						PMID: 8631731
				Phage mEp167	S	PMID: 18516490
				Enterobacteria phage mEp237	S	PMID: 30043202
				Enterobacterial phage mEp213	S	PMID: 23103976
				Escherichia virus T1	S	PMID: 29523063
						PMID: 15109731
						PMID: 18516490
				Enterobacteria phage vB_EcoS_IME18	S	PMID: 31105661
				Escherichia virus T7	P	PMID: 21033576
		NP_459196.1	Salmonella enterica	Enterobacteria phage ES18	S	PMID: 11344156
						PMID: 15659686
				Enterobacteria phage phi80	S	PMID: 11344156
fliC	Flagellin	NP_460912.1	Salmonella enterica	Salmonella virus SPN19	S	PMID: 22927964

fliK	Flagellar hook-length control protein	NP_460927.1	Salmonella enterica	Salmonella virus iEPS5	S	PMID: 23747700 PMID: 31105661 PMID: 25191318
fljB	Phase 2 flagellin	NP_461698.1	Salmonella enterica	Salmonella virus SPN19	S	PMID: 22927964
lamB	Maltoporin	NP_418460.1	Escherichia coli	Escherichia virus P1	M	PMID: 6237098
				Enterobacteria phage Tula	M	PMID: 7002910 PMID: 22135098
				Enterobacteria phage Bp7	M	D01083026
				Escherichia virus TP1	M	PMID: 23202520
				Escherichia virus Lambda	S	PMID: 22282803 PMID: 23202520
				Phage 21	S	PMID: 3881596
				Escherichia virus HK97	S	PMID: 25532427
nfrA	Bacteriophage adsorption protein A	NP_415100.1	Escherichia coli	Escherichia virus N4	P	PMID: 29523063 PMID: 22297513 PMID: 8226649
ompA	Outer membrane protein A	NP_415477.1	Escherichia coli	Enterobacteria phage K3	M	PMID: 16341340
				Enterobacteria phage M1	M	PMID: 26755501 PMID: 8027994
				Enterobacteria phage Ox2	M	PMID: 26755501 PMID: 6378883
				Enterobacteria phage Mi	M	PMID: 3000773
				Enterobacteria phage vB_EcoM_IME340	M	PMID: 31105661
		NP_460044.1	Salmonella enterica	Enterobacteria phage Ox2	M	doi: 10.1111/j.1574-6968.1984.tb00190.x
		NP_706879.2	Shigella flexneri	Shigella virus Sf6	P	PMID: 24673644 PMID: 30171001
ompC	Outer membrane porin C	NP_416719.1	Escherichia coli	Escherichia virus T4	M	PMID: 29523063 PMID: 24673644
				Enterobacteria phage Tula	M	PMID: 7002910 PMID: 22135098
				Enterobacteria phage Tulb	M	PMID: 26755501 PMID: 330500

ompF	Outer membrane porin F	NP_461210.1	Salmonella enterica	Enterobacteria phage Ox2	M	PMID: 22135098 PMID: 6378883
				Enterobacteria phage Bp7	M	D01083026
				Escherichia virus Lambda	S	PMID: 29523063
				Phage 434	S	PMID: 26755501 PMID: 360042
				Enterobacteria phage PA-2	S	PMID: 2428812
				Salmonella phage vB_SenMS16	M	PMID: 29523063 PMID: 23289425
				Shigella virus Sf6	P	PMID: 24673644 PMID: 30171001
				Enterobacteria phage T2	M	PMID: 26755501 PMID: 360042
				Enterobacteria phage Tula	M	PMID: 26755501 PMID: 330500
				Enterobacteria vB_EcoM_IME281 phage	M	PMID: 31105661
ompF	Outer membrane porin protein F	WP_002211302.1	Yersinia pestis	Escherichia virus Lambda	S	PMID: 29599247 PMID: 22282803
				Yersinia phage Yep-phi	P	PMID: 27722870 PMID: 24006436
				Yersinia virus R1RT	M	PMID: 26965493 PMID: 27342557
				Yersinia virus TG1	M	PMID: 27342557
phoE	Outer membrane porin PhoE	NP_414776.1	Escherichia coli	Enterobacteria phage K3	M	PMID: 16341340 PMID: 368014
pip	Phage infection protein	NP_267826.1	Lactococcus lactis	Salmonella phage 5	S	PMID: 16349376
				Lactococcus virus c2	S	PMID: 26755501 PMID: 16349376
				Enterococcus phage VPE25	S	PMID: 26735798 PMID: 30104266
				Enterococcus phage VFW	S	PMID: 27578757 PMID: 30104266 PMID: 27578757

		WP_015064319.1	Bordetella bronchiseptica	Bordetella virus BPP1	P	PMID: 26104433 PMID: 11896279
pSARA30_16	Type IV protein secretion/mating pair formation complex protein	WP_000820909.1	Salmonella enterica	Pseudomonas virus PRD1	T	PMID: 9260964
sap	S-layer protein sap	NP_843397.1	Bacillus anthracis	Bacillus virus AP50	T	PMID: 29523063 PMID: 24363347
tolC	Outer membrane protein TolC	NP_232066.1	Vibrio cholerae	Vibrio phage VP3	P	PMID: 29259138
		NP_417507.2	Escherichia coli	Escherichia virus TLS	S	PMID: 29523063 PMID: 11350161
tonB	Protein TonB	NP_415768.1	Escherichia coli	Enterobacteria phage phi80	S	PMID: 26755501
				Escherichia virus H8	S	PMID: 17526714
				Escherichia virus T1	S	PMID: 26755501 PMID: 29523063 PMID: 15109731
				Escherichia virus T7	P	PMID: 21033576
tsx	Nucleoside-specific channel-forming protein Tsx	NP_414945.1	Escherichia coli	Enterobacteria phage vB_EcoM_IME339	M	PMID: 31105661
				Enterobacteria phage T6	M	doi: 10.1007/bf00267199?
VAA_RS04305	Outer membrane protein ompK	WP_013856162.1	Vibrio anguillarum	Vibrio phage KVP40	M	PMID: 7867914 PMID: 26081633 PMID: 27412881
VAS14_16696	Outer membrane protein OmpK	WP_005370112.1	Photobacterium angustum	Vibrio phage KVP40	M	PMID: 7867914 PMID: 27412881
VC0633	Outer membrane protein U	NP_230282.1	Vibrio cholerae	Vibrio phage ICP2	P	PMID: 27412881
VCA0867	Outer membrane protein W	NP_233253.1	Vibrio cholerae	Vibrio phage VP5	P	PMID: 29523063 PMID: 24719419
yjaE	Uncharacterized protein	NP_267038.1	Lactococcus lactis	Lactococcus virus bIL67	S	PMID: 27389474
yncD	Probable TonB-dependent receptor YncD	NP_415968.1	Escherichia coli	Enterobacteria phage vB_EcoS_IME347	S	PMID: 31105661 PMID: 30146706
yueB	ESX secretion system protein YueB	NP_391064.1	Bacillus subtilis	Bacillus phage SPP1	S	PMID: 29523063 PMID: 23222721 PMID: 22171743

Fam : virus family ; M : Myoviridae ; T : Tectiviridae ; S : Siphoviridae ; P : Podoviridae.

Table S 5.4 Genome statistics.

Attribute	Value	% of total
Genome		
Size (bp)	35870	100
DNA Coding (bp)	33501	93
DNA G+C (bp)	23470	65
Genes		
Total	45	100
<u>Strand</u>		0
Sense	35	78
Antisense	10	22
<u>Functions</u>		
Predicted	27	60
Hypothetical	18	40
RNA genes		
tRNA	0	0
tmRNA	0	0

Table S 5.5 PvEP4A ORFs and BLAST hits.

ORF	AA length	Predicted product	Best phage hit	E-value	% ID	Cov.
1	234	terminase small subunit	Rhodobacter phage RcRhea	0	100	100
2	702	terminase large subunit	Rhodobacter phage RcSaxon	0	100	100
3	184	endolysin	Rhodobacter phage RcRhea	0	100	75
4	78	hypothetical protein RCCRONUS_4	Rhodobacter phage RcCronus	0	100	48
5	192	head-tail joining	Rhodobacter phage RcCronus	0	99	100
6	527	portal protein	Rhodobacter phage RcRhea	0	100	94
7	620	major head protein	Rhodobacter phage RcRhea	0	100	100
8	80	hypothetical protein RCRHEA_8	Rhodobacter phage RcRhea	4E-45	100	100
9	103	head closure Hc1	Rhodobacter phage RcCronus	0	99	100
10	130	major tail protein	Rhodobacter phage RcRhea	0	100	100
11	115	tail assembly chaperone	Rhodobacter phage RcRhea	0	100	100
12	46	tail assembly chaperone	Rhodobacter phage RcRhea	7E-23	100	28
13	752	tail length tape measure protein	Rhodobacter phage RcRhea	0	100	100
14	217	glycoside hydrolase	Rhodobacter phage RcRhea	0	100	100
15	36	hypothetical protein RCRHEA_15	Rhodobacter phage RcRhea	7E-15	100	100
16	276	tail assembly protein	Rhodobacter phage RcRhea	0	100	100
17	139	minor tail protein	Rhodobacter phage RcRhea	0	100	100
18	1302	tail protein	Rhodobacter phage RcCronus	0	100	100
19	444	hypothetical protein RCCRONUS_19	Rhodobacter phage RcCronus	0	85	55
20	136	minor tail protein	Rhodobacter phage RcSaxon	0	98	100
21	123	hypothetical protein RCRHEA_21	Rhodobacter phage RcRhea	0	100	86
22	537	DNA transfer protein	Rhodobacter phage RcRhea	0	100	100
23	155	single-stranded DNA- binding protein	Rhodobacter phage RcCronus	0	100	100
24	260	ABC transporter	Rhodobacter phage RcSaxon	0	100	98
25	160	hypothetical protein RCCRONUS_25	Rhodobacter phage RcCronus	0	100	100
26	90	hypothetical protein RCSAXON_27	Rhodobacter phage RcSaxon	0	97	100
27	86	hypothetical protein RCSAXON_28	Rhodobacter phage RcSaxon	0	98	100
28*	88	hypothetical protein	Breoghanian sp.	5E-16	79	98
29	54	hypothetical protein RCRHEA_29	Rhodobacter phage RcRhea	8E-13	94	98
30	79	hypothetical protein PM403_gp38	Pseudomonas phage BUCT-PX-5	9E-13	81	86
31	80	Arc-like repressor	Rhodobacter phage RcRhea	5E-23	100	100
32	217	plasmid partitioning protein	Rhodobacter phage RcSaxon	0	99	100
33	99	hypothetical protein RCSAXON_34	Rhodobacter phage RcSaxon	0	100	100
34	340	RepA-like replication initiator	Rhodobacter phage RcSaxon	0	100	100
35	199	hypothetical protein RCSAXON_36	Rhodobacter phage RcSaxon	0	99	100
36	411	DNA methyltransferase	Rhodobacter phage RcCronus	0	100	100
37	94	hypothetical protein RCRHEA_37	Rhodobacter phage RcRhea	0	98	100

38	202	deoxyribonucleoside 5' monophosphate phosphatase	Rhodobacter phage RcRhea	0	100	100
39	69	hypothetical protein RCRHEA_39	Rhodobacter phage RcRhea	2E-42	100	100
40	56	hypothetical protein RCCRONUS_39	Rhodobacter phage RcCronus	4E-13	100	100
41	367	DNA primase	Rhodobacter phage RcRhea	0	100	100
42	573	DNA primase	Rhodobacter phage RcSaxon	0	100	100
43	142	hypothetical protein RCRHEA_43	Rhodobacter phage RcRhea	0	100	100
44	173	hypothetical protein RCCRONUS_43	Rhodobacter phage RcCronus	0	100	100
45	200	minor tail protein	Rhodobacter phage RcSaxon	0	99	100

ID: Percentage identity; Cov.: Percentage coverage. *no phage hit was recovered from the blast so the best bacterial species hit is displayed.

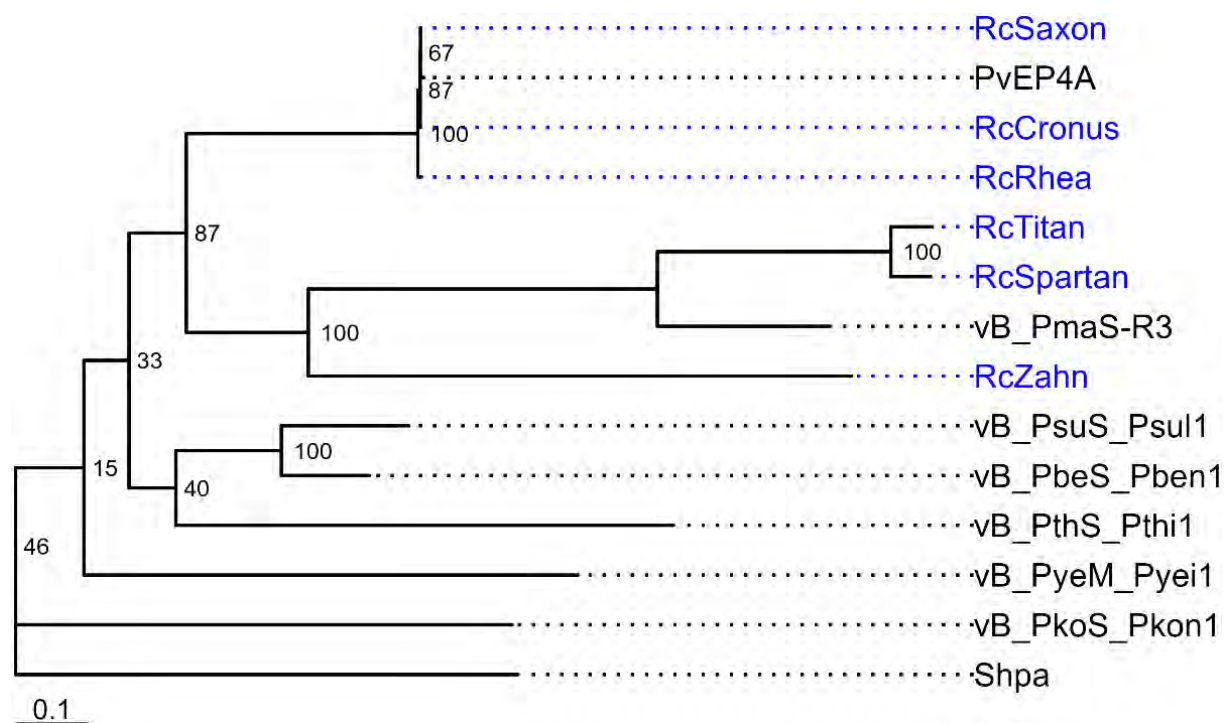


Figure S 5.1 Phylogenetic tree constructed with terminase large subunit genes of PvEP4A, some *Rhodobacter* phages (blue), and *Paracoccus* phages (black). A tree based on the whole genome sequence can be found in Figure 5.3. Scale bar: average number of substitutions per site. Numbers at the nodes are support values (percent of generated trees supporting the branch).

Table S 5.6 PvEP4A host range is narrow.

Species	Strain	Phage susceptibility
<i>Paracoccus versutus</i>	DSM582	susceptible
<i>Paracoccus versutus</i>	P2	resistant
<i>Paracoccus versutus</i>	P45	resistant
<i>Paracoccus versutus</i>	P46	partially susceptible
<i>Paracoccus pantotrophus</i>	DSM1404	resistant
<i>Paracoccus pantotrophus</i>	DSM65	resistant
<i>Paracoccus denitrificans</i>	DSM1407	resistant
<i>Paracoccus aminophilus</i>	DSM8538	resistant
<i>Paracoccus aestuarii</i>	DSM19484	partially susceptible
<i>Paracoccus alkenifer</i>	DSM11593	resistant
<i>Paracoccus siganidrum</i>	DSM26381	resistant
<i>Paracoccus yeei</i>	KB 2	resistant
<i>Paracoccus yeei</i>	KB 3	resistant
<i>Paracoccus yeei</i>	KB 4	resistant
<i>Rhodobacter capsulatus</i>	DSM1710	resistant

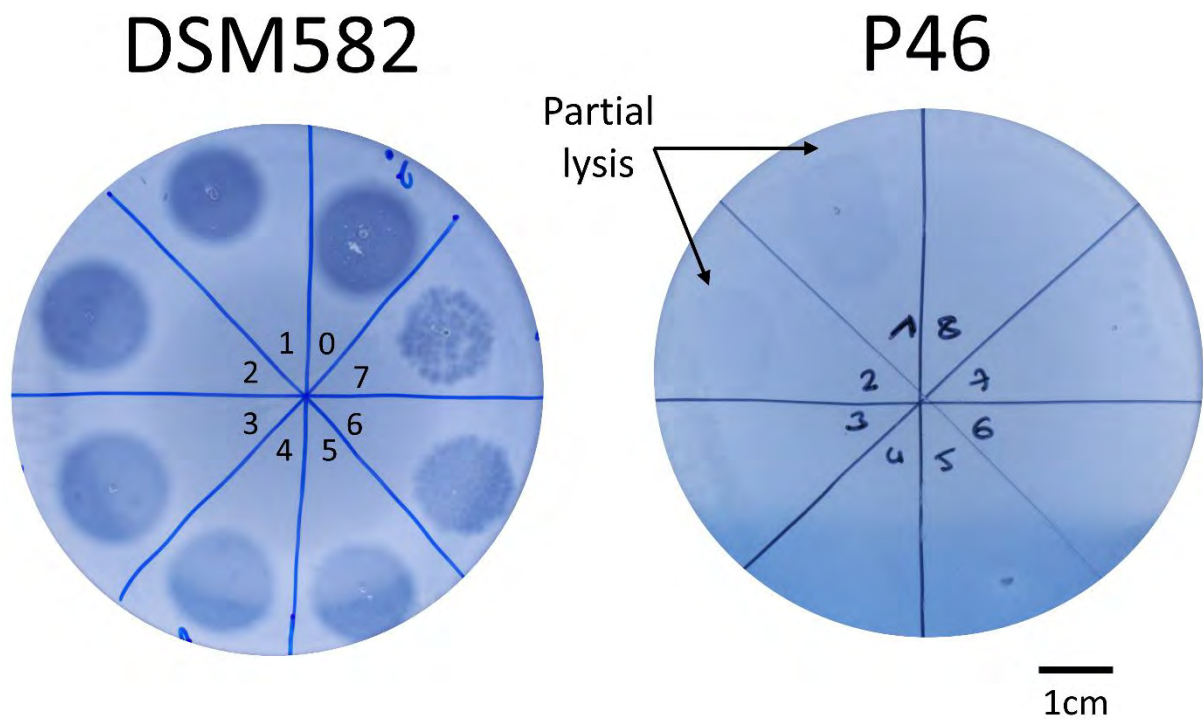


Figure S 5.2 Double-layer agar plates with *P. versutus* DSM582 (left) and P46 (right) in the soft agar overlay. Numbers indicate phage dilutions with 0 being the most concentrated and 8 the less. DSM582 exhibited lysis zones (darker zones) at all concentrations, indicating high susceptibility. In contrast, P46 showed partial resistance with faint spots at high phage concentrations (1, 2) and no visible lysis at lower concentrations. Images taken after 48h incubation at 30°C.

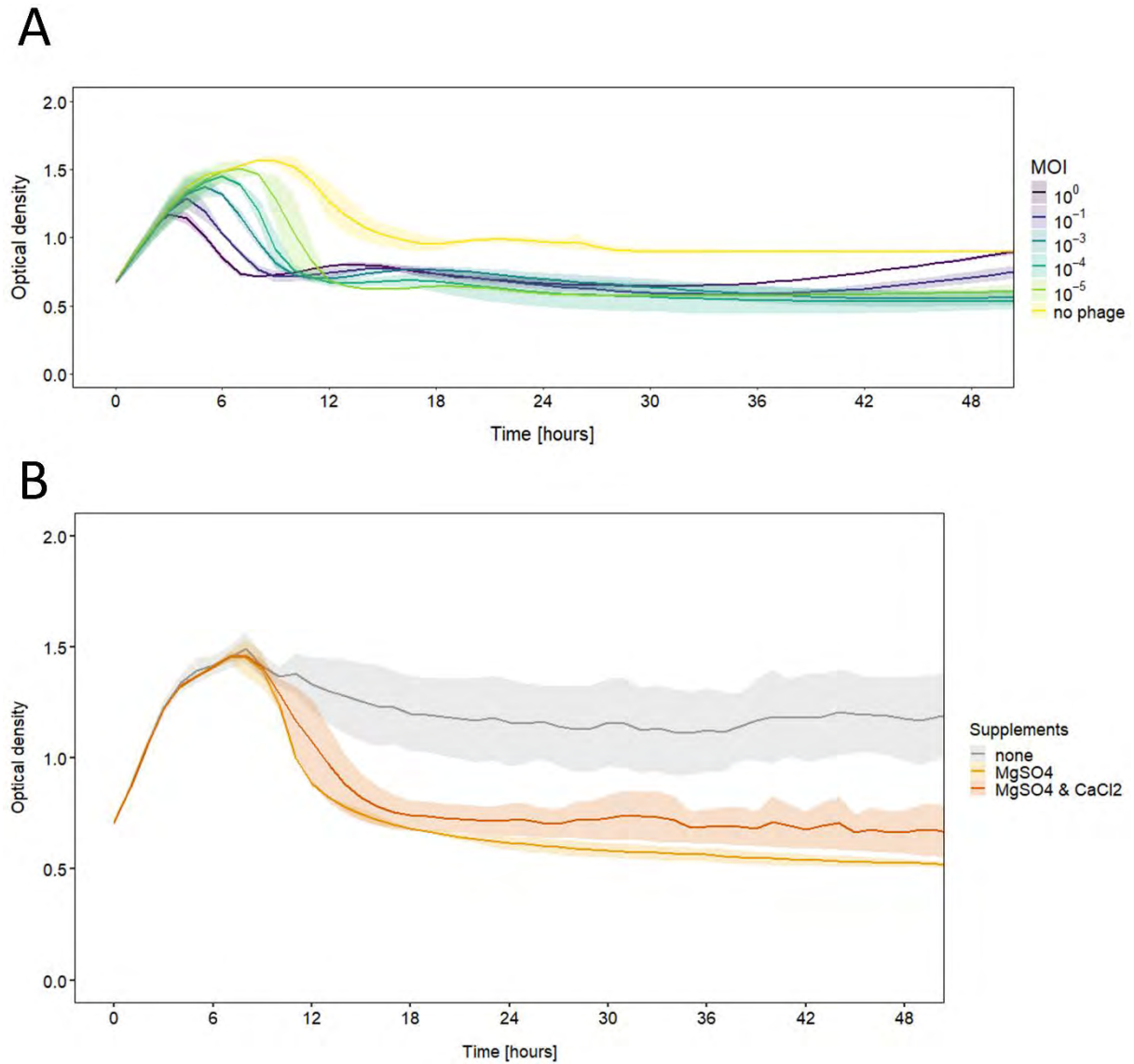


Figure S 5.3 Infection dynamics of the phage PvEP4A with its host *P. versutus* DSM582. (A) Growth curve of the host *P. versutus* with various concentrations of PvEP4A (MOI: multiplicity of infection or ratio of phages to bacteria). (B) The addition of divalent magnesium cation increased the aggregation tendency of *P. versutus*. Growth curve of the host *P. versutus* in LB with no supplement (grey), with 10 mM MgSO₄ (yellow) and with with 10 mM MgSO₄ and 1mM CaCl₂. Solid lines represent the mean, and shaded areas indicate the standard error of (A) three or (B) two replicates.

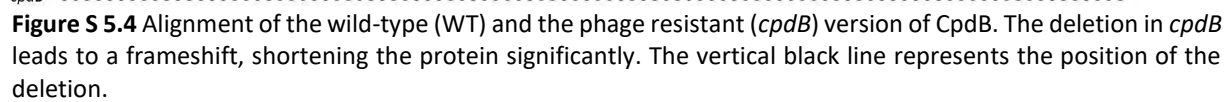


Table S 5.7 Concentration of cyclic di-nucleotides and their metabolites in the *P. versutus* wild-type and *cpdB* mutant. The concentrations of compounds were normalized by the protein concentration present in the cells. Means of 3 biological replicates and 2 technical replicates $\pm$ standard deviations are given in pmol per mg of proteins. No data were obtained for the following compounds: 3'3' c-di-AMP, 3'3' cGAMP, 2'3'-cGAMP, 3',5'-cGMP, pGpG. A Student's *t*-test highlights the significant difference for c-di-GMP.

	3'3' c-di-GMP	3',5'-cAMP	pApA
WT	1.79 $\pm$ 0.55	1.78 $\pm$ 1.39	1.32 $\pm$ 0.70
<i>cpdB</i>	14.93 $\pm$ 2.99	2.79 $\pm$ 2.30	1.72 $\pm$ 0.88
T-test p-value	9 x 10 ⁻⁷	0.38	0.41

Table S 5.8 Table of loci whose description suggests their involvement in production or degradation of c-di-GMP. Grey highlights loci whose expression was significantly dysregulated. Bold highlights a regulator.

Locus	L2FC	Padj	Gene	Description
LOCUS_43320	-4.07	1.3E-79		diguanylate cyclase response regulator
LOCUS_15400	-1.60	9.4E-13		diguanylate cyclase
LOCUS_18000	-0.09	8.4E-01		diguanylate cyclase
LOCUS_21630	-3.60	5.8E-58		diguanylate cyclase
LOCUS_45580	-4.64	3.4E-82		diguanylate cyclase
LOCUS_10330	-3.85	1.2E-78		diguanylate phosphodiesterase
LOCUS_15720	-0.24	4.7E-01		phosphoribosyl 1,2-cyclic phosphodiesterase
LOCUS_21160	-0.15	7.1E-01		phosphodiesterase
LOCUS_26430	-0.14	7.9E-01		phosphodiesterase
LOCUS_29410	-0.74	1.4E-03		multifunctional 2',3'-cyclic-nucleotide 2'-phosphodiesterase/5'-nucleotidase/3'-nucleotidase
LOCUS_38800	0.27	3.3E-01	<i>cpdB</i>	2',3'-cyclic-nucleotide 2'-phosphodiesterase

Transcripts analysis with DESeq2 p-value adjusted for multiple testing (padj) and logarithm in base 2 of the expression ratio of the wild-type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant) (L2FC).

Table S 5.9. Genes associated to cell motility in *P. versutus* based on functional annotation (Huerta-Cepas et al., 2019). The genes whose expression did not significantly vary between the wild-type and *cpdB* mutant are highlighted in grey.

Locus	L2FC	Padj	Gene	Description
LOCUS_02030	-5.12	6E-48	<i>flgH</i>	flagellar L-ring protein
LOCUS_02040	-4.10	7E-23	<i>flgA</i>	flagella basal body P-ring formation protein FlgA
LOCUS_02050	-3.16	5E-37	<i>flgG</i>	flagellar basal-body rod protein FlgG
LOCUS_02060	-5.04	1E-42	<i>flgF</i>	flagellar basal-body rod protein FlgF
LOCUS_02070	-6.17	1E-28	<i>fliQ</i>	flagellar biosynthetic protein FliQ
LOCUS_02080	-4.83	2E-27	<i>fliE</i>	flagellar hook-basal body complex protein FliE
LOCUS_02090	-5.06	3E-04	<i>flgC</i>	flagellar basal body rod protein FlgC
LOCUS_02100	-5.38	6E-39	<i>flgB</i>	flagellar biosynthesis protein FlgB
LOCUS_02110	-3.02	2E-25	<i>fliI</i>	ATP synthase
LOCUS_04800	0.20	6E-01		hypothetical protein
LOCUS_05560	-0.28	5E-01	<i>motB</i>	flagellar motor protein
LOCUS_16190	1.49	2E-01		addiction module toxin, HicA family protein
LOCUS_20450	-4.46	8E-84	<i>flgI</i>	flagellar P-ring protein 2
LOCUS_20460	-4.90	4E-90	<i>flgL</i>	flagellar hook protein FlgL
LOCUS_20470	-5.33	4E-109	<i>flgK</i>	flagellar hook-associated protein
LOCUS_20480	-5.79	4E-95	<i>flgE</i>	flagellar basal body protein
LOCUS_20490	-4.58	1E-48	<i>motB</i>	chemotaxis MotB protein
LOCUS_20920	-3.98	5E-46	<i>fliP</i>	flagellar biosynthetic protein FliP
LOCUS_20930	-4.21	7E-49	<i>fliN</i>	hypothetical protein
LOCUS_20950	-4.80	1E-89	<i>fliF</i>	flagellar M-ring protein
LOCUS_20960	-5.49	2E-72	<i>fliL</i>	hypothetical protein
LOCUS_20990	-5.19	2E-90	<i>motA</i>	flagellar motor protein MotA
LOCUS_21010	-4.67	4E-80	<i>flhA</i>	flagellar biosynthesis protein FlhA
LOCUS_21020	-3.70	5E-29	<i>fliR</i>	flagellar biosynthetic protein FliR
LOCUS_21030	-3.55	2E-31	<i>flhB</i>	flagellar biosynthetic protein FlhB
LOCUS_26030	-0.14	9E-01		membrane protein
LOCUS_27920	-4.23	5E-42	<i>flgF</i>	flagellar basal body rod protein FlgF
LOCUS_27930	-3.86	5E-39	<i>flbT</i>	flagellar biosynthesis repressor FlbT
LOCUS_27940	-3.99	3E-43	<i>flaF</i>	flagellar biosynthesis regulatory protein FlaF
LOCUS_27950	-5.35	2E-71	<i>flaA</i>	flagellin
LOCUS_27970	-5.27	1E-75		hypothetical protein
LOCUS_27980	-3.80	3E-52		hypothetical protein
LOCUS_27990	-3.17	5E-32	<i>flgD</i>	basal-body rod modification protein FlgD
LOCUS_28620	-1.60	6E-16	<i>fliG</i>	flagellar motor switch protein FliG
LOCUS_39100	-2.79	3E-28	<i>fliM</i>	flagellar switch protein FliM
LOCUS_47440	0.28	6E-01		hypothetical protein
LOCUS_51760	-0.16	8E-01		hypothetical protein
LOCUS_55480	0.11	9E-01		hypothetical protein
LOCUS_55570	0.25	7E-01		hypothetical protein

Abbreviations: padj: DESeq2 p-value adjusted for multiple testing; L2FC, logarithm in base 2 of the expression ratio of the wild type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant).

Table S 5.10. Selected genes associated to “Cell wall/membrane/envelop biogenesis” category in *P. versutus* based on functional annotation (Huerta-Cepas et al., 2019). Only genes whose expression significantly vary between the wild-type and *cpdB* mutant are shown. Bold rows highlight loci matching known phage receptors (Table S 5.16), capsule-associated loci are highlighted in grey (Table S 5.15).

Locus	L2FC	Padj	Gene	Description
LOCUS_05150*	-5.03	1E-103		hypothetical protein
LOCUS_29350*	-4.46	1E-22		peptidase
LOCUS_15920	-4.04	4E-79		glycosyl transferase
LOCUS_37280	-3.25	2E-65		hypothetical protein
LOCUS_02710	-2.01	8E-11	<i>galE1</i>	dTDP-4-dehydrorhamnose 3,5-epimerase
LOCUS_17500	-1.83	8E-13		hypothetical protein
LOCUS_20280*	-1.35	1E-12	<i>acrA_1</i>	RND transporter
LOCUS_01810	-1.32	7E-06		membrane protein
LOCUS_20740	-1.19	4E-10		outer-membrane lipoprotein carrier protein
LOCUS_01740	-1.14	8E-04		sugar transferase
LOCUS_20910*	-1.14	7E-07		RND transporter
LOCUS_20900*	-1.11	3E-07		RND transporter
LOCUS_00270	-1.06	7E-06	<i>pal</i>	peptidoglycan-associated lipoprotein
LOCUS_00440	-0.92	2E-05		lytic transglycosylase
LOCUS_10450	-0.88	3E-04		HlyD family type I secretion periplasmic adaptor subunit
LOCUS_20410	-0.82	8E-03		hypothetical protein
LOCUS_54030	-0.81	1E-02		hypothetical protein
LOCUS_16120	-0.78	3E-03		hypothetical protein
LOCUS_02140	-0.77	2E-02		hypothetical protein
LOCUS_05270	-0.74	2E-03	<i>kpsS</i>	capsule biosynthesis protein CapA
LOCUS_39040	-0.67	5E-03	<i>MltB_2</i>	murein transglycosylase
LOCUS_21570	-0.66	3E-02	<i>lpsB</i>	glycosyl transferase family 1
LOCUS_19050	-0.60	1E-02	<i>MltB_1</i>	murein transglycosylase
LOCUS_29480	-0.60	2E-02		membrane protein
LOCUS_25010	-0.59	2E-02		hypothetical protein
LOCUS_15680	-0.59	1E-02		D-alanyl-D-alanine carboxypeptidase
LOCUS_40820	-0.56	2E-02	<i>kpsE</i>	capsule polysaccharide transporter
LOCUS_06540	-0.54	1E-02		hypothetical protein
LOCUS_10030	-0.52	4E-02		hypothetical protein
LOCUS_20840	-0.50	5E-02	<i>mepA</i>	penicillin-insensitive murein endopeptidase
LOCUS_42380	0.63	1E-02		hypothetical protein
LOCUS_43410	0.68	1E-02	<i>murl</i>	glutamate racemase
LOCUS_18780	0.84	1E-02	<i>nosX_1</i>	FAD:protein FMN transferase
LOCUS_51270	0.84	1E-02		hypothetical protein
LOCUS_53220	1.38	2E-05		membrane protein
LOCUS_46880	1.78	2E-06		BCCT family transporter

Abbreviations: padj: DESeq2 p-value adjusted for multiple testing; L2FC, logarithm in base 2 of the expression ratio of the wild type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant). *Despite being the two most downregulated loci from that category, LOCUS_05150 and LOCUS_29350, are not promising candidates. LOCUS_05150 belongs to glycosyl hydrolase family 28 known for their ability to degrade pectin (Bradley et al., 2022), while LOCUS_29350 is part of the main GTA operon and annotated as a peptidase. Among the four most dysregulated RND transporter loci, three were downregulated (LOCUS_20280, LOCUS_20910, LOCUS_20900) and one upregulated (LOCUS_53220). RND transporters are primarily known as drug efflux pumps with no known involvement in phage resistance (Nikaido, 2018).

Table S 5.11. Genes whose expression significantly differs between the wild-type and *cpdB* mutant and which are associated to signal transduction in *P. versutus* based on functional annotation (Huerta-Cepas et al., 2019). The genes coding either for enzymes involved in the synthesis of c-di-GMP from two Guanosine-5'-triphosphate (GTP) molecules, diguanylate cyclases, or for degradation of c-di-GMP are highlighted in grey.

Locus	L2FC	padj	Name	Description
LOCUS_01580	-2.51	5.3 E-15		hypothetical protein
LOCUS_01590	-3.34	3.6 E-33		anti-sigma factor antagonist
LOCUS_05030	-1.48	6.2 E-09		sensor histidine kinase
LOCUS_05040	-1.37	3.8 E-07		DNA-binding response regulator
LOCUS_09220	-2.52	6.5 E-14		hypothetical protein
LOCUS_10330	-3.85	1.2 E-78		diguanylate phosphodiesterase
LOCUS_12160	-1.75	4.0 E-15		two-component sensor histidine kinase
LOCUS_13860	0.62	2.8 E-02		cyclic nucleotide-binding protein
LOCUS_14610	0.61	2.8 E-02		hypothetical protein
LOCUS_15340	-0.48	4.5 E-02		phosphoglycerate mutase
LOCUS_15400	-1.60	9.4 E-13		diguanylate cyclase
LOCUS_15450	0.57	1.5 E-02		ATPase
LOCUS_16060	-5.34	7.6 E-111		response regulator
LOCUS_17480	-4.05	5.6 E-83		hypothetical protein (homolog to CckA)
LOCUS_18500	-1.83	2.0 E-13		nickel transporter
LOCUS_18510	-2.58	1.1 E-29		fused response regulator/phosphatase
LOCUS_21630	-3.60	5.8 E-58		diguanylate cyclase
LOCUS_22070	-0.50	3.3 E-02		two-component response regulator
LOCUS_26370	-0.79	6.0 E-04		DNA-binding response regulator
LOCUS_38300	0.65	8.3 E-03		hypothetical protein
LOCUS_42040	-0.74	1.3 E-04	ntrX	sigma-54-dependent Fis family transcriptional regulator
LOCUS_42780	0.85	2.9 E-03	phoB_2	response regulator
LOCUS_43320	-4.07	1.3 E-79		diguanylate cyclase response regulator
LOCUS_45580	-4.64	3.4 E-82		GGDEF domain-containing protein

Abbreviations: padj: DESeq2 p value adjusted for multiple testing; L2FC, logarithm in base 2 of the expression ratio of the wild type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant).

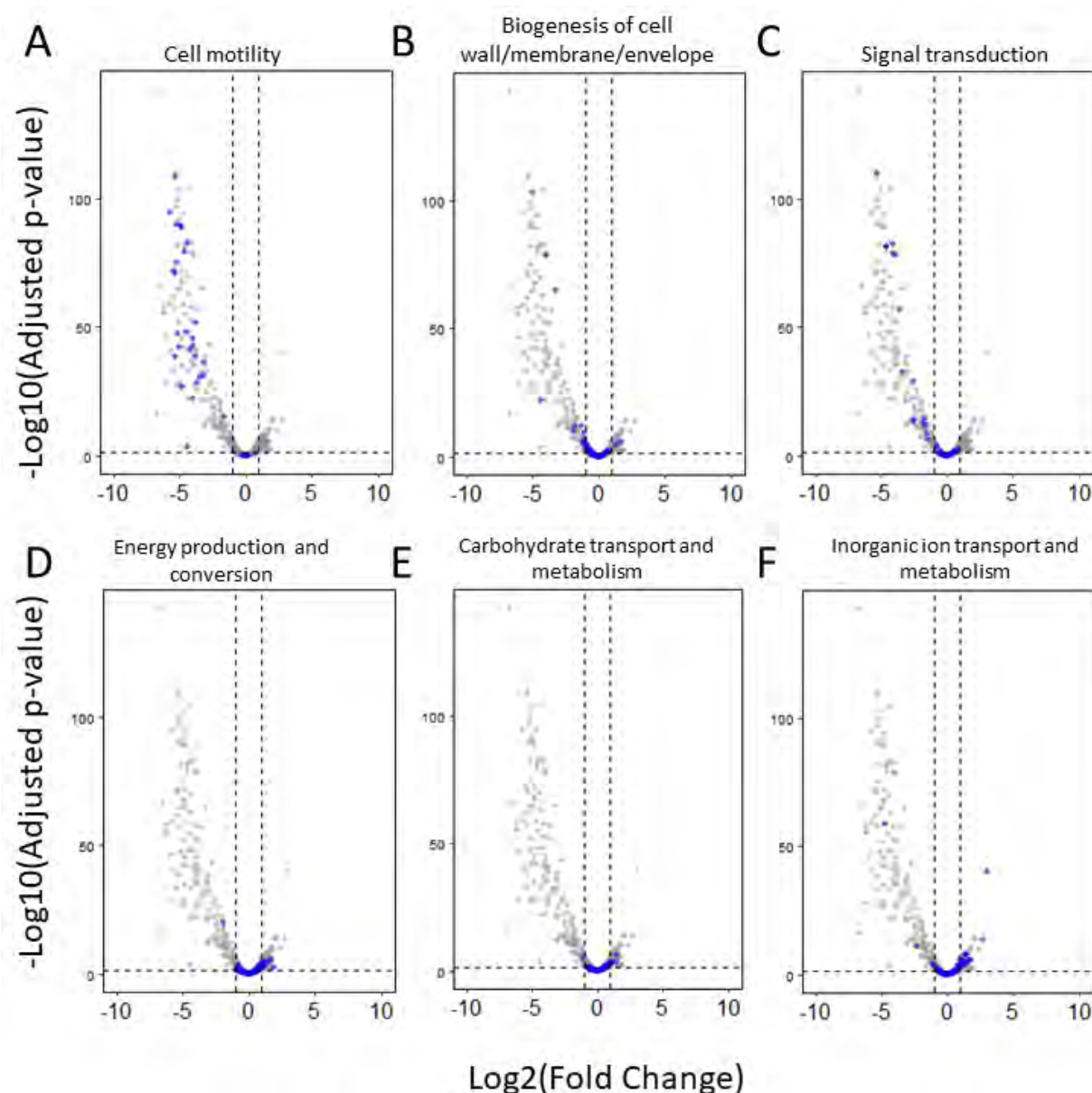


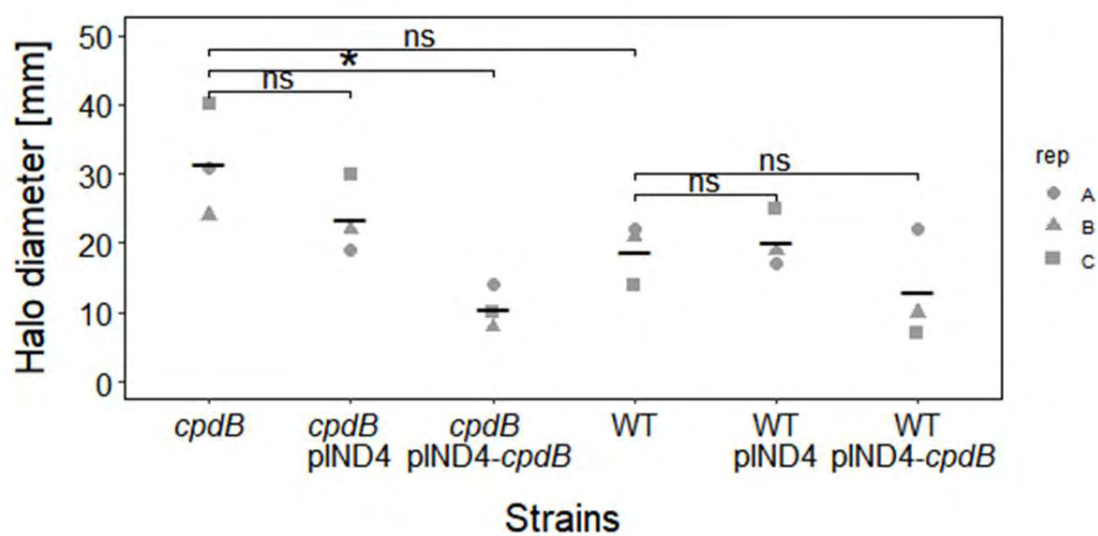
Figure S 5.5 Volcano Plots displaying differential gene expression analysis from RNA-seq data separated by COG categories containing mainly downregulated (A, B, C) and upregulated genes (D, E, F). Each dot represents a locus, with blue dots indicating loci within a specific COG category. The x-axis denotes the log₂ fold change (FC) between conditions, and the y-axis represents the -log₁₀ p-values. The horizontal line marks the p-value threshold ($p = 0.05$), and the vertical lines indicate the fold change thresholds ($\text{Log}_2(\text{Fold Change}) = \pm 1$).

Table S 5.12. Genes whose expression significantly differs between the wild-type and *cpdB* mutant and which are classified to carbohydrate transport and metabolism COG category based on functional annotation (Huerta-Cepas et al., 2019).

Locus	L2FC	Padj	Gene	Description
LOCUS_44320	1.63	3E-04	thuG	trehalose/maltose ABC transporter permease
LOCUS_44340	1.57	4E-07	thuE	ABC transporter substrate-binding protein
LOCUS_15840	1.19	9E-04	rbsC	ribose ABC transporter permease
LOCUS_18430	1.18	8E-04		ABC transporter permease
LOCUS_07840	1.17	7E-05		Bcr/CflA family drug resistance efflux transporter
LOCUS_15850	1.16	2E-04	rbsA	ribonucleotide-diphosphate reductase subunit alpha
LOCUS_46680	1.06	2E-02		sugar ABC transporter substrate-binding protein
LOCUS_18440	1.05	6E-04		ABC transporter ATP-binding protein
LOCUS_43790	1.03	2E-02		sugar ABC transporter permease
LOCUS_15860	0.98	3E-03	rbsB	ribose ABC transporter substrate-binding protein
LOCUS_03180	0.97	3E-03		MFS transporter
LOCUS_43800	0.97	2E-02		ABC transporter permease
LOCUS_18420	0.92	3E-03		sugar ABC transporter
LOCUS_43820	0.88	4E-03		ABC transporter ATP-binding protein
LOCUS_42570	0.88	6E-03		sugar ABC transporter permease
LOCUS_28900	0.87	7E-04		GntP protein
LOCUS_42580	0.87	9E-03	smoE	sugar ABC transporter substrate-binding protein
LOCUS_33340	0.81	4E-02		MFS transporter
LOCUS_28890	0.79	8E-03		gluconokinase
LOCUS_43770	0.79	5E-02		ABC transporter substrate-binding protein
LOCUS_48810	0.66	4E-02	uxuA	mannonate dehydratase
LOCUS_48770	0.65	3E-02		hypothetical protein
LOCUS_53760	0.64	3E-02		ABC transporter permease
LOCUS_53770	0.62	1E-02		ABC transporter ATP-binding protein
LOCUS_10130	-0.55	4E-02		hypothetical protein
LOCUS_25010	-0.59	2E-02		hypothetical protein
LOCUS_54570	-0.75	4E-02		C4-dicarboxylate ABC transporter
LOCUS_21880	-0.83	3E-02		2-methylisocitrate lyase
LOCUS_21310	-0.87	4E-02	eda_1	2-dehydro-3-deoxy-6-phosphogalactonate aldolase
LOCUS_48260	-0.95	2E-02	manB	phosphomannomutase
LOCUS_36960	-1.19	4E-06		hydroxypyruvate isomerase

Abbreviations: padj: DESeq2 p value adjusted for multiple testing; L2FC, logarithm in base 2 of the expression ratio of the wild type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant).

A



B

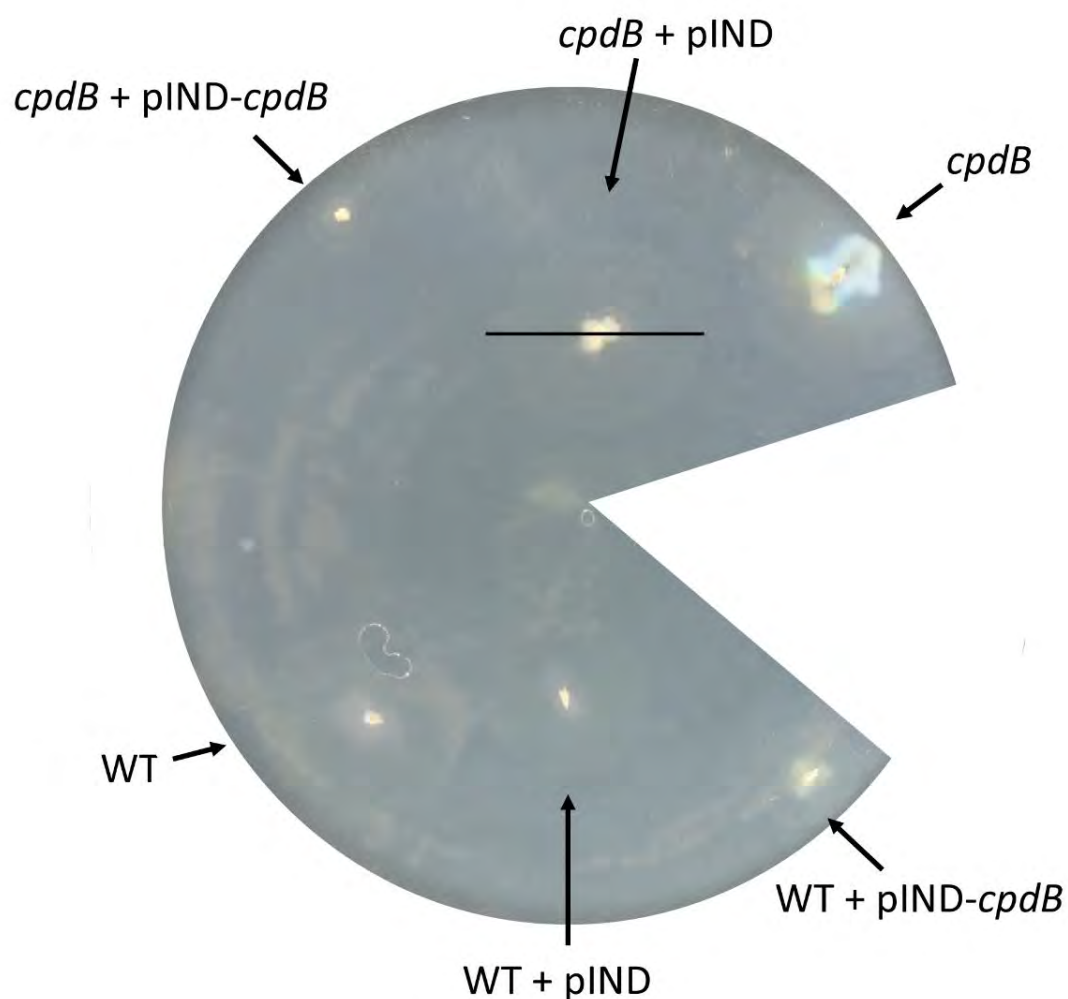


Figure S 5.6 Increased concentration of c-di-GMP due to nucleotides deletion in *cpdB* in *P. versutus* led to a light increase in motility. The swimming assay was performed with wild-type strain (WT), the phage resistant mutant (*cpdB*) and complemented strains with the inducible-expression plasmid pIND4 empty or carrying *cpdB*. (A) Dots represent replicates with a specific shape for each plate, bars the means. (B) Partial picture of the plate of replicate B (triangle) with the black bar corresponding to 22mm.

Table S 5.13 This table shows the top 3 BLASTp matches of T4P proteins from *Pseudomonas aeruginosa* and Tad pilus, a type of T4P, proteins from *Caulobacter vibrioides* against *P. versutus* DSM582 proteins. *P. aeruginosa* proteins were chosen for their role in phage reception, while *C. vibrioides* proteins were selected due to their phylogenetic proximity to *P. versutus*. The matches exhibit weak homology to *P. versutus* proteins, indicating no significant findings.

Protein	Locus	BlastP		RNAseq		Gene	Product
		%ID	E-val.	L2FC	Padj		
P. aeruginosa T4P							
PilA	LOCUS_39110	32	1E+00	-0.15	7E-01		sodium:solute symporter
	LOCUS_09980	28	1E+00	-0.35	2E-01		hypothetical protein
	LOCUS_44080	25	2E+00	0.06	9E-01	lepA	elongation factor 4
PilB	LOCUS_31550	30	8E-07	0.09	NA	trbB	P-type conjugative transfer ATPase TrbB
	LOCUS_12670	31	2E-06	0.04	9E-01	trbB	P-type conjugative transfer ATPase TrbB
	LOCUS_13900	31	4E-06	0.10	9E-01	trbB	P-type conjugative transfer ATPase TrbB
PilC	LOCUS_16690	39	9E-01	0.06	9E-01		hypothetical protein
	LOCUS_53280	32	1E+00	0.67	1E-02		twin-arginine translocation pathway signal
	LOCUS_44450	24	2E+00	-0.35	NA		MFS transporter
PilD	LOCUS_05380	44	4E-01	0.29	4E-01		membrane protein
	LOCUS_54450	27	6E-01	1.36	2E-04	pcaC	4-carboxymuconolactone decarboxylase
	LOCUS_23530	32	1E+00	0.01	1E+00	livM	branched-chain amino acid ABC transporter permease
PilE	LOCUS_11070	36	4E-01	0.10	NA		hypothetical protein
	LOCUS_34390	37	6E-01	0.56	5E-02	arcA	arginine deiminase
	LOCUS_09130	33	1E+00	0.17	7E-01	cbbX	CbbX protein
PilT	LOCUS_55570	30	5E-06	0.19	7E-01		hypothetical protein
	LOCUS_51460	47	1E-02	-0.16	NA		hypothetical protein
	LOCUS_23300	29	4E-02	-0.25	5E-01		ABC transporter ATP-binding protein
C. vibrioides Tad pilus (T4cP)							
pilA	LOCUS_18210	40	2E-01	-0.01	1E+00	deoD	purine nucleoside phosphorylase
	LOCUS_25360	44	8E-01	0.20	6E-01	dapF	diaminopimelate epimerase
	LOCUS_26800	40	3E+00	-1.12	5E-05		heavy metal-transporting ATPase
CpaA	LOCUS_40520	34	2E+00	-0.16	7E-01		hypothetical protein
	LOCUS_42790	32	2E+00	0.58	5E-02	actP	cation acetate symporter
	LOCUS_29030	35	5E+00	0.12	8E-01		hypothetical protein
CpaB	LOCUS_16610	43	4E-01	-0.12	7E-01		MBL fold hydrolase
	LOCUS_00690	45	8E-01	-0.39	1E-01		ABC transporter permease
	LOCUS_45410	26	9E-01	0.42	2E-01		30S ribosomal protein S1
CpaC	LOCUS_13900	34	5E-01	0.10	9E-01	trbB	P-type conjugative transfer ATPase TrbB
	LOCUS_32820	37	5E-01	0.51	2E-01	trbB	P-type conjugative transfer ATPase TrbB
	LOCUS_14780	41	2E+00	-0.30	4E-01	cls	cardiolipin synthase
CpaD	LOCUS_49550	42	3E-01	0.02	1E+00		oxidoreductase
	LOCUS_41040	42	2E+00	0.10	8E-01		integrase
	LOCUS_22080	37	2E+00	-0.19	5E-01		2-nitropropane dioxygenase
CpaE	LOCUS_53010	21	5E-02	-0.30	3E-01	minD	site-determining protein
	LOCUS_22220	44	6E-02	0.19	6E-01	parA	chromosome partitioning protein ParA
	LOCUS_31260	43	5E-01	-0.26	NA		cobyrinic acid a,c-diamide synthase
CpaF	LOCUS_13900	34	7E-40	0.10	9E-01	trbB	P-type conjugative transfer ATPase TrbB
	LOCUS_12670	34	3E-39	0.04	9E-01	trbB	P-type conjugative transfer ATPase TrbB
	LOCUS_31550	37	9E-39	0.09	NA	trbB	P-type conjugative transfer ATPase TrbB

CpaG	LOCUS_13540	38	1E+00	0.11	8E-01	vdh_1	salicylaldehyde dehydrogenase
	LOCUS_07640	37	2E+00	-0.33	5E-01		hypothetical protein
	LOCUS_46130	45	2E+00	-0.26	3E-01	ftsZ1	cell division protein FtsZ
CpaH	LOCUS_53510	41	3E-01	0.09	NA	phoD	phosphonate ABC transporter substrate-binding protein
	LOCUS_52430	40	3E+00	-0.32	NA		HlyD family type I secretion periplasmic adaptor subunit
	LOCUS_24250	27	5E+00	0.41	1E-01	coxI_2	cytochrome c oxidase subunit 1
CpaI	LOCUS_16430	35	7E-01	-0.33	3E-01		laccase domain protein
	LOCUS_31820	31	1E+00	0.18	7E-01		N5,N10-methylene tetrahydromethanopterin reductase
	LOCUS_08050	25	2E+00	1.15	1E-05		TonB-dependent receptor
CpaJ	LOCUS_09080	30	3E+00	0.10	9E-01	prkA	phosphoribulokinase
	LOCUS_10090	29	6E+00	0.19	6E-01		ABC transporter
	LOCUS_20110	30	5E+00	1.56	3E-06		TonB-dependent receptor
CpaK	LOCUS_15800	36	2E+00	0.45	2E-01		hypothetical protein
	LOCUS_25410	31	5E+00	0.08	8E-01	dxs2	1-deoxy-D-xylulose-5-phosphate synthase 2
	LOCUS_26620	32	9E+00	0.34	3E-01	pqqC	pyrroloquinoline-quinone synthase

Abbreviations: padj: DESeq2 p value adjusted for multiple testing; L2FC, logarithm in base 2 of the expression ratio of the wild type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant).

Table S 5.14 Loci whose annotation (description and/or KEGG pathway) suggested a link with LPS, with results from the transcriptomics study. Grey highlights the only significantly dysregulated gene of this table.

Locus	L2FC	Padj	Gene	Description	KEGG pathways
LOCUS_02250	-0.10	0.84	<i>lpxK</i>	Transfers the gamma-phosphate of ATP to the 4'-position of a tetraacyldisaccharide 1-phosphate intermediate to form tetraacyldisaccharide 1,4'-bis-phosphate (lipid IVA)	LPS biosynthesis proteins
LOCUS_15150	0.10	0.82	<i>lptF</i>	LPS export ABC transporter permease LptF	Transporters
LOCUS_15160	0.08	0.83	<i>lptG</i>	LPS export ABC transporter permease LptG	Transporters
LOCUS_15170	0.03	0.93	<i>lptD</i>	Involved in the assembly of LPS at the surface of the outer membrane	Transporters
LOCUS_19700	-0.15	0.70	<i>kdtA</i>	3-deoxy-D-manno-octulosonic acid transferase	LPS biosynthesis proteins
LOCUS_20070	0.06	0.89	<i>lptA</i>	Involved in the assembly of LPS. Required for the translocation of LPS from the inner membrane to the outer membrane	Transporters
LOCUS_20090	0.13	0.76	<i>gutQ</i>	KpsF/GutQ protein	LPS biosynthesis proteins
LOCUS_21390	0.24	0.46		Glycosyltransferase family 25 (LPS biosynthesis protein)	Glycan metabolism
LOCUS_21420	0.13	0.76		glycosyltransferase involved in LPS biosynthesis	Glycan metabolism
LOCUS_21510	-0.16	0.71	<i>lptE</i>	LPS-assembly	Transporters
LOCUS_21570	-0.59	0.03	<i>lpsB</i>	Glycosyl transferases group 1	LPS biosynthesis proteins
LOCUS_37040	-0.16	0.67	<i>lpxA</i>	Involved in the biosynthesis of lipid A, a phosphorylated glycolipid that anchors the LPS to the outer membrane of the cell	NA
LOCUS_37060	-0.01	0.99	<i>lpxB</i>	Condensation of UDP-2,3-diacylglucosamine and 2,3-diacylglucosamine-1-phosphate to form lipid A disaccharide, a precursor of lipid A, a phosphorylated glycolipid that anchors the LPS to the outer membrane of the cell	NA
LOCUS_40760	0.31	0.30	<i>lpxD</i>	Catalyzes the N-acylation of UDP-3-O-acylglucosamine using 3-hydroxyacyl-ACP as the acyl donor. Is involved in the biosynthesis of lipid A, a phosphorylated glycolipid that anchors the LPS to the outer membrane of the cell	NA

Abbreviations: padj: DESeq2 p value adjusted for multiple testing; L2FC, logarithm in base 2 of the expression ratio of the wild type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant).

Table S 5.15 Significantly differentially expressed genes predicted belonging to capsule synthetic pathways by CapsuleFinder (Rendueles *et al.*, 2017). In grey are highlighted genes upregulated in the *cpdB* mutant. Bold highlights loci discussed in text.

System	HMM profile	Stat.	%cov		Locus	L2FC	Padj	Gene	Product
			Profile	Seq.					
ABC	Acetyltransf_1	A	95	31	LOCUS_16070	-0.51	5E-02		acetyltransferase
	Acetyltransf_1	A	93	55	LOCUS_25690	-0.74	3E-02	mobC	molybdopterin-guanine dinucleotide biosynthesis protein MobC
	Epimerase	A	100	69	LOCUS_02710	-1.93	8E-11	galE1	dTDP-4-dehydrorhamnose 3,5-epimerase
	Epimerase	A	99	32	LOCUS_25010	-0.54	2E-02		hypothetical protein
	Epimerase	A	69	68	LOCUS_26400	-0.80	9E-04		oxidoreductase
	Glycos_transf_1	A	99	50	LOCUS_21570	-0.59	3E-02	lpsB	glycosyl transferase family 1
	WchA	A	99	82	LOCUS_01740	-1.04	8E-04		sugar transferase
	KpsE	M	100	60	LOCUS_40820	-0.51	2E-02	kpsE	capsule polysaccharide transporter
	KpsS	A	99	82	LOCUS_05270	-0.68	2E-03	kpsS	capsule biosynthesis protein CapA
	KpsM	M	93	87	LOCUS_22450	1.51	4E-06		transport permease protein
	KpsT	M	91	43	LOCUS_15850	1.07	2E-04	rbsA	ribonucleotide-diphosphate reductase subunit alpha
	KpsT	M	70	66	LOCUS_22440	1.38	2E-04		2-phenylethanol ABC transporter ATP-binding protein
	KpsT	M	92	67	LOCUS_42850	0.97	6E-04	nosF	ABC transporter ATP-binding protein
	KpsT	M	76	49	LOCUS_43820	0.79	4E-03		ABC transporter ATP-binding protein
	KpsT	M	95	64	LOCUS_44310	0.91	1E-02	thuK	sugar ABC transporter ATP-binding protein
GroupIV_s_stricte	YihO	M	53	55	LOCUS_07840	1.08	7E-05		Bcr/CflA family drug resistance efflux transporter
PGA	CapA	A	72	31	LOCUS_29410	-0.74	1E-03		multifunctional 2',3'-cyclic-nucleotide phosphodiesterase/5'-nucleotidase/ 3'-nucleotidase
Syn_cps3	Cps3M	A	97	96	LOCUS_48260	-0.81	2E-02	manB	phosphomannomutase
	Cps3S	M	88	63	LOCUS_15920	-4.02	4E-79		glycosyl transferase

Abbreviations: Status (Stat.) refers to the requirement of the genes for the specific system with mandatory (M) and accessory (A); Coverage percentage (%cov) of profile and sequence (sequence) refer to the percentage of the HMM profile or the input sequence that is covered by the match. padj: DESeq2 p-value adjusted for multiple testing; L2FC, logarithm in base 2 of the expression ratio of the wild type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant).

Table S 5.16 Significantly upregulated homologs of known phage receptors upregulated in *cpdB* mutant.

Gene	Refseq	Locus	%ID	E-val	L2FC	Padj	Product	COG
<i>fhuA</i>	NP_459196.1	LOCUS_36650	23	4E-24	3.06	9E-41	iron transport outer membrane receptor	P
		LOCUS_50160	29	5E-80	0.57	5E-02	TonB-dependent receptor	P
	NP_414692.1	LOCUS_46070	25	2E-33	1.00	1E-04	ligand-gated channel	P
		LOCUS_17310	24	5E-22	1.35	1E-08	ferric-rhodotorulic acid/ferric-coproden receptor FhuE	P
<i>fepA</i>	NP_415116.1	LOCUS_09230	44	0E+00	0.94	3E-06	outer membrane receptor FepA	P
	NP_415116.1	LOCUS_02300	29	2E-68	1.07	5E-05	outer membrane protein	P
	NP_459577.1	LOCUS_20110	23	6E-05	1.56	3E-06	TonB-dependent receptor	P
<i>yncD</i>	NP_415968.1	LOCUS_08050	23	2E-08	1.15	1E-05	TonB-dependent receptor	P
<i>yncD</i>	NP_415968.1	LOCUS_42500	23	1E-20	0.84	5E-03	ligand-gated channel	P
<i>ompC</i>	NP_416719.1	LOCUS_49750	28	9E-01	1.53	4E-05	quinol dehydrogenase ferredoxin subunit NapH	C
<i>btuB</i>	NP_418401.1	LOCUS_40890	44	3E-01	0.59	2E-02	hypothetical protein	NA
		LOCUS_28400	44	3E-01	0.55	4E-02	hypothetical protein	NA
pSARA 30_16	WP_000820909.1	LOCUS_09940	36	4E-01	0.87	5E-04	branched-chain amino acid ABC transporter substrate-binding protein	E

BLASTp of reference sequence (Refseq) on *P. versutus* proteins with percentage identity (%ID) et e-value (E-val) indicated. Transcripts analysis with DESeq2 p-value adjusted for multiple testing (padj) and logarithm in base 2 of the expression ratio of the wild-type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant) (L2FC). The Clusters of Orthologous Genes (COG) categories are listed as follows: C: Energy production and conversion; E: Amino Acid transport and metabolism; P: Inorganic ion transport and metabolism; NA: not available.

Chapter 6 - Study of a conserved Gene Transfer Agent in *Paracoccus* species

6.1 Foreword

In this chapter, I present a new bacterial model for investigating horizontal gene transfer mediated by gene transfer agents (GTAs)—*Paracoccus versutus* DSM582—and examine aspects of this particular HGT mechanism.

The key findings of this chapter are the following:

- Homologues of the main RcGTA operon are present in multiple species belonging to the *Paracoccus* genus.
- *P. versutus* DSM582 is the only among a set of 14 *Paracoccus* strains to be able to perform horizontal gene transfer mediated by GTAs at detectable levels.
- We assembled the full genomic sequence of *P. versutus* DSM582, which contains four circular replicons.
- The packaging of genomic DNA into GTAs is biased towards two regions close to the origin of replication of two secondary replicons.
- The PvGTA operon is mainly expressed during the exponential phase.
- High intracellular cyclic di-GMP concentration decreases GTA production and reception in *P. versutus*.

6.2 Abstract

Gene transfer agents (GTAs) are virus-like particles encoded in many bacterial genomes. Unlike viruses, GTAs mainly package fragments of the genome of their host, typically from 4 to 14 kilobase pairs in size. These fragments can be subsequently transferred to related recipient cells, enabling horizontal gene transfer. While the *Rhodobacter capsulatus* GTA (RcGTA) is the best-characterized model, the operon encoding RcGTA structural components is widespread in various orders of alphaproteobacteria, including Rhizobiales, Rhodobacterales, Caulobacterales, and others, suggesting that it was already present in the common ancestor of the main alphaproteobacteria clades. However, research on GTA-mediated gene transfer has largely centered on *R. capsulatus* as the primary model organism. New models would enable comparative analyses of genetic and regulatory mechanisms, revealing both conserved core elements and species-specific adaptations. In this study, we introduce *Paracoccus versutus* as a novel model organism for investigating horizontal gene transfer (HGT) mediated by GTAs. Our comparative analysis reveals both similarities and differences between *P. versutus* and *R. capsulatus* GTAs. A negative correlation between intracellular cyclic di-GMP concentrations and GTA transfer rates is conserved in both species, and GTA particles package DNA fragments of similar sizes. However, significant differences exist. In *P. versutus*, the main GTA operon exhibits maximal induction during the exponential growth phase, contrasting with the stationary phase expression observed in *R. capsulatus*. Furthermore, *P. versutus* GTAs demonstrate a preference for encapsulating genomic regions proximal to replication origins, suggesting a potential link between genome replication and GTA production. These distinct characteristics imply that the evolutionary advantages conferred by GTAs may vary between *P. versutus* and *R. capsulatus*. The introduction of *P. versutus* as an alternative model organism for GTA studies offers an opportunity to expand our understanding of these intriguing DNA transfer vehicles and their roles in bacterial evolution and adaptation across diverse species.

6.3 Introduction

Horizontal gene transfer is a major driver of bacterial evolution, facilitating the acquisition of novel genetic material that can rapidly alter cellular physiology and survival. This process influences both short-term bacterial adaptability and long-term evolutionary trajectories (Daubin and Szöllősi, 2016). While natural transformation, bacteriophage-mediated transduction, and conjugation are well-established mechanisms of HGT, gene transfer agents (GTAs) represent a less studied but increasingly important fourth mechanism.

GTAs are hypothesized to have evolved from defective, “domesticated” prophages, mediating gene transfer through phage-like protein capsids (Lang *et al.*, 2017). Unlike bacteriophages, GTAs preferentially package host DNA fragments. GTAs have garnered increasing interest due to their role in bacterial gene flow, ecological interactions, and evolutionary dynamics, especially in the marine ecosystem (McDaniel *et al.*, 2010, 2012; Von Wintersdorff *et al.*, 2016).

Initially discovered in the alphaproteobacterium *Rhodobacter capsulatus* (Marrs, 1974), GTAs have since been identified in various species, including *Desulfovibrio desulfuricans*, *Methanococcus voltae*, *Brachyspira hyodysenteriae* and *Bartonella* spp. (Chapter 1, Table 2) (Banks and Le, 2024; Craske *et al.*, 2024). The GTAs from these species exhibit independent evolutionary origins (Redfield *et al.*, 2019). Nevertheless, RcGTA-like elements are widely distributed among alphaproteobacteria, particularly within the Rhodobacterales order (Lang and Beatty, 2007; Shakya *et al.*, 2017; Redfield *et al.*, 2019). Experimental evidence has demonstrated functional GTA production in species such as *Ruegeria pomeroyi*, *Roseovarius nubinhibens*, and *Dinoroseobacter shibae* (Biers *et al.*, 2008; McDaniel *et al.*, 2010; Tomasch *et al.*, 2018). However, variations in GTA characteristics across species underscore the importance of identifying and studying new models to elucidate shared and species-specific features, especially because the majority of our current understanding stems from the *R. capsulatus* GTA (RcGTA) model.

The production of RcGTAs involves a tightly regulated and intricate process within donor cells. It begins with the expression of phage-like genes, leading to the assembly of GTA particles and the packaging of approximately 4 kb fragments of random host genomic DNA (Yen *et al.*, 1979; Hynes *et al.*, 2012). Once assembled, the donor cell undergoes lysis, releasing the GTA

particles into the environment (Hynes *et al.*, 2012; Westbye *et al.*, 2013). This process is governed by a complex regulatory network involving quorum sensing and the CtrA phosphorelay system (Chapter 1, Figure 5) (Sherlock and Fogg, 2022b). Central to this regulation is the transcriptional activator GafA, which binds to the promoter region of the GTA operon to initiate the synthesis of structural components essential for GTA assembly (Fogg, 2019). Additionally, cyclic di-GMP, a secondary messenger, modulates CtrA phosphorylation, further influencing the transcription of GTA-related genes (Farrera-Calderon *et al.*, 2020). Environmental cues such as high cell density or nutrient limitation trigger this regulatory cascade (Marrs, 1974; Lang and Beatty, 2000; Westbye *et al.*, 2017), ensuring that GTA production is optimized for conditions favoring gene transfer within bacterial communities.

The reception of RcGTA particles is equally complex and highly regulated. Once released, GTA particles bind to recipient cells through interactions with extracellular polysaccharides and inject their DNA into the periplasmic space (Brimacombe *et al.*, 2013; Alim *et al.*, 2023). The DNA is subsequently transported into the cytoplasm via natural transformation-like machinery, where it undergoes homologous recombination mediated by RecA, integrating into the recipient genome (Brimacombe *et al.*, 2014, 2015). This process is also influenced by quorum sensing and CtrA-dependent regulatory mechanisms (Brimacombe *et al.*, 2014, 2015), ensuring efficient DNA uptake and integration. The persistence of HGT mediated by GTAs suggests that the population-level benefits outweigh the loss of producer cells. However, uncovering the full scope of GTA evolutionary significance and ecological roles requires the investigation of additional models beyond the well-characterized RcGTA system.

The *Paracoccus* genus, belonging to the Rhodobacterales order, presents a promising clade for discovering novel RcGTA-like models. Members of this genus have demonstrated remarkable adaptability to diverse ecological niches and lifestyles (Kelly *et al.*, 2006; Lasek *et al.*, 2018; Hollensteiner *et al.*, 2023). The potential for GTA-mediated HGT in *Paracoccus* species remains unexplored, offering a unique opportunity to expand our understanding of this important genetic transfer mechanism within the alphaproteobacteria.

This study investigates the occurrence and regulation of GTA-mediated HGT in the *Paracoccus* genus, introducing *Paracoccus versutus* DSM582 as a new model organism. Our findings reveal significant differences between *P. versutus* and the established *R. capsulatus* model. Notably, *P. versutus* primarily expresses the main GTA operon during the exponential phase, in contrast

to *R. capsulatus*, which shows stationary phase induction. Additionally, DNA fragments encapsulated in *P. versutus* GTA particles exhibit a bias toward genomic regions near the replication origins of two out of four replicons in its multipartite genome, differing from the random DNA packaging typically observed in *R. capsulatus*. However, in both *P. versutus* and in *R. capsulatus*, there is a negative correlation between elevated cyclic di-GMP concentrations and GTA transfer rates. The identification of *P. versutus* as a producer and recipient of functional GTAs offers a novel and convenient organism for further investigation of the functions, biogenesis, and regulation of GTAs. This expands our understanding of these intriguing genetic elements across diverse bacterial species and provides new insights into the complexities of GTA-mediated HGT.

6.4 Materials & Methods

6.4.1 Bacterial strains and culture conditions

The bacterial strains, plasmids, and primers used in this chapter are listed in Table S 6.1, Table S 6.2, and Table S 6.3. The bacteria were routinely grown in LB broth (tryptone, 10 g/L; yeast extract 5 g/L; NaCl, 5 g/L; tryptone and yeast extract from Oxoid) or on LB supplemented with 1.5% agar. Where necessary, the media were supplemented with kanamycin (Km, 50 µg/mL), gentamicin, (Gm, µg/mL), tetracyclin (Tet, 12 µg/mL), and isopropyl-β-d-thiogalactopyranoside (IPTG, 1mM) to trigger expression of genes carried by the (IPTG)-inducible-expression plasmid pIND4 (Ind *et al.*, 2009). Spontaneous rifampicin resistant colonies were isolated by plating the strain of interest onto selective LB agar. *E. coli* strains were grown at 37°C; *Paracoccus* strains were grown at 30°C; for liquid cultures, an agitation of 180 rpm was applied.

6.4.2 Strains and plasmids construction

Plasmid construction and amplification were done in *E. coli* DH5α-λpir. Chemocompetent cells were prepared and transformed according to standard CaCl₂-based protocols (Sambrook and Russell, 2001). Electroporation into *P. versutus* was done at 25 µF, 400 Ω, and 2.5 kV after two washes of an overnight culture with a 15% glycerol, 1 mM MOPS buffer. Stocks of all strains were conserved at –80° C in 20% glycerol.

The plasmid pIND4-*gafA* was constructed by Gibson assembly with the NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs) (Gibson *et al.*, 2009). Briefly, sequences of

plasmid pIND4 and *gafA* were amplified separately with overlapping regions by PCR with the Q5® High-Fidelity DNA Polymerase (New England Biolabs). After purification of PCR products with NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel), they were assembled by Gibson assembly and the obtained plasmids were transformed into chemocompetent *E. coli*. The transformants were selected on kanamycin. The constructs were confirmed by Sanger sequencing (Microsynth) before being transformed into *P. versutus* strains. The plasmid pIND4 was also transformed in *P. versutus* strains as a control.

The transcriptional reporter plasmid pSEVA631-Pgta-sfGFP to monitor the main GTA operon expression in *P. versutus* was constructed by Céline Terrettaz in multiple steps. She first amplified a tetracycline resistance cassette by PCR and cloned it in pSEVA631 at BglII position by digesting first the plasmid and the insert, and then ligating them together before transformation into *E. coli*. To construct transcriptional fusions, the gene coding for sfGFP and the 326 bps upstream of the GTA operon were amplified by PCR individually before being united by PCR and cloned between the EcoRI and HindIII sites of the pSEVA631-tetAR. Finally, due to insufficient fluorescence, a Shine-Dalgarno sequence was added before the sfGFP gene by PCR.

P. versutus deletion mutants were constructed by forced recombination between homologous DNA sequences according to the method of Martínez-García and de Lorenzo (Martínez-García and de Lorenzo, 2011). Briefly, the 400-600 bps flanking the gene of interest (GTA operon and LOCUS_17550) upstream and downstream were amplified by PCR separately and purified directly from the PCR mix or from an agarose gel (Gel and PCR clean-up kit, Macherey-Nagel). Then, both fragments were fused together thanks to a primer-encoded 20–30 bps overlap during a second PCR amplification round. The resulting fragment was purified, digested with EcoRI and BamHI restriction enzymes, and inserted into the pEMG suicide vector. The integrity of the inserted DNA fragment was verified by Sanger sequencing (Microsynth). Then, *P. versutus* was electroporated with the plasmid pEMG-ΔGTA and pEMG-ΔRecep. The plasmid inserts itself in the host genome by homologous recombination of the flanking regions insert. To excise the pEMG from the host genome, pME6000 modified plasmid encoding I-SceI restriction enzyme was electroporated in a *P. versutus* clone obtained from the first electroporation. The loss of the pEMG was first evaluated by streaking Tet-resistant colonies from the second electroporation on LB-Km plates. The susceptibility to Km indicates the

removal of pEMG from the host genome. However, in the case of the removal of the whole GTA operon, we obtained only revertant to the wild-type type. Therefore, a chloramphenicol resistance cassette was inserted in the pEMG-ΔGTA plasmid between the upstream and downstream regions by Gibson assembly with the NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs) (Gibson *et al.*, 2009) to build the pEMG-ΔGTA-CmR plasmid. It allowed the selection of the deletion mutants only. Then, to have a clean mutant, the chloramphenicol resistance cassette was removed with the pEMG-ΔGTA_{bis} plasmid. The deletions were finally confirmed by PCR.

6.4.3 GTA transfer experiments

Paracoccus spp. donor and recipient cultures were grown in LB overnight. Cells were cleared from donor cultures by centrifugation (3min at 13'000 rpm) and the supernatant was filtered through a 0.22 μm syringe filter. Recipient cell density was quantified by spectrophotometry with Thermo Scientific™ GENESYS™ 10S UV-Vis Spectrophotometer (Thermo Fisher Scientific, United-states) and, based on the results, cells were diluted or concentrated to an optical density (OD) of 0.2 at 600nm. The cells were washed in NaCl 0,9% and resuspended in same volume of G-Buffer (10 mM Tris-HCl (pH 7.8), 1 mM MgCl₂, 1 mM CaCl₂, 1 mM NaCl) supplemented with 500 μg/mL of bovine serum albumin. Reactions were carried out in 1.5 mL Eppendorf by mixing 500 μL of recipient cell solution with 500 μL of GTA solution and briefly vortexing. After 24h of incubation at 30°C with 180 rpm shaking, 20 μL from the mix were diluted in NaCl 0,9% and plated on a LB agar plate to determine the total number of cells. The remaining 980μL (only 100μL for *Paracoccus versutus* DSM582 in case of expected transfer) were inoculated on LB agar plates supplemented with 50 μg/mL of rifampicin to determine the number of rifampicin-resistant cells. The transformation rate, also named transformability or transfer frequency, was calculated by dividing the number of resistant in CFU/mL by the total number of CFU/mL. Each experiment consisted of two treatments (GTAs from rifampicin susceptible and resistant strains), a statistically significant difference in transformation rates between the two treatments was considered as gene transfers.

For the residual activity assay, a step was added before the actual gene transfer experiment. GTA solutions were mixed either with bacterial suspensions or G-buffer ("no cell" control) as for gene transfer experiment and incubated for 24 hours at 30°C with 180 rpm shaking. After

24h of incubation, the solutions were centrifuged and filtered to remove any cell. The flowthrough was used as a GTA solution for a regular transfer assay.

6.4.4 DNA encapsulated in GTAs: extraction, sequencing, and analysis

To extract and sequence the DNA fragments encapsulated in GTAs, the GTA particles produced by *P. versutus* DSM582 were first purified and concentrated. Hence, a bacterial culture of 40 mL was grown at 30°C for 72 hours in LB. The culture was then centrifuged twice for 10 min at 8000G at 4°C. The supernatant was filtered with a 0.22 µm syringe filter. The flowthrough was concentrated with Amicon® Pro Purification System with 50kDa Amicon® Ultra-0.5 Device. Approximately 30 mL was concentrated before plugging the filter. The concentrate was resuspended in 2 mL of G-buffer by thoroughly washing the filter membrane.

Following GTA particle concentration, encapsulated DNA was extracted using a phenol:chloroform:isoamyl alcohol (25:24:1) method. Equal volumes of the extraction mixture and sample were combined, vortexed for 20 seconds, and centrifuged at 8000 G for 15 min at room temperature. The DNA-containing aqueous phase was transferred to a new tube and mixed with 0.1 volume of 3M sodium acetate. DNA precipitation was achieved by adding 2.5 volumes of 100% ethanol, inverting the tube, and incubating on ice for 15 min. The precipitated DNA was pelleted by centrifugation for 3 min, washed twice with 1 mL 70% ethanol, and air-dried under a laminar flow hood for at least 20 min. The DNA pellet was resuspended in 100 µL of Milli-Q® Type 1 Ultrapure Water and quantified using a Thermo Scientific™ NanoDrop 2000 spectrophotometer. DNA fragment size and integrity were assessed by gel electrophoresis, revealing two distinct bands corresponding to genomic DNA and GTA content (data not shown). The 4 kb band, representing DNA encapsulated in GTAs, was excised and purified using the Macherey-Nagel™ NucleoSpin™ Gel and PCR Clean-up Kit. The purified DNA was subsequently submitted to Novogene (UK) for Illumina sequencing.

Reads were trimmed and filtered with fastP (Chen *et al.*, 2018), mapped with bowtie2 (Langmead and Salzberg, 2012), and counted with featureCounts (Liao *et al.*, 2014). Genome coverage of encapsulated DNA in GTAs was determined by the number of reads at a specific position of the genome. The mean coverage by PvGTA-DNA was calculated for sliding windows of 10kb with the *rollapply* function of the *zoo* package (Zeileis and Grothendieck, 2005). In addition, the genome was examined according to the nucleobase composition with the *Biostrings* package (Pagès *et al.*, 2022) using a sliding window of 10kb with 1kb steps. GC

content was calculated as the sum of G and C divided by the sum of all nucleotides. GC skew was calculated as follows: $(C - G)/(G + C)$ (Arakawa and Tomita, 2007). The cumulative GC skew was calculated in the same way, except that it takes advantage of the sum of neighboring windows from an arbitrary start. The origin and terminus of replication were predicted using the cumulative GC skew results whereby the minimum value of the cumulative skew corresponded to the terminus, and the maximum value corresponded to the origin of the replication (Grigoriev, 1998).

6.4.5 Flow-cytometry and reporter assays

Determination of *P. versutus* GTA transcriptional activity by flow cytometry was conducted as follows: overnight cultures of *P. versutus* strains were grown in LB medium supplemented with appropriate antibiotics at 30°C. The following day, cultures were washed twice with 0.9% NaCl and diluted to a final OD₆₀₀ of 0.001 in fresh LB containing appropriate antibiotics and IPTG to induce expression of the pIND4 plasmid cargo. The diluted cultures were distributed into 96-well plates and incubated in a plate reader (Cytation5, BioTek) for regular optical density measurements. At predetermined time points, samples were collected and analyzed using a flow cytometer (BD Accuri C6 Plus, BD Biosciences), with 50'000 events typically recorded per sample.

Data analysis was performed using RStudio software. To determine the proportion of highly fluorescent cells in each sample, we employed a dynamic thresholding method based on the relationship between Forward Scatter Area (FSC-A) and fluorescence intensity (FL1-A). First, the FSC-A and FL1-A values were log-transformed to stabilize variance and normalize the data distribution. A linear regression model was then fitted to the log-transformed data to capture the relationship between FSC-A and FL1-A. The residuals from the linear regression model, representing the difference between observed and predicted fluorescence values, were calculated. A dynamic threshold was defined as the regression-predicted fluorescence plus two times the standard deviation of the residuals. This threshold accounts for variations in the data and identifies cells with significantly higher fluorescence. Finally, the proportion of fluorescent cells was determined by calculating the fraction of cells with fluorescence value exceeding the dynamic threshold. This method allows for a robust and adaptive determination of fluorescent cell populations, accounting for variations in sample distributions and the inherent relationship between FSC-A and fluorescence.

To study the effect of quorum sensing molecules, a modified assay was performed. An exponential phase culture of the reporter strain was diluted to an OD₆₀₀ of 0.1 in fresh LB with antibiotics and distributed into multiple tubes. Various concentrations of hexadecanoyl-L-homoserine lactone (C16-HSL) diluted in DMSO were added to the cultures. The final DMSO concentration was 5 % (v/v), also in the control without C16-HSL. The cultures were then distributed in 96-well plates and incubated in a plate reader (Cytation5, BioTek) for regular optical density measurements. After 15min incubation at 30°C, samples were collected and analyzed using a flow cytometer (BD Accuri C6 Plus, BD Biosciences), with 50'000 events typically recorded per sample.

6.4.6 Phylogenetic analysis

GyrA and GyrB sequences were retrieved from the National Center for Biotechnology Information (NCBI) GenBank database. A CLUSTAL Multiple sequence alignment was performed using the MUSCLE alignment software (Edgar, 2004). Maximum Likelihood phylogenies were generated with the best model for each partition in the IQ-TREE web server (Nguyen *et al.*, 2015), with 10'000 ultrafast bootstraps (UFBoot) trees to assess node support values. The phylogenies were visualized and annotated using RStudio (RStudio Team, 2020) with the help of package *ggtree* (Yu *et al.*, 2017).

6.4.7 Sequencing and annotation of *P. versutus* DSM582 genome

Genomic DNA from an exponential culture was extracted using the Qiagen Genomic-tip 20/G kit and sequenced on a PacBio Sequel instrument. Hifi reads were generated using PacBio CCS 4.1.0 and the genome was assembled and circularized using Flye 2.6 (Kolmogorov *et al.*, 2020). Genome annotation was performed using DFAST, with additional functional classification of genomic content carried out using the eggNOG (Tanizawa *et al.*, 2018; Hernández-Plaza *et al.*, 2023).

The investigation of the proportion of putative essential genes was done as follows. We first made a list of strict homologues between *P. versutus* and *R. sphaeroides* using reciprocal blast and retrained only hits with $\geq 70\%$ percent identity and $\geq 50\%$ coverage. Following this, the number of essential genes among these homologues, as determined by Tn-seq (Burger *et al.*, 2017), was counted. The percentage of putative essential homologues was calculated by dividing the number of identified essential homologues by the total number of homologues for each replicon in *P. versutus*.

6.4.8 Visual representation and statistical analysis

The statistical analyses and graphics were performed with *RStudio* software (RStudio Team, 2020). Graphics were designed with *ggplot2* and base R functions. The circular plots were created with the package *circlize* (Gu *et al.*, 2014). Regarding statistical analyses, the normality of the data distribution was tested with Shapiro-Wilk normality test, the equality of variances with F-test (when two groups), or Levene's test (more than two groups). Finally, means were compared either with Student's t-test/Wilcoxon rank test or ANOVA with post-hoc Tukey or Dunnett tests. The results of the statistical tests are indicated in the figures accordingly: *p-value <0.05; **p-value <0.01; ***p-value <0.001.

6.5 Results

6.5.1 *P. versutus* DSM582 is the only among a set of *Paracoccus* strains to be able to perform horizontal gene transfer mediated by GTA.

To identify potential GTA-mediated transfer systems in the *Paracoccus* genus, we performed homology searches using BLAST with the RcGTA operon as the query sequence against publicly available *Paracoccus* genomes (Altschul *et al.*, 1990). All analyzed *Paracoccus* genomes contained an operon homologous to the main RcGTA cluster, with $\geq 90\%$ of query coverage and $\geq 67\%$ identity (Figure 6.1). Notably, *Paracoccus* species possess an additional GTA-related gene at the operon 3' end (highlighted in red in Figure 6.1). This gene shows homology to *rcc00171* of *R. capsulatus*, which encodes tail fibers and is located in a distinct locus separate from the main GTA operon (Hynes *et al.*, 2016).

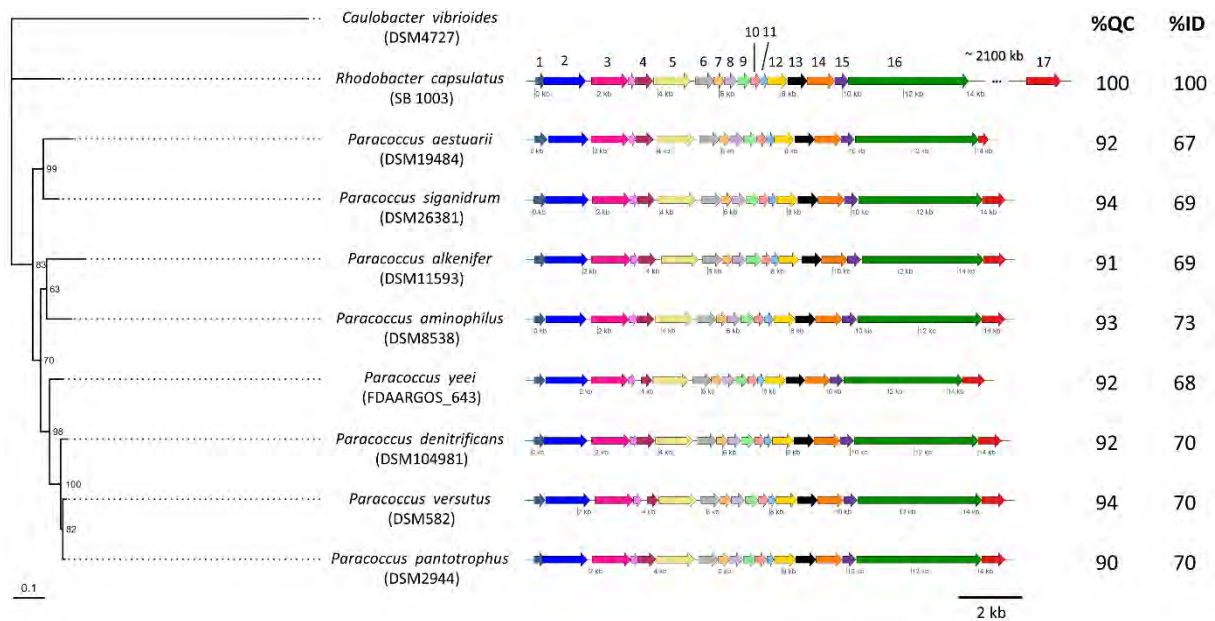


Figure 6.1. The whole cluster of the GTA operon is present in many *Paracoccus* species. On the left side, the result of a non-exhaustive *Paracoccus* tree is based on the multiple sequence alignment of the GyrA and GyrB and maximum likelihood-based construction. Bootstrap values are indicated on each branch and the distance scale is at the bottom of the figure. The tree was rooted with *Caulobacter vibrioides*. In the middle, the main GTA operon from *R. capsulatus* and several species of *Paracoccus* are represented with labels according to Lang and colleagues (2017): 1. Small terminase; 2. Large terminase; 3. Portal; 4. Prohead Protease; 5. Major Capsid protein; 6. Adaptator; 7. Stopper; 8. Tail terminator; 9. Tail tube; 10. Tail chaperone 1; 11. Tail chaperone 2; 12. Tape measure; 13. Distal tail; 14. Hub; 15. Peptidase; 16. Megatron; 17. Tail fiber. The arrows are proportional to gene size. On the right side, the results overview (%QC: percentage of query cover; %ID: percentage identity) of the BLASTn of the whole main GTA operon in *Paracoccus* species based on the RcGTA operon.

Given the presence of the main GTA operon in *Paracoccus* spp., we investigated their capacity for GTA-mediated HGT. Transfer experiments were conducted by mixing recipient cells with cell-free supernatants potentially containing GTA particles. Acquisition of rifampicin resistance, resulting from mutations in *rpoB*, served as a genetic marker (Lang *et al.*, 2012). Supernatants from rifampicin-susceptible strains served as controls for the occurrence of spontaneous resistance. While spontaneous rifampicin resistance rates varied among strains and species, *P. versutus* DSM582 exhibited a significantly higher frequency of rifampicin resistance when exposed to supernatant from a rifampicin-resistant strain compared to a susceptible strain (Figure 6.2A), indicating that HGT occurred in this strain.

To confirm that genetic transfer in *P. versutus* DSM582 was dependent on GTAs, we deleted the operon responsible for GTA production and repeated the experiment. The Δ GTA strain could receive but not donate the marker (Figure 6.2B), confirming that deletion of the main GTA operon impaired GTA production. Transfer experiments with newly isolated environmental *P. versutus* strains showed no detectable genetic marker transfer (Figure 6.2C), suggesting that HGT may be limited to specific strains under the tested conditions.

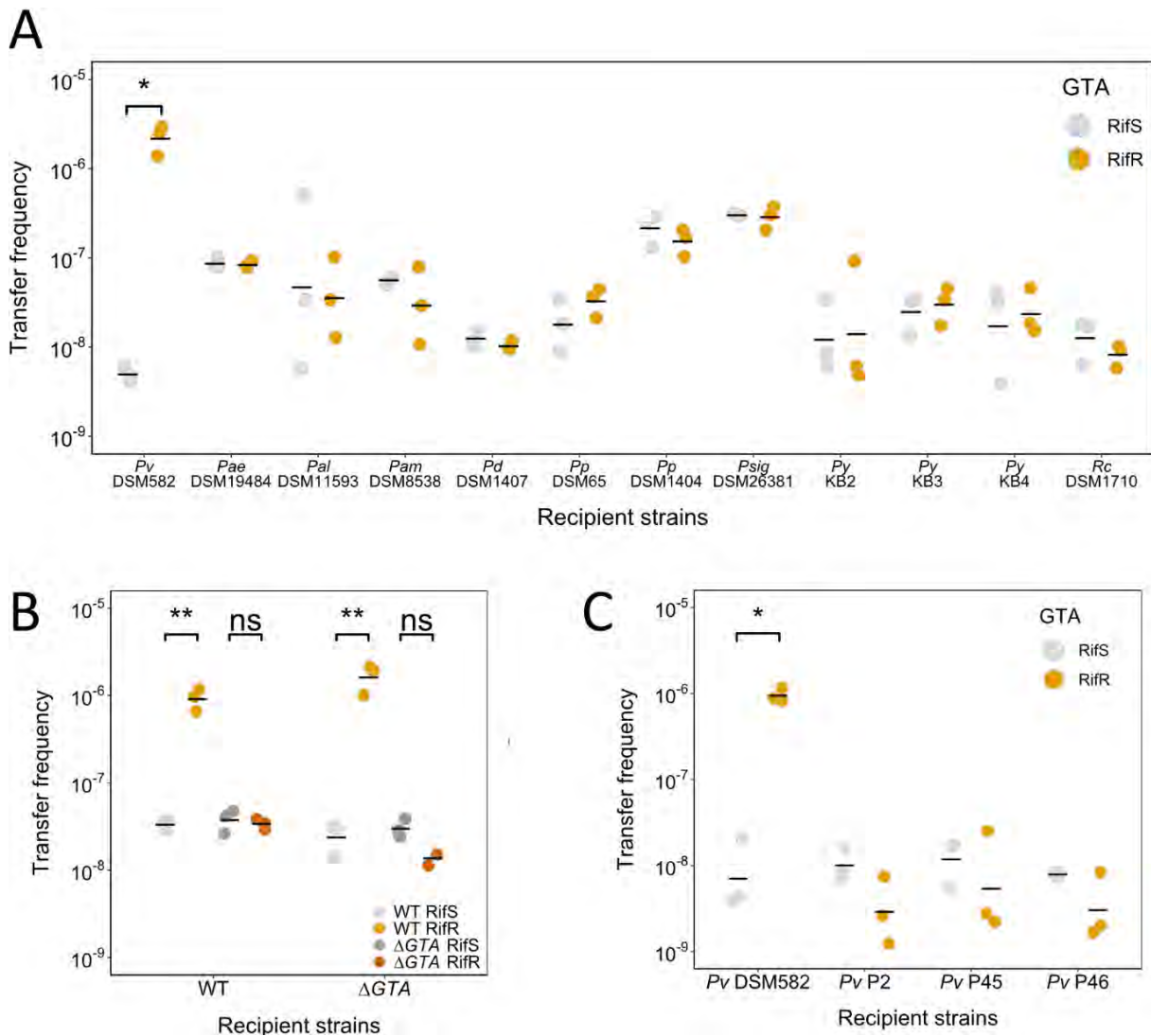


Figure 6.2. *Paracoccus versutus* DSM582 produces GTAs and can acquire a genetic marker via GTAs. (A) Among various species of *Paracoccus*, horizontal gene transfer was only observed for *Paracoccus versutus* DSM582. We monitored the resistance to rifampicin in the population exposed to the supernatant from a rifampicin-resistant strain (RifR, orange) compared to one from a rifampicin-susceptible strain (RifS, grey). Abbreviation for the strains' name: *Pv* for *P. versutus*; *Pae* for *P. aestuari*; *Pal* for *P. alkenifer*; *Pam* for *P. aminophilus*; *Pd* for *P. denitrificans*; *Pp* for *P. pantotrophus*; *Psig*, *P. siganidrum*; *Py* for *P. yeei*; *Rc* for *R. capsulatus*. (B) Deletion of the main GTA operon led to the inability of *P. versutus* DSM582 to produce GTAs but not to receive it. (C) Among various strains of *Paracoccus versutus*, HGT was only observed in strain DSM582. Dots show independent replicates (three) and bars the means. The experiments were performed at least two times with similar results.

Additionally, we conducted cross-transfer experiments using *P. versutus* DSM582 as the GTA donor and other strains as recipients, to investigate their capacity to receive GTAs. No significant transfer was detected although a higher frequency of rifampicin resistance was observed in *P. versutus* P46 (Figure 6.3). Sequence divergence is unlikely to prevent recombination between *P. versutus* strains given the conservation of *rpoB* sequence (Table S 6.4) (Lang *et al.*, 2012). Among all tested *Paracoccus* strains, GTA-mediated HGT was only clearly observed in *P. versutus* DSM582.

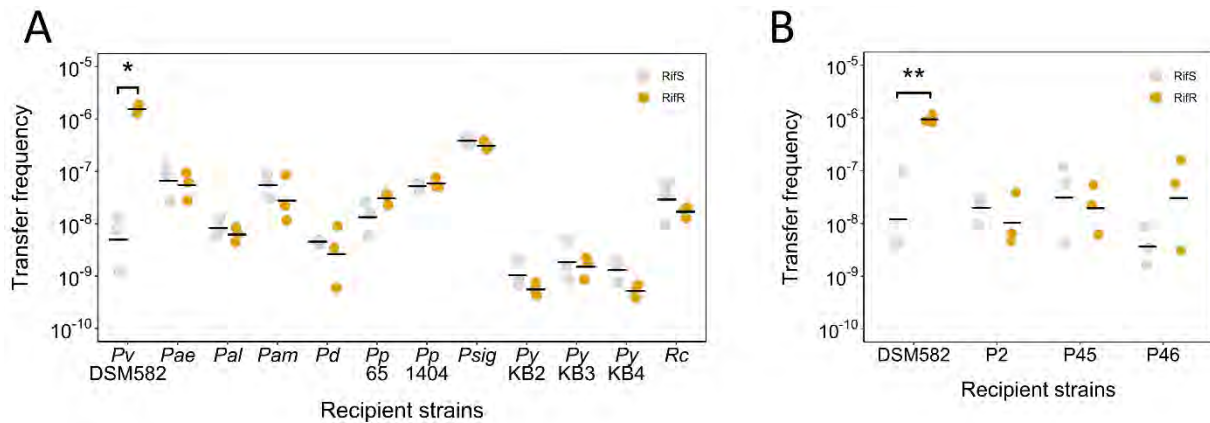


Figure 6.3 Although *P. versutus* DSM582 produces GTAs, none of the other *Paracoccus* spp. strains (A) or other *P. versutus* strains (B) seem to be able to receive the resistance marker efficiently from *P. versutus* DSM582. We monitored the resistance to rifampicin in a recipient population exposed to the supernatant from the rifampicin-resistant *P. versutus* DSM582 (RifR, orange) compared to one from the rifampicin-susceptible *P. versutus* DSM582 (RifS, grey). Abbreviation for strains' name: Pv for *P. versutus* with respective strains names; Pae for *P. aestuari* DSM19484; Pal for *P. alkenifer* DSM11593; Pam for *P. aminophilus* DSM8538; Pd for *P. denitrificans* DSM1407; Pp for *P. pantotrophus* with respective strains names; Psig, *P. siganidrum* DSM26381; Py for *P. yeei* with respective strains names; Rc for *R. capsulatus* DSM1710. Dots show independent replicates (three) and bars the means. The experiments were performed at least two times with similar results.

6.5.2 The genome of *P. versutus* DSM582 contains four replicons.

To gain a better understanding of the genetics behind GTA-mediated HGT in *P. versutus*, we sequenced the genome of *P. versutus* DSM582. Long-read sequencing facilitated the assembly of a complete genome, revealing a multipartite structure comprising four distinct circular replicons (Table 6.1, Figure 6.4). Multipartite genomes are not uncommon in the *Paracoccus* genus (Table S 6.5, Figure S 6.1) (Winterstein and Ludwig, 1998; Kelly *et al.*, 2006; Dziewit *et al.*, 2014; Lasek *et al.*, 2018).

Table 6.1 General features of *P. versutus* DSM582 genome.

General features of genome	Chromosome	Chromid	Megaplasmid	Plasmid
Size (bp)	2'799'769	1'904'959	889'277	131'523
GC content (%)	67.7	67.4	67.8	62.5
Coding density (%)	89.4	89.4	90.1	78.8
Number of ORFs	2765	1876	832	149
Number of rRNA genes	6	3	0	0
Number of tRNA genes	40	18	4	0
% Putative essential ORFs ^a	35.1 (174/496)	29.0 (88/303)	6.5 (2/31)	0 (0/4)
Position of the Origin of replication ^b	2'189'000	144'000	398'000	-
Position of Terminus ^b	634'000	1'179'000	821'000	-

^a Initially, strict homologues between *P. versutus* and *Rhodobacter sphaeroides* were identified through reciprocal BLAST analysis, retaining only matches with at least 70% sequence identity and 50% coverage. Following this, the number of essential genes among these homologues, as determined by Tn-seq (Burger *et al.*, 2017), was counted. The percentage of putative essential homologues was calculated by dividing the number of identified essential homologues by the total number of homologues for each replicon in *P. versutus*.

^b Origins of replication and termini were identified using GC skew analysis. For the chromosome, chromid, and megaplasmid, a 10'000 bp sliding window with 1000 bp increments was employed. Due to its smaller size, the plasmid was not analyzed to avoid unreliable results (Hubert, 2022).

The two largest replicons in *P. versutus* DSM582 likely represent the main chromosome and a chromid. As expected, the main chromosome is the largest replicon and contains the majority of predicted essential genes (Table 6.1). The second-largest replicon also harbors homologues of many putative essential genes. Its classification was unclear. Several authors consider the second largest replicons carrying essential genes as chromids (Harrison *et al.*, 2010; Mergeay and Houdt, 2018). Based on the genomes they investigated, they noticed that those replicons gather the following criteria: (i) chromids have plasmid-type maintenance and replication systems, (ii) chromids have a nucleotide composition close to that of the chromosome, and (iii) chromids carry core genes that are found on the chromosome in other species. Other authors such as DiCenzo and Finan present two categories for the second largest replicon carrying essential core genes: chromid and secondary chromosomes (DiCenzo and Finan, 2017). The difference between a secondary chromosome and a chromid is that the secondary chromosome results from a split of the chromosome into two (DiCenzo and Finan, 2017). Hence, second chromosomes are expected to exhibit high synteny with the chromosomes of related species, with the degree of similarity dependent on the time since their evolutionary divergence (DiCenzo and Finan, 2017).

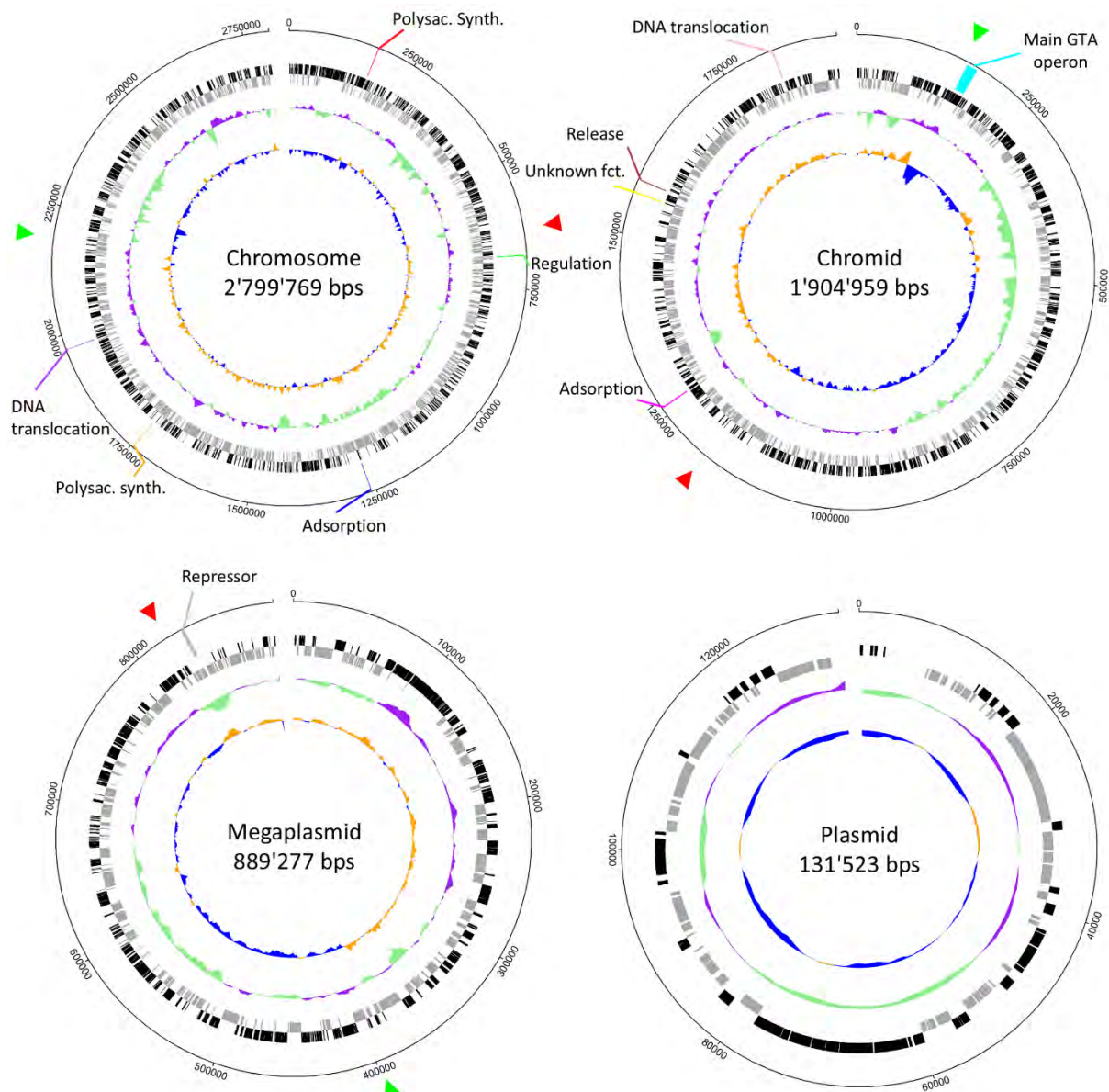


Figure 6.4 Schematic overview of the genome of *P. versutus* DSM582. The concentric circles represent, from inside out: (i) the GC skew (positive, orange; negative, blue), (ii) the GC content (below average green, above, violet), (iii) genes encoded on the reverse and forward strand (gray and black), (iv) homologues of *R. capsulatus* genes involved in genetic transfer mediated by GTAs grouped by category (Table S 6.6), (v) genome coordinates using the coding sequence of *dnaA* as origin for the chromid, (vi) predicted position of origin of replication (green triangle) and terminus (red triangle) based on cumulative GC skew.

To distinguish between a chromid and a secondary chromosome, we first attempted to evaluate genomic synteny with closely related species, preferably those with a single chromosome. All available genomes of species closely related to *P. versutus*, such as *P. pantotrophus* and *P. denitrificans*, also possess two replicons larger than 1 Mb (Table S 6.5, Figure S 6.1). The genomes of the closest strains with only one large replicon, *P. aminovorans* JCM7685 and *Paracoccus* sp. SMMA_5, display some synteny with both largest replicons of *P. versutus* DSM582, but lots of genome rearrangements occurred as well (Figure S 6.2).

Secondly, we investigated the replication systems. Interestingly, the second largest replicon contains replication/partitioning modules composed of a replicase and the partitioning operon (*parAB*), as such associated with plasmids in Rhodobacterales (Figure S 6.3, Table S 6.7) (Petersen, 2011; Petersen *et al.*, 2011). Therefore, we considered the second-largest replicon as a chromid.

Comparative genomic analyses of various *P. versutus* strains revealed that the two smallest replicons are absent in some strains, indicating their dispensable nature (Table S 6.5, Figure S 6.4, Figure S 6.5). These replicons carry replication and partitioning modules typical of plasmids (Table S 6.7, Figure S 6.3). Furthermore, the scarcity of homologues essential in *R. sphaeroides* supports their classification as plasmids (Table 6.1). Based on these characteristics, both replicons were classified as plasmids. One of these replicons, with a size of almost 900 kb, may be further categorized as a megaplasmid (DiCenzo and Finan, 2017), while the other remains classified as a standard plasmid. In conclusion, the genome of *P. versutus* DSM582 showcases a multipartite structure, with a main chromosome, a chromid, a megaplasmid, and a plasmid.

6.5.3 Packaging of DNA into GTAs is biased towards two regions close to the origin of replication.

Gene transfer agents and bacteriophages present a key difference : while bacteriophages package almost exclusively their own genetic material, GTAs encapsulate random segments of the host organism's genome (Hynes *et al.*, 2012). To analyze whether the packaging of DNA into *P. versutus* GTA particles was random or showed preferences for distinct genomic loci as highlighted in several species (Berglund *et al.*, 2009; Québatte *et al.*, 2017; Tomasch *et al.*, 2018; Gozzi *et al.*, 2022), we extracted the DNA from isolated GTA particles and sent it for high-throughput sequencing. Mapping of reads onto the *P. versutus* genome allowed us to identify regions encapsulated in GTAs, but also to determine biases in encapsulated fragments' frequency based on the relative read counts at specific positions along the chromosome.

With a total of 4.6 billion reads from the sequencing of DNA encapsulated in GTA particles, the genome was well covered: all positions of the three largest replicons as well as about 98% for the small plasmid had matching reads (Table 6.2, Figure 6.5). In addition, the sequencing depth of the four replicons was similar, with means ranging from 698 to 855 reads per

position (Table 6.2). Interestingly, plasmid fragments were also encapsulated in GTAs, indicating the possibility for both chromosomal and plasmid DNA transfer via GTAs in *P. versutus*. This contrasts with observations in *Bartonella henselae* where exchange via BaGTA was restricted to chromosomal DNA (Québatte *et al.*, 2017), and in *D. shibae* for which the low coverage of the plasmids led the authors to conclude that plasmids are virtually excluded from encapsulation into GTA particles (Tomasch *et al.*, 2018).

Table 6.2 General features of DNA sequencing of GTA content.

General features	Chromosome	Chromid	Megaplasmid	Plasmid
Total reads	2'180'473'677	1'628'751'498	725'938'080	91'758'659
Mean reads per position	779	855	816	698
Median reads per position	780	808	812	792
Coverage (%)	100	100	100	97.95
Position with highest reads coverage	2'500'877	147'078	393'438	34'935

Although the reads were distributed evenly, one region was much more frequently packaged than the rest of the genome in GTAs (Figure 6.5, Figure S 6.6). This region is located on the chromid, approximately 50 kb to 300 kb from the chromosomal replication initiator *dnaA*, with a peak at position 147'078. It exhibited a coverage of 2'640 reads per position, considerably higher than the overall median frequency (808 reads per position). Additionally, a second region on the megaplasmid, spanning from 350 kb to 450 kb, showed slight enrichment, peaking at position 393'438. This peak had a frequency of 1'571 reads per position, roughly double the median coverage.

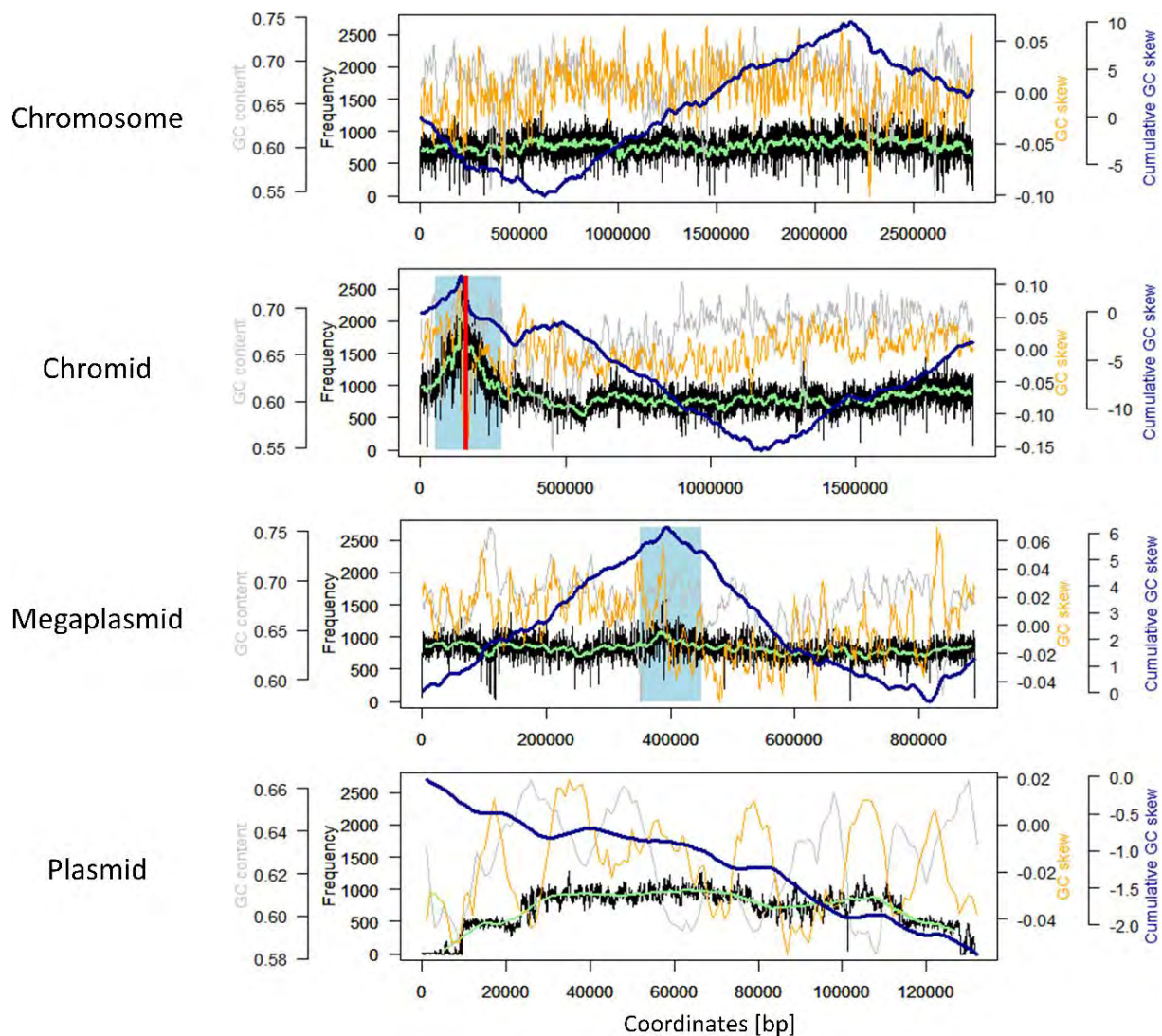


Figure 6.5 Two regions (blue) located close to the origins of replication were more frequently found in GTAs than the rest of the genome. Each replicon is represented with DNA coverage (black), mean coverage (green), GC content (grey), GC skew (orange), and cumulative GC skew (dark blue). The maximum and the minimum of the cumulative GC skew line shows the predicted position of the origin of replication and of the terminus of replication respectively. The red area on the chromid corresponds to the position of the main GTA operon.

Biases in genome coverage of GTA-packaged DNA have been observed in several GTA-producing bacteria. Multiple hypotheses have been proposed to explain these biases. These include the spatial structure of the chromosome influencing DNA accessibility, specific DNA motifs being preferentially recognized by the packaging machinery, and differential copy numbers with some genomic regions present multiple times in the cell due to run-off replication (Berglund *et al.*, 2009; Québatte *et al.*, 2017; Tomasch *et al.*, 2018; Gozzi *et al.*, 2022). The copy-number hypothesis was formulated following the observation that DNA fragments from a high plasticity zone were more frequently encapsulated in BaGTA than the rest of the genome (Berglund *et al.*, 2009). This high plasticity zone contains several short clusters of phage genes, with the peak of encapsulated fragments corresponding to a phage

homologous helicase. This observation supports the hypothesis of run-off replication due to a phage-derived extra origin of replication (Berglund *et al.*, 2009).

To investigate the potential causes of packaging bias in *P. versutus* GTAs, we first examined whether the presence of predicted prophages might explain the observed pattern, as seen in BaGTA (Berglund *et al.*, 2009; Québatte *et al.*, 2017). Using PHASTEST, we identified prophages on the chromosome and the chromid (Table S 6.8) (Wishart *et al.*, 2023). Interestingly, the main GTA operon is situated near the enrichment region on the chromid (Figure 6.5). However, no prophage was predicted on the megaplasmid, which does not account for the observed enrichment on this replicon. We then calculated the GC skew for all replicons of *P. versutus* to determine the position of the replication origin and terminus because during exponential growth, bacteria can initiate new rounds of replication before the previous ones are completed, resulting in a higher copy number of origin-proximal regions (Trojanowski *et al.*, 2018). Interestingly, both enriched regions were found to be located close to the predicted origin of replication of their respective replicons (Figure 6.5, Figure S 6.6), suggesting a potential link between replication and GTA content. However, it is puzzling that no encapsulation bias was observed for the main chromosome close to its predicted origin of replication.

We also searched for shared motifs between both enriched regions that might explain the observed packaging bias, but our results were inconclusive. Additionally, data regarding the spatial organization of the replicons are currently not available for *P. versutus*, limiting our ability to assess the impact of chromosomal structure on DNA accessibility for GTA packaging.

6.5.4 GTA operon is induced in a subpopulation and its expression is highest during exponential phase.

GTA particle release requires cell lysis (Hynes *et al.*, 2012), a costly behavior limited to a small fraction of the population in known models. Previous studies estimated maximal GTA production at approximately 3% in *R. capsulatus* and 14% in *B. henselae* (Hynes *et al.*, 2012; Québatte *et al.*, 2017). To investigate GTA expression in *P. versutus*, we constructed a reporter plasmid containing the putative promoter region of the GTA main operon (approximately 400 bp upstream of the small terminase gene, *LOCUS_29200*) fused to a promoterless *sfGFP* reporter gene. Single-cell GTA gene expression was monitored using flow cytometry, revealing that only a small subset of the population exhibited high fluorescence (Figure 6.6A). To

quantify this fluorescent subpopulation, we employed a linear regression model based on cell size and fluorescence intensity. A dynamic threshold was established using the regression-predicted fluorescence value plus two standard deviations as a limit (Figure 6.6B). Cells exceeding this threshold were considered to be actively expressing the GTA main operon.

To determine the proportion of the population expressing the main GTA operon and the timing of maximal expression, we conducted growth curves and collected samples at each growth phase for analysis using a flow cytometer. Despite a slightly delayed growth pattern in the reporter strain compared to the controls, samples were collected at comparable growth phases (Figure 6.6C). The maximum proportion of cells expressing the construct was approximately 3% and was observed during exponential phase (Figure 6.6D). Notably, this proportion decreased progressively along the growth curve. Nevertheless, almost 1% of cells still expressed the construct after approximately 5 hours in stationary phase. Researchers conducted a comparable study using *R. capsulatus*, with a translational reporter for the GTA particle's portal protein (*rcc01684*). They observed a highly fluorescent subgroup only when the lysis gene was mutated, indicating that cells producing GTAs typically undergo rapid lysis (Hynes *et al.*, 2012). It suggests that the *P. versutus* fluorescent cells observed at one moment may not be the same as those seen at another time point. If cells expressing the PvGTA main operon indeed lyse quickly, it implies that different subsets of the population are expressing the main GTA operon throughout various stages of growth.

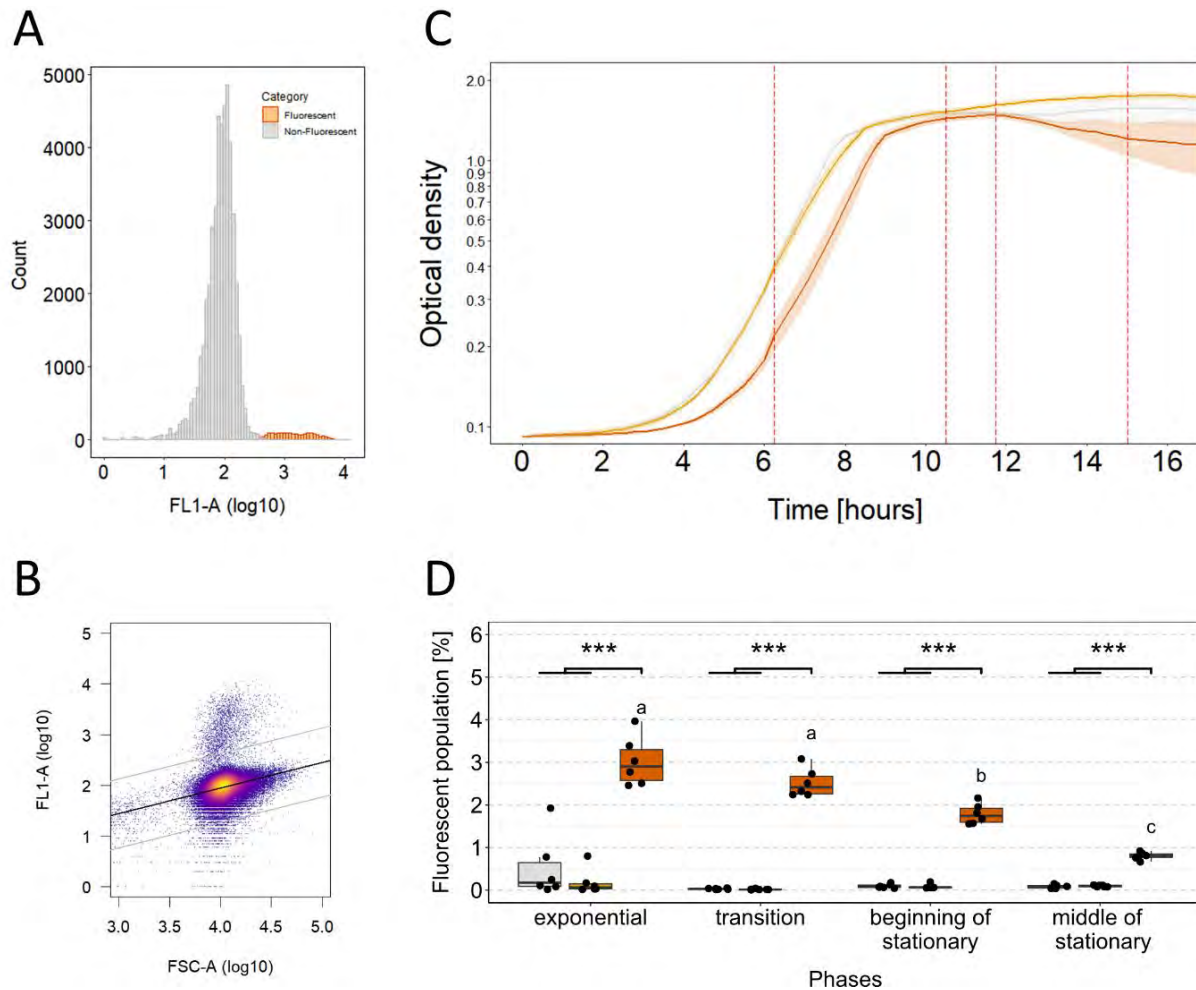


Figure 6.6. The expression of the main GTA operon in *P. versutus* DSM582 peaks during the exponential growth phase and diminishes as the population transitions to the stationary phase. (A) Histogram of green fluorescence intensity (FL1-A) showing fluorescent (orange) and non-fluorescent (grey) populations. (B) Density plot of forward scatter (FSC-A) vs. green fluorescence (FL1-A). Each point represents a single cell. Color intensity indicates event density. The central line shows the linear regression model, and grey lines represent the dynamic threshold for fluorescence. Events exceeding the upper threshold were considered fluorescent. (C) Growth curves for the wild-type strain (grey), and strains with pSEVA631-*tetR* (light orange, control) and pSEVA631-Pgta-*sfGFP-tetR* (dark orange). Lines show means, shaded areas indicate standard deviation of 6 replicates. Red lines mark sampling times. (D) GTA operon expression across growth phases. Boxplots show median (central line), first and third quartiles (box limits), and data points within 1.5 times the interquartile range (whiskers). Individual replicates (n=6) are shown as dots. Different letters indicate significant differences between phases ($p < 0.05$).

The inverse correlation between GTA expression levels and population density observed in Figure 6.6 suggested that GTA production might be influenced by quorum sensing. In *R. capsulatus*, the GtaI/R quorum sensing system is known to regulate RcGTA production and reception, with GtaI producing two long-chain acyl-homoserine lactones (HSLs), C14- and C16-acyl-HSL, which trigger RcGTA production (Leung *et al.*, 2012; Brimacombe *et al.*, 2013). The C16-acyl-HSL molecule is the most abundant acyl-HSL produced not only by *R. capsulatus* but also by various species of *Paracoccus* (Schaefer *et al.*, 2002; Morinaga *et al.*, 2020).

Therefore, we hypothesized that quorum sensing might also affect GTA production in *P. versutus*.

To investigate this hypothesis, we examined GTA operon expression at the single-cell level using the transcriptional reporter strain exposed to various concentrations of C16-HSL. *P. denitrificans* has been shown to produce C16-HSL at a concentration of up to 14 μM with a maximal production in the late exponential phase (Morinaga *et al.*, 2020). Our experiments revealed that the addition of C16-HSL rapidly stimulated the reporter expression, with increasing concentrations of quorum sensing molecules promoting transcription until a plateau was reached at approximately 17 μM (Figure 6.7). Further experiments with knock-out mutants of *gtal* and *gtar* homologues are required to confirm the stimulating role of HSL on GTA expression in *P. versutus*.

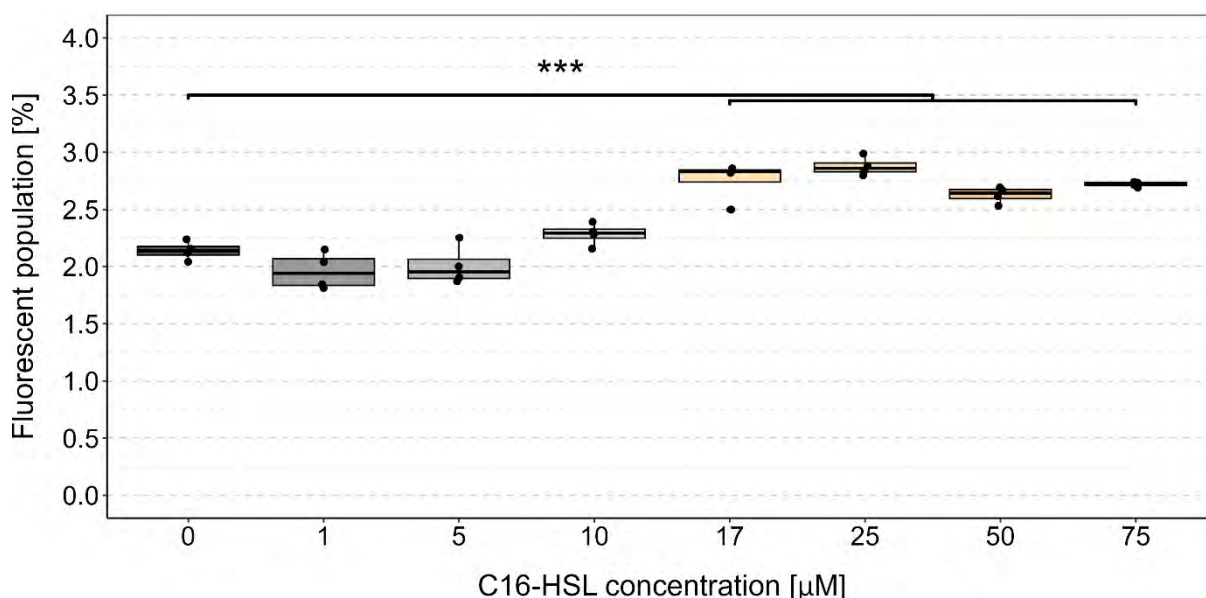


Figure 6.7. The expression of the GTA operon in *P. versutus* DSM582 is affected by the addition of C16-HSL. Percentages of cells expressing the GTA operon at high levels 15min after the addition of HSL. The gene expression was quantified based on fluorescence intensities using flow cytometry.

6.5.5 Overexpression of GafA, a possible master regulator of GTA production, did not lead to a higher transfer rate in *P. versutus* or GTA expression.

The transcriptional regulator GafA plays a crucial regulatory role in the expression of the main GTA operon in *R. capsulatus*. It binds to the promoter region of the GTA operon and interacts directly with the RNA polymerase omega subunit, thereby initiating the transcription of both structural and DNA-packaging elements necessary for GTA production (Fogg, 2019; Sherlock and Fogg, 2022b). BLASTp analysis revealed significant homology between GafA and the

protein encoded by *LOCUS_7010* in *P. versutus*, with a query coverage of 95% and sequence identity of 52.62%. Based on this homology, we henceforth refer to *LOCUS_7010* as the *gafA* gene in *P. versutus*.

Previous studies have shown that *R. capsulatus* GTA overproducer mutant exhibited increased GTA production due to a mutation in *rcc00280*, a transcriptional repressor of *gafA* (Hynes *et al.*, 2016; Ding *et al.*, 2019). Furthermore, overexpression of *gafA* in *R. capsulatus* wild-type led to higher transfer rates (Fogg, 2019). Based on this, we hypothesized that overexpressing the *gafA* homologue in *P. versutus* might similarly enhance GTA production. Therefore, we constructed a strain harboring an inducible plasmid carrying the *gafA* homologue of *P. versutus*. Contrary to our expectations, overexpression of *gafA* did not result in a significant increase in either the expression of the main GTA operon or GTA-mediated transfer (Figure 6.8). These findings suggest that the endogenous expression levels of GafA may already be sufficient for maximal GTA production in *P. versutus*. It is noteworthy that the presence of the pIND plasmid resulted in a slight increase in transfer frequency, though the underlying mechanism for this effect remains unclear. Additionally, the effect of *gafA* overexpression on GTA reception efficiency was not evaluated in the current experimental design.

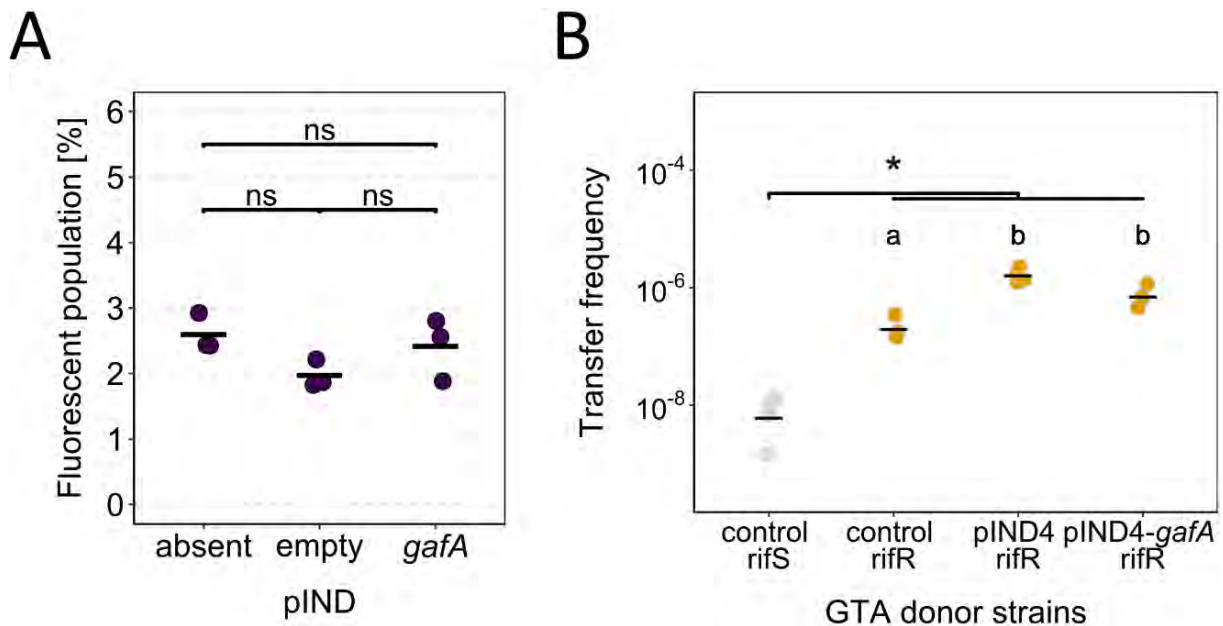


Figure 6.8. Overexpression of GafA did not increase GTA expression and transfer in *P. versutus* DSM582. (A) Percentage of the population that was highly fluorescent in strains carrying the transcriptional reporter pSEVA631-*tetAR*-pGTA-*sfGFP*. Overexpression of *gafA* bore in pIND4 plasmid did not increase the proportion of populations. (B) Transfer frequencies of *P. versutus* DSM582 wild type receiving GTAs from wild type strain (control) and complemented strains containing plasmids pIND4 or pIND4-*gafA* with 1mM IPTG induction. The donor strains were either rifampicin-susceptible (rifS, grey) as a control for spontaneous rifampicin-susceptible mutants, or rifampicin-resistant (rifR, orange). All transfers with GTAs from rifR strains were significantly different from transfers with GTA from rifS strain. In addition, the presence of plasmid pIND4 also increased the transfer rate relative to control rifR as highlighted with letters “a” and “b”. Dots show independent replicates and bars the means. The experiments were performed at least two times.

As shown previously, several strains of *P. versutus* exhibited no detectable HGT mediated by GTAs (Figure 6.2A). The gene *gafA* of those strains displays a few missense mutations (Figure S 6.7). To investigate whether the absence of detectable transfer resulted from regulatory issues, we overexpressed the *gafA* gene from the transformable strain DSM582 in these strains and conducted GTAs transfer experiments using strain DSM582 as a recipient or as a donor, given its demonstrated ability to receive and transfer genetic material via GTA. Although GTAs from strains overexpressing *gafA* led to marginally higher transformability, the transfer rates were not statistically significant compared to the control (Figure 6.9A). Additionally, overexpression of *gafA* did not improve the ability of these strains to receive a selective marker via GTAs (Figure 6.9B). This indicates that the inability of the strains to perform HGT mediated by GTAs was unlikely due to issues with only GafA.

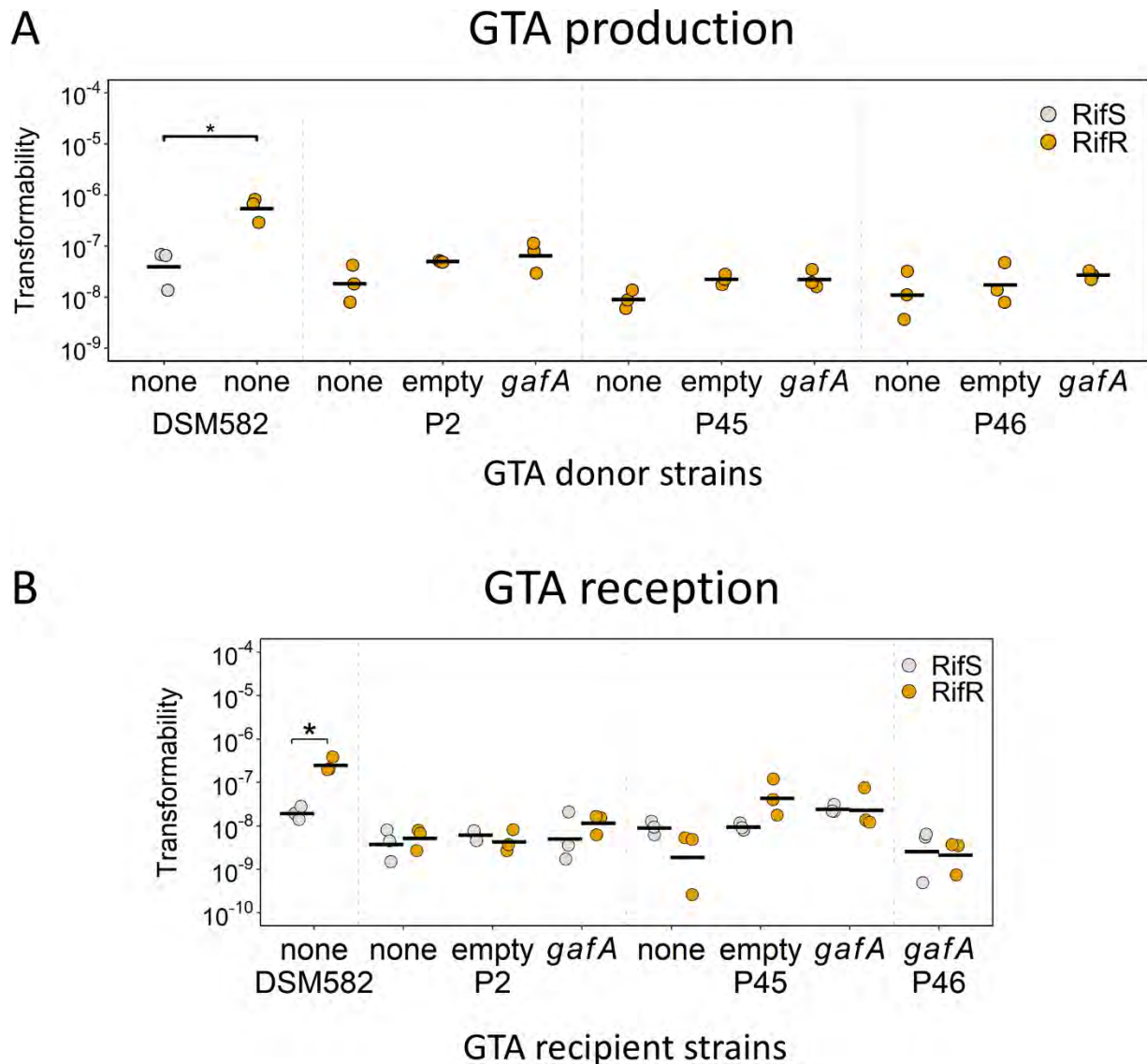


Figure 6.9 Overexpression of *gafA* in the other *P. versutus* strains did not lead to significant HGT mediated by GTAs. (A) Transfer frequencies of *P. versutus* DSM582 with GTA solutions from DSM582 strain and environmental strains (P2, P45, P46) without plasmid (none), or complemented with either pIND4 (empty) or pIND4-*gafA* (*gafA*). Donor strains were either rifampicin-susceptible (RifS, grey) as a control for spontaneous rifampicin-resistant mutants, or rifampicin-resistant (RifR, orange). Transfer experiments with GTA solution from each RifS strain were not performed for simplification. (B) Transfer frequencies of various *P. versutus* wild-type (none) and complemented with pIND4 (empty) or pIND4-*gafA* (*gafA*) in the presence of GTA solution from markerless (RifS, grey) or rifampicin-resistant (RifR, orange) *P. versutus* DSM582. Controls (none and empty) were missing for strain P46 due to repeated contamination. Dots show independent replicates (usually three) and bars the means. The experiments were performed at least two times.

6.5.6 High intracellular cyclic di-GMP concentrations decreased *P. versutus* ability to perform HGT mediated by GTAs.

We previously isolated a phage-resistant mutant of *P. versutus* DMS582 (Chapter 5). This strain harbors a mutation in *cpdB*, resulting in an elevated intracellular concentration of 3',5'-cyclic diguanylate (c-di-GMP). Transcriptomic analysis of the phage-resistant mutant during exponential growth revealed that elevated c-di-GMP levels influence various

phenotypes, including motility and metabolism (Chapter 5), consistent with c-di-GMP established role as a secondary messenger (Hengge, 2009; Jenal *et al.*, 2017).

Since high c-di-GMP concentrations decreases GTA production and main RcGTA cluster expression (Pallegar, Canuti, *et al.*, 2020; Pallegar, Langille, *et al.*, 2020), we investigated the impact of the *cpdB* mutation on GTA-mediated HGT in *P. versutus*. Transfer assays revealed a reduction in transformability when the *cpdB* mutant served as GTA donor (Figure 6.10A), suggesting lower GTA production. This was corroborated by flow cytometry analysis, which showed significantly reduced GTA operon expression in the *cpdB* mutant compared to the wild-type strain (Figure 6.10B). While complementation with *gafA* restored expression of the main GTA operon, transfer assays demonstrated no recovery of GTA production to wild-type levels (Figure S 6.8). This indicates that, despite *gafA*-mediated restoration of main GTA operon expression, other processes, such as GTA particles maturation or release, remain negatively impacted by elevated c-di-GMP concentrations.

The *cpdB* mutant also exhibited an impaired ability to receive DNA from GTAs (Figure 6.10C), a defect that *gafA* complementation also failed to rescue (Figure S 6.8B). To determine if the impaired recipient capability was due to an adsorption issue, we assessed the residual activity of a GTA solution after contact with wild-type or mutant cells. GTA transfer was detected only in the control condition, where the GTA solution was first in contact with cell-free medium (Figure 6.10D). This suggests that the GTA particles attach to mutants at levels similar to wild-type cells. However, the experimental design was not optimal due to the absence of a bacterial control, and some GTAs might have been removed or damaged during the additional filtration step with cells. In addition, we cannot exclude the presence of extracellular modifications in the *cpdB* mutant, as evidenced by the reduced adsorption of phage PvEP4A (Chapter 5). In summary, the alteration of *cpdB* affected both GTA production and reception, underscoring its global importance in HGT mediated by GTAs.

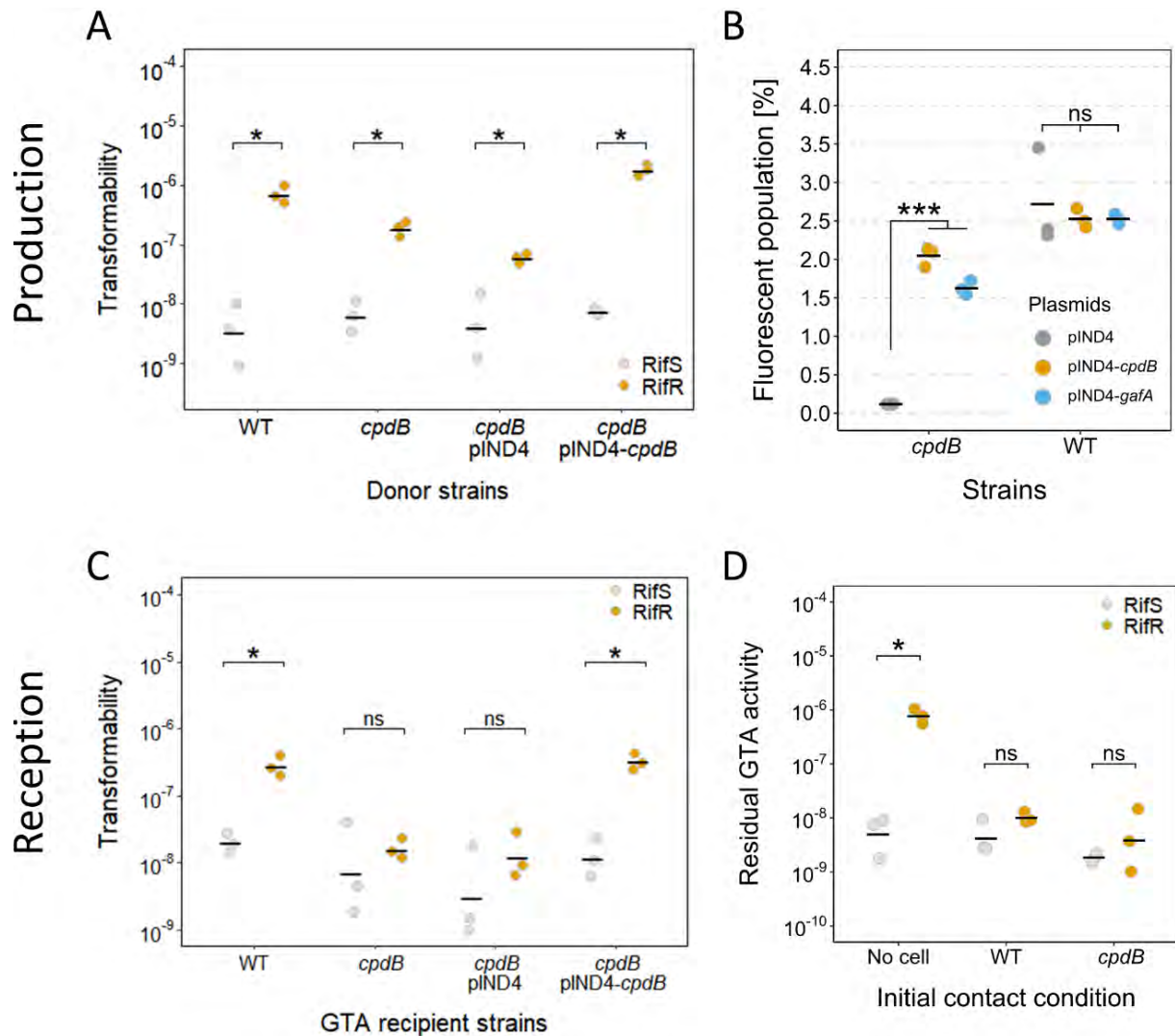


Figure 6.10. The concentration of c-di-GMP impacts not only GTA production but also the transformability of recipients. (A) Production of GTAs was affected by the *cpdB* mutation as shown by a reduced transformability of the wild-type recipient with GTAs from the mutant. (B) The *cpdB* mutation impaired the expression of the main GTA operon which was partially recovered by complementation with *cpdB* or *gafA*. For details of the method, see section Material and Method or legend of Figure 6.6. (C) Nucleotide deletion in *cpdB* also reduced the ability of the mutant to transform with GTAs. The GTA solutions were obtained from strain with the original *cpdB* either susceptible (grey) or resistant (orange) to rifampicin. Rifampicin resistance was used as a marker of transfer. (D) The residual activity of GTA solution was estimated by performing transfer experiments with GTA solutions that have been in contact with bacterial cells. Dots show independent replicates and bars the means.

Since elevated c-di-GMP level decreases both GTA production and reception, we analyzed transcriptomic data for *P. versutus* homologues of *R. capsulatus* GTA-mediated HGT genes. In the phage-resistant mutant, multiple HGT-associated genes exhibited significant differential expression (Table 6.3).

For GTA production, genes encoding structural components and DNA packaging machinery of GTAs were generally downregulated, possibly due to the downregulation of the *gafA* given the role of its homologue in RcGTA operon expression (Fogg, 2019). However, *LOCUS_29310*, situated in the middle of the main GTA operon, showed no significant change in expression

despite elevated c-di-GMP concentration. Its homologue in *R. capsulatus*, *rcc01693*, is annotated as a phage tail assembly chaperone (Kogay *et al.*, 2022), but its specific role in RcGTA remains unclear. The downregulation of the main GTA operon in *cpdB* mutant is consistent with the lower expression of the GTA operon observed with the transcriptional reporter strains (Figure 6.10B). Genes coding for head spikes, the components of GTA particles allowing the binding of GTA to recipient cells (Alim *et al.*, 2023), showed no differential expression in the mutant, but their identification remains uncertain due to disrupted genomic colocalization and low sequence identity of *LOCUS_40350* with *R. capsulatus* homologue. Interestingly, the gene encoding the holin (*LOCUS_43380*) was downregulated, while the expression of the adjacent gene (*LOCUS_43390*), coding for the putative endolysin, remained unaffected. In *R. capsulatus*, both the endolysin (Rcc00555) and holin (Rcc00556) are required for complete release of GTAs (Westbye *et al.*, 2013). This differential expression pattern suggests reduced GTA release efficiency in *P. versutus*. Furthermore, reduced expression was also observed for *LOCUS_07020*, a *P. versutus* gene homologous to the regulatory gene *rcc01866* in *R. capsulatus*, which is known to play a role in GTA release and potentially maturation (Hynes *et al.*, 2016; Sherlock and Fogg, 2022b). This suggests that, in addition to the observed downregulation of the main GTA operon, GTA release was also likely compromised in the *cpdB* mutant.

For GTA reception, it occurs through three main steps: adsorption, DNA injection and translocation, and finally DNA recombination (Chapter 1, Figure 4). In *R. capsulatus*, the bacterial capsule was found to be essential for GTA adsorption. Genetic studies identified three key contributors: two polysaccharide biosynthesis genes (*rcc01081* and *rcc01932*) and a serine acetyltransferase gene (*rcc01700*) (Brimacombe *et al.*, 2013; Sherlock and Fogg, 2022a). In *P. versutus*, the expression of the homologous genes was not affected by increased c-di-GMP concentrations (Table 6.3). Interestingly, the deletion of *LOCUS_17550*, the *P. versutus* homologue of *rcc01932*, did not impair GTA reception (Figure S 6.9), suggesting that *LOCUS_17550* is not involved in GTA-mediated HGT in this species. However, the CapsuleFinder tool identified several genes in *P. versutus* potentially involved in capsule biogenesis (Rendueles *et al.*, 2017). The expression of some of these genes was significantly altered in the mutant strain (Table S 6.9), particularly *LOCUS_15920*, which showed a more than 16-fold reduction in expression. Moreover, *LOCUS_15920* is annotated as a glycosyl

transferase, like the *rcc01081* and *rcc01932* genes in *R. capsulatus* (Brimacombe *et al.*, 2013). This similarity suggests a potential role for *LOCUS_15920* in PvGTA reception, but further investigation is needed to confirm this hypothesis. Regarding the last two steps of GTA reception, homologues involved in DNA translocation and DNA recombination exhibited a reduced expression (Table 6.3). As for GTA production, multiple steps of GTA reception might be disturbed by the higher concentration of c-di-GMP.

Taken together, our results demonstrate that high intracellular c-di-GMP levels, caused by a *cpdB* mutation, have a multifaceted negative impact on GTA-mediated HGT in *P. versutus*. The *cpdB* mutant showed reduced GTA production, possibly stemming from the downregulation of the main GTA operon and genes associated with lysis. Furthermore, the mutant ability to receive GTA was also compromised, potentially due to alterations in capsule biogenesis and decreased expression of genes required for DNA translocation and recombination. While *gafA* complementation partially restored GTA operon expression, it did not fully restore GTA production or reception to wild-type levels, suggesting that c-di-GMP influence extends beyond *gafA*-mediated regulation. It is noteworthy that, alongside *gafA*, several homologues of genes associated with RcGTA regulation were downregulated. This included *LOCUS_07020*, a homologue of *R. capsulatus* regulatory gene *rcc01866* known to mediate GTA release and potentially maturation, as well as *LOCUS_17480* and *LOCUS_21630*, which encode homologues of the CtrA phosphorylation key players *cckA* and *divL* (Hynes *et al.*, 2016; Westbye *et al.*, 2018; Sherlock and Fogg, 2022b). These data emphasize the central role of c-di-GMP in regulating GTA-mediated HGT, affecting both donor and recipient cells and revealing a complex relationship between c-di-GMP signaling and bacterial genetic exchange.

Table 6.3. Homologues of *R. capsulatus* involved in HGT with GTAs and their expression change between the wild-type and the *cpdB* mutant. In bold, significant L2FC.

Function	Locus	L2FC	Padj	Name	Description	Rc ID	Rc name*	% ID	Evalue	COG category
adsorption	LOCUS_12550	0.26	0.55	LysR family	transcriptional regulator	rcc01079	ghsA	41	7E-01	Transcription
adsorption	LOCUS_40350	-0.18	0.63	hypothetical protein		rcc01080	ghsB	26	8E+00	NA
DNA packaging	LOCUS_29200	-3.04	0.00	hypothetical protein		rcc01682		49	1E-16	NA
DNA packaging	LOCUS_29210	-3.83	0.00	large terminase		rcc01683		79	0E+00	Function Unknown
structural protein	LOCUS_29220	-5.75	0.00	portal protein		rcc01684		77	0E+00	Function Unknown
unknown	LOCUS_29230	-6.74	0.00	hypothetical protein		rcc01685		77	2E-37	NA
virion maturation	LOCUS_29240	-6.15	0.00	peptidase U35		rcc01686		70	7E-70	Function Unknown
structural protein	LOCUS_29250	-6.73	0.00	phage capsid protein		rcc01687		60	8E-173	Function Unknown
structural protein	LOCUS_29260	-6.40	0.00	hypothetical protein		rcc01688		60	2E-56	NA
structural protein	LOCUS_29270	-5.97	0.00	hypothetical protein		rcc01689		53	1E-31	Function Unknown
structural protein	LOCUS_29280	-5.42	0.00	hypothetical protein		rcc01690		61	4E-44	Function Unknown
structural protein	LOCUS_29290	-4.89	0.00	tail protein		rcc01691		77	6E-79	Function Unknown
structural protein	LOCUS_29300	-3.19	0.00	hypothetical protein		rcc01692		63	1E-29	Function Unknown
structural protein	LOCUS_29310	-0.31	NA	hypothetical protein		rcc01693		65	6E-23	Function Unknown
structural protein	LOCUS_29320	-4.79	0.00	tail protein		rcc01694		62	9E-88	Function Unknown
structural protein	LOCUS_29330	-5.30	0.00	glycoside hydrolase family 24		rcc01695		74	4E-110	Function Unknown
structural protein	LOCUS_29340	-4.66	0.00	hypothetical protein		rcc01696		60	1E-112	Function Unknown
DNA entry	LOCUS_29350	-4.37	0.00	peptidase		rcc01697		64	2E-52	Cell wall/ membrane/ envelop biogenesis
structural protein	LOCUS_29360	-4.10	0.00	phage host specificity protein		rcc01698		60	0E+00	Function Unknown
attachment	LOCUS_29370	-5.44	0.00	ribonuclease III		rcc00171	Tail fiber	41	7E-16	Function Unknown
release	LOCUS_43380	-2.48	0.00	hypothetical protein		rcc00556		55	3E-50	Function Unknown
release	LOCUS_43390	-0.59	0.01	hypothetical protein		rcc00555		65	2E-88	Function Unknown
Regulation										
master regulator	LOCUS_07010	-5.02	0.00	helix-turn-helix containing protein	domain	rcc01865	gafA	53	7E-112	Transcription
virion maturation	LOCUS_07020	-5.59	0.00	hypothetical protein		rcc01866		49	4E-12	NA

quorum sensing	LOCUS_01240	0.33	0.35		LuxR family regulator	transcriptional	rcc00328	gtaR	69	3E-98	Transcription
quorum sensing	LOCUS_01250	-0.12	0.79		N-acyl-L-homoserine synthetase	lactone	rcc00329	gtal	67	9E-101	Secondary metabolites biosynthesis, transport and catabolism
extracellular repressor of RcGTA gene expression	LOCUS_54220	0.16	0.77		hypothetical protein		rcc00280		49	9E-32	Secondary metabolites biosynthesis, transport and catabolism
phosphorelay pathway	LOCUS_15670	-0.71	0.00		LexA repressor		rcc02361	lexA	71	3E-120	Transcription
phosphorelay pathway	LOCUS_17480	-4.05	0.00		hypothetical protein		rcc01749	cckA	64	0E+00	Signal transduction mechanisms
phosphorelay pathway	LOCUS_21630	-3.60	0.00		diguanylate cyclase		rcc00042	divL	54	5E-143	Signal transduction mechanisms
phosphorelay pathway	LOCUS_40490	-0.27	0.36		ATP-dependent ATP-binding subunit	Clp protease	rcc02608	clpX	88	0E+00	Post-translational modification, protein turnover, chaperones
phosphorelay pathway	LOCUS_40500	0.00	0.99	clpP	ATP-dependent proteolytic subunit	Clp protease	rcc02609	clpP	79	6E-128	Post-translational modification, protein turnover, chaperones
phosphorelay pathway	LOCUS_41780	-0.43	0.15	ctrA	DNA-binding response regulator		rcc01663	ctrA	84	2E-129	Signal transduction mechanisms
phosphorelay pathway	LOCUS_46310	0.05	0.91		histidine phosphotransferase		rcc03000	chpT	48	1E-54	Function Unknown
GTA reception											
polysaccharide biosynthesis genes	LOCUS_01720	0.24	0.42		colanic acid biosynthesis glycosyltransferase	WcaL	rcc01081		31	5E-05	Cell wall/membrane/envelop biogenesis
polysaccharide biosynthesis genes	LOCUS_17550	-0.21	0.52	mraY	phospho-N-acetylmuramoyl-pentapeptide-transferase		rcc01932		29	2E-09	Cell wall/membrane/envelop biogenesis
reception	LOCUS_29380	-0.28	0.29	cysE	serine O-acetyltransferase		rcc01700	cysE1	80	1E-149	Amino Acid transport and metabolism
Transformation											
DNA transport from periplasm to cytoplasm	LOCUS_19850	-4.04	0.00	comF	amidophosphoribosyltransferase		rcc00197	comF	53	3E-66	Function Unknown
DNA transport from periplasm to cytoplasm	LOCUS_45600	-4.80	0.00		hypothetical protein		rcc02362	comEC	51	1E-168	Function Unknown

DNA transport from periplasm to cytoplasm	LOCUS_43210	-4.98	0.00		ATPase AAA	rcc00460	comM	69	0E+00	Post-translational modification, protein turnover, chaperones
recombination	LOCUS_37290	-5.25	0.00		DNA processing protein DprA	rcc03098	drpA	57	5E-118	Replication, recombination and repair
recombination	LOCUS_17460	-2.30	0.00	recA	protein RecA	rcc01751	recA	87	2E-09	Replication, recombination and repair

Abbreviations: padj: DESeq2 p value adjusted for multiple testing; L2FC, logarithm in base 2 of the expression ratio of the wild type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant); Rc name: cursive names correspond to genes, others are according to the model from Sherlock and Fogg (Sherlock and Fogg, 2022b); %ID refers to percentage identity shared between *R. capsulatus* genes and homologues in *P. versutus* with e-value indicated.

6.6 Discussion

Here, we investigated GTA-mediated HGT in the *Paracoccus* genus, belonging to the Rhodobacterales order like the classic GTA model, *R. capsulatus*. Despite being discovered half a century ago, HGT mediated by GTAs is still poorly understood compared to other HGT mechanisms. In addition, most of our knowledge stems from a narrow range of model organisms, particularly *R. capsulatus*, hindering a comprehensive understanding of GTAs and their implications across diverse bacterial taxa. Although we investigated multiple *Paracoccus* species and strains carrying the main GTA operon in their genomes, we only detected transfer mediated by GTAs in one strain, *P. versutus* DSM582. Therefore, we studied different aspects of GTAs in this model. While GTA particles encapsulated DNA fragments from the entire genome, regions near the origin of replication of the chromid and megaplasmid were more frequently represented in GTA packages. Consistent with the burden carried by GTA producers leading to cell death, we showed that only a small subset of the population expresses GTA structural component genes but mainly during the exponential phase. Elevated concentrations of c-di-GMP, resulting from a mutation conferring phage resistance, negatively affected both the production and reception of GTAs. This observation suggests a crucial regulatory role for this second messenger in GTA-mediated gene transfer.

6.6.1 In pursuit of a model to study GTA-mediated transfer in *Paracoccus* spp.

Given the widespread presence of the main GTA operon in *Paracoccus* spp. (Figure 6.1), we expected to find multiple strains capable of significant levels of GTA-mediated HGT. This expectation was further supported by previous investigations in *R. capsulatus*, where over 80% of 33 strains examined could either produce and/or receive GTAs (Wall et al., 1975). However, we observed HGT mediated by GTAs in only one strain of *Paracoccus* in this work (Figure 6.2, Figure 6.3). The absence of detected gene transfer in most *Paracoccus* strains may be attributed to several factors. On the one hand, GTA production and reception are tightly regulated processes that may require specific conditions for optimal transfer. While gene transfers in *R. capsulatus* are primarily induced at high cell density or under nutrient limitations (Solioz et al., 1975; Lang and Beatty, 2000; Westbye et al., 2017), they are favored in unstressed, exponentially-growing *B. henselae* cells (Québatte et al., 2017). Our study involved *Paracoccus* species isolated from diverse environments, each potentially requiring unique conditions for GTA-mediated HGT. Consequently, our experimental setup may not

have adequately addressed the specific requirements of each species. On the other hand, the sensitivity of our transfer detection method may have been a limiting factor. The choice of rifampicin resistance as a genetic marker, while convenient and compatible with GTA packaging limitations, may have lacked the resolution to identify transfer events occurring at frequencies below the spontaneous mutation rate. Future investigations should explore more sensitive markers, such as fluorescence-based reporters or antibiotics with lower spontaneous resistance frequencies, to enhance detection sensitivity. Despite these limitations, our approach successfully identified one strain, *P. versutus* DSM582, capable of GTA-mediated HGT. This finding underscores the potential for GTA-mediated gene transfer within the *Paracoccus* genus and highlights the need for further investigation into the specific conditions and mechanisms governing this process across different species.

6.6.2 Bias in DNA fragments encapsulated in GTA particles

One of the most surprising results was the bias in genome coverage of DNA fragments encapsulated in GTA particles, with distinct peaks observed on two out of four replicons (Figure 6.5). This pattern contrasts sharply with the even genome coverage seen in RcGTA, which encapsulates the entire genome uniformly, except for the RcGTA locus itself (Hynes *et al.*, 2012). Nevertheless, enrichment patterns have been observed in GTAs from other bacterial species, albeit with considerable variation in peak number and genomic position, suggesting diverse underlying mechanisms may be at play.

Several bacterial species exhibit unique GTA packaging patterns that provide valuable context for our findings. In *Bartonella* spp., for instance, the packaging bias of GTAs is linked to a run-off replication (RoR) origin derived from a defective prophage, leading to the amplification of genomic regions surrounding the RoR origin (Berglund *et al.*, 2009; Québatte *et al.*, 2017). *Caulobacter crescentus* GTAs, on the other hand, exhibit underpackaging at the origin and terminus of replication, with two prominent peaks midway between them (Gozzi *et al.*, 2022). Intriguingly, the GTA operon of *C. crescentus* is also located in one of these peaks. Researchers have hypothesized that this pattern resulted from chromosomal arm cohesion, localization of those regions to mid-cell, and proximity of the GTA cluster (Gozzi *et al.*, 2022). *D. shibae* presents an even more complex GTA packaging pattern with seven distinct peaks of overrepresented genomic regions, also showing lower packaging of the terminus and origin of replication (Tomasch *et al.*, 2018).

The various hypotheses proposed to explain these patterns can be broadly categorized into three main mechanisms: higher genome copy numbers (as seen in *Bartonella* spp. GTAs), chromosome spatial structures (as proposed for *C. crescentus* GTA and *D. shibae* GTA), or recognition of specific sites (suggested for *D. shibae* GTA as well) (Figure 6.11). These diverse explanations highlight the complexity and species-specific nature of GTA packaging mechanisms.

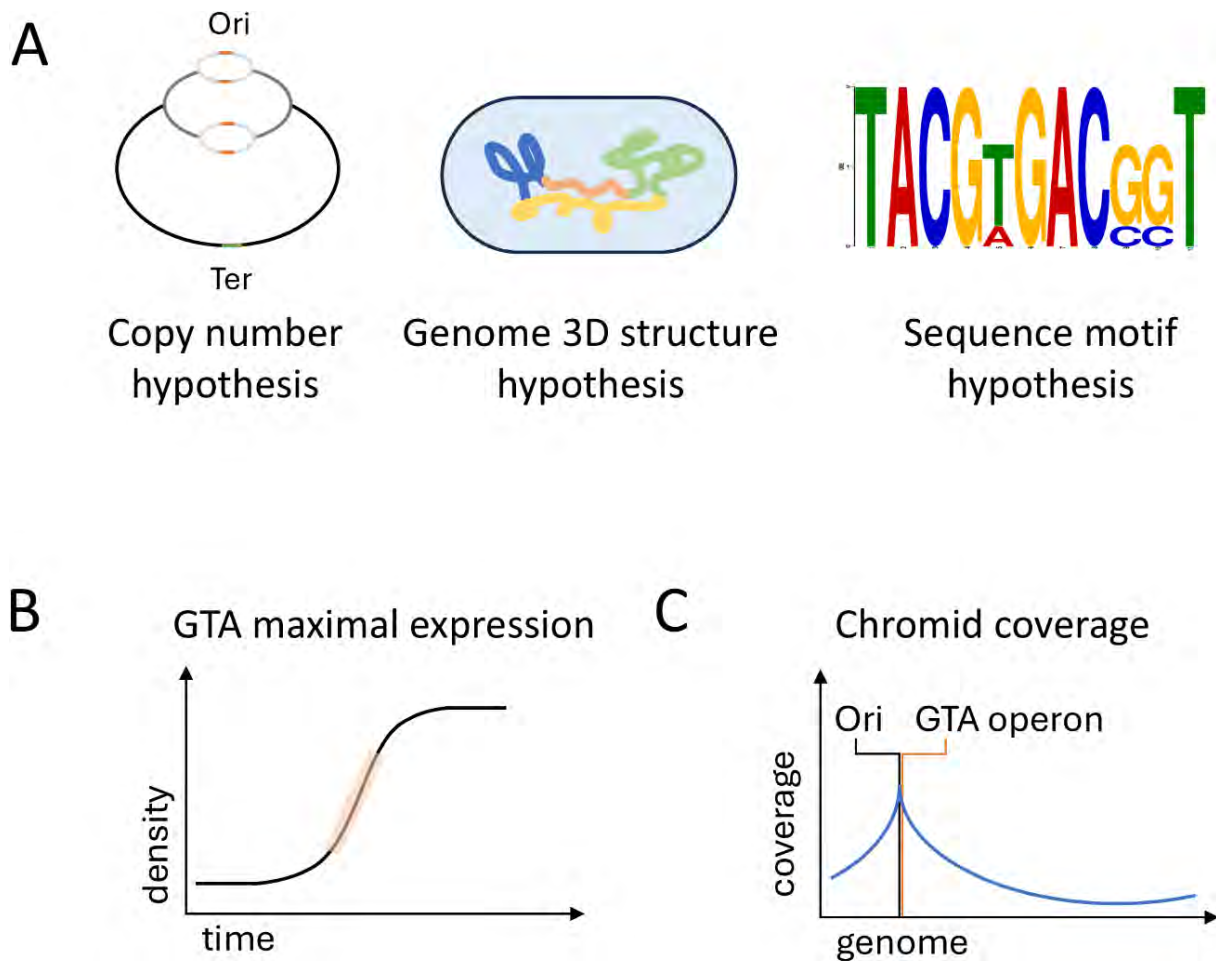


Figure 6.11 Hypotheses (A) to explain the observed bias in DNA encapsulated in GTAs and illustrations of results obtained in *P. versutus* regarding (B) timing of the main GTA operon expression and (C) genome coverage of DNA encapsulated in GTAs.

In the case of *P. versutus*, the hypothesis that certain genome regions are overrepresented in GTA particles due to higher copy numbers is plausible. Notably, the maximal expression of the main GTA operon occurs during the exponential growth phase (Figure 6.6, Figure 6.11B), a period characterized by rapid cell division (Ughy *et al.*, 2023). This hypothesis is further supported by the observation that the peaks are located near the origins of replication of the respective replicons (Figure 6.5, Figure 6.11C). However, this would suggest an

unconventional replication order, starting with the chromid, followed by the megaplasmid, and then the chromosome and plasmid. Indeed, replication of multipartite genomes such for *Vibrio cholerae* and *Sinorhizobium meliloti* starts with replication of the chromosome followed by the other replicons (Rasmussen *et al.*, 2007; De Nisco *et al.*, 2014; DiCenzo and Finan, 2017). Currently, no experimental data are available to corroborate the replication order in *P. versutus*. An alternative explanation for the observed enrichment pattern could be run-off replication, as proposed for *Bartonella* (Berglund *et al.*, 2009). In that study, the authors found no sequence similarity between the region close to the peak of the amplification and typical plasmid origin, providing no evidence for a megaplasmid-derived origin of replication. Therefore, they concluded that replication likely initiated from an origin associated with an inactivated prophage. In *P. versutus*, the typical *repABC* operon, involved in plasmid replication, was only found in the plasmid and no other prophage genes—aside from the GTA operon—were predicted within the enriched regions (Table S 6.8). Overall, while copy number variation remains a compelling explanation, the exact mechanisms underlying the observed enrichment patterns in *P. versutus* GTA particles would require further investigation. We also searched for some shared motifs between both enriched regions that might explain the observed packaging bias, but our results were inconclusive. Additionally, data regarding the spatial organization of the replicons are currently not available for *P. versutus*. However, chromosomal macrodomains might be predicted using techniques to capture chromatin conformation (Song and Loparo, 2015), which could provide valuable insights in future studies. An alternative hypothesis might be that GTAs preferentially package DNA near the main GTA operon due to physical proximity during assembly which might be consistent with the observations in *C. crescentus* (Gozzi *et al.*, 2022). However, in *P. versutus* and in *C. crescentus*, the peaks were not perfectly centered on the main GTA operon, weakening this hypothesis.

Given the current evidence, particularly the position of the peaks near replication origins and the timing of main GTA expression, we believe that the copy number hypothesis is the most likely explanation for the observed packaging bias in *P. versutus* GTAs. However, to fully understand this phenomenon, the replication order of the various replicons in *P. versutus* needs to be elucidated through further research. Alternatively, finding another *Paracoccus* species capable of GTA-mediated gene transfer might provide valuable insights, as the genome architecture varies within the genus (Table S 6.5).

The unique aspects of GTA packaging raise intriguing questions about the evolutionary pressures that have shaped this system. For instance, the preferential packaging of certain genomic regions could reflect an adaptive strategy to transfer specific genes or functional modules that are particularly beneficial in the ecological niches occupied by those species. This possibility opens up exciting avenues for future research into the ecological and evolutionary significance of GTA-mediated gene transfer in bacterial populations.

6.6.3 Regulation of GTA in *P. versutus*—what a phage-resistant mutant teaches us.

The regulatory network of GTA has been studied mainly in *R. capsulatus*, leading to the development of a refined model over time. This model suggests that RcGTA production is intricately regulated by a complex network involving quorum sensing, the CtrA phosphorelay system, and the direct activator GafA (Chapter 1, Figure 5) (Fogg, 2019; Sherlock and Fogg, 2022b). Similar regulatory mechanisms have been observed in *D. shibae*, where the main GTA operon is controlled by the CtrA phosphorelay and quorum sensing (Wang *et al.*, 2014; Koppenhöfer *et al.*, 2019). These similarities between *R. capsulatus* and *D. shibae* raise the question of whether the regulatory network of GTAs is conserved across other RcGTA-like models.

Our investigation indicates that quorum sensing plays a positive role in the expression of the main GTA operon (Figure 6.7), consistent with findings in *R. capsulatus* and *D. shibae* (Leung *et al.*, 2012; Koppenhöfer *et al.*, 2019). Nevertheless, the timing of expression of the main GTA operon differ between the species: PvGTA is mainly expressed during the exponential phase (Figure 6.6) while RcGTA production is maximal during stationary phase (Solioz *et al.*, 1975; Lang and Beatty, 2000). Therefore, confirmatory tests with deletion mutants for *gtaI* and *gtaR* homologues are needed to clarify the role of quorum sensing in PvGTA.

In this study, we did not provide direct experimental evidence regarding the role of CtrA phosphorelay in the GTA regulation of *P. versutus* but focused our investigations on the impact of elevated concentration of c-di-GMP on PvGTA (Figure 6.12). Since CtrA activity is modulated by c-di-GMP levels in *R. capsulatus* (Farrera-Calderon *et al.*, 2020; Pallegar, Langille, *et al.*, 2020), the results with the *cpdB* mutant might provide valuable insights.

Increased intracellular c-di-GMP levels downregulated multiple genes associated with GTA-mediated HGT (Table 6.3, Figure 6.12). Transcriptomic analysis revealed decreased expression of *gafA*, the main PvGTA operon, and the holin-encoding gene *LOCUS_43380* in the

cpdB mutant. Additionally, reduced expression was observed for *LOCUS_07020*, a *P. versutus* homologue of the *R. capsulatus* regulatory gene *rcc01866*, which has been implicated in GTA particle release and maturation (Hynes *et al.*, 2016; Sherlock and Fogg, 2022b). This coordinated downregulation of genes essential for GTA production and release is consistent with the observed reduction in GTA-mediated transfer efficiency (Figure 6.10A). Interestingly, these findings parallel previous observations in *R. capsulatus*, where elevated c-di-GMP levels were shown to decrease RcGTA operon expression and capsid protein production (Pallegar, Langille, *et al.*, 2020).

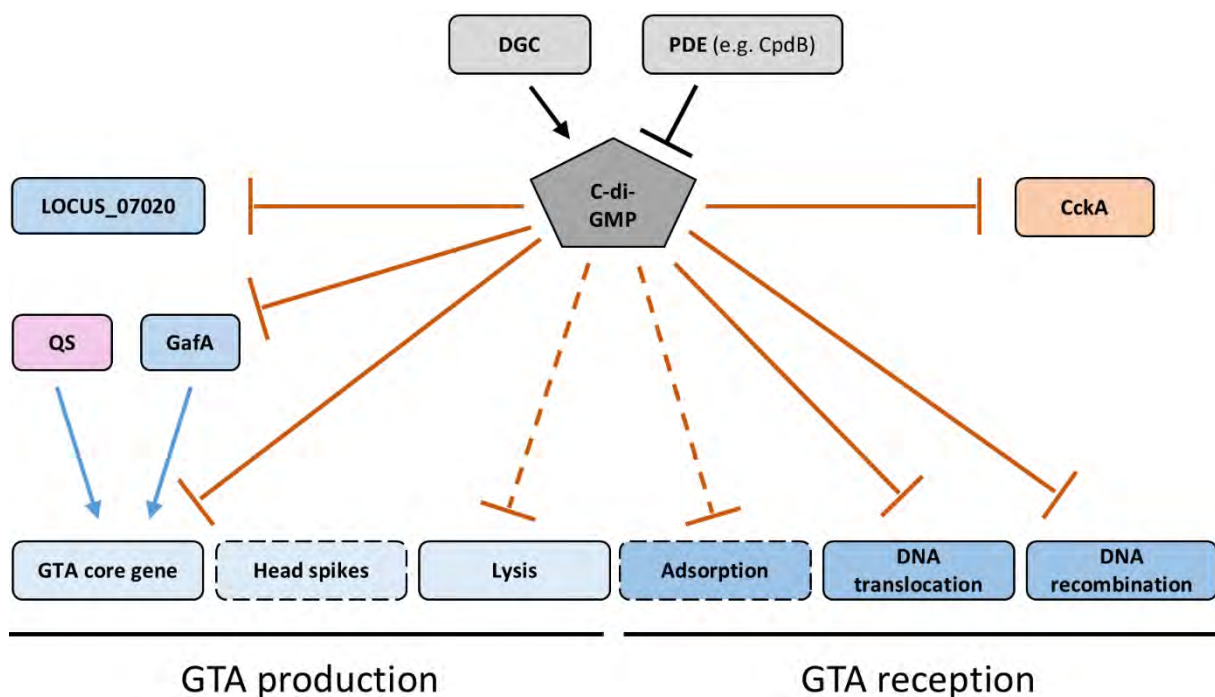


Figure 6.12 Visual summary of the findings regarding the regulation of GTAs in *P. versutus*. Arrows indicate activation (standard arrows) or repression (flat-head arrows), with dashed arrows representing uncertainty about the effect. Arrow colors: orange (transcriptomics data), black (literature), and blue (experiments). Box colors: orange (*cckA-chpT-ctrA* phosphorelay), light/medium/dark blue (GTA production/regulation/reception), pink (quorum sensing), and grey (c-di-GMP metabolism). Dashed boxes indicate categories where homolog identification remains uncertain.

The overexpression of *gafA* restored the main PvGTA operon expression, but failed to fully reinstate transfer to wild-type levels (Figure 6.10B, Figure S 6.8). This observation suggests a two-step GTA production process, similar to *R. capsulatus* (Sherlock and Fogg, 2022b). In this model, the main GTA operon is under the direct control of *gafA*, while other crucial components such as lysis and head spikes are also regulated by additional partners including phosphorylated CtrA and *rcc01866* (Chapter 1, Figure) (Mercer *et al.*, 2012; Hynes *et al.*, 2016; Sherlock and Fogg, 2022b).

GTA reception was also impaired in the *cpdB* mutant (Figure 6.10C). The transformation-related genes *comEC*, *comF*, *comM*, *dprA* and *recA* showed reduced expression in *cpdB* mutant (Table 6.3), consistent with the reduction in GTA reception as observed in transfer assay (Figure 6.10C). These genes are known targets of CtrA regulation in *D. shibae* and are co-regulated with quorum sensing in *R. capsulatus* (Brimacombe *et al.*, 2014, 2015; Koppenhöfer *et al.*, 2019). However, the *ctrA* and quorum sensing genes expression remained unchanged in the *cpdB* mutant (Table 6.3), suggesting a different regulation of GTA reception in *P. versutus*.

Collectively, these data show that elevated c-di-GMP levels in *P. versutus* impair both GTA production and reception, by downregulating multiple genes involved in these processes. While the mechanism by which c-di-GMP influences these genes remains unclear, a potential role for the CckA-ChpT-CtrA phosphorelay—analogous to that described in *R. capsulatus*—cannot be ruled out. However, several observations suggest that the RcGTA regulatory model may not fully apply to PvGTA. In *R. capsulatus*, the CckA-ChpT-CtrA phosphorelay is essential for CtrA phosphorylation (Westbye *et al.*, 2018; Farrera-Calderon *et al.*, 2020). Given the reduced expression of *cckA* and *divL* in the *cpdB* mutant (Table 6.3), CtrA would theoretically remain rather unphosphorylated, if *P. versutus* follows the same regulatory cascade than *R. capsulatus*. However, our findings in the *cpdB* mutant appear inconsistent with the RcGTA model, as unphosphorylated CtrA activates *gafA* expression and triggers recipient ability to a greater extent than the phosphorylated form (Brimacombe *et al.*, 2014; Fogg, 2019), contrasting with the observed decrease of GTA-related functions. This highlights the need for additional investigations to elucidate the function of the CckA-ChpT-CtrA phosphorelay in PvGTA regulation. Future studies should also incorporate quantitative analysis of intracellular and extracellular capsid protein concentrations during GTA production, along with direct assessment of CtrA phosphorylation states under varying c-di-GMP conditions. Such approaches would provide key information to gain a better understanding of GTA regulation in *P. versutus*.

6.7 Conclusion

The prevalence and conservation of GTAs, particularly in alphaproteobacteria, have long suggested that they provide cells with a benefit. However, research conducted on only a few

models has provided a limited and potentially biased view of this complex mechanism. Our study highlights a new model, *P. versutus* DSM582, closely related to *R. capsulatus* but with its own specific features, such as the timing of GTA expression and bias in encapsulated DNA, potentially impacting the benefits conferred by GTAs. While the effects of c-di-GMP seem similar on *P. versutus* and *R. capsulatus* GTAs, the role of CckA-ChpT-CtrA phosphorelay in PvGTA remains to be determined. Our discovery that *P. versutus* produces functional GTAs now offers a new, easy-to-handle organism to further dissect the functions, biogenesis, and regulation of GTAs, expanding our understanding of these fascinating genetic elements across diverse bacterial species.

6.8 References

- Alim, N.T.B., Koppenhöfer, S., Lang, A.S., and Beatty, J.T. (2023) Extracellular Polysaccharide Receptor and Receptor-Binding Proteins of the *Rhodobacter capsulatus* Bacteriophage-like Gene Transfer Agent RcGTA. *Genes (Basel)* **14**: 1–16.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.
- Arakawa, K. and Tomita, M. (2007) The GC skew index: A measure of genomic compositional asymmetry and the degree of replicational selection. *Evol Bioinforma* **3**: 159–168.
- Banks, E.J. and Le, T.B.K. (2024) Co-opting bacterial viruses for DNA exchange: structure and regulation of gene transfer agents. *Curr Opin Microbiol* **78**: 1–10.
- Berglund, E.C., Frank, A.C., Calteau, A., Pettersson, O.V., Granberg, F., Eriksson, A.S., et al. (2009) Run-off replication of host-adaptability genes is associated with gene transfer agents in the genome of mouse-infecting *Bartonella grahamii*. *PLoS Genet* **5**: 1–16.
- Biers, E.J., Wang, K., Pennington, C., Belas, R., Chen, F., and Moran, M.A. (2008) Occurrence and expression of gene transfer agent genes in marine bacterioplankton. *Appl Environ Microbiol* **74**: 2933–2939.
- Brimacombe, C.A., Ding, H., and Beatty, J.T. (2014) *Rhodobacter capsulatus* DprA is essential for RecA-mediated gene transfer agent (RcGTA) recipient capability regulated by quorum-sensing and the CtrA response regulator. *Mol Microbiol* **92**: 1260–1278.
- Brimacombe, C.A., Ding, H., Johnson, J.A., Thomas Beatty, J., and Beatty, J.T. (2015) Homologues of genetic transformation DNA import genes are required for *Rhodobacter capsulatus* gene transfer agent recipient capability regulated by the response regulator CtrA. *J Bacteriol* **197**: 2653–2663.
- Brimacombe, C.A., Stevens, A., Jun, D., Mercer, R.G., Lang, A.S., and Beatty, T.J. (2013) Quorum-sensing regulation of a capsular polysaccharide receptor for the *Rhodobacter capsulatus* gene transfer agent (RcGTA). *Mol Microbiol* **87**: 802–817.
- Burger, B.T., Imam, S., Scarborough, M.J., Noguera, D.R., and Donohue, T.J. (2017) Combining Genome-Scale Experimental and Computational Methods To Identify Essential Genes in *Rhodobacter sphaeroides*. *mSystems* **2**: 1–18.
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018) Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: 884–890.
- Craske, M.W., Wilson, J.S., and Fogg, P.C.M. (2024) Gene transfer agents: structural and functional properties of domesticated viruses. *Trends Microbiol* **32**: 1–12.
- Daubin, V. and Szöllősi, G.J. (2016) Horizontal Gene Transfer and the History of Life. *Cold Spring Harb Perspect Biol* **8**: 1–11.
- DiCenzo, G.C. and Finan, T.M. (2017) The Divided Bacterial Genome: Structure, Function, and Evolution. *Microbiol Mol Biol Rev* **81**: 1–37.
- Ding, H., Grüll, M.P., Mulligan, M.E., Lang, A.S., and Thomas Beatty, J. (2019) Induction of *Rhodobacter capsulatus* gene transfer agent gene expression is a bistable stochastic process repressed by an

- extracellular calcium-binding rtx protein homologue. *J Bacteriol* **201**: 1–18.
- Dziewit, L., Czarnecki, J., Wibberg, D., Radlinska, M., Mrozek, P., Szymczak, M., et al. (2014) Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids. *BMC Genomics* **15**: 1–16.
- Edgar, R.C. (2004) MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 1–19.
- Farrera-Calderon, R.G., Pallegar, P., Westbye, A.B., Wiesmann, C., Lang, A.S., and Beatty, J.T. (2020) The CckA-ChpT-CtrA phosphorelay controlling *Rhodobacter capsulatus* gene transfer agent production is bidirectional and regulated by cyclic di-GMP. *J Bacteriol* **203**: 1–17.
- Fogg, P.C.M. (2019) Identification and characterization of a direct activator of a gene transfer agent. *Nat Commun* **10**: 1–10.
- Gao, F. and Zhang, C.T. (2008) Ori-Finder: A web-based system for finding oriCs in unannotated bacterial genomes. *BMC Bioinformatics* **9**: 1–6.
- Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**: 343–345.
- Gozzi, K., Tran, N.T., Modell, J.W., Le, T.B.K.K., and Laub, M.T. (2022) Prophage-like gene transfer agents promote *Caulobacter crescentus* survival and DNA repair during stationary phase. *PLoS Biol* **20**: 1–34.
- Grigoriev, A. (1998) Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res* **26**: 2286–2290.
- Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014) Circlize implements and enhances circular visualization in R. *Bioinformatics* **30**: 2811–2812.
- Harrison, P.W., Lower, R.P.J., Kim, N.K.D., and Young, J.P.W. (2010) Introducing the bacterial “chromid”: Not a chromosome, not a plasmid. *Trends Microbiol* **18**: 141–148.
- Hengge, R. (2009) Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol* **7**: 263–273.
- Hernández-Plaza, A., Szklarczyk, D., Botas, J., Cantalapiedra, C.P., Giner-Lamia, J., Mende, D.R., et al. (2023) eggNOG 6.0: enabling comparative genomics across 12 535 organisms. *Nucleic Acids Res* **51**: 389–394.
- Hollensteiner, J., Schneider, D., Poehlein, A., Brinkhoff, T., and Daniel, R. (2023) Pan-genome analysis of six *Paracoccus* type strain genomes reveal lifestyle traits. *PLoS One* **18**: 1–24.
- Hubert, B. (2022) SkewDB, a comprehensive database of GC and 10 other skews for over 30,000 chromosomes and plasmids. *Sci Data* **9**: 1–9.
- Hynes, A.P., Mercer, R.G., Watton, D.E., Buckley, C.B., and Lang, A.S. (2012) DNA packaging bias and differential expression of gene transfer agent genes within a population during production and release of the *Rhodobacter capsulatus* gene transfer agent, RcGTA. *Mol Microbiol* **85**: 314–325.
- Hynes, A.P., Shakya, M., Mercer, R.G., Gröll, M.P., Bown, L., Davidson, F., et al. (2016) Functional and Evolutionary Characterization of a Gene Transfer Agent’s Multilocus “genome.” *Mol Biol Evol* **33**: 2530–2543.

- Ind, A.C., Porter, S.L., Brown, M.T., Byles, E.D., De Beyer, J.A., Godfrey, S.A., and Armitage, J.P. (2009) Inducible-expression plasmid for *Rhodobacter sphaeroides* and *Paracoccus denitrificans*. *Appl Environ Microbiol* **75**: 6613–6615.
- Jenal, U., Reinders, A., and Lori, C. (2017) Cyclic di-GMP: Second messenger extraordinaire. *Nat Rev Microbiol* **15**: 271–284.
- Kelly, D.P., Rainey, F.A., and Wood, A.N.N.P. (2006) The Genus *Paracoccus*. *The Prokaryotes* **5**: 232–249.
- Kogay, R., Koppenhöfer, S., Beatty, J.T., Kuhn, J.H., Lang, A.S., and Zhaxybayeva, O. (2022) Formal recognition and classification of gene transfer agents as viriforms. *Virus Evol* **8**: 1–13.
- Kolmogorov, M., Bickhart, D.M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S.B., et al. (2020) metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat Methods* **17**: 1103–1110.
- Koppenhöfer, S., Wang, H., Scharfe, M., Kaever, V., Wagner-Döbler, I., and Tomasch, J. (2019) Integrated Transcriptional Regulatory Network of Quorum Sensing, Replication Control, and SOS Response in *Dinoroseobacter shibae*. *Front Microbiol* **10**: 1–15.
- Lang, A.S. and Beatty, J.T. (2000) Genetic analysis of a bacterial genetic exchange element: The gene transfer agent of *Rhodobacter capsulatus*. *Proc Natl Acad Sci U S A* **97**: 859–864.
- Lang, A.S. and Beatty, J.T. (2007) Importance of widespread gene transfer agent genes in alpha-proteobacteria. *Trends Microbiol* **15**: 1–9.
- Lang, A.S., Westbye, A.B., and Beatty, J.T. (2017) The Distribution, Evolution, and Roles of Gene Transfer Agents in Prokaryotic Genetic Exchange. *Annu Rev Virol* **4**: 87–104.
- Lang, A.S., Zhaxybayeva, O., and Beatty, J.T. (2012) Gene transfer agents: Phage-like elements of genetic exchange. *Nat Rev Microbiol* **10**: 472–482.
- Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.
- Lasek, R., Szuplewska, M., Mitura, M., Decewicz, P., Chmielowska, C., Pawlot, A., et al. (2018) Genome structure of the opportunistic pathogen *Paracoccus yeei* (Alphaproteobacteria) and identification of putative virulence factors. *Front Microbiol* **9**: 1–17.
- Leung, M.M., Brimacombe, C.A., Spiegelman, G.B., and Beatty, J.T. (2012) The GtaR protein negatively regulates transcription of the gtaRI operon and modulates gene transfer agent (RcGTA) expression in *Rhodobacter capsulatus*. *Mol Microbiol* **83**: 759–774.
- Liao, Y., Smyth, G.K., and Shi, W. (2014) FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.
- Marrs, B. (1974) Genetic recombination in *Rhodopseudomonas capsulata*. *Proc Natl Acad Sci U S A* **71**: 971–973.
- Martínez-García, E. and de Lorenzo, V. (2011) Engineering multiple genomic deletions in Gram-negative bacteria: Analysis of the multi-resistant antibiotic profile of *Pseudomonas putida* KT2440. *Environ Microbiol* **13**: 2702–2716.
- McDaniel, L.D., Young, E., Delaney, J., Ruhnau, F., Ritchie, K.B., and Paul, J.H. (2010) High frequency of horizontal gene transfer in the oceans. *Science (80-)* **330**: 50.

- McDaniel, L.D., Young, E.C., Ritchie, K.B., and Paul, J.H. (2012) Environmental factors influencing gene transfer agent (GTA) mediated transduction in the subtropical ocean. *PLoS One* **7**: 1–9.
- Mercer, R.G., Quinlan, M., Rose, A.R., Noll, S., Beatty, J.T., and Lang, A.S. (2012) Regulatory systems controlling motility and gene transfer agent production and release in *Rhodobacter capsulatus*. *FEMS Microbiol Lett* **331**: 53–62.
- Mergeay, M. and Houdt, R. Van (2018) Plasmids as Secondary Chromosomes. *Mol Life Sci* 1–5.
- Morinaga, K., Nagakubo, T., Nomura, N., and Toyofuku, M. (2020) Involvement of membrane vesicles in long-chain-AHL delivery in *Paracoccus* species. *Environ Microbiol Rep* **12**: 355–360.
- Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., and Minh, B.Q. (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**: 268–274.
- De Nisco, N.J., Abo, R.P., Wu, C.M., Penterman, J., and Walker, G.C. (2014) Global analysis of cell cycle gene expression of the legume symbiont *Sinorhizobium meliloti*. *Proc Natl Acad Sci U S A* **111**: 3217–3224.
- Noé, L. and Kucherov, G. (2005) YASS: Enhancing the sensitivity of DNA similarity search. *Nucleic Acids Res* **33**: 540–543.
- Pagès, H., Aboyoun, P., Gentleman, R., and Debroy, S. (2022) Biostrings: Efficient manipulation of biological strings. *R Packag* 1–32.
- Pallegar, P., Canuti, M., Langille, E., Peña-Castillo, L., and Lang, A.S. (2020) A Two-Component System Acquired by Horizontal Gene Transfer Modulates Gene Transfer and Motility via Cyclic Dimeric GMP. *J Mol Biol* **432**: 4840–4855.
- Pallegar, P., Langille, E., Gomelsky, M., and Lang, A.S. (2020) Cyclic di-GMP-Mediated Regulation of Gene Transfer and Motility in *Rhodobacter capsulatus*. **202**: 1–17.
- Petersen, J. (2011) Phylogeny and compatibility: Plasmid classification in the genomics era. *Arch Microbiol* **193**: 313–321.
- Petersen, J., Brinkmann, H., Berger, M., Brinkhoff, T., Päucker, O., and Pradella, S. (2011) Origin and evolution of a novel DnaA-like plasmid replication type in rhodobacterales. *Mol Biol Evol* **28**: 1229–1240.
- Québatte, M., Christen, M., Harms, A., Körner, J., Christen, B., and Dehio, C. (2017) Gene Transfer Agent Promotes Evolvability within the Fittest Subpopulation of a Bacterial Pathogen. *Cell Syst* **4**: 611–621.
- Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., and Lopez, R. (2005) InterProScan: Protein domains identifier. *Nucleic Acids Res* **33**: 116–120.
- Rasmussen, T., Jensen, R.B., and Skovgaard, O. (2007) The two chromosomes of *Vibrio cholerae* are initiated at different time points in the cell cycle. *EMBO J* **26**: 3124–3131.
- Redfield, R.J., Beatty, J.T., and Lang, A.S. (2019) Gene transfer agents. *Encycl Microbiol* **1**: 370–377.
- Rendueles, O., Garcia-Garcerà, M., Néron, B., Touchon, M., and Rocha, E.P. (2017) Abundance and co-occurrence of extracellular capsules increase environmental breadth: Implications for the emergence of pathogens. *PLoS Pathog* **13**: 1–30.

- RStudio Team (2020) RStudio: Integrated Development for R. RStudio. <http://www.rstudio.com/>.
- Sambrook, J. and Russell, D.W. (2001) Molecular Cloning: A Laboratory Manual. Third Edition.
- Schaefer, A.L., Taylor, T.A., Beatty, J.T., and Greenberg, E.P. (2002) Long-chain acyl-homoserine lactone quorum-sensing regulation of *Rhodobacter capsulatus* gene transfer agent production. *J Bacteriol* **184**: 6515–6521.
- Shakya, M., Soucy, S.M., and Zhaxybayeva, O. (2017) Insights into origin and evolution of α -proteobacterial gene transfer agents. *Virus Evol* **3**: 1–13.
- Sherlock, D. and Fogg, P.C.M. (2022a) Loss of the *Rhodobacter capsulatus* Serine Acetyl Transferase. *Appl Environ Microbiol* **88**: 1–18.
- Sherlock, D. and Fogg, P.C.M. (2022b) The archetypal gene transfer agent RcGTA is regulated via direct interaction with the enigmatic RNA polymerase omega subunit. *Cell Rep* **40**: 1–19.
- Solioz, M., Yen, H.-C.C., and Marrs, B. (1975) Release and uptake of gene transfer agent by *Rhodopseudomonas capsulata*. *J Bacteriol* **123**: 651–657.
- Song, D. and Loparo, J.J. (2015) Building bridges within the bacterial chromosome. *Trends Genet* **31**: 164–173.
- Tanizawa, Y., Fujisawa, T., and Nakamura, Y. (2018) DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* **34**: 1037–1039.
- Tomasch, J., Wang, H., Hall, A.T.K., Patzelt, D., Preusse, M., Petersen, J., et al. (2018) Packaging of *Dinoroseobacter shibae* DNA into gene transfer agent particles is not random. *Genome Biol Evol* **10**: 356–369.
- Trojanowski, D., Hołowka, J., and Zakrzewska-Czerwińska, J. (2018) Where and when bacterial chromosome replication starts: A single cell perspective. *Front Microbiol* **9**: 1–9.
- Ughy, B., Nagyapati, S., Lajko, D.B., Letoha, T., Prohaszka, A., Deeb, D., et al. (2023) Reconsidering Dogmas about the Growth of Bacterial Populations. *Cells* **12**: 1–15.
- Wall, J.D., Weaver, P.F., and Gest, H. (1975) Gene transfer agents, bacteriophages, and bacteriocins of *Rhodopseudomonas capsulata*. *Arch Microbiol* **105**: 217–224.
- Wang, H., Ziesche, L., Frank, O., Michael, V., Martin, M., Petersen, J., et al. (2014) The CtrA phosphorelay integrates differentiation and communication in the marine alphaproteobacterium *Dinoroseobacter shibae*. *BMC Genomics* **15**: 1–17.
- Westbye, A., O'Neill, Z., Schellenberg-Beaver, T., and Beatty, T.J. (2017) The *Rhodobacter capsulatus* gene transfer agent is induced by nutrient depletion and the RNAP omega subunit. *Microbiology* **163**: 1–9.
- Westbye, A.B., Kater, L., Wiesmann, C., Ding, H., Yip, C.K., and Beatty, J.T. (2018) The protease ClpXP and the PAS domain protein DivL regulate CtrA and gene transfer agent production in *Rhodobacter capsulatus*. *Appl Environ Microbiol* **84**: 1–17.
- Westbye, A.B., Leung, M.M., Florizone, S.M., Taylor, T.A., Johnson, J.A., Fogg, P.C., and Beatty, J.T. (2013) Phosphate concentration and the putative sensor kinase protein CckA modulate cell lysis and release of the *Rhodobacter capsulatus* gene transfer agent. *J Bacteriol* **195**: 5025–5040.

- Von Wintersdorff, C.J.H., Penders, J., Van Niekerk, J.M., Mills, N.D., Majumder, S., Van Alphen, L.B., et al. (2016) Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Front Microbiol* **7**: 1–11.
- Winterstein, C. and Ludwig, B. (1998) Genes coding for respiratory complexes map on all three chromosomes of the *Paracoccus denitrificans* genome. *Arch Microbiol* **169**: 275–281.
- Wishart, D.S., Han, S., Saha, S., Oler, E., Peters, H., Grant, J.R., et al. (2023) PHASTEST: Faster than PHASTER, better than PHAST. *Nucleic Acids Res* **51**: 443–450.
- Yen, H.C., Hu, N.T., and Marrs, B.L. (1979) Characterization of the gene transfer agent made by an overproducer mutant of *Rhodopseudomonas capsulata*. *J Mol Biol* **131**: 157–168.
- Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T.T.Y. (2017) Ggtree: an R Package for Visualization and Annotation of Phylogenetic Trees With Their Covariates and Other Associated Data. *Methods Ecol Evol* **8**: 28–36.
- Zeileis, A. and Grothendieck, G. (2005) Zoo: S3 infrastructure for regular and irregular time series. *J Stat Softw* **14**: 1–27.

6.9 Supplementary Material

Table S 6.1 Strains used in this study.

Strains		Relevant characteristics	Reference
<i>P. versutus</i>	DSM582	PvGTA producer	DSMZ
<i>P. versutus</i>	DSM582 rifR	PvGTA producer, spontaneous rifampicin-resistant mutant	This study
<i>P. versutus</i>	DSM582 ΔGTA	Mutant with deletion of the main GTA operon	This study
<i>P. versutus</i>	DSM582 ΔGTA rifR	Mutant with deletion of the main GTA operon, spontaneous rifampicin-resistant mutant	This study
<i>P. versutus</i>	DSM582 ΔRecep	Mutant with deletion of LOCUS_17550	This study
<i>P. versutus</i>	DSM582 ΔRecep rifR	Mutant with deletion of LOCUS_17550, spontaneous rifampicin-resistant mutant	This study
<i>P. versutus</i>	DSM582 pSEVA631- <i>tetAR</i>		This study
<i>P. versutus</i>	DSM582 pSEVA631-Pgta- <i>sfGFP</i>	RifS DSM582 strain with pSEVA631-Pgta- <i>sfGFP</i> plasmid; reporter strain for GTA main operon	This study
<i>P. versutus</i>	DSM582 pIND4	control for presence of pIND4	This study
<i>P. versutus</i>	DSM582 pIND4 rifR	control for presence of pIND4	This study
<i>P. versutus</i>	DSM582 pIND4 pSEVA631-Pgta- <i>sfGFP</i>	control for presence of pIND4	This study
<i>P. versutus</i>	DSM582 pIND4- <i>cpdB</i>		This study
<i>P. versutus</i>	DSM582 pIND4- <i>cpdB</i> rifR		This study
<i>P. versutus</i>	DSM582 pIND4- <i>cpdB</i> pSEVA631-Pgta- <i>sfGFP</i>		This study
<i>P. versutus</i>	DSM582 pIND4- <i>gafA</i>		This study
<i>P. versutus</i>	DSM582 pIND4- <i>gafA</i> rifR		This study
<i>P. versutus</i>	DSM582 pIND4- <i>gafA</i> pSEVA631-Pgta- <i>sfGFP</i>		This study
<i>P. versutus</i>	<i>cpdB</i> mutant	Spontaneous PvEP4A phage resistant strain DSM582 due to a mutation in <i>cpdB</i> , RifS	Chapter 5
<i>P. versutus</i>	<i>cpdB</i> mutant rifR	Spontaneous rifampicin-resistant mutant from the phage resistant strain	This study
<i>P. versutus</i>	<i>cpdB</i> mutant pIND4	control for presence of pIND4	Chapter 5
<i>P. versutus</i>	<i>cpdB</i> mutant pIND4 rifR	control for presence of pIND4	This study
<i>P. versutus</i>	<i>cpdB</i> mutant pIND4 pSEVA631-Pgta- <i>sfGFP</i>	control for presence of pIND4	This study
<i>P. versutus</i>	<i>cpdB</i> mutant pIND4- <i>cpdB</i>		Chapter 5
<i>P. versutus</i>	<i>cpdB</i> mutant pIND4- <i>cpdB</i> rifR		This study
<i>P. versutus</i>	<i>cpdB</i> mutant pIND4- <i>cpdB</i> pSEVA631-Pgta- <i>sfGFP</i>		This study
<i>P. versutus</i>	<i>cpdB</i> mutant pIND4- <i>gafA</i>		Chapter 5
<i>P. versutus</i>	<i>cpdB</i> mutant pIND4- <i>gafA</i> rifR		This study
<i>P. versutus</i>	<i>cpdB</i> mutant pIND4- <i>gafA</i> pSEVA631-Pgta- <i>sfGFP</i>		This study
<i>P. versutus</i>	P2	Wild-type phenotype	Chapter 5
<i>P. versutus</i>	P2 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>P. versutus</i>	P2 pIND4	control for presence of pIND4	This study
<i>P. versutus</i>	P2 pIND4 rifR	control for presence of pIND4	This study
<i>P. versutus</i>	P2 pIND4- <i>cpdB</i>		This study

<i>P. versutus</i>	P2 pIND4- <i>cpdB</i> rifR		This study
<i>P. versutus</i>	P2 pIND4- <i>gafA</i>		This study
<i>P. versutus</i>	P2 pIND4- <i>gafA</i> rifR		This study
<i>P. versutus</i>	P45	Wild-type phenotype	Chapter 5
<i>P. versutus</i>	P45 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>P. versutus</i>	P45 pIND4	control for presence of pIND4	This study
<i>P. versutus</i>	P45 pIND4 rifR	control for presence of pIND4	This study
<i>P. versutus</i>	P45 pIND4- <i>cpdB</i>		This study
<i>P. versutus</i>	P45 pIND4- <i>cpdB</i> rifR		This study
<i>P. versutus</i>	P45 pIND4- <i>gafA</i>		This study
<i>P. versutus</i>	P45 pIND4- <i>gafA</i> rifR		This study
<i>P. versutus</i>	P46	Wild-type phenotype	Chapter 5
<i>P. versutus</i>	P46 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>P. versutus</i>	P46 pIND4	control for presence of pIND4	This study
<i>P. versutus</i>	P46 pIND4 rifR	control for presence of pIND4	This study
<i>P. versutus</i>	P46 pIND4- <i>cpdB</i>		This study
<i>P. versutus</i>	P46 pIND4- <i>cpdB</i> rifR		This study
<i>P. versutus</i>	P46 pIND4- <i>gafA</i>		This study
<i>P. versutus</i>	P46 pIND4- <i>gafA</i> rifR		This study
<i>P. denitrificans</i>	DSM1407	Wild-type phenotype	DSMZ
<i>P. denitrificans</i>	DSM1407 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>P. pantotrophus</i>	DSM1404	Wild-type phenotype	DSMZ
<i>P. pantotrophus</i>	DSM1404 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>P. pantotrophus</i>	DSM65	Wild-type phenotype	DSMZ
<i>P. pantotrophus</i>	DSM65 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>P. aminophilus</i>	DSM8538	Wild-type phenotype	DSMZ
<i>P. aminophilus</i>	DSM8538 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>P. aestuarii</i>	DSM19484	Wild-type phenotype	DSMZ
<i>P. aestuarii</i>	DSM19484 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>P. alkenifer</i>	DSM11593	Wild-type phenotype	DSMZ
<i>P. alkenifer</i>	DSM11593 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>P. siganidrum</i>	DSM26381	Wild-type phenotype	DSMZ
<i>P. siganidrum</i>	DSM26381 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>R. capsulatus</i>	DSM1710	Wild-type phenotype	DSMZ
<i>R. capsulatus</i>	DSM1710 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>P. yeei</i>	KB 2	Wild-type phenotype	P. Junier
<i>P. yeei</i>	KB 2 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>P. yeei</i>	KB 3	Wild-type phenotype	P. Junier
<i>P. yeei</i>	KB 3 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>P. yeei</i>	KB 4	Wild-type phenotype	P. Junier
<i>P. yeei</i>	KB 4 rifR	Spontaneous rifampicin-resistant mutant	This study
<i>E. coli</i>	DH5 α - λ pir	Laboratory cloning strain	Lab strain

Table S 6.2 Plasmids used in this study.

Plasmids	Relevant characteristics	Reference
pEMG	Km ^R , oriR6K, MCS with two flanking I-SceI sites, used for markerless deletions	(Martínez-García and de Lorenzo, 2011)
pEMG-ΔGTA	Initial plasmid used for deletion of the main GTA operon	This study
pEMG-ΔGTA-CmR	pEMG-ΔGTA with the addition of a chloramphenicol resistance cassette	This study
pEMG-ΔGTA_bis	Plasmid used to remove the chloramphenicol resistance cassette after the removal of the main GTA operon	This study
pEMG-ΔRecep	Plasmid used for deletion of LOCUS_17550 in DSM582	This study
pME6000-IscI	Helper plasmid to remove pEMG chromosomal insertion	C. Terrettaz
pIND4	IPTG-inducible expression vector designed for use in alphaproteobacteria, carrying a kanamycin resistance marker	(Ind <i>et al.</i> , 2009)
pIND4-cpdB	pIND4 containing <i>cpdB</i> from DSM582	Chapter 5
pIND4-gafA	pIND4 containing <i>gafA</i> from DSM582	This study
pSEVA631-tetAR	pSEVA631 having tetAR cassette inserted at BglII site	C. Terrettaz
pSEVA631-Pgta-sfGFP	pSEVA631-tetAR with main GTA promotor and sfGFP gene	C. Terrettaz

Table S 6.3 Oligonucleotides used in this study

Oligonucleotides	Sequence (5'-3')	Purpose
ΔGTA		
GTA_UFR_F	AAA GAA TTC ATG GGC ATG GTC GAT CAT CA	Amplification of region upstream of the main GTA operon
GTA_UFR_R	CCA AGA GTT TCG GAA AGA CAT GAT CGA GCC GC	
GTA_DFR_F	ATC ATG TCT TTC CGA AAC TCT TGG TCA TTC TTG ATC G	Amplification of region downstream of the main GTA operon
GTA_DFR_R	AAA GGA TCC GTC ATG GCC TCG ATC ATC G	
pEMG-ΔGTA_F	CGA AAC TCT TGG TCA TTC TTG ATC G	Amplification of the pEMG-ΔGTA for insertion of chloramphenicol cassette
pEMG-ΔGTA_R	GAA AGA CAT GAT CGA GCC GC	
CmR_pEMG-ΔGTA_F	AAG AAT GAC CAA GAG TTT CGT TGC CAT GTT TTA CGG CAG T	Amplification of the chloramphenicol cassette for insertion in pEMG-ΔGTA
CmR_pEMG-ΔGTA_R	GCG GCT CGA TCA TGT CTT TCT TCA CTT ACA CCG CTT CTC AAC	
GTA_UFR_R_bis	ACC AAG AGT TTC G GCA AGG TGC GCA TCG TGC	Used with GTA_UFR_F for amplification of upstream region, to build the pEMG-ΔGTA_bis
GTA_DFR_F_bis	ATG CGC ACC TTG C C GAA ACT CTT GGT CAT TCT TGA TCG	Used with GTA_DFR_R for amplification of downstream region, to build the pEMG-ΔGTA_bis
ΔRecep		
GTAREcep_UFR_F	AAA GAA TTC GAC TCC GGG CAC GCA TTT CG	Amplification of region upstream of LOCUS_17550
GTAREcep_UFR_R	AAG CCT GTC AGC GCT ACG CTG TCC TTT CGC CAG GG	
GTAREcep_DFR_F	GAA AGG ACA GCG TAG CGC TGA CAG GCT TTC TCC	Amplification of region downstream of LOCUS_17550
GTAREcep_DFR_R	AAA GGA TCC CCA GCA CCA CCA CCT CGC	
pIND4-gafA		
gib_gafA_with_pIND4_F	GTG ATG AGA TCT ATG ACC ATG ACG ATT TTG ACG GGA GAT TTC GCG C	Amplification of gafA in DSM582 for gibbon assembly with pIND4
gib_gafA_with_pIND4_Rbis	ATG GGA GGA TCC TCA CCC GAT CAG CGG CGC	
gib_pIND4_with_gafA_Fbis	CTG ATC GGG TGA GGA TCC TCC CAT GGT TAA TTT C	Amplification of pIND4 for gibbon assembly with gafA
gib_pIND4_with_gafA_R	CGT CAT GGT CAT AGA TCT CAT CAC CAT CAC C-3	
Main GTA operon reporter		
tetAR_BglII_F	AAA GAT CTA TGT TGT CTA CAT GGC TCT GC	Amplification of tetracycline resistance to replace the gentamicin cassette of pSEVA631
tetAR_BglII_R	AAA GAT CTC CTA TCG TTT CCA CGA TCA GC	
gtaP_FP_UF	AAG AAT TCG CGA CTG ATC GGC CTT GC	Amplification of GTA promotor of DSM582 strain
gtaP_FP_UR	GCC TTT ACG CAT TCT TGA TCG CCC TGC CCT	
gtaP_FP_DF	GCG ATC AAG AAT GCG TAA AGG CGA AGA GC	Amplification of sfGFP
gtaP_FP_DR	CCA AGC TTT TAT TTG TAC AGT TCA TCC ATA CC	
add_RBS_sfGFP_R	AGG ATC CTC TTG ATC GCC CTG CCC TC	Modification of sfGFP RBS
add_RBS_sfGFP_F	AGG ATC CTA AGG AGG TAA AAA AAA AAT GCG TAA AGG CGA AGA GC	

Table S 6.4 Sequence identity between *P. versutus* DSM582 with the other strains/species for the *rpoB* gene.

Species	Strain	% Identity	Nr. Mismatch	Nr. Gap
<i>Paracoccus versutus</i>	DSM582	100.0	0	0
<i>Paracoccus versutus</i>	P2	99.4	23	0
<i>Paracoccus versutus</i>	P46	99.4	26	0
<i>Paracoccus versutus</i>	P45	99.4	26	0
<i>Paracoccus denitrificans</i>	DSM104981	95.0	198	4
<i>Paracoccus pantotrophus</i>	DSM2944	91.9	321	13
<i>Paracoccus yeei</i>	FDAARGOS643	91.9	321	13
<i>Paracoccus siganidrum</i>	DSM26381	89.8	400	20
<i>Paracoccus aestuarii</i>	DSM19484	88.7	442	25
<i>Paracoccus alkenifer</i>	DSM11593	87.9	453	41
<i>Paracoccus aminophilus</i>	DSM8538	86.9	527	15

Table S 6.5 Genomic characteristics of complete *Paracoccus* genome and one *R. capsulatus* genome used in the phylogenetic analysis (Figure S 6.1). Rep: total number of replicons; large rep: number of replicons larger than 1 Mb; NA: data not available; Bold text: Genomes containing two replicons larger than 1Mb; Shading in “Genome size” column: Darker shades indicate larger genomes. Genomes are presented in the same order as in the phylogenetic tree (Figure S 6.1).

Species	Strain	Genome ID	Rep.	Large rep.	Genome size [bp]	Size largest rep. [bp]	Size 2 nd largest rep. [bp]
<i>P. versutus</i>	DSM582	NA	4	2	5 725 528	2 799 769	1 904 959
<i>P. versutus</i>	UW1	GCF_029203935.1	4	2	5 502 442	2 768 327	1 740 310
<i>P. versutus</i>	DSM17099	GCF_029909795.1	2	2	4 969 440	2 731 001	2 238 439
<i>P. pantotrophus</i>	DSM 11072	GCF_029910495.1	4	2	4 259 988	2 219 897	1 499 165
<i>P. pantotrophus</i>	ACCC10489	GCF_013414285.1	4	2	4 626 003	2 391 084	1 604 287
<i>P. pantotrophus</i>	DSM2944	GCF_008824185.1	4	2	4 409 379	2 248 278	1 525 820
<i>P. denitrificans</i>	HAMBI_2443	GCF_034627565.1	3	2	5 237 024	2 853 118	1 730 091
<i>P. denitrificans</i>	PD1222	GCF_000203895.1	3	2	5 236 194	2 852 282	1 730 097
<i>P. denitrificans</i>	DYTN-1	GCF_023238405.1	3	2	5 237 032	2 853 092	1 730 123
<i>P. denitrificans</i>	ATCC19367	GCF_004063735.1	3	2	5 242 327	2 858 426	1 730 087
<i>P. denitrificans</i>	XW11	GCF_040409775.1	3	2	5 078 518	3 025 368	1 537 841
<i>P. sp.</i>	TOH	GCF_030388245.1	5	2	4 506 955	2 149 639	1 315 351
<i>P. sp.</i>	MA	GCF_020990385.1	3	2	4 473 736	2 389 614	1 412 597
<i>P. kondratievae</i>	BJQ0001	GCF_009301415.1	3	2	4 228 383	2 255 404	1 513 387
<i>P. methylovorus</i>	H4-D09	GCF_016919705.1	5	2	4 381 421	2 171 951	1 499 640
<i>P. sp.</i>	SMMA_5_TC	GCF_009696685.2	3	1	3 122 888	2 655 856	325 864
<i>P. sp.</i>	SMMA_5	GCF_009696665.2	3	1	3 123 001	2 655 986	325 864
<i>P. aminovorans</i>	JCM7685	GCF_900005615.1	4	1	4 038 130	3 108 335	740 381
<i>P. aminovorans</i>	BM109	GCF_033555915.1	4	2	4 297 217	2 932 317	1 155 326
<i>P. yeei</i>	CCUG 32053	GCF_003612915.1	9	1	4 668 852	3 421 679	482 273
<i>P. yeei</i>	FDAARGOS_643	GCF_008693865.1	5	1	4 689 804	3 637 766	363 819
<i>P. yeei</i>	FDAARGOS_252	GCF_002073635.2	8	1	4 829 607	3 622 127	295 841
<i>P. yeei</i>	TT13	GCF_002749495.1	7	1	4 620 225	3 588 087	327 671
<i>P. aminophilus</i>	JCM7686	GCF_000444995.1	6	1	4 873 118	3 613 807	438 126
<i>P. contaminans</i>	RKI 16-01929T=LMG 29738T=CCM 8701T=CIP 111112T	GCF_002105555.1	2	1	3 033 494	2 939 512	93 982
<i>P. jeotgali</i>	CBA4604	GCF_002865605.1	3	1	3 299 166	3 121 136	149 225

<i>P. mutanolyticus</i>	RSP-02	GCF_003285265.1	1	1	3 592 357	3 592 357	NA
<i>P. sanguinis</i>	OM2164	GCF_012689545.1	4	1	3 373 772	3 204 673	119 507
<i>P. sp.</i>	MC1862	GCF_016617715.1	7	1	3 633 757	3 307 010	113 686
<i>P. suum</i>	SC2-6	GCF_003324675.1	1	1	3 250 998	3 250 998	NA
<i>P. marcusii</i>	3501	GCF_041879345.1	11	1	3 999 597	3 059 608	424 132
<i>P. sp.</i>	Arc7-R13	GCF_004010775.1	9	1	4 040 012	3 231 929	199 685
<i>P. marcusii</i>	MBLB0836	GCF_028621715.1	1	1	3 175 811	3 175 811	NA
<i>P. marcusii</i>	CP157	GCF_019141545.1	23	1	4 077 279	3 065 501	254 393
<i>P. marcusii</i>	KGP	GCF_009493695.2	5	1	3 936 521	3 114 772	328 249
<i>P. marcusii</i>	HNIBRBA1908	GCF_039905895.1	6	1	3 914 663	3 078 589	368 501
<i>P. liaowanqingii</i>	2251	GCF_004683865.2	11	1	4 925 228	3 271 357	310 431
<i>P. aestuarii</i>	DSM19484	GCF_028553885.1	10	1	3 675 213	3 089 821	296 221
<i>P. stylophorae</i>	LMG25392	GCF_028553765.1	2	1	3 642 172	3 634 760	7 412
<i>P. seriniphilus</i>	DSM14827	GCF_028553745.1	5	1	4 261 722	2 779 753	599 725
<i>P. saliphilus</i>	DSM18447	GCF_028553805.1	2	1	4 621 331	4 378 448	242 883
<i>P. sp.</i>	M09	GCF_030863885.1	1	1	4 566 970	4 566 970	NA
<i>P. everestensis</i>	S8-55	GCF_021491915.1	1	1	2 957 008	2 957 008	NA
<i>P. sp.</i>	AK26	GCF_011801365.1	5	1	3 663 216	2 787 975	471 727
<i>P. sp.</i>	S3-43	GCF_029027965.1	1	1	3 221 824	3 221 824	NA
<i>P. alcaliphilus</i>	DSM 8512	GCF_028553725.1	5	1	4 772 712	3 557 425	429 813
<i>P. fistulariae</i>	KCTC22803	GCF_028553785.1	4	1	3 723 810	3 621 100	90 204
<i>P. zhejiangensis</i>	J6	GCF_002847445.1	5	1	4 637 247	3 985 112	267 361
<i>P. tegillarcae</i>	BM15	GCF_002847305.1	4	1	3 954 736	3 759 886	92 522
<i>P. aerodenitrificans</i>	SCSIO 75817	GCF_027913215.1	5	1	3 558 095	3 318 305	198 273
<i>P. sp.</i>	SCSIO 75233	GCF_027912675.1	11	1	3 892 551	3 267 188	220 952
<i>P. sediminicola</i>	SCSIO 76264	GCF_027912835.1	7	1	3 534 024	3 175 232	146 119
<i>P. albus</i>	SCSIO 80058	GCF_027913035.1	5	1	3 761 358	3 318 811	210 876
<i>R. capsulatus</i>	SB1003	GCF_000021865.1	2	1	3 871 920	3 738 958	132 962

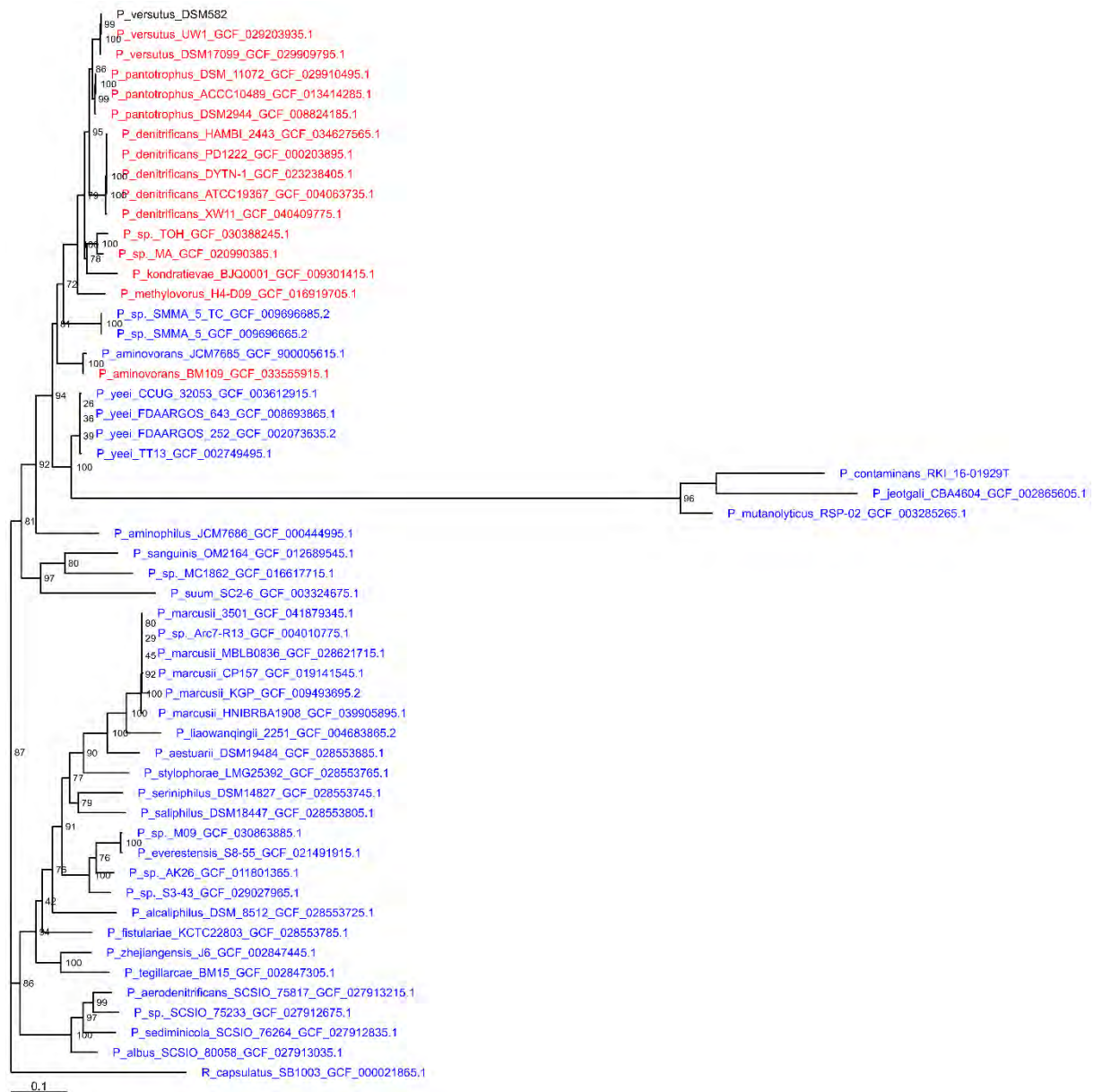


Figure S 6.1 Maximum likelihood phylogenetic tree of *Paracoccus* species based on concatenated GyrA and GyrB protein sequences. The tree was constructed using all complete *Paracoccus* genomes available in the NCBI database as of October 8, 2024. Bootstrap values (%) are shown on each branch, and the scale bar represents the number of substitutions per site. *R. capsulatus* was used as an outgroup to root the tree. The strain of interest, *P. versutus* DSM582, is highlighted in black. Red labels indicate strains with two replicons larger than 1 Mb, while blue labels represent genomes with only one replicon larger than 1 Mb.

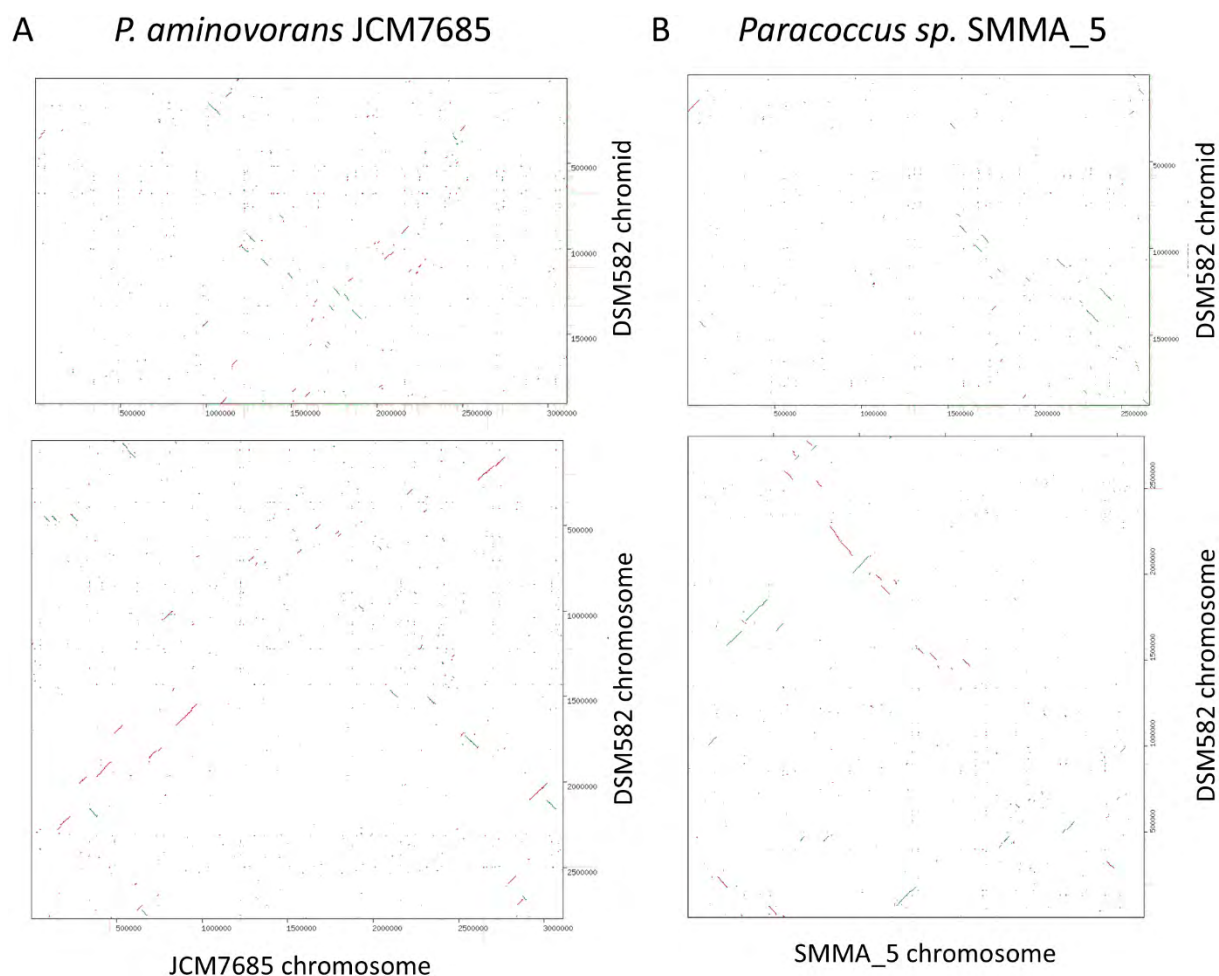


Figure S 6.2 Dot plots comparing the two largest replicons of *P. versutus* DSM582 with those of (A) *P. aminovorans* JCM7685 and (B) *Paracoccus* sp. SMMA_5. Plots were generated using the YASS online tool with an e-value threshold of 1×10^{-30} and default settings for other parameters (Noé and Kuchero, 2005). Green diagonal lines represent regions of sequence similarity in the same orientation, while red diagonal lines indicate inverted regions or alignments between the reverse complement of one sequence and the forward sequence of the other.

Table S 6.6 *P. versutus* (Pv) loci homologous to *R. capsulatus* (Rc) loci involved in GTA-mediated HGT and their functional annotations in *R. capsulatus*. More detailed information about the role of each gene from the main GTA operon are available in Chapter 1. Abbreviations: Rep.: Replicons; % ID; percentage identity of *P. versutus* locus relative to *R. capsulatus* locus.

Rep.	Pv locus	Start	End	Rc locus	% ID	Description	Category
1	LOCUS_01720	174044	175249	rcc01081	31	Glycosyl transferase, group 1	Polysaccharide biosynthesis
	LOCUS_07010	684277	685464	rcc01865	53	GafA transcriptional regulator	Regulation
	LOCUS_07020	685678	686010	rcc01866	49		
	LOCUS_12550	1258879	1259796	rcc01079	41	GhsA head spike protein	Adsorption
	LOCUS_17550	1744948	1746030	rcc01932	29	Glycosyl transferase, family 4	Polysaccharide biosynthesis
	LOCUS_19850	1970778	1971503	rcc00197	53	ComF	DNA translocation
2	LOCUS_29200	150597	150956	rcc01682	49	Small terminase	Main GTA operon
	LOCUS_29210	150838	152247	rcc01683	80	Large terminase	
	LOCUS_29220	152396	153595	rcc01684	77	Portal	
	LOCUS_29230	153597	153830	rcc01685	76	CDS	
	LOCUS_29240	153832	154368	rcc01686	70	Prohead Protease	
	LOCUS_29250	154392	155558	rcc01687	61	Major Capsid protein	
	LOCUS_29260	155686	156309	rcc01688	60	Adaptator	
	LOCUS_29270	156306	156656	rcc01689	53	Stopper	
	LOCUS_29280	156653	157072	rcc01690	61	Tail terminator	
	LOCUS_29290	157085	157498	rcc01691	77	Tail tube	
	LOCUS_29300	157500	157817	rcc01692	63	Tail chaperone 1	
	LOCUS_29310	157814	158023	rcc01693	65	Tail chaperone 3	
	LOCUS_29320	158048	158710	rcc01694	62	Tape measure	
	LOCUS_29330	158714	159346	rcc01695	74	Distal tail	
	LOCUS_29340	159343	160161	rcc01696	60	Hub	
	LOCUS_29350	160158	160580	rcc01697	64	Peptidase	
	LOCUS_29360	160596	164483	rcc01698	60	Megatron	
	LOCUS_29370	164476	165198	rcc00171	41	Tail fiber	
	LOCUS_40350	1260677	1262299	rcc01080	26	GhsB head spike protein	Adsorption
	LOCUS_43210	1554389	1555903	rcc00460	69	ComM	Unkown function
	LOCUS_43380	1572127	1572660	rcc00556	55	Holin family protein	Release
	LOCUS_43390	1572651	1573262	rcc00555	65	Holin-associated N-acetylmuramidase	
	LOCUS_45600	1818065	1820245	rcc02362	51	ComEC	DNA translocation
3	LOCUS_54220	830475	832811	rcc00280		Calcium binding, extracellular, RTX family protein	Repressor
					49		

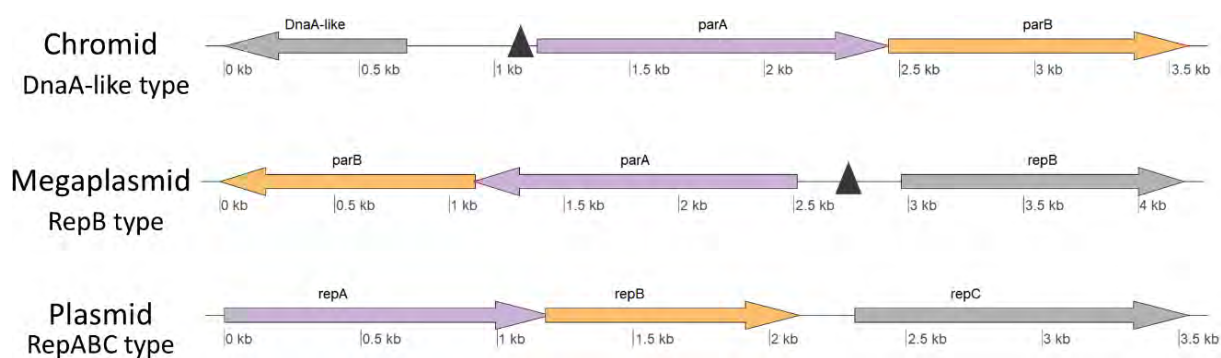


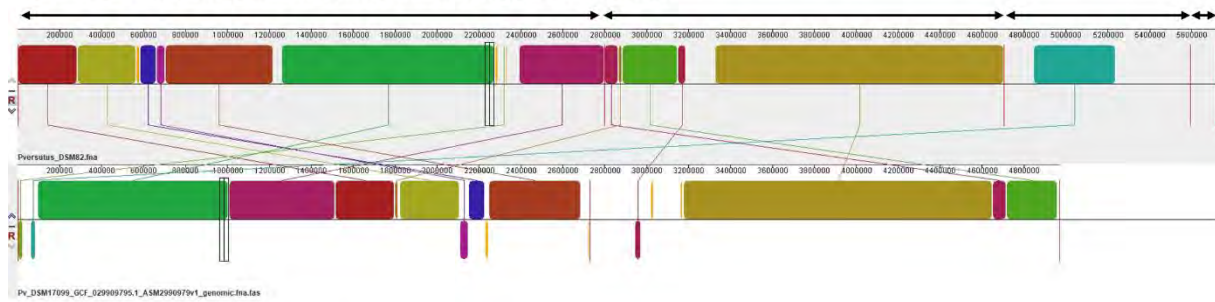
Figure S 6.3 Genetic organization of replication/partitioning modules from the three *P. versutus* DSM582 extrachromosomal replicons. Partitioning genes, *parA* and *parB*, are shown in colors, replicases in grey. The *parAB* genes of the repABC operon are referred to as *repAB* (Petersen, 2011). Black triangles indicate the putative origin of replication located within AT- rich regions of the extrachromosomal replication modules as predicted more precisely by Ori-Finder (Gao and Zhang, 2008). The corresponding loci can be found in Table S 6.7.

Table S 6.7 Loci coding for the replication/partitioning modules from the three *P. versutus* DSM582 extrachromosomal replicons

Replicon	Locus	5'	3'	Strand	Name	Product
Chromid	LOCUS_29080	136'381	137'055	-	DnaA-like*	hypothetical protein
	LOCUS_29090	137'538	138'842	+	parA	Chrom. partitioning protein ParA
	LOCUS_29100	138'839	139'948	+	parB	replication protein ParB
Megaplasmid	LOCUS_50010	390'710	391'822	-	parB	Chrom. partitioning protein ParB
	LOCUS_50020	391'815	393'224	-	parA	plasmid partitioning ATPase ParA
	LOCUS_50030	393'678	394'925	+	repB	replication protein RepB
Plasmid	LOCUS_55110	33'674	34'897	-	repC	replication initiation protein
	LOCUS_55120	35'098	36'030	-	repB	plasmid partitioning protein RepB
	LOCUS_55130	36'014	37'207	-	repA	plasmid partitioning protein RepA

* This locus was named DnaA-like based on the annotation for the replication type given for the second largest replicon of *Paracoccus denitrificans* PD1222 (Petersen, 2011). Moreover, InterProScan tool identified one domain as DnaA N-terminal domain and the other domain as belonging to the Winged helix-like DNA-binding domain superfamily (Quevillon et al., 2005).

A DSM582 versus DSM17099



B DSM582 versus UW1

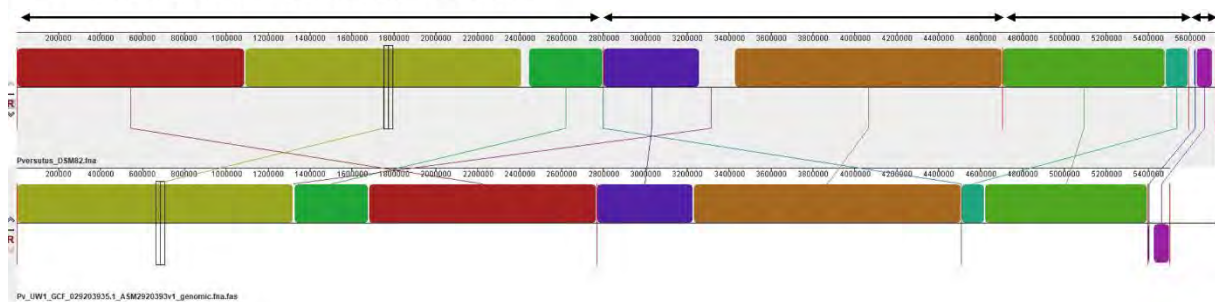
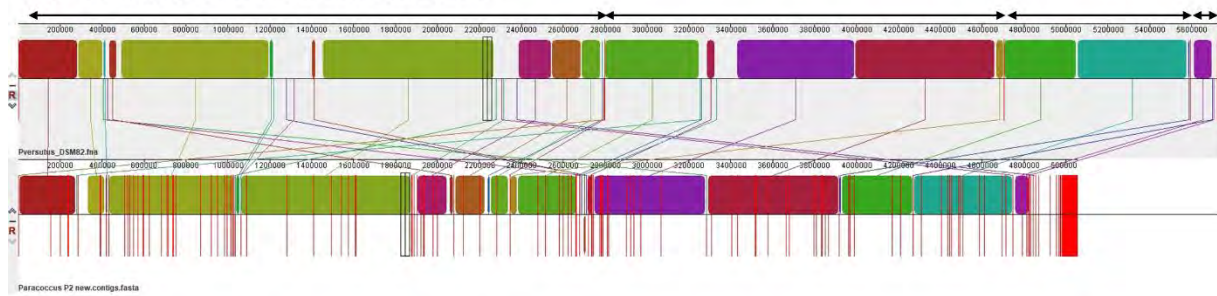
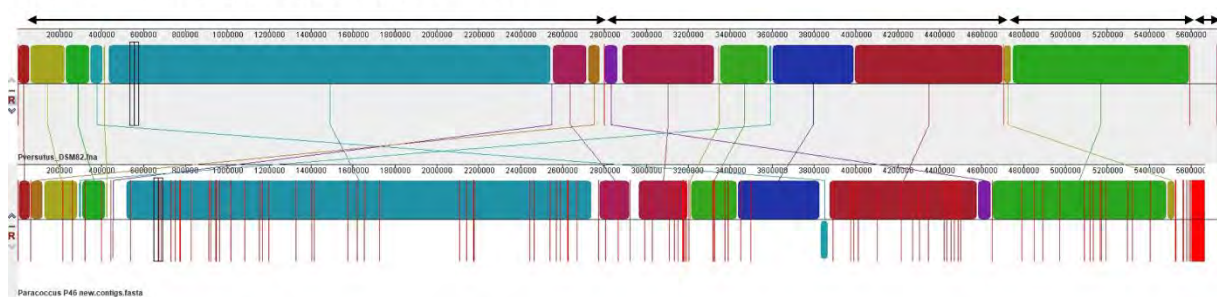


Figure S 6.4 Mauve alignment of *P. versutus* DSM582 (top) against complete genomes of *P. versutus* (bottom) (A) DSM17099 (GCF_029909795.1) and (B) UW1 (GCF_029203935.1). Replicons of *P. versutus* DSM582 are highlighted with arrows with biggest replicon on the left side. The Mauve software divides the genome into several locally collinear blocks, with homologous regions of the genome colored identically and connected by lines. Blocks below the genome's center line are inverted relative to the reference sequence.

A DSM582 versus P2



B DSM582 versus P45



C DSM582 versus P46

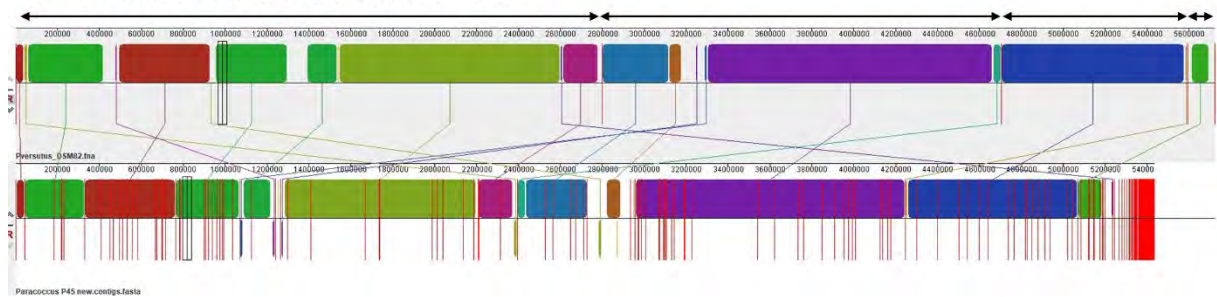


Figure S 6.5 Mauve alignment of *P. versutus* DSM582 (top) against three other *P. versutus* genomes from environmental strains (bottom). Genomes from the environmental strains were composed of many contigs separated by vertical red lines. Replicons of *P. versutus* DSM582 are highlighted with arrows with biggest replicon on the left side. The Mauve software divides the genome into several locally collinear blocks, with homologous regions of the genome colored identically and connected by lines. Blocks below the genome's center line are inverted relative to the reference sequence.

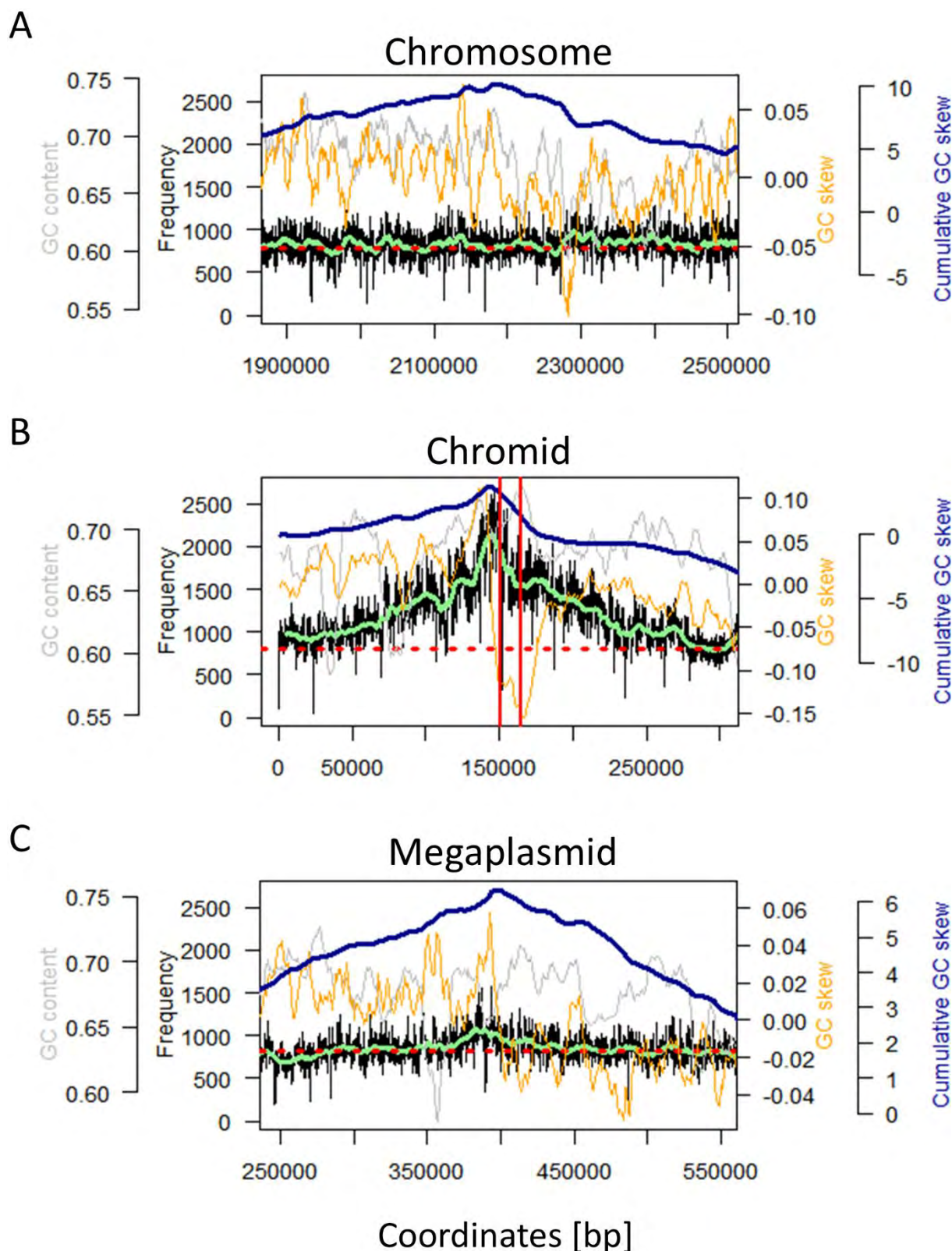


Figure S 6.6. Gene fragments enrichment are located very close to the origins of replication. Only a subset of the replicon is represented with DNA coverage (black), local mean coverage (green), median coverage for the whole chromosome (red) GC content (grey), GC skew (orange), and cumulative GC skew (dark blue). We calculated the mean coverage, GC content, GC skew and the cumulative GC skew with sliding windows of 10kb, every 1kb. The maximum of the cumulative GC skew line shows the predicted position of the origin of replication. The vertical red lines on chromid delimit the positions of the main GTA operon.

Table S 6.8 *P. versutus* has multiple predicted prophages regions in its genome, based on PHASTEST web tool (Wishart *et al.*, 2023).

Replicon	Prophage region	Positions
Chromosome	1	449'923-472'422
Chromid	3	66'276-83'585 150'838-166'100 (GTA main operon) 1'302'693-1'329'507
Megaplasmid	0	-
Plasmid	0	-

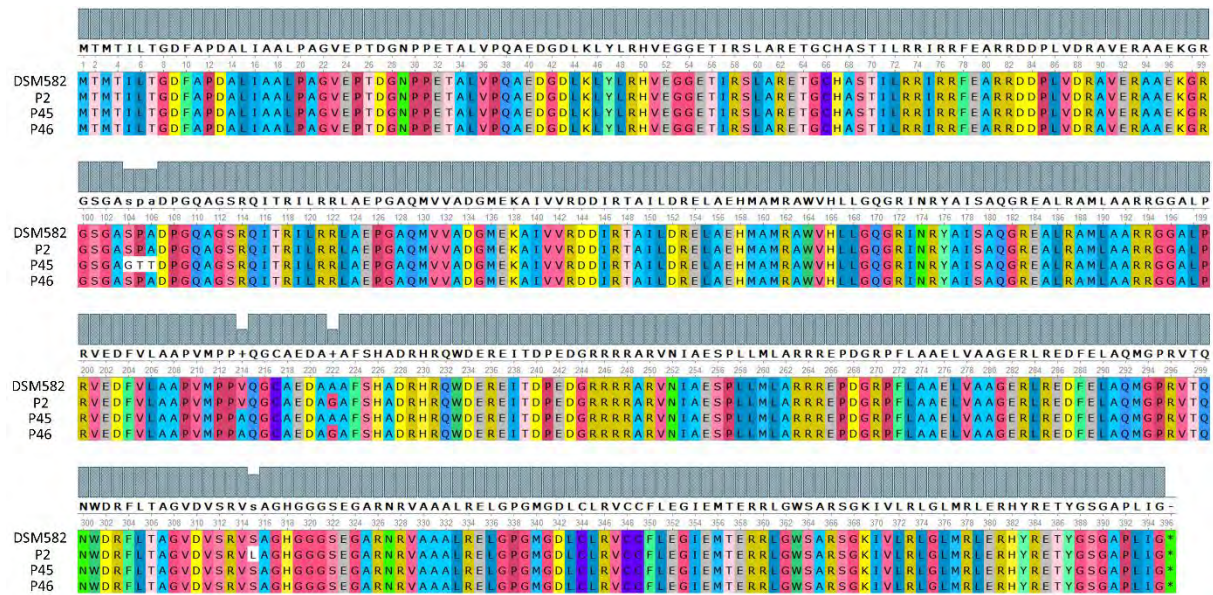


Figure S 6.7. Alignment of GafA from several strains of *P. versutus* (DSM582, P2, P45, and P46) highlights few missense mutations.

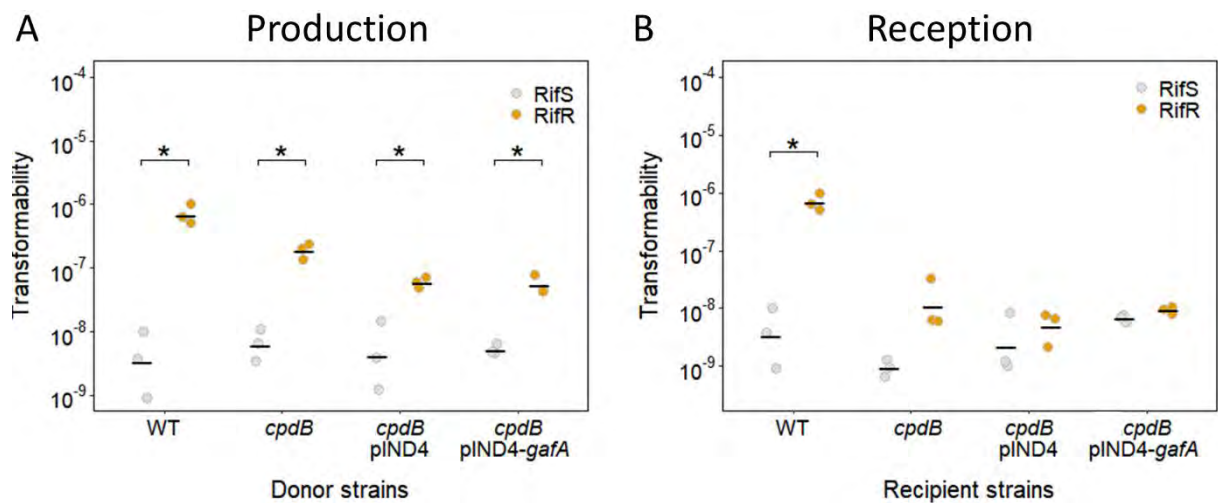


Figure S 6.8 Complementation of the *cpdB* mutant with *gafA* did not restore (A) GTA production and (B) reception to levels like the wild-type strain. Dots show independent replicates and bars the means.

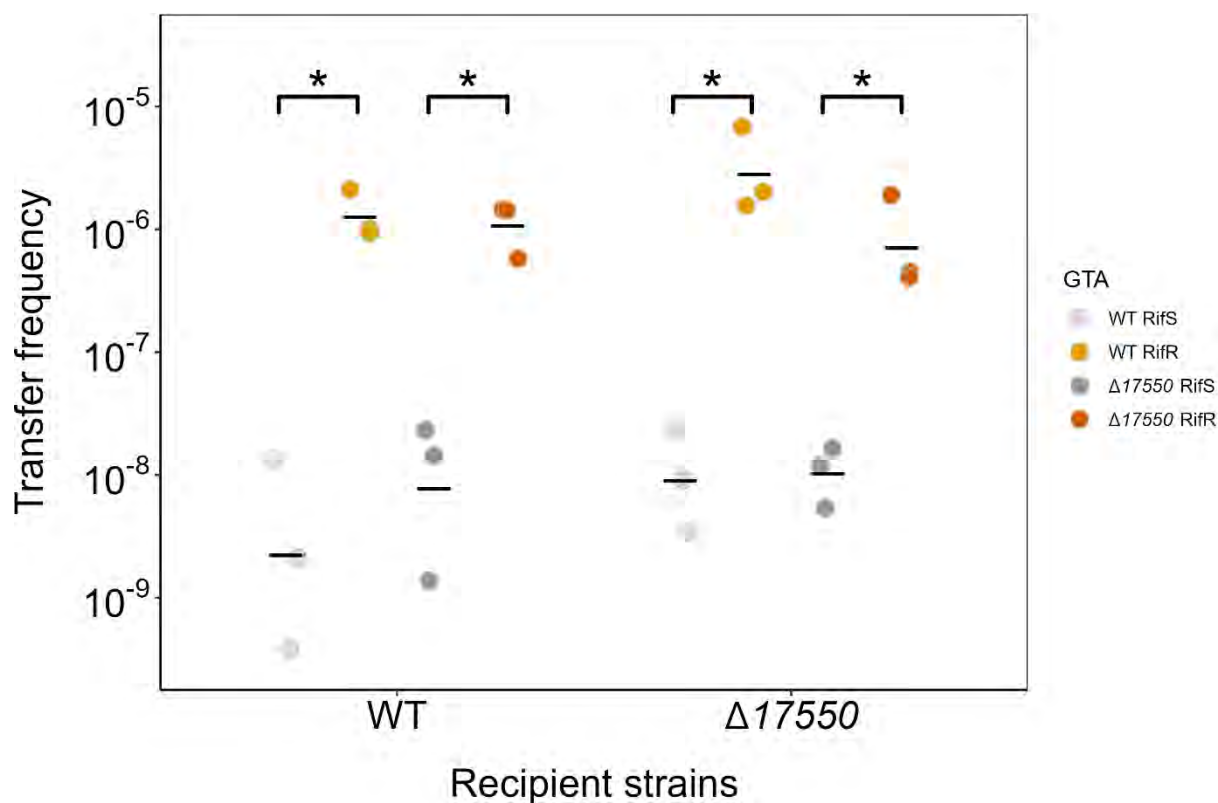


Figure S 6.9 Deletion of *LOCUS_17550*, a homologue of *rcc01932*, did not impair the ability of the strain to produce and receive GTA, as it happens in *R. capsulatus* for reception (Brimacombe *et al.*, 2013). GTA solutions were obtained from *P. versutus* DSM582 wild-type (WT) and $\Delta 17550$. Both strains were either rifampicin-susceptible (RifS, grey), or rifampicin-resistant (RifR, orange). Dots show independent replicates and bars the means. The experiment was performed only one time.

Table S 6.9 Loci predicted by CapsuleFinder (Rendueles *et al.*, 2017) to be involved in capsule biogenesis pathways and with transcript levels higher than 2-fold between the wild-type and *cpdB* mutant of *P. versutus*. A more complete table containing all significantly different loci is presented in Chapter 5 (table S15).

Locus	L2FC	padj	Product	Predicted system	Gene status
LOCUS_01740	-1.04	7.5E-04	sugar transferase	ABC	accessory
LOCUS_02710	-1.93	7.7E-11	dTDP-4-dehydrorhamnose 3,5-epimerase	ABC	accessory
LOCUS_07840	1.08	6.7E-05	Bcr/CflA family drug resistance efflux transporter	GroupIV_s_stricte	mandatory
LOCUS_15850	1.07	2.5E-04	ribonucleotide-diphosphate reductase subunit alpha	ABC	mandatory
LOCUS_15920	-4.02	3.6E-79	glycosyl transferase	Syn_cps3	mandatory
LOCUS_22440	1.38	1.9E-04	2-phenylethanol ABC transporter ATP-binding protein	ABC	mandatory
LOCUS_22450	1.51	4.0E-06	transport permease protein	ABC	mandatory

Abbreviations: padj: DESeq2 p-value adjusted for multiple testing ; L2FC, logarithm in base 2 of the expression ratio of the wild type over the *cpdB* mutant (a negative value means that the expression is reduced in the *cpdB* mutant).

Chapter 7 - Synthesis, general discussion, perspectives and conclusion

7.1 Research background

Horizontal gene transfer is a major driver of evolution and adaptation in prokaryotes. It occurs through several mechanisms, with the canonical ones being conjugation, natural transformation, and transduction. However, DNA can also be transferred via less common systems, including GTAs (Solioz *et al.*, 1975), DNA-containing membrane vesicles (Rumbo *et al.*, 2011) and nanotubes between cells (Dubey and Ben-yehuda, 2011). These understudied systems may play a significant role in gene transfer, such as GTAs in the open ocean environment, which might lead to transfer frequencies much higher than those of transformation and transduction (McDaniel *et al.*, 2010). In addition to introducing genetic diversity, the transferred genetic material is hypothesized, particularly in the context of natural transformation, to offer benefits by serving as a template for DNA repair and as a source of nutrients.

7.2 Synthesis

During my doctoral thesis, I explored the potential benefits and the regulation of natural transformation in *P. stutzeri* and HGT mediated by GTAs in *P. versutus*.

7.2.1 Part I – Natural transformation of *P. stutzeri*

In the first part of this work, I aimed to investigate the benefits of natural competence for nutrient acquisition. While *P. stutzeri* was able to use DNA as a source of carbon and nitrogen, this ability was not primarily dependent on the competence apparatus. Given the absence of known alternative mechanisms for importing long stretches of DNA in *P. stutzeri*, I hypothesized that DNA degradation should occur extracellularly. Bioinformatic screening allowed me to identify three potential extracellular nucleases, of which I chose to experimentally investigate in priority EndA, given its homology to a nuclease with proven activity in various species such as *S. oneidensis* (Heun *et al.*, 2012) and *V. cholerae* (Seper *et*

al., 2011). However, deletion of the *endA* gene did not significantly impact the ability of *P. stutzeri* to use DNA as a source of nutrients, suggesting that the two other putative nucleases warrant further investigation. Overall, our study did not conclusively demonstrate that the competence apparatus was necessary for DNA-based nutrition in *P. stutzeri*.

In the second part, I explored three aspects of competence regulation in *P. stutzeri*: competence induction factors, single-cell expression of competence genes, and the role of the *tfoX* gene in transformability. High concentrations of extracellular DNA were found to downregulate the expression of competence genes *comA* and *pilA*, while mild ultraviolet radiation upregulated their expression. At the single-cell level, most bacterial cells expressed competence genes during the transition to stationary phase, with *pilA* exhibiting more heterogeneous expression compared to *comA*. Deletion of two *tfoX* homologues in *P. stutzeri* did not significantly affect its transformability. Collectively, these findings highlight the distinct characteristics of *P. stutzeri* compared to well-studied Gram-negative models and emphasize the importance of investigating additional models to expand our understanding of bacterial competence regulation across diverse species.

The primary research goal of Chapter 4 was to identify genes involved in DNA uptake and cell-cell interactions during transformation in *P. stutzeri* on a large scale. Our observations revealed that transformation rates using donor cells are comparable to those achieved with saturating concentrations of purified DNA. This suggests that donor cells may play a significant role in natural transformation, through active DNA secretion or the presence of specific proteins that enhance DNA binding and transport into recipient cells. Given the multitude of potential hypotheses, a high-throughput approach was necessary to refine the investigation and expand the repertoire of genes involved in natural transformation in *P. stutzeri*. Consequently, I adapted a transposon insertion sequencing (TIS) method to conduct a comprehensive genetic study of natural transformation in this species, examining the roles played by the recipient and the donor cells.

However, the study faced significant technical challenges. The TIS library preparation proved suboptimal, resulting in insufficient sequencing depth. Overamplification of some genomic regions led to disproportionate read mapping, which compromised the ability to draw meaningful conclusions about gene functions in natural transformation. Nevertheless, the datasets allowed to identify essential genes for *P. stutzeri* growth under standard conditions.

Despite the experimental setbacks, the research highlights the complexity of studying natural transformation and the need for methodological refinement in bacterial genetic studies. The study underscores the potential of TIS in microbial genetics while demonstrating the challenges of applying this technique to specific bacterial species like *P. stutzeri*. The experience gained may inform future research strategies and protocol optimizations for similar investigations.

7.2.2 Part II – Phage and Phage-like elements of *Paracoccus versutus*

Chapter 5 presents a new bacteriophage, PvEP4A, infecting *P. versutus* DSM582 and its characterization. PvEP4A was identified as a dsDNA, tailed bacteriophage with a narrow host range, infecting only its original *P. versutus* host. Interestingly, it is more closely genetically related to *R. capsulatus* phages than to known phages infecting *Paracoccus*. The research revealed that the host rapidly developed resistance through a deletion in the *cpdB* gene, which increased intracellular c-di-GMP concentration. This genetic modification led to decreased expression of flagella, capsular polysaccharides, and outer membrane receptors, potentially explaining the phage-resistance phenotype. The research expanded the number of known phages infecting *Paracoccus* species and revealed complex phage-host interactions, demonstrating how bacteria can rapidly develop resistance mechanisms that alter multiple aspects of cellular physiology.

Chapter 6 introduces *P. versutus* DSM582 as a new model organism for investigating GTA, expanding our understanding of HGT mechanisms beyond the well-established *R. capsulatus* model. The GTA system of *P. versutus* DSM582 demonstrates unique features that distinguish this species from previous GTA models. Notably, DNA packaging into GTAs exhibited a strong bias near the origin of replication on the second replicon and a weaker one on the third replicon, with no observable bias on the largest replicon. This suggests a potential link between genome replication and GTA production in *P. versutus*. The research also revealed significant differences in GTA expression and regulation between *P. versutus* and *R. capsulatus*. While *R. capsulatus* typically induces GTA production during the stationary growth phase, *P. versutus* maximally expresses its main GTA operon during the exponential growth phase. This fundamental difference implies that the evolutionary advantages of GTAs may vary across bacterial species. Intriguingly, the study uncovered unexpected regulatory dynamics. Overexpression of a homologue of *gafA*, a master regulator of RcGTA production,

did not increase transfer rates or GTA expression. Moreover, high intracellular cyclic di-GMP concentrations unexpectedly decreased *P. versutus* ability to perform HGT. The introduction of *P. versutus* as an alternative GTA model organism provides an opportunity to broaden scientific understanding of these fascinating DNA transfer vehicles. By exploring the unique characteristics of GTAs across diverse bacterial species, researchers could gain deeper insights into bacterial evolution, adaptation, and genetic exchange mechanisms.

7.3 Discussion

7.3.1 Alternative benefits of HGT mechanisms?

HGT mechanisms are often studied for their role in the acquisition of new genes. While the acquired genes might be beneficial in themselves, HGT also promotes new combinations of genetic material across individuals in a population, thereby increasing genetic diversity for selection to act upon. This potential for allele recombination would, for instance, enable bacteria to break linkage between beneficial and deleterious mutations, rescue genotypes from Muller's ratchet, and explore novel adaptive combinations (Vos, 2009). However, HGT mechanisms might provide additional benefits to the host and population, particularly in the contexts of natural transformation and GTA-mediated transfer, both of which are host-controlled processes.

7.3.1.1 Natural transformation: alternative benefits

Natural competence is traditionally viewed as advantageous for nutrient acquisition, genome repair, and evolutionary adaptation (Chapter 1) (Nielsen *et al.*, 2007; Seitz and Blokesch, 2012). Some researchers suggests additional benefits, proposing that natural transformation may enable bacteria to eliminate costly integrated mobile genetic elements such as prophages or conjugative elements and so participate in genome maintenance (Ambur *et al.*, 2016).

In this thesis, we specifically investigated the potential role of natural competence in DNA-based nutrition using *P. stutzeri* as a model system (Chapter 2). Our results demonstrated that, under the tested conditions, disruption of two key competence genes (*pilA* and *comA*) did not impair growth when DNA was provided as the sole carbon or nitrogen source, contrasting with previous findings in *E. coli* (Finkel and Kolter, 2001; Palchevskiy and Finkel, 2006). This indicates that the competence apparatus may not be essential or have a subtle contribution for DNA-based nutrition in *P. stutzeri*.

Emerging evidence suggests that the competent state itself may confer benefits unrelated to DNA uptake. For instance, in *S. aureus*, host-induced respiratory stress triggers competence development, which enhances glucose uptake and glycolysis, thereby supporting bacterial proliferation during infection (Cordero *et al.*, 2022). Similarly, the growth arrest associated with competence in *B. subtilis* may be advantageous when the population encounters conditions lethal to dividing bacteria (Haijema *et al.*, 2001; Johnsen *et al.*, 2009; Yüksel *et al.*, 2016). Furthermore, Type IV pili, which are important for DNA uptake during natural transformation, have well-documented roles in surface adhesion, twitching motility, and biofilm formation (Giltner *et al.*, 2012). Given these multifunctional properties, it would be intriguing to explore whether T4P contributes to additional functions in competent cells apart from facilitating DNA internalization. Taken together, competence seems to provide the bacterium with advantages extending beyond DNA internalization.

7.3.1.2 Gene Transfer Agents: alternative benefits

The benefits of GTA production and the acquisition of GTA-packaged DNA remain incompletely understood. As a mechanism of HGT, GTAs enable the acquisition of new genetic material. However, given that the DNA fragments packaged in GTAs are often limited to ~4-14 kb in size (Table 1.2), GTAs are less likely to mediate the transfer of complete operons. However, they might be well suited to facilitating the exchange of allelic variants, which may fine-tune adaptive traits. Such allele shuffling may be especially relevant for traits under strong diversifying selection. In *Bartonella*, for instance, BaGTA is thought to promote recombination-driven diversification of type IV secretion systems and surface autotransporters involved in host interaction, enhancing *Bartonella* adaptability to diverse hosts (Korotaev *et al.*, 2025). Allele shuffling may also play a key role in enabling bacteria to survive under diverse environmental conditions, particularly by enhancing their ability to resist infection from multiple bacteriophages.

Beyond the acquisition of new genes or allele, GTAs may serve additional biological functions. Early work by Marrs and colleagues hypothesized a role in DNA repair (Marrs *et al.*, 1977), a hypothesis recently substantiated in *C. crescentus*, where GTA-mediated homologous recombination was shown to support genome maintenance (Gozzi *et al.*, 2022). GTAs might also provide benefits analogous to those proposed for natural transformation, including elimination of deleterious mutations or parasitic DNA elements (Ambur *et al.*, 2016).

However, the benefits of GTAs might also be unrelated to the genetic material they carry. The lysis mediated by GTA particles could provide nutrients to surrounding cells, as the programmed cell death phenomenon does (Allocati *et al.*, 2015; Kogay *et al.*, 2020). Additionally, if GTA production is triggered by phage infection, the resulting cell lysis might function as an abortive infection mechanism during phage attack, limiting viral spread in population through self-sacrifice. Nevertheless, GTAs might also offer protection to the host similar than to prophages that render the host resistant to phage superinfection. Prophages interfere with subsequent phage infections through diverse mechanisms, such as blocking phage access to cell receptors, preventing genome injection across the cell envelope, inhibiting replication after phage DNA entry, and disrupting virion assembly pathways, among other strategies (Bondy-Denomy *et al.*, 2016; Patel *et al.*, 2024). To date, research on GTA-associated genes has primarily focused on those encoding structural components, DNA encapsulation machinery, and a limited number of genes involved in reception and regulation (Chapter 1). However, a recent study identified 59 gene families that likely co-evolved with GTA genes (Kogay and Zhaxybayeva, 2024). While most of these families are linked to stress response, DNA repair, and biofilm formation, half exhibits no known functional connection to the GTA production cycle. Investigating these gene families could reveal novel insights into potential benefits of GTAs.

To resolve these hypotheses, experimental validation of GTA-associated benefits—such as nutrient release, infection abortion and superinfection—is critical. A more complete understanding of the GTA production and reception cycle, coupled with functional analysis of co-evolving gene families, will be essential for elucidating the ecological role of GTAs in microbial communities and the evolutionary forces that maintain this mechanism.

7.3.2 Who gives? Who receives?

HGT involves the exchange of genetic material between donor and recipient cells. In mechanisms like conjugation, the roles are clear: the donor carries a conjugative plasmid or integrative conjugative element. While some HGT mechanisms, like conjugation, have clearly defined roles for donor and recipient cells based on the presence of conjugative plasmids or integrative conjugative elements, the determinants for donor and recipient status in other processes, such as natural transformation and GTA-mediated transfer, remain less clear. Interestingly, even within a genetically homogenous population, cells can adopt distinct roles,

with some acting as donors and others as recipients. This raises fundamental questions about the biological and environmental cues that govern these roles.

7.3.2.1 Natural transformation: donors and recipients determinants

In natural transformation, recipient cells internalize extracellular DNA, which typically originates from lysed donor cells. This lysis can be triggered by various factors, including natural cell death, predation, or autolysis (Steinmoen *et al.*, 2002; Hamilton and Dillard, 2006; Veening and Blokesch, 2017). For example, certain competent bacteria, like *V. cholerae* and *S. pneumoniae*, prey on neighboring cells, coupling competence regulation with the production of killing factors (Veening and Blokesch, 2017). *V. cholerae* utilizes a type VI secretion system (T6SS) to inject effector proteins into neighboring cells, leading to lysis of non-kin cells, while kin cells are protected by immunity proteins. In contrast, *S. pneumoniae* employs bacteriocin secretion to eliminate neighboring cells, with producing cells possessing specific immunity mechanisms to prevent self-lysis (Veening and Blokesch, 2017). Alternatively, active secretion of ssDNA via type IV secretion systems (T4SS) has been observed in *N. gonorrhoeae*, promoting HGT and biofilm formation (Hamilton *et al.*, 2005; Callaghan *et al.*, 2017, 2021). These examples highlight multiple factors, including environmental conditions, intercellular interactions, and intrinsic cellular factors, that can determine whether a cell becomes a DNA donor.

The importance of donor cells in natural transformation is often underestimated and remains poorly studied. Nevertheless, evidence suggests that donors may have an important role in natural transformation of *P. stutzeri* (Stewart *et al.*, 1983). Our findings align with this hypothesis, as transformation efficiencies were comparable when using either donor cells or saturating concentrations of extracellular DNA (Chapter 4). The underlying mechanism of DNA donation remains elusive in *P. stutzeri*. The absence of observed cell lysis suggests that DNA transfer may occur via secretion of nuclease-sensitive DNA (Stewart *et al.*, 1983). Furthermore, deletion of *clpV* homologue, a T6SS component involved in sheath disassembly in *P. aeruginosa* (LeRoux *et al.*, 2012), had no impact on transformation with donor cells (Chapter 3), indicating that *P. stutzeri* does not prey on donors like *V. cholerae*. Potential parallels with *N. gonorrhoeae* were initially considered, but genomic analysis via SecReT4 revealed dispersed, non-clustered T4SS-related genes in *P. stutzeri* (Bi *et al.*, 2013), which may represent remnants of a degraded system or false positives. Additionally, while *P. stutzeri* can

incorporate ssDNA, this occurs at a much lower rate than with dsDNA (Meier *et al.*, 2002). This suggests a different mechanism than *N. gonorrhea* and *V. cholerae*.

Besides the mechanisms that enable DNA acquisition from donor cells, another area of investigation is the potential difference between donor-derived DNA and exogenous DNA provided in experimental assays. Notably, transformation efficiency in *P. stutzeri* is significantly enhanced with DNA fragments ≥ 10 kb (Carlson *et al.*, 1983), raising the possibility that donor cells may supply DNA in a more optimal form (e.g., length, protection, or structural state) than purified genomic DNA. To systematically dissect donor-specific contributions, future studies could employ genome-wide approaches such as TIS. As proposed in Chapter 4, a high-throughput screen using a donor mutant library coupled with transformation assays might allow the identification of genetic determinants essential for DNA donation. Beyond uncovering mechanistic factors (e.g., secretion systems, DNA processing enzymes), this strategy may also reveal the regulatory networks that govern donor-cell determination, providing insights into the physiological conditions promoting donor activity.

Natural transformation requires recipient cells to develop competence. In *P. stutzeri*, competence seems to arise globally across populations (Chapter 3), with most cells expressing *pilA*—encoding the type IV pilus subunit for DNA uptake—albeit at variable levels. In contrast, *comA*, essential for DNA translocation into the cytoplasm, exhibits more uniform expression. This suggests widespread but heterogeneous competence, raising questions about its functional consequences.

This heterogeneity aligns with observations in *V. cholerae* and *S. pneumoniae* (Lo Scrudato and Blokesch, 2012; Moreno-Gómez *et al.*, 2017), contrasting with the bistable competence of *B. subtilis*, with only 10–20% of the population expressing competence (Haijema *et al.*, 2001; Maamar and Dubnau, 2005; Maamar *et al.*, 2007). Phenotypic heterogeneity within a clonal population has been postulated to be advantageous because it mitigates the extinction risks of a population in fluctuating environments (Veening *et al.*, 2008; Grimbergen *et al.*, 2015). For natural transformation, this might imply that more competent cells are more likely to acquire DNA, beneficial or deleterious. If the DNA reduces fitness, these individuals are purged without threatening the population stability; if advantageous, they proliferate. However, the direct link between *pilA* expression heterogeneity and transformability remains to be tested, before speculating on the mechanisms underlying this phenotype. Approaches

like FACS coupled with transformation assays could resolve whether higher *pilA* levels correlate with increased DNA uptake.

7.3.2.2 *Transfer mediated by GTA: donors and recipients determinants*

Factors determining which cells become donors or recipients are not well-studied for transfer mediated by GTAs. Nevertheless, one common characteristic shared by GTA systems is that GTA production is limited to a small subset of the population with maximal estimates for *R. capsulatus* at around 3% (Hynes *et al.*, 2012) and at about 14% in *B. henselae* (Québatte *et al.*, 2017). In the case of *P. versutus*, our results indicate that up to 3 % of the population expressed the main GTA operon, with a maximum during exponential growth (Chapter 6).

Interestingly, the stringent response regulates GTA production differently in two GTA model species. In *R. capsulatus*, nutrient depletion triggers (p)ppGpp, which seems to promote RcGTA production (Westbye *et al.*, 2017). In contrast, (p)ppGpp decreases BaGTA production in *B. henselae* (Québatte *et al.*, 2017). As a result, actively replicating cells—with low (p)ppGpp—are more likely to produce BaGTA. In *P. versutus*, the factors determining donor cell status are unknown, but we showed that PvGTA encapsulated more frequently some regions close to the origin of replication of two replicons (Chapter 6). Among multiple hypotheses, this pattern might be explained by the fact that these regions were in several copies when the encapsulation process started. Additionally, PvGTA main operon was mainly expressed during exponential phase (Chapter 6). These observations lead us to hypothesize that GTA production in *P. versutus* could take place in replicating cells, but further evidence is needed.

Factors determining recipient cell status during GTA-mediated transfer are also not well understood. Analysis indicates that BaGTA preferentially dock to actively replicating cells and DNA uptake follows a competence-related mechanism (Québatte *et al.*, 2017). Factors leading to reception of RcGTA have been identified like presence of a capsule and being in stationary phase (Brimacombe *et al.*, 2013; Alim *et al.*, 2023), but mechanisms determining which cells become a recipient are unclear. Regarding PvGTA recipient cells, our current understanding is limited to the observation that nucleotide deletion in the *cpdB* gene reduces recipient efficiency (Chapter 6). Despite efforts to identify a specific PvGTA receptor, these have been unsuccessful thus far. Therefore, rather than focusing solely on receptors, monitoring

expression levels of competence-associated genes could provide alternative insights into which cells are predisposed to act as recipients.

7.3.3 Conclusive remarks

Although this thesis has explored some aspects of natural transformation in *P. stutzeri* and GTAs in *P. versutus*, many fundamental questions remain unresolved. In particular, the factors governing cellular roles in these processes—such as the determinants of *pilA* expression heterogeneity or the regulatory mechanisms driving GTA production—are still poorly understood. Further research is needed to elucidate how genetic and environmental cues coordinate these behaviors at the single-cell and population levels. Moreover, while some insights have been gained into the ecological and evolutionary benefits of these processes, their full impact on microbial populations and communities remains to be elucidated.

7.4 Perspectives

The interplay between donor and recipient cells during HGT remains a rich area for exploration, with many fundamental questions still unanswered. Future research should aim to address several key aspects:

1. How does heterogeneity in competence expression and GTA production affect HGT efficiency and fitness? During this thesis, we observed that competence in *P. stutzeri* and GTA production in *P. versutus* were not uniformly expressed across populations. For natural transformation, a first step would be to determine whether higher *pilA* expression correlates with increased DNA uptake. Techniques such as FACS coupled with multi-marker transformation assays could clarify this relationship. If competence heterogeneity is confirmed, follow-up studies could assess its adaptive benefits—compared to populations with intermediate, uniform competence—using experimental evolution and computational modeling. Similarly, only a subset of *P. versutus* cells expressed the GTA operon. Investigating the trade-offs behind this proportion is important. Indeed, higher number of producers may not be so advantageous if recipients are already overwhelmed by the number of GTAs released, not to mention the fact that a larger portion of the population would lyse. Evolution assays with tunable levels of GTA production or competence levels could help resolve these questions.

2. Are there specific genetic or structural factors that limit which cells can successfully receive DNA from a specific donor? For GTAs, recipient range might be constrained by compatibility between receptor and tail fiber or restriction-modification systems. Identifying the PvGTA receptor—for instance, by testing candidate genes suggested in this work (Chapter 6) or using TIS-based essentiality screens—would clarify host specificity. A complementary approach could adapt phage receptor-binding assays to map how mutations in putative receptors affect GTA attachment (Farquharson *et al.*, 2021). Meanwhile, restriction-modification system might also be a barrier to GTA-mediated transfer; its role might be more tractable via plasmid transformation assays than GTA transfers assays.

Understanding these dynamics not only deepens our knowledge of microbial evolution but also has practical implications. For instance, elucidating the factors that govern donor-recipient interactions may inform strategies to control gene flow in clinical settings or in environmental contexts where HGT shapes microbial communities. By addressing these open questions, researchers can uncover new principles governing horizontal gene transfer and its role in microbial adaptability.

7.5 Conclusion

This thesis explored aspects of HGT mechanisms, focusing on natural transformation in *P. stutzeri* and GTA-mediated HGT in *P. versutus*. The research aimed to elucidate the potential benefits and regulation of natural competence, as well as to introduce *P. versutus* as a new model for studying GTA-mediated HGT. While the study on *P. stutzeri* did not conclusively demonstrate that the competence apparatus was necessary for DNA-based nutrition, it revealed the distinct characteristics of this species compared to well-studied Gram-negative models, particularly regarding competence regulation. The investigation into *P. versutus* uncovered unique features of its GTA system, including biased DNA packaging, differences in GTA expression and regulation compared to *R. capsulatus*, suggesting that the evolutionary advantages of GTAs may vary across bacterial species. Taken together, this thesis contributes to a broader understanding of bacterial competence regulation, phage-host interactions, and the GTAs in *Paracoccus* genus, emphasizing the importance of investigating diverse models to expand our knowledge of HGT. The introduction of *P. versutus* as an alternative GTA model

organism provides a crucial opportunity to broaden scientific understanding of these fascinating DNA transfer particles.

7.6 References

- Alim, N.T.B., Koppenhöfer, S., Lang, A.S., and Beatty, J.T. (2023) Extracellular Polysaccharide Receptor and Receptor-Binding Proteins of the *Rhodobacter capsulatus* Bacteriophage-like Gene Transfer Agent RcGTA. *Genes (Basel)* **14**: 1–16.
- Allocati, N., Masulli, M., Di Ilio, C., and De Laurenzi, V. (2015) Die for the community: An overview of programmed cell death in bacteria. *Cell Death Dis* **6**: 1–10.
- Ambur, O.H., Engelstädter, J., Johnsen, P.J., Miller, E.L., and Rozen, D.E. (2016) Steady at the wheel: Conservative sex and the benefits of bacterial transformation. *Philos Trans R Soc B Biol Sci* **371**: 1–11.
- Bi, D., Liu, L., Tai, C., Deng, Z., Rajakumar, K., and Ou, H.Y. (2013) SecReT4: A web-based bacterial type IV secretion system resource. *Nucleic Acids Res* **41**: 660–665.
- Bondy-Denomy, J., Qian, J., Westra, E.R., Buckling, A., Guttman, D.S., Davidson, A.R., and Maxwell, K.L. (2016) Prophages mediate defense against phage infection through diverse mechanisms. *ISME J* **10**: 2854–2866.
- Brimacombe, C.A., Stevens, A., Jun, D., Mercer, R.G., Lang, A.S., and Beatty, T.J. (2013) Quorum-sensing regulation of a capsular polysaccharide receptor for the *Rhodobacter capsulatus* gene transfer agent (RcGTA). *Mol Microbiol* **87**: 802–817.
- Callaghan, M., Heilers, J.-H., van der Does, C., and Dillard, J.P. (2017) Secretion of chromosomal DNA by the *Neisseria gonorrhoeae* type IV secretion system. *Curr Top Microbiol Immunol* **413**: 323–345.
- Callaghan, M.M., Klimowicz, A.K., Shockey, A.C., Kane, J., Pepperell, C.S., and Dillard, J.P. (2021) Transcriptional and translational responsiveness of the *neisseria gonorrhoeae* type IV secretion system to conditions of host infections. *Infect Immun* **89**: 1–18.
- Carlson, C.A., Pierson, L.S., Rosen, J.J., and Ingraham, J.L. (1983) *Pseudomonas stutzeri* and related species undergo natural transformation. *J Bacteriol* **153**: 93–99.
- Cordero, M., García-Fernández, J., Acosta, I.C., Yepes, A., Avendano-Ortiz, J., Lisowski, C., et al. (2022) The induction of natural competence adapts staphylococcal metabolism to infection. *Nat Commun* **13**: 1–17.
- Dubey, G.P. and Ben-yehuda, S. (2011) Intercellular Nanotubes Mediate Bacterial Communication. *Cell* **144**: 590–600.
- Farquharson, E.L., Lightbown, A., Pulkkinen, E., Russell, T., Werner, B., and Nugen, S.R. (2021) Evaluating Phage Tail Fiber Receptor-Binding Proteins Using a Luminescent Flow-Through 96-Well Plate Assay. *Front Microbiol* **12**: 1–12.
- Finkel, S.E. and Kolter, R. (2001) DNA as a nutrient: Novel role for bacterial competence gene homologs. *J Bacteriol* **183**: 6288–6293.

- Giltner, C.L., Nguyen, Y., and Burrows, L.L. (2012) Type IV Pilin Proteins: Versatile Molecular Modules. *Microbiol Mol Biol Rev* **76**: 740–772.
- Gozzi, K., Tran, N.T., Modell, J.W., Le, T.B.K.K., and Laub, M.T. (2022) Prophage-like gene transfer agents promote *Caulobacter crescentus* survival and DNA repair during stationary phase. *PLoS Biol* **20**: 1–34.
- Grimbergen, A.J., Siebring, J., Solopova, A., and Kuipers, O.P. (2015) Microbial bet-hedging: The power of being different. *Curr Opin Microbiol* **25**: 67–72.
- Haijema, B.J., Hahn, J., Haynes, J., and Dubnau, D. (2001) A ComGA-dependent checkpoint limits growth during the escape from competence. *Mol Microbiol* **40**: 52–64.
- Hamilton, H.L. and Dillard, J.P. (2006) Natural transformation of *Neisseria gonorrhoeae*: From DNA donation to homologous recombination. *Mol Microbiol* **59**: 376–385.
- Hamilton, H.L., Domínguez, N.M., Schwartz, K.J., Hackett, K.T., and Dillard, J.P. (2005) *Neisseria gonorrhoeae* secretes chromosomal DNA via a novel type IV secretion system. *Mol Microbiol* **55**: 1704–1721.
- Heun, M., Binnenkade, L., Kreienbaum, M., and Thormann, K.M. (2012) Functional specificity of extracellular nucleases of *Shewanella oneidensis* MR-1. *Appl Environ Microbiol* **78**: 4400–4411.
- Hynes, A.P., Mercer, R.G., Watton, D.E., Buckley, C.B., and Lang, A.S. (2012) DNA packaging bias and differential expression of gene transfer agent genes within a population during production and release of the *Rhodobacter capsulatus* gene transfer agent, RcGTA. *Mol Microbiol* **85**: 314–325.
- Johnsen, P.J., Dubnau, D., and Levin, B.R. (2009) Episodic selection and the maintenance of competence and natural transformation in *Bacillus subtilis*. *Genetics* **181**: 1521–1533.
- Kogay, R., Wolf, Y.I., Koonin, E. V., and Zhaxybayeva, O. (2020) Selection for reducing energy cost of protein production drives the gc content and amino acid composition bias in gene transfer agents. *MBio* **11**: 1–13.
- Kogay, R. and Zhaxybayeva, O. (2024) Co-evolution of gene transfer agents and their alphaproteobacterial hosts. *J Bacteriol* **206**: 4–7.
- Korotaev, A., Niggli, Q., Congedi, V., and Dehio, C. (2025) Transcriptional antitermination integrates the expression of loci of diverse phage origin in the chimeric *Bartonella* Gene Transfer Agent BaGTA. *bioRxiv* 1–37.
- LeRoux, M., De Leon, J.A., Kuwada, N.J., Russell, A.B., Pinto-Santini, D., Hood, R.D., et al. (2012) Quantitative single-cell characterization of bacterial interactions reveals type VI secretion is a double-edged sword. *Proc Natl Acad Sci U S A* **109**: 19804–19809.
- Maamar, H. and Dubnau, D. (2005) Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop. *Mol Microbiol* **56**: 615–624.

- Maamar, H., Raj, A., and Dubnau, D. (2007) Noise in gene expression determines cell fate in *Bacillus subtilis*. *Science* (80-) **317**: 526–529.
- Marrs, B., Wall, J.D., and Gest, H. (1977) Emergence of the biochemical genetics and molecular biology of photosynthetic bacteria. *Trends Biochem Sci* **2**: 105–108.
- McDaniel, L.D., Young, E., Delaney, J., Ruhnau, F., Ritchie, K.B., and Paul, J.H. (2010) High frequency of horizontal gene transfer in the oceans. *Science* (80-) **330**: 50.
- Meier, P., Berndt, C., Weger, N., and Wackernagel, W. (2002) Natural transformation of *Pseudomonas stutzeri* by single-stranded DNA requires type IV pili, competence state and *comA*. *FEMS Microbiol Lett* **207**: 75–80.
- Moreno-Gómez, S., Sorg, R.A., Domenech, A., Kjos, M., Weissing, F.J., Van Doorn, G.S., and Veening, J.W. (2017) Quorum sensing integrates environmental cues, cell density and cell history to control bacterial competence. *Nat Commun* **8**: 1–12.
- Nielsen, K.M., Johnsen, P.J., Bensasson, D., and Daffonchio, D. (2007) Release and persistence of extracellular DNA in the environment. *Environ Biosafety Res* **6**: 37–53.
- Palchevskiy, V. and Finkel, S.E. (2006) *Escherichia coli* competence gene homologs are essential for competitive fitness and the use of DNA as a nutrient. *J Bacteriol* **188**: 3902–3910.
- Patel, P.H., Taylor, V.L., Zhang, C., Getz, L.J., Fitzpatrick, A.D., Davidson, A.R., and Maxwell, K.L. (2024) Anti-phage defence through inhibition of virion assembly. *Nat Commun* **15**: 1–11.
- Québatte, M., Christen, M., Harms, A., Körner, J., Christen, B., and Dehio, C. (2017) Gene Transfer Agent Promotes Evolvability within the Fittest Subpopulation of a Bacterial Pathogen. *Cell Syst* **4**: 611–621.
- Rumbo, C., Fernandez-Moreira, E., Merino, M., Poza, M., Mendez, J.A., Soares, N.C., et al. (2011) Horizontal Transfer of the OXA-24 Carbapenemase Gene via Outer Membrane Vesicles : a New Mechanism of Dissemination of Carbapenem Resistance Genes in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* **55**: 3084–3090.
- Lo Scrudato, M. and Blokesch, M. (2012) The Regulatory Network of Natural Competence and Transformation of *Vibrio cholerae*. *PLoS Genet* **8**: 1–17.
- Seitz, P. and Blokesch, M. (2012) Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. *FEMS Microbiol Rev* **37**: 336–363.
- Seper, A., Fengler, V.H.I., Roier, S., Wolinski, H., Kohlwein, S.D., Bishop, A.L., et al. (2011) Extracellular nucleases and extracellular DNA play important roles in *Vibrio cholerae* biofilm formation. *Mol Microbiol* **82**: 1015–1037.
- Solioz, M., Yen, H.-C.C., and Marrs, B. (1975) Release and uptake of gene transfer agent by *Rhodopseudomonas capsulata*. *J Bacteriol* **123**: 651–657.

- Steinmoen, H., Knutsen, E., and Håvarstein, L.S. (2002) Induction of natural competence in *Streptococcus pneumoniae* triggers lysis and DNA release from a subfraction of the cell population. *Proc Natl Acad Sci U S A* **99**: 7681–7686.
- Stewart, G.J., Carlson, C.A., and Ingraham, J.L. (1983) Evidence for an active role of donor cells in natural transformation of *Pseudomonas stutzeri*. *J Bacteriol* **156**: 30–35.
- Veening, J.W. and Blokesch, M. (2017) Interbacterial predation as a strategy for DNA acquisition in naturally competent bacteria. *Nat Rev Microbiol* **15**: 621–629.
- Veening, J.W., Smits, W.K., and Kuipers, O.P. (2008) Bistability, epigenetics, and bet-hedging in bacteria. *Annu Rev Microbiol* **62**: 193–210.
- Vos, M. (2009) Why do bacteria engage in homologous recombination? *Trends Microbiol* **17**: 226–232.
- Westbye, A., O'Neill, Z., Schellenberg-Beaver, T., and Beatty, T.J. (2017) The *Rhodobacter capsulatus* gene transfer agent is induced by nutrient depletion and the RNAP omega subunit. *Microbiology* **163**: 1–9.
- Yüksel, M., Power, J.J., Ribbe, J., Volkmann, T., and Maier, B. (2016) Fitness trade-offs in competence differentiation of *Bacillus subtilis*. *Front Microbiol* **7**: 1–15.

Chapter 8 - Appendices

8.1 Conference presentations

1. Poster: 18th International Symposium on Microbial Ecology (ISME18), Lausanne, Switzerland, 14-19 August 2022 and Swiss Society for Microbiology meeting (SSM2022), Lausanne, Switzerland, 30 Aug. – 1 Sep. 2022.



Global approach of the regulation of natural transformation in *Pseudomonas stutzeri*

Camille Tinguely, Pilar Junier & Diego Gonzalez

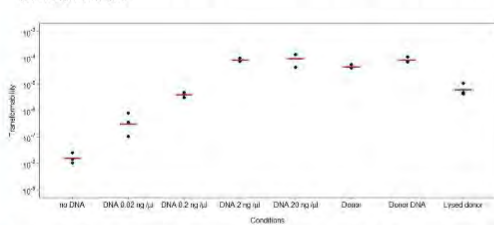
Laboratory of Microbiology, Biology Institute, University of Neuchâtel, Switzerland



unine
Université de Neuchâtel

BACKGROUND

- Natural transformation is a horizontal gene transfer (HGT) mechanism by which naked DNA present in the environment is internalized and integrated into the genome.
- As a costly and risky process, strong selective forces drive bacteria to regulate natural transformation and competence to maximize the benefits and reduce the costs of the process. One of this optimization is to select the uptaken DNA.
- The soil bacteria *Pseudomonas stutzeri* expresses various levels of transformability and can be transformed with free DNA but also in presence of donor cells, making it an interesting model to study natural transformation.



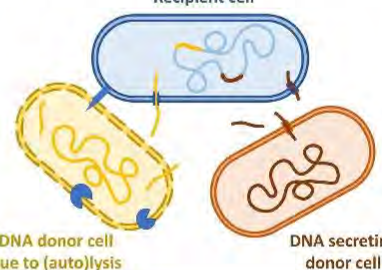
Transformability of the recipient strain *P. stutzeri* ATCC 17587 with no DNA (control), with various concentration of extracellular DNA, with donor bacteria *Acinetobacter*, with extracted DNA from the donors fraction and with lysed donor cells. The transformability is calculated as the number of rifampicin-resistant cells per total recipient cell. The three replicates are indicated with the black dots, the mean is in red.

AIM

Understand recipient and donor cells strategies to optimize natural transformation

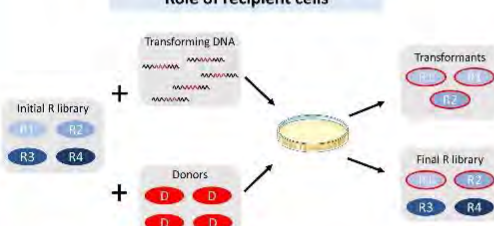
QUESTIONS

- Which genes are required for *Pseudomonas stutzeri* to acquire DNA by natural transformation ?
- Does the donor play an active role in natural transformation ?

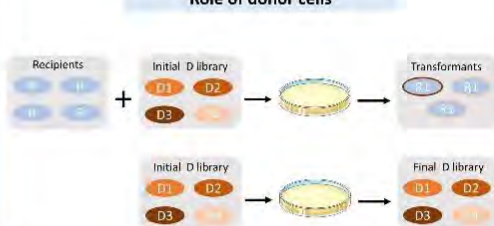


EXPERIMENTAL DESIGN

Role of recipient cells

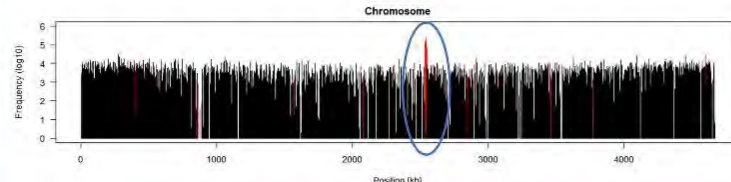


Role of donor cells



PRELIMINARY RESULTS

- The sequencing of initial recipient library indicated 13% saturation with a bias in reads distribution



Bias in reads distribution

Protocol adaptation

Resequencing

→ Results are coming



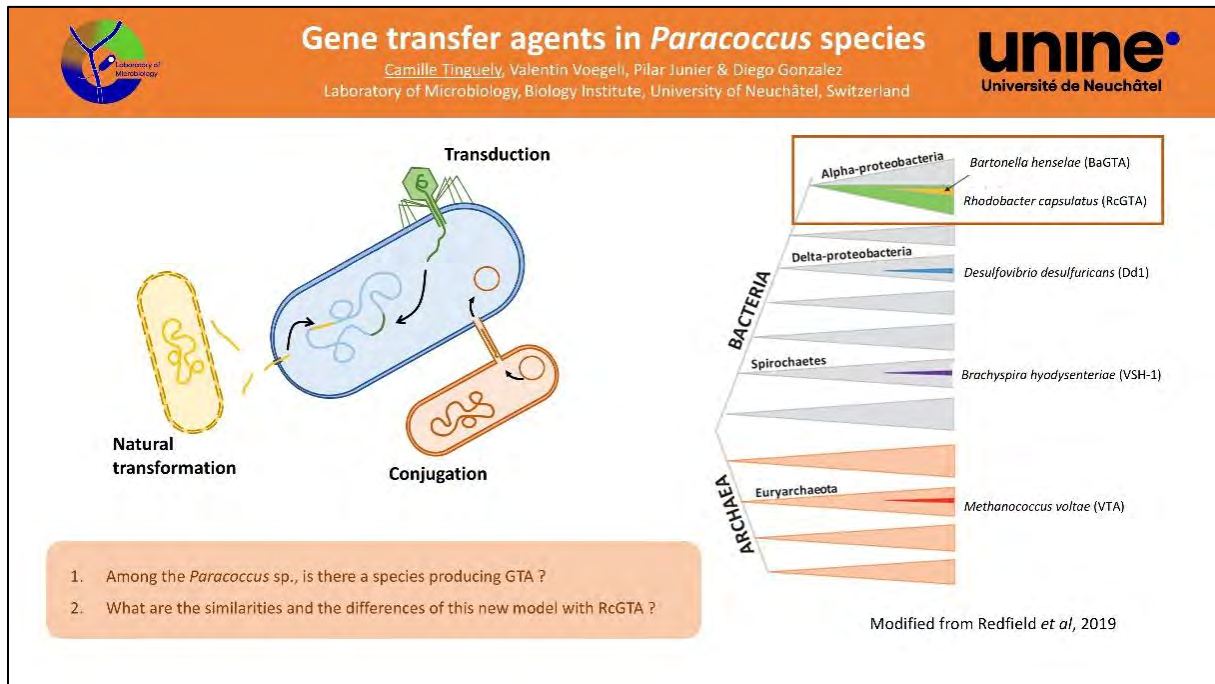
Camille Tinguely
camille.tinguely@unine.ch



Acknowledgements

This research is supported by the Fondation Pierre Mercier pour la Science.

2. Poster pitch: SWISS Young Microbiologists Symposium 2023, Zurich, Switzerland, 2 June 2023.



4. Oral presentation at the Swiss Society for Microbiology meeting (SSM2023), Lausanne, Switzerland, 30-31 August 2023.



Gene transfer agents (GTA) in *Paracoccus* species

Camille Tinguely, Valentin Voegeli, Pilar Junier & Diego Gonzalez

Laboratory of Microbiology, Biology Institute, University of Neuchâtel, Switzerland

31.08.23

8.2 Other projects

Gamification as a tool to teach key concepts in microbiology to bachelor-level students in biology: a case study using microbial interactions and soil functioning

Aislinn Estoppey¹, Camille Tinguely¹, Melissa Cravero¹, Estelle Blandenier¹, Lucie Brand¹, Julien Court¹, Marie Zaninetti¹, Saskia Bindschedler^{1*}, Pilar Junier^{*}

¹ Laboratory of Microbiology, University of Neuchâtel, Rue Emile-Argand 11, CH-2000, Neuchâtel, Switzerland.

*Corresponding author

Abstract: Microbiology is a difficult topic to teach given that the objects of study are mostly invisible to the learner. The majority of university students beginning their training in biology are more interested in natural objects that can be seen with the naked eye. Nonetheless, micro-organisms are key components of the biosphere and a good microbiological background is required for a thorough training in natural sciences. Lectures are still a common teaching format in universities. However, it is a passive learning format and no longer considered the most adequate approach in most teaching situations. Instead, alternatives consisting of more active teaching formats have been recognized to better motivate students to acquire and consolidate knowledge. In addition, transferable skills, such as effective communication, critical thinking and time management, are acquired simultaneously. A similar engagement can be obtained using games as part of the teaching experience. In this study, we designed a card game to teach key concepts in basic bacteriology and mycology to bachelor-level students. The first task consists of creating and designing microbial characters based on a list of species. This proved very useful for second-year bachelor students in terms of grasping concepts such as cell morphologies, taxonomy and life cycles. In the second task, third-year students used the characters created in the second-year class to develop a game based on an ecological function, namely forest litter degradation. In addition, they also considered experimental validation of the microbial activities and incorporated knowledge acquired in other fields.

Diurnal cycles drive rhythmic physiology and promote survival in facultative phototrophic bacteria

Camille Tinguely^{1#}, Mélanie Paulméry^{1#}, Céline Terrettaz^{1#}, Diego Gonzalez^{1*}

¹Laboratory of Microbiology, Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland.

[#]Contributed equally.

^{*}Corresponding author

Abstract: Bacteria have evolved many strategies to spare energy when nutrients become scarce. One widespread such strategy is facultative phototrophy, which helps heterotrophs supplement their energy supply using light. Our knowledge of the impact that such behaviors have on bacterial fitness and physiology is, however, still limited. Here, we study how a representative of the genus *Porphyrobacter*, in which aerobic anoxygenic phototrophy is ancestral, responds to different light regimes under nutrient limitation. We show that bacterial survival in stationary phase relies on functional reaction centers and varies depending on the light regime. Under dark-light alternance, our bacterial model presents a diphasic life history dependent on phototrophy: during dark phases, the cells inhibit DNA replication and part of the population lyses and releases nutrients, while subsequent light phases allow for the recovery and renewed growth of the surviving cells. We correlate these cyclic variations with a pervasive pattern of rhythmic transcription which reflects global changes in diurnal metabolic activity. Finally, we demonstrate that, compared to either a phototrophy mutant or a bacteriochlorophyll *a* overproducer, the wild type strain is better adapted to natural environments, where regular dark-light cycles are interspersed with additional accidental dark episodes. Overall, our results highlight the importance of light-induced biological rhythms in a new model of aerobic anoxygenic phototroph representative of an ecologically important group of environmental bacteria.

8.3 Curriculum Vitae

CAMILLE TINGUELY

PhD candidate in Microbiology

Route du Moulin 24, 1646 Echarlens, Switzerland

Mobile: +41 79 901 26 07

Email: camille.tinguel@yahoo.com

Research Interests

My general interest lies in microbial ecology, with a dual focus on applied and fundamental research. On one side, I am passionate about harnessing microorganisms to develop innovative solutions, such as exploring biological control strategies. On the other hand, I am deeply interested in understanding the evolutionary and ecological mechanisms that drive microbial adaptation, including processes like horizontal gene transfer and bacterium-phage interactions. These complementary interests allow me to bridge practical applications with fundamental scientific inquiry.

Education

2020 – Present	PhD in Biology. University of Neuchâtel, Switzerland
2019 – 2020	Diploma in Teaching for Secondary Schools (DEEM) with a complementary certificate in professional pedagogy. University of Fribourg, Switzerland
2015 – 2017	Master of Science in Biology, ETH Zurich, Switzerland.
2012 – 2015	Bachelor of Science in Biology, ETH Zurich, Switzerland.

Research experiences

May 2025- Present	Research Assistant, Biotechnology Group and Cheese Quality, Cultures and Terroir Group, Agroscope of Liebefeld. Topics: <i>Development of an optimized culture medium for the growth of Brevibacterium spp.; YeastGuard: Evaluating the impact of yeasts on the proliferation of Listeria spp. in a mini-cheese model.</i>
Oct. 2020– Sep. 2024	PhD thesis in microbiology, Laboratory of microbiology, University of Neuchâtel. Title: <i>Natural Transformation and Gene Transfer Agents: New Perspectives on Horizontal Gene Transfer in environmental bacteria.</i>
Feb. 2018 – July 2019	Research Assistant, Medical Microbiology Group, Department of Medicine, University of Fribourg. Topic: <i>Fighting the resistance: development of rapid diagnostic tests for detection of antibiotic resistance in clinically significant Gram-negative bacteria.</i>

March – Aug. 2017	Scientific Assistant, Plant Pathology Group, ETH Zurich. Topic: <i>Isolation and characterization of insecticidal Pseudomonas protegens</i> .
Aug. 2016 – March 2017	Master thesis, Group of Pathogen Ecology, ETH Zurich. Title: <i>Impact of modification of quorum-sensing on the cost of rifampicin resistance and virulence factor production of Pseudomonas aeruginosa</i> .
Feb. – July 2016	Semester project, Plant Pathology Group, ETH Zurich. Title: <i>Oral toxicity of Pseudomonas protegens CHAO to different larval stages of Pieris brassicae, effect of CHAO feeding on phenotype of adults and comparison of qPCR and colony-plating for monitoring of CHAO in insects</i> .
Sep.- Dec. 2015	Semester project, Group of Biocommunication, ETH Zurich. Title: <i>Expression of Anthocyanidins in leaves of two color morphs of Althaea rosea</i> .

Teaching experiences

Nov. – Dec. 2023	PhD student assistant for the practical of molecular biology for 1st year Master students in biogeoscience (1day/week), University of Neuchâtel.
March – April 2023	PhD student assistant for the Problem-based learning in Microbiology for 3rd year Bachelor students in biology (4 weeks), University of Neuchâtel. Topic: <i>Roles and importance of microbial interactions as an ecosystem function: development of a game as a teaching and communication tool</i> .
Nov. – Dec. 2022	PhD student assistant for the practical of molecular biology for 1st year Master students in biogeoscience (1day/week), University of Neuchâtel.
May 2022	Service-learning project in 3 primary schools (15h lessons preparation, 15h teaching), University of Neuchâtel.
March – April 2022	PhD student assistant for the Problem-based learning in Microbiology for 3rd year Bachelor students in biology (4 weeks), University of Neuchâtel. Topic: <i>Tiny Earth Switzerland: Searching for new antifungal compounds in the environment of honeybees</i> .
Nov. – Dec. 2021	PhD student assistant for the practical of molecular biology for 1st year Master students in biogeoscience (1day/week), University of Neuchâtel.
March – April 2021	PhD student assistant for the Problem-based learning in Microbiology for 3rd year Bachelor students in biology (4 weeks), University of Neuchâtel. Topic: <i>Tiny Earth Switzerland: Searching for new antibiotics</i> .
Aug. 2019 – June 2020	Teaching Intern for anatomy and biology lessons (4h/week). École de Culture Générale de Fribourg, Switzerland.

Conferences

30-31 August 2023	Swiss Society for Microbiology meeting (SSM2023) , Lausanne, Switzerland. Presentation title: <i>Gene transfer agents in Paracoccus species</i> .
9-13 July 2023	Federation of European Microbiological Societies meeting (FEMS2023) , Hamburg, Germany. Poster title: <i>Gene transfer agents in Paracoccus species: a new bacterial model</i> .
2 June 2023	SWISS Young Microbiologists Symposium 2023 , Zurich, Switzerland. Poster title: <i>Gene transfer agents in Paracoccus species: a new bacterial model</i> .
30 Aug. – 1 Sep. 2022	Swiss Society for Microbiology meeting (SSM2022) , Lausanne, Switzerland. Poster title: <i>Global approach of the regulation of natural transformation in Pseudomonas stutzeri</i> .
14-19 August 2022	18th International Symposium on Microbial Ecology (ISME18) , Lausanne, Switzerland. Poster title: <i>Global approach of the regulation of natural transformation in Pseudomonas stutzeri</i> .

Supervision experiences

July 2023 – Dec. 2023	Co-advising of Lorenzo Bisanti (Biology technician trainee), Laboratory of Microbiology, University of Neuchâtel
Nov. 2022 – March 2023	Co-advising of Matys Constantino (Biology technician trainee), Laboratory of Microbiology, University of Neuchâtel
Sep. 2022 – Nov. 2022	Advising of Lea Ripoll Odermatt (Bachelor student), Laboratory of Microbiology, University of Neuchâtel.
May 2022– Oct. 2022	Co-advising of Naïma Mangia (Biology technician trainee), Laboratory of Microbiology, University of Neuchâtel
Sep. 2021 – Aug. 2022	Co-advising of Valentin Voegeli (Master student) for Master thesis, Laboratory of Microbiology, University of Neuchâtel

Other professional experiences

June 2022-2023-2024	Certificate Examination Expert. École de Culture Générale de Fribourg.
Oct. 2017 – Jan. 2018	Quality Assurance Assistant. Mifroma SA, Switzerland

Publications

1. Nordmann P, Helsen N, Kieffer N, **Tinguely C**, Greub G, Poirel L. Rapid detection of β -lactamase activity using the rapid Amp NP test. Microbiol Spectr. 2025 Apr;13(4):e0078224. doi: 10.1128/spectrum.00782-24. Epub 2025 Mar 6. PMID: 40047444; PMCID: PMC11960058.
2. Nordmann P, Bouvier M, Delaval A, **Tinguely C**, Poirel L, Sadek M. Rapid Detection of Ceftazidime/Avibactam Susceptibility/Resistance in Enterobacterales by Rapid

- CAZ/AVI NP Test. *Emerg Infect Dis.* 2024 Feb;30(2):255-261. doi: 10.3201/eid3002.221398.
3. Estoppey A, **Tinguely C**, Cravero M, Blandenier E, Brand L, Court J, Zaninetti M, Bindschedler S, Junier P. Gamification as a tool to teach key concepts in microbiology to bachelor-level students in biology: a case study using microbial interactions and soil functioning. *Access Microbiol.* 2024 Feb 6;6(2):000699.v3. doi: 10.1099/acmi.0.000699.v3.
 4. **Tinguely C**, Paulméry M, Terrettaz C, Gonzalez D. Diurnal cycles drive rhythmic physiology and promote survival in facultative phototrophic bacteria. *ISME Commun.* 2023 Nov 24;3(1):125. doi: 10.1038/s43705-023-00334-5.
 5. Nordmann P, Sadek M, **Tinguely C**, Poirel L. Rapid Resalmipenem/Acinetobacter NP Test for Detection of Carbapenem Susceptibility/Resistance in *Acinetobacter baumannii*. *J Clin Microbiol.* 2021 May 19;59(6):e03025-20. doi: 10.1128/JCM.03025-20.
 6. Sadek M, **Tinguely C**, Poirel L, Nordmann P. Rapid Polymyxin/Pseudomonas NP test for rapid detection of polymyxin susceptibility/resistance in *Pseudomonas aeruginosa*. *Eur J Clin Microbiol Infect Dis.* 2020 Sep;39(9):1657-1662. doi: 10.1007/s10096-020-03884-x.
 7. Sadek M, Poirel L, **Tinguely C**, Nordmann P. A Selective Culture Medium for Screening Ceftazidime-Avibactam Resistance in Enterobacterales and *Pseudomonas aeruginosa*. *J Clin Microbiol.* 2020 Aug 24;58(9):e00965-20. doi: 10.1128/JCM.00965-20.
 8. Flury P, Vesga P, Dominguez-Ferreras A, **Tinguely C**, Ullrich CI, Kleespies RG, Keel C, Maurhofer M. Persistence of root-colonizing *Pseudomonas protegens* in herbivorous insects throughout different developmental stages and dispersal to new host plants. *ISME J.* 2019 Apr;13(4):860-872. doi: 10.1038/s41396-018-0317-4.
 9. Lescat M, Poirel L, **Tinguely C**, Nordmann P. A Resazurin Reduction-Based Assay for Rapid Detection of Polymyxin Resistance in *Acinetobacter baumannii* and *Pseudomonas aeruginosa*. *J Clin Microbiol.* 2019 Feb 27;57(3):e01563-18. doi: 10.1128/JCM.01563-18.

Skills

Lab skills	Microbiological Techniques (bacteria and phage isolation and culture, biofilm formation, motility assays, MIC testing, etc.), molecular biology (Plasmid/DNA/RNA extraction and quantification, PCR, cloning, transposon sequencing, and mutant library construction), biochemistry methods basics (protein expression induction, enzymatic activity measurements, double hybrid BACTH system), microscopy and flow cytometry.
IT	Microsoft Office Suite, RStudio, Image J, various molecular biology softwares (Snapgene, BioEdit, GeneiousPrime)
Languages	French (native), English (advanced, MSc and Ph.D. in English, public presentations), German (intermediate, Bachelor 1st year in German)

References

Prof. Pilar Junier
Laboratory of microbiology University of Neuchâtel
Rue Emile-Argand 11 CH-2000 Neuchâtel
+41 32 718 22 44
pilar.junier@unine.ch

Dr. Diego Gonzalez
Laboratory of microbiology University of Neuchâtel
Rue Emile-Argand 11 CH-2000 Neuchâtel
+41 78 605 48 19
diego.gonzalez@unine.ch