



**HAL**  
open science

# Natural competence for genetic transformation in *Staphylococcus aureus*

Shiyuan Feng

► **To cite this version:**

Shiyuan Feng. Natural competence for genetic transformation in *Staphylococcus aureus*. *Genetics*.  
Université Paris-Saclay, 2021. English. NNT : 2021UPASL095 . tel-04368867

**HAL Id: tel-04368867**

**<https://theses.hal.science/tel-04368867>**

Submitted on 2 Jan 2024

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

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

# Natural competence for genetic transformation in *Staphylococcus aureus*

*Compétence naturelle pour la transformation génétique chez  
Staphylococcus aureus*

## Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 577, Structure et dynamique des systèmes vivants (SDSV)  
Spécialité de doctorat : Sciences de la vie et de la santé  
Graduate School : Life Science and Health. Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche **Institut de Biologie Intégrative de la  
Cellule - I2BC** (CEA, CNRS, Université Paris-Saclay), sous la direction de  
**Nicolas MIROUZE**, Chargé de Recherche INSERM, I2BC

Thèse soutenue à Paris-Saclay, le 10 décembre 2021, par

**Shiyuan FENG**

## Composition du Jury

<b>Nicolas BAYAN</b> Professeur, Université Paris-Saclay, I2BC, Gif-sur-Yvette	Président
<b>Nathalie CAMPO</b> Chargée de Recherche CNRS, Laboratoire de Microbiologie et Génétique Moléculaires, Toulouse	Rapporteur & Examinatrice
<b>Xavier CHARPENTIER</b> Directeur de Recherche INSERM, Centre International de Recherche en Infectiologie, Lyon	Rapporteur & Examineur
<b>Angelika GRUNDLING</b> Professeure, Imperial College London, London	Examinatrice
<b>Philippe BOULOC</b> Directeur de Recherche CNRS, I2BC, Gif-sur-Yvette	Examineur
<b>Nicolas MIROUZE</b> Chargé de Recherche INSERM, I2BC, Gif-sur-Yvette	Directeur de thèse

**Titre :** Compétence naturelle pour la transformation génétique chez *Staphylococcus aureus*

**Mots clés :** *Staphylococcus aureus*, compétence naturelle, transformation génétique

**Résumé :** Le transfert horizontal de gènes (THG) contribue de façon importante à l'évolution des espèces bactériennes. L'acquisition de certains gènes peut permettre à une bactérie d'utiliser de nouveaux substrats ou de survivre dans des environnements toxiques. L'analyse de nombreux génomes a également révélé qu'un nombre important de gènes augmentant la virulence et la résistance aux antibiotiques de pathogènes humains majeurs, tel que *Staphylococcus aureus*, a été acquis par THG.

Pendant ma thèse, le but de mes projets était d'étudier, chez *S. aureus*, le développement de la compétence naturelle, une adaptation physiologique nécessaire pour la transformation génétique, l'un des mécanismes de THG chez les bactéries. Pour atteindre cet objectif, j'ai pu identifier certains stress sentis par les cellules, déchiffrer les complexes voies de régulations induites en réponse à ces signaux et enfin, chercher à élucider le rôle joué par les multiples régulateurs centraux de la compétence.

**Title :** Natural competence for genetic transformation in *Staphylococcus aureus*

**Keywords :** *Staphylococcus aureus*, natural competence, genetic transformation

**Abstract :** Horizontal gene transfer (HGT) greatly contributed to the evolution of bacterial species. Acquisition of certain genes can completely change the fitness of a bacterium allowing the use of new substrates or the survival in environments otherwise considered as toxic. Analysis of sequence and functions in numerous genomes recently revealed that an important number of genes increasing virulence and antibiotic resistance in major human pathogens, such as *Staphylococcus aureus*, have been acquired through HGT.

During my PhD, the aim of my projects was to investigate, in *S. aureus*, the development of natural competence, a physiological adaptation necessary to perform genetic transformation, an important mechanism used by bacteria to perform HGT. To reach such goal, I have particularly identified environmental stimuli sensed by the cells, deciphered complex regulatory pathways triggered in response to these signals and characterized the role of multiple competence central regulators.

## Acknowledgments

First and foremost, my appreciation goes to my supervisor, Dr. Nicolas Mirouze, without whom this thesis would not be possible. I will remember Nicolas for his continued patience, guidance and support throughout my years of study in his laboratory. His example of integrity, enthusiasm and dedication to research will remain with me as a clear base on which I hope to establish my own scientific career.

I would like to acknowledge my PhD tutors, Dr. Jean-luc Pernodet and Dr. Etienne Dervyn, for their insightful discussion and guidance in the completion of this thesis.

The time spent in the team of Nicolas Mirouze had not only been of immense scientific growth for me but also a memorable one. I would like to thank my fellow lab-mates, Yolande Hauck, Fedy Morgene, Marine Rigot, Macha Dussouchaud, Roza Mohammedi, Paul Budisavljevic. It was my pleasure to work with them. I particularly would like to thank Yolande for helping me adjust to the lab environment during the first year of my PhD, and thank Fedy for his kindly help and discussion.

I would like to express my deepest gratitude to my family, particularly my girlfriend, who had always believed in my capabilities and supported my choice of career. Thank you with all my heart and soul. Looking forward to see you soon. Thank you, my dear grandfather, and miss you every day.

I would like also to acknowledge the funding sourced for my thesis. To CSC, thank for supporting for 4 years. In addition, thank my supervisor again for providing the financial needs for the last 3 months.

Last but not the least, I want to express my sincere thanks to all persons who have been involved all along these years of my PhD.





### Summary of the thesis in English

During evolution, the ability of bacteria to adapt to evolving environments often resulted from the acquisition of new genes, and therefore new functions, through Horizontal Gene Transfer (HGT). HGT, defined as the transmission of genetic material between organisms that are not in a parent-offspring relationship, is an important means by which bacteria ensure genomic plasticity. HGT may occur via three main mechanisms: transduction, conjugation and genetic transformation. The latter mechanism ensures the internalization and homologous recombination within the chromosome of high molecular-weight exogenous DNA. Genetic transformation requires that bacterial cells enter a regulated differentiated state called competence. Competence regulation varies considerably among species. Central competence regulatory pathways include dedicated alternative sigma factors, transcription activators and transcription co-regulators. Importantly, expression, stability and activity of these central regulators are highly modulated by a large variety of regulatory cascades. These regulatory pathways trigger competence development in response to diverse environmental and cellular cues ranging from quorum sensing, cellular nutritional status and stressful conditions. Interestingly, despite global similarities between these pathways, competence regulation has probably evolved independently in different species in order to adjust to a large diversity of ecological niches and stimuli. Competence has been thoroughly studied in a number of different bacteria, particularly *Bacillus subtilis* and *Streptococcus pneumoniae*. Interestingly, new bacteria displaying such aptitude are often discovered, and one of the latest is the human pathogen, *Staphylococcus aureus*.

1. During my PhD, the main objective was to characterize the development of competence in a new important model organism in this field, *S. aureus*. To achieve such objective, my aim was to answer several essential questions:

1- What are the environmental signals sensed by the cells to induce the development of competence in *S. aureus*?

## Summary

2- What are the regulatory pathways integrating these signals and transmitting the information to the central competence regulators?

3- What is the exact role of the three potential central competence regulators present in *S. aureus*: SigH, ComK1 and ComK2?

First, through the construction of a *S. aureus* genomic library in yeast, screened with the yeast two-hybrid technique, I have identified partners of essential proteins involved in competence development and genetic transformation. Then, using a genetic approach and transcriptional analysis techniques (reporter genes and RNA-sequencing), I have particularly shown that:

- *S. aureus* can induce competence efficiently thanks to optimized protocols
- The three central competence regulators are all important for a complete development of competence.
- SigH and ComK1 are both essential for the expression of the genes involved in genetic transformation
- ComK2 controls the expression of other genes involved in other processes, such as the general stress response, toxin/antitoxin systems and amino acid / nucleic acid metabolism.
- Nutrient limitation/cell density and oxygen availability are two important signals sensed by *S. aureus* to induce competence.
- Different oxygen concentrations can activate or inhibit competence through different oxygen-sensing two-component systems
- A protein of unknown function, found interacting with ComK2 in Y2H, work alongside ComK2 to inhibit competence under anaerobic conditions

### Summary of the thesis in French

Pendant l'évolution, la capacité des bactéries à s'adapter à des environnements fluctuants a souvent été le résultat de l'acquisition de nouveaux gènes, et donc de nouvelles fonctions, par Transfer Horizontal de Gènes (THG). Le THG, défini comme la transmission de matériel génétique entre organismes sans lien de parenté, est un moyen important utilisé par les bactéries pour assurer leur plasticité génétique. Ce THG peut être réalisé par trois principaux mécanismes : la transduction, la conjugaison et la transformation génétique (TG). Ce dernier mécanisme assure l'internalisation et la recombinaison homologue dans le chromosome receveur d'importantes quantités d'ADN exogène. La TG nécessite que les cellules basculent dans un état physiologique différencié appelé compétence naturelle (CN). La régulation de la CN varie considérablement d'un organisme à l'autre. Les voies centrales de régulation incluent des facteurs sigma alternatifs, des activateurs transcriptionnels et des co-régulateurs. Il est important de préciser que l'expression, la stabilité et l'activité de ces régulateurs centraux sont largement modulées par un grand nombre de cascades de régulation. Ces voies de régulation induisent le développement de la CN en réponse à divers signaux environnementaux tels que la forte densité cellulaire, le statut nutritionnel de la cellule et des conditions stressantes. Il est intéressant de remarquer que malgré des grandes similarités entre ces mécanismes, la régulation de la CN a probablement évolué de façon indépendante pour chaque espèce dans le but de s'ajuster à un large éventail de niches écologiques et de stimuli. La CN a été très étudiée chez différentes bactéries modèles, telles que *Bacillus subtilis* et *Streptococcus pneumoniae*. De nouvelles espèces bactériennes avec de telles aptitudes sont souvent découvertes et l'une des dernières en date n'est autre que le pathogène humain *Staphylococcus aureus*.

Pendant ma thèse, mon principal objectif était de caractériser le développement de la CN chez un nouvel organisme modèle dans le domaine, *S. aureus*. Pour atteindre un tel objectif, mon but était de répondre à différentes questions :

- 1- Quels sont les signaux environnementaux perçus par les cellules pour induire le

## Summary

développement de la CN chez *S. aureus* ?

- 2- Quelles sont les voies de transduction du signal qui intègrent ces stimuli et transmettent l'information aux régulateurs centraux de la CN ?
- 3- Quel est le rôle exact des trois potentiels régulateurs centraux de la CN chez *S. aureus* : SigH, ComK1 et ComK2 ?

Tout d'abord, par la construction d'une banque d'ADN génomique de *S. aureus* dans la levure, criblée par la technique du double hybride, j'ai pu identifier les partenaires de protéines impliquées dans le développement de la CN et de la TG. Ensuite, en utilisant une approche de génétique, associée à des techniques d'analyse de la transcription (gènes rapporteurs et RNA-sequencing), j'ai plus particulièrement démontré que :

- *S. aureus* peut induire efficacement la CN grâce à des protocoles optimisés
- Les trois régulateurs centraux sont tous essentiels au développement complet de la CN
- SigH et ComK1 sont tous les deux essentiels à l'expression des gènes dédiés à la TG
- ComK2 contrôle l'expression d'autres gènes impliqués dans d'autres processus tels que la réponse générale au stress, des systèmes toxine/antitoxine et le métabolisme des acides aminés et acides nucléiques.
- La limitation en nutriment/densité cellulaire et la concentration en oxygène sont des signaux environnementaux importants perçus par *S. aureus* pour induire la CN
- Différentes concentrations en oxygène peuvent activer ou inhiber le développement de la CN à travers différents systèmes à deux composants
- Une protéine de fonction inconnue, interagissant avec ComK2, travaille aux côtés de son partenaire pour inhiber la CN dans des conditions d'anaérobiose

## Version française de la thèse

### 1 Introduction

#### 1.1 Brève introduction de *Staphylococcus aureus*

*Staphylococcus aureus* appartient au phylum des Firmicutes, à la classe des bacilles, à l'ordre des Bacilliales et à la famille des Staphylococcaceae. *S. aureus* est une bactérie coccale Gram-positif et anaérobie facultative. La particularité d'un anaérobie facultatif est sa capacité à convertir l'ATP par respiration aérobie en présence d'oxygène et de basculer vers la respiration anaérobie (fermentation) en l'absence d'oxygène 1. *S. aureus* produit également l'enzyme catalase qui transforme le peroxyde d'hydrogène (H<sub>2</sub>O<sub>2</sub>) en oxygène et en eau. Cette activité catalase est utilisée pour distinguer les staphylocoques des streptocoques et des entérocoques.

*S. aureus* colonise une part importante de la population humaine (30 % de porteurs persistants et 30 % de porteurs intermittents). Ce pathogène a le potentiel de provoquer une grande variété de complications, allant de simples infections cutanées bénignes jusqu'à de graves maladies telles que la septicémie, l'ostéomyélite et l'endocardite 2. De plus, la mortalité associée aux infections à *S. aureus* est souvent corrélée à la présence de gènes de virulence et de résistance aux antibiotiques. Malheureusement, au cours des cinquante dernières années, certaines souches de *S. aureus* ont acquis une résistance à un large spectre d'antibiotiques pour devenir une cause majeure d'infections nosocomiales mais aussi d'épidémies chez les individus sains en dehors de l'hôpital.

Jusqu'à peu, les scientifiques pensaient que *S. aureus* utilisait principalement la conjugaison et la transduction, deux importants modes de transfert horizontal de gènes (THG), pour acquérir de nouvelles séquences génétiques, notamment associées à des gènes de résistance aux antibiotiques. Or la récente démonstration de la capacité de *S. aureus* à induire la compétence naturelle pour la transformation génétique, le troisième mode de THG chez les bactéries, a quelque peu changé notre façon d'appréhender le THG chez ce pathogène majeur.

#### 1.2 État de l'art actuel

Pour pouvoir être génétiquement transformable, les bactéries doivent basculer dans un état physiologique particulier appelé « compétence naturelle ». Cette différenciation cellulaire est

induite en réponse à des stimuli environnementaux. Ces signaux, souvent associés à des stress, sont intégrés par la cellule, qui induit des cascades complexes de régulation qui aboutissent à l'activation de régulateurs centraux de la compétence. Ces régulateurs, décrits et étudiés chez un grand nombre de bactéries modèles (#REF), contrôlent à leur tour l'expression d'une centaine de gènes, participant à l'établissement de l'état de compétence. Parmi ces gènes induits se trouvent les gènes essentiels à la transformation génétique.

La capacité de *S. aureus* à induire la compétence naturelle pour la transformation génétique n'a été décrite que récemment. En conséquence, les deux processus restent mal compris. Cependant, le peu d'informations disponibles soulignent déjà l'importance d'étudier cette différenciation physiologique chez cet important agent pathogène humain. En effet, chez *S. aureus* la régulation centrale de la compétence combine des régulateurs précoces trouvés à l'origine dans les deux autres organismes modèles : un facteur sigma alternatif cryptique, SigH (comparable à ComX chez *S. pneumoniae* 3), et deux gènes codant pour des homologues putatifs du facteur de transcription ComK (présent chez *B. subtilis* 4). De façon intéressante, l'utilisation de plusieurs régulateurs pourrait permettre à *S. aureus* d'être en mesure d'utiliser un plus large éventail de stimuli environnementaux pour décider de devenir ou non compétent pour la transformation génétique.

De plus, même si les publications originales peuvent être considérées comme des pierres angulaires dans le domaine, les conditions expérimentales décrites pour induire naturellement la compétence n'étaient clairement pas optimales et ont forcé les auteurs à surexprimer les régulateurs centraux (#REF). En contournant les signaux environnementaux naturels et les voies de transduction des signaux, cette stratégie a clairement limité la quantité d'informations disponibles.

## 2 Objectifs et stratégie

Au cours de ma thèse, j'ai profité de protocoles optimisés, développés dans le laboratoire d'accueil, pour induire efficacement cette adaptation environnementale (jusqu'à 80% de cellules compétentes dans une culture planctonique donnée) afin d'étudier le développement NATUREL de la compétence chez le pathogène humain *S. aureus*. Pour atteindre cet objectif, j'ai consacré mon travail de thèse à étudier trois questions principales :

- 1/ Quels sont les signaux environnementaux conduisant au développement NATUREL de la compétence chez *S. aureus* ?
- 2/ Quels sont les étapes et les acteurs intégrant ces signaux et transmettant l'information aux régulateurs centraux de la compétence ?
- 3/ Quel est le rôle exact de chaque régulateur central de la compétence ?

Pour répondre à ces questions, ma thèse a été divisée en deux principales parties :

- 1) tout d'abord, en utilisant la technique du double-hybride dans la levure, j'ai construit un réseau d'interaction protéine-protéine centré sur le développement de la compétence et la transformation génétique naturelle.
  - 2) Ensuite, dans la seconde moitié de mon doctorat, j'ai effectué des études génétiques, étudiant l'impact de la suppression de gènes candidats sur le développement de la compétence.
- 
- 1) Au cours de la première moitié de ma thèse, j'ai personnellement construit une banque génomique de *S. aureus* chez la levure. Ensuite, j'ai criblé cette banque, en utilisant la technique du double hybride dans la levure, pour identifier des protéines ou des peptides capables d'interagir avec les trois régulateurs de compétence centrale et les protéines d'intérêt impliquées lors de la transformation génétique. Les partenaires potentiels identifiés dans cette première partie de ma thèse ont été essentiels pour proposer des hypothèses et répondre aux trois questions citées ci-dessus.
  - 2) Au cours de la seconde moitié de mon doctorat, j'ai mené des études génétiques pour étudier, au niveau transcriptionnel, la régulation de la compétence naturelle chez *S. aureus*. Pour ce faire, j'ai construit une série de fusions transcriptionnelles entre certains promoteurs



d'intérêt et un gène rapporteur : le gène codant pour la « Green Fluorescent Protein » (GFP). Ces fusions transcriptionnelles ont été utilisées pour étudier la régulation transcriptionnelle des trois régulateurs de compétence centraux et de plusieurs gènes de compétence tardive dans différents contextes génétiques. C'est pourquoi j'ai également construit une série de délétions de gènes afin de comprendre l'impact de régulateurs candidats sur le développement de la compétence.

En utilisant cette approche, j'ai investigué deux projets principaux : d'abord, j'ai montré comment lors d'une croissance de le milieu CS<sub>2</sub>, tous les régulateurs centraux de compétence sont impliqués dans un développement complet et naturel de la compétence alors que seuls SigH et ComK1 activent l'expression des gènes de compétence tardive (i.e. impliqués dans la transformation génétique). Ensuite, j'ai montré comment en milieu GS (conditions anaérobies), ComK2 peut inhiber l'expression des gènes de transformation génétique, probablement via une interaction avec une protéine de fonction inconnue identifiée dans l'étude Y2H (1<sup>ère</sup> partie de ma thèse). Enfin, j'ai montré comment la concentration en oxygène peut moduler le développement naturel de la compétence. J'ai notamment démontré comment différentes conditions d'oxygène peuvent être intégrées par différents systèmes de détection d'oxygène pour activer ou inhiber le développement de la compétence.

### 3. Résultats

#### 3.1 Construction d'un réseau d'interaction protéine-protéine centré sur la compétence et la transformation génétique chez *S. aureus*

Afin de générer notre réseau d'interactions protéine/protéine, nous avons tout d'abord construit des bibliothèques d'ADN génomique de *S. aureus* dans la levure et criblé ces librairies avec plusieurs protéines impliquées à différents niveaux au cours du développement de la compétence et de la transformation génétique chez *S. aureus*: 3 régulateurs centraux de la compétence : SigH, ComK1 et ComK2 ; ainsi que les protéines impliquées dans la transformation génétique, permettant la fixation et le transport d'ADN exogène (ComGA, ComGB, ComC, ComEA, ComEB, ComEC, ComFA, ComFC et ComC), et contribuant à la recombinaison homologue dans le chromosome hôte (DprA et RecA).

Au total, sur l'ensemble de ces criblages, 896 candidats aux interactions ont été isolés et identifiés. Parmi ces candidats, 56 % viennent de la bibliothèque-N1, 19 % viennent de la bibliothèque-N2 et 25 % viennent de la bibliothèque-N3. D'après la séquence des inserts d'ADN, 30,36 % de ces clones (c'est-à-dire 272) ont exprimé une protéine de fusion en phase. Parmi ces clones en phase, 117 clones correspondaient à ComEB (dans les criblages ComEB), et 19 clones à SA2107 (dans les criblages ComK2). En fait, nous avons obtenu un total de 109 interactions potentielles. Elles sont toutes représentées dans notre réseau (Figure 13).

En plus des criblages de nos bibliothèques d'ADNg, nous avons également testé la capacité de nos protéines d'intérêt à interagir directement les unes avec les autres (une approche que nous appelons « matrice »). Ainsi, les 12 gènes entiers ainsi que les domaines indépendants présents en N- ou C-terminal de ces mêmes gènes ont été testés. Une telle stratégie nous a permis de tester l'hypothèse simple selon laquelle les protéines impliquées dans les mêmes voies ont une forte probabilité d'interagir directement. Un total de 23 constructions impliquant les protéines entières ou les domaines protéiques ont été testés pour leur capacité à interagir les uns avec les autres. Cette approche a fourni d'excellentes informations et également un soutien pour notre réseau. En effet, certaines interactions identifiées à partir de notre bibliothèque ont également été détectées dans la matrice, comme l'auto-interaction ComEB (tableau 11 et figure 12F). Au total, 6 interactions ont ainsi été identifiées, comme le montre la figure 12.

### **3.2 Régulation du transfert horizontal de gènes médié par la compétence naturelle chez *S. aureus***

Ici, en utilisant un protocole optimisé pour induire naturellement la compétence dans les cellules planctoniques, conduisant à une efficacité de transformation améliorée (jusqu'à  $1,5 \times 10^{-7}$ ), nous avons montré que les trois potentiels régulateurs centraux de la compétence chez *S. aureus* sont tous essentiels pour un développement complet de la compétence. En effet, nous avons constaté que les gènes impliqués dans la transformation naturelle peuvent être divisés en plusieurs classes en fonction des régulateurs contrôlant leur expression. De plus, nous avons montré que la disponibilité de l'oxygène est un signal important menant au développement de la compétence via l'induction de l'un des trois régulateurs centraux (probablement SigH). Nos résultats démontrent la complexité du développement des compétences chez *S. aureus*, par rapport à d'autres organismes modèles historiques. Nous prévoyons que nos résultats seront un point de départ pour prouver l'importance de la compétence naturelle et de la transformation génétique pour la plasticité génomique de *S. aureus*. De plus, nous pensons également que nos données nous permettront d'élucider les conditions environnementales conduisant à des acquisitions de virulence et de résistance aux antibiotiques in vivo, chez cet important agent pathogène humain.

### **3.3 Développement naturel de la compétence chez *S. aureus* dans des conditions d'anaérobie**

*S. aureus* favorise l'hypoxie dans les tissus, ce qui est un signal clé pour la formation de biofilm et une virulence accrue. *S. aureus* possède trois systèmes de détection de la concentration en oxygène qui se trouvent être trois systèmes de transduction à deux composants (TCS) : NreBC 7, SrrAB 8 et AirSR 9 . Il est important de noter que la publication originale démontrant la capacité de *S. aureus* à être naturellement transformable a déjà montré que la compétence naturelle pouvait être induite sous différentes disponibilités d'O<sub>2</sub> 10. C'est pourquoi nous avons décidé d'étudier comment l'absence d'O<sub>2</sub> pourrait affecter le développement de la compétence lorsque *S. aureus* est cultivé dans des conditions anaérobies en milieu GS.

Fait intéressant, nous avons montré que, comme dans le milieu CS2, SigH et ComK1 activent l'expression des gènes de transformation génétique lors du développement de la compétence dans le milieu GS. De plus, nous montrons que, dans des conditions d'anaérobiose, ComK2 est également impliqué dans le développement de la compétence, mais en tant qu'inhibiteur de l'expression de certains gènes de transformation génétique. Nous avons également proposé que cette activité d'inhibition pourrait être liée à une interaction entre ComK2 et une protéine à fonction non prédite, SA2107. Ensuite, nous avons démontré comment le manque d'O<sub>2</sub> dans le milieu GS inhibe le développement de la compétence via le TCS NreBC. Enfin, nous avons rassemblé des preuves suggérant que le TCS à détection d'oxygène NreBC et le régulateur de compétence central ComK2 pourraient fonctionner dans la même voie pour inhiber le développement de la compétence en réponse à une faible concentration en O<sub>2</sub>.

Tous ces résultats indiquent clairement que la disponibilité de l'O<sub>2</sub> régule le développement de la compétence chez *S. aureus* mais compte tenu de toutes les différentes conditions décrites ci-dessus, ces régulations doivent être complexes et intégrées à plusieurs niveaux.

## 4. Discussion

### 4.1 Le réseau d'interaction protéine à protéine centré sur la compétence naturelle et la transformation génétique

Au cours de la première année de ma thèse, j'ai généré une banque génomique de *S. aureus* chez la levure. Au cours de l'année suivante, j'ai criblé cette librairie, en utilisant la technique du double hybride dans la levure, avec différents appâts impliqués dans le développement de la compétence et de la transformation naturelle chez *S. aureus*. En parallèle, j'ai également testé, dans une expérience « matricielle », si les protéines choisies comme appâts pouvaient interagir directement entre elles. Au final, les interactions détectées lors des criblages et dans les expériences « matrices » sont toutes résumées dans notre réseau d'interaction protéine-protéine. Surtout, dans les différentes expériences que nous avons menées, des interactions très intéressantes ont été détectées qui pourraient être classées comme :

- interactions déjà validées dans d'autres organismes (comme l'interaction ComGA-RecA),
- des interactions jamais vues auparavant mais qui ont du sens car elles impliquent des protéines travaillant dans la même voie (comme l'interaction ComEA-ComEB)
- interactions complètement nouvelles avec des partenaires qui n'étaient pas censés interagir les uns avec les autres avant cette étude (comme l'interaction ComK2-SA2107)

Même si nous avons détecté des interactions très intéressantes, nous sommes un peu déçus par nos résultats de « screening de librairie ». En effet, le nombre total d'interactions potentielles semble faible par rapport aux études précédentes. Deux stratégies différentes pourraient être menées par le laboratoire d'accueil à l'avenir afin d'améliorer nos résultats :

1/ Par rapport aux précédentes librairies construites (*B. subtilis* ou *S. pneumoniae*, INRA, Jouy en Josas), la taille la plus abondante des fragments génomiques présents dans nos librairies semble un peu élevée (200 à 400pb contre 100-200pb dans bibliothèques précédentes). Par conséquent, une solution pourrait être de reconstruire nos bibliothèques en augmentant la quantité d'enzymes de restriction utilisées dans les digestions partielles afin de réduire la taille moyenne des fragments générés. Cette approche pourrait produire des fragments codant pour des peptides plus petits qui pourraient mieux détecter une interaction. Comme ce « problème » de taille n'est pour l'instant qu'une hypothèse, il faudrait le démontrer.

2/ A ce stade, il est également important de mentionner que nous collaborons avec une autre équipe de l'institut, dirigée par F.X. Barre, avec qui nous avons travaillé sur un nouveau

protocole pour générer des bibliothèques génomiques. Les banques de proies sont ainsi construites sur la base de protocoles de biologie moléculaire associés au séquençage à haut débit. En résumé, l'ADN génomique est fragmenté par sonication. Les extrémités des fragments sont réparées par les mêmes techniques que celles utilisées pour l'analyse de fréquence de marqueurs (MFA) ou la capture de conformation chromosomique (3C). Les fragments génomiques sont ensuite clonés dans nos trois vecteurs afin que chaque fragment soit exprimé dans la bonne phase.

Enfin, de nombreuses interactions attendues, car déjà décrites dans la littérature, n'ont pas été confirmées dans notre étude, ni dans le criblage ni dans les expériences « matrice » 11. Plusieurs explications peuvent être proposées pour élucider cela :

(i) Les fragments générés par digestion partielle ne codent pas pour des peptides qui pourraient révéler une interaction importante

(ii) Les protéines homologues d'espèces différentes, fusionnées au domaine activateur de GAL4, n'adoptent pas exactement la même conformation tridimensionnelle, révélant ou non une interaction

(iii) L'interaction attendue n'existe pas *in vivo* dans le nouvel organisme modèle considéré

Pour vérifier quel scénario explique la non-détection d'une interaction attendue, nous devons tester à nouveau l'interaction dans de nouvelles bibliothèques chez la levure (voir ci-dessus) ou directement *in vivo*, chez *S. aureus*, via des techniques alternatives telles que la microscopie.

## **4.2 Régulation du transfert horizontal de gènes médié par la compétence naturelle chez *S. aureus***

Au cours de ma thèse, j'ai montré dans différents projets comment le développement de la compétence chez *S. aureus* peut être influencé (activé ou inhibé) par différentes conditions de croissance associées à différentes disponibilités en oxygène.

Dans un premier temps, j'ai proposé qu'au cours de la croissance en CS<sub>2</sub>, la concentration en oxygène diminue progressivement. Ce changement dans la disponibilité de l'oxygène est détecté par le TCS SrrAB, qui favorise (potentiellement par le biais de SigH) le développement de la compétence.

En parallèle, j'ai montré que dans des conditions d'anaérobie claires (c'est-à-dire croissance statique en milieu GS), le manque d'oxygène est détecté par un deuxième TCS, NreBC, qui inhibe la compétence, potentiellement via le troisième régulateur central de la compétence, ComK2.

Enfin, des chercheurs de l'équipe d'accueil ont réalisé des expériences en utilisant des souches de *S. aureus* exprimant la GFP en tant que rapporteur transcriptionnel et cultivées dans des plaques à 96 puits à l'intérieur de notre lecteur de plaques. Ce montage expérimental nous a permis de moduler la disponibilité en oxygène en (i) appliquant, ou non, un film au sommet des puits empêchant tout échange gazeux et (ii) en modifiant le volume de la culture dans le puits, ce qui affecte le volume d'air (et d'oxygène) disponible, au moins en présence du film. Fait intéressant, lorsque le film était absent, le développement de la compétence est resté relativement faible et quasiment pas affecté par le volume de la culture. Au contraire, lorsque le film était présent, le développement de la compétence était fortement influencé par le volume de la culture : les puits avec un volume de culture important (250-300  $\mu$ l) induiraient mal la compétence (1% de cellules compétentes), tandis que les puits avec un faible volume (25 à 50  $\mu$ l) induirait pleinement la compétence (jusqu'à 25 % de cellules compétentes).

On peut donc imaginer un modèle dans lequel les conditions environnementales optimales pour le développement de la compétence nécessitent une concentration en oxygène spécifique :

(I) À haute disponibilité en oxygène, le développement de la compétence est inhibé par un mécanisme inconnu

(II) Lorsque la culture en CS2 atteint la phase stationnaire (tube falcon fermé ou puits avec film), la concentration en oxygène commence à diminuer et la compétence est induite

(III) Les conditions d'anaérobie (culture statique en milieu GS) inhibent le développement de la compétence, potentiellement via le TCS NreBC et le régulateur central ComK2

L'équipe de Morikawa a récemment décrit un nouveau protocole pour induire efficacement la compétence lors de la formation de biofilm <sup>12</sup>. Fait intéressant, lors de la croissance en biofilm, des gradients (en nutriment et en oxygène) sont créés parce que les cellules proches de la surface du biofilm consomment les nutriments et l'oxygène disponibles qui se raréfient ainsi dans les couches les plus profondes du biofilm <sup>13</sup>. Selon notre modèle, les cellules compétentes détectées par Morikawa feraient partie de couche spécifiques à l'intérieur du

biofilm, là où la concentration en oxygène diminue suffisamment mais pas trop pour être au stade II (Fig. 3).

Pour tester une telle hypothèse, des mesures précises de la concentration en oxygène en corrélation avec le développement du programme transcriptionnel de compétence tout au long de la croissance dans diverses conditions environnementales seront essentielles. Pour ce faire, nous proposons d'utiliser un nouveau système développé par la société Presens : en effet, cette entreprise a développé des capteurs d'oxygène optiques, non invasifs, capables de mesurer la pression partielle d'oxygène dissous et gazeux. Ces capteurs sont fixés sur la surface intérieure du verre transparent ou de la matière plastique. La concentration en oxygène peut alors être mesurée sans contact et non destructive depuis l'extérieur, à travers la paroi du contenant.

Un tel équipement serait essentiel pour nos futures expériences car il nous permettrait de mesurer la concentration en oxygène tout au long de la croissance en CS2 dans des tubes fermés de type Falcon, en GS dans des tubes Eppendorf fermés et statiques mais également en CS2 et GS dans des plaques 96 puits filmées. Ces mesures d'oxygène seront également compatibles avec les expériences GFP et Luciférase.



## 5 Références

1. C. Ryan, K.R., *Sherris medical microbiology*. McGraw Hill, 2004.
2. Lowy, F.D., *Medical progress - Staphylococcus aureus infections*. . New England Journal of Medicine 1998. **339**: p. 520–532
3. Kuroda, M., et al., *Whole genome sequencing of meticillin-resistant Staphylococcus aureus*. The Lancet, 2001. **357**(9264): p. 1225-1240.
4. Annette Fagerlund, P.E.G.a.L.S.H., *Staphylococcus aureus competence genes: mapping of the SigH, ComK1 and ComK2 regulons by transcriptome sequencing*. Molecular Microbiology, 2014. **94**(3): p. 557–579.
5. Carreau, A., et al., *Why is the partial oxygen pressure of human tissues a crucial parameter? Small molecules and hypoxia*. Journal of cellular and molecular medicine, 2011. **15**(6): p. 1239-1253.
6. Haag, A. and F. Bagnoli, *The Role of Two-Component Signal Transduction Systems in Staphylococcus aureus Virulence Regulation*. Current topics in microbiology and immunology, 2016. **409**.
7. Müllner, M., et al., *A PAS Domain with an Oxygen Labile [4Fe-4S]<sub>2</sub><sup>+</sup> Cluster in the Oxygen Sensor Kinase NreB of Staphylococcus carnosus*. Biochemistry, 2008. **47**(52): p. 13921-13932.
8. Ulrich, M., et al., *The staphylococcal respiratory response regulator SrrAB induces ica gene transcription and polysaccharide intercellular adhesin expression, protecting Staphylococcus aureus from neutrophil killing under anaerobic growth conditions*. Molecular Microbiology, 2007. **65**(5): p. 1276-1287.
9. Yan, M., et al., *The Essential Two-Component System YhcSR Is Involved in Regulation of the Nitrate Respiratory Pathway of Staphylococcus aureus*. Journal of Bacteriology, 2011. **193**(8): p. 1799-1805.
10. Kazuya Morikawa, A.J.T., Yumiko Inose, Melody Tsai, Le Thuy Nguyen Thi, Toshiko Ohta, Tarek Msadek, *Expression of a Cryptic Secondary Sigma Factor Gene Unveils Natural Competence for DNA Transformation in Staphylococcus aureus*. PLoS Pathog, 2012. **8**(11): p. e1003003.
11. Mortier-Barrière, I., et al., *A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA*. Cell, 2007. **130**(5): p. 824-836.
12. Maree, M., et al., *MRSA emerges through natural transformation*. bioRxiv, 2020: p. 2020.12.08.415893.
13. Stewart, P.S. and M.J. Franklin, *Physiological heterogeneity in biofilms*. Nature Reviews Microbiology, 2008. **6**(3): p. 199-210.

## Table of Contents

Introduction .....	25
<b>1. <i>Staphylococcus aureus</i></b> .....	<b>27</b>
<b>1.1. General presentation</b> .....	<b>27</b>
<b>1.2. Pathogenicity and virulence of <i>S. aureus</i></b> .....	<b>29</b>
<b>1.2.1. <i>S. aureus</i> virulence mechanisms</b> .....	<b>29</b>
<b>1.2.2. <i>S. aureus</i> virulence factors</b> .....	<b>33</b>
<b>1.3. <i>S. aureus</i> environmental adaptation and regulatory networks</b> .....	<b>41</b>
<b>1.4. <i>S. aureus</i> resistance to antibiotics</b> .....	<b>51</b>
<b>2. Genetic transformation and competence development</b> .....	<b>61</b>
<b>2.1. Genetic transformation</b> .....	<b>61</b>
<b>2.2. Development of natural competence in important model organisms</b> .....	<b>67</b>
<b>2.3. Competence development in <i>S. pneumoniae</i></b> .....	<b>69</b>
<b>2.3.1. Presentation of the bacterium</b> .....	<b>69</b>
<b>2.3.2. Griffith's historical experiment</b> .....	<b>69</b>
<b>2.3.3. Competence regulatory pathways in <i>S. pneumoniae</i></b> .....	<b>71</b>
<b>2.4. Competence development in <i>Bacillus subtilis</i></b> .....	<b>77</b>
<b>2.4.1. Presentation of the bacterium</b> .....	<b>77</b>
<b>2.4.2. Competence regulatory pathways in <i>B. subtilis</i></b> .....	<b>77</b>
<b>2.5. Competence development in <i>Vibrio Cholerae</i></b> .....	<b>85</b>
<b>2.5.1. Presentation of the bacterium</b> .....	<b>85</b>
<b>2.5.2. Competence regulatory pathways in <i>V. cholerae</i></b> .....	<b>85</b>
<b>2.6. Competence development in <i>S. aureus</i></b> .....	<b>89</b>
<b>2.6.1. Current state of the art</b> .....	<b>89</b>
2.6.1.1. <i>S. aureus</i> has all the genes required for competence development and natural transformation .....	89
2.6.1.2. <i>S. aureus</i> has three potential central competence regulators.....	89
2.6.1.3. Competence in <i>S. aureus</i> is induced in CS2 and GS medium .....	93
2.6.1.4. Global transcriptional analysis of SigH, ComK1 and ComK2 regulons .....	95
2.6.1.5. TCS VraSR, AgrCA and BraSR are involved in the expression of the <i>comG</i> promoter..	97
2.6.1.6. <i>S. aureus</i> is able to accomplish genetic transformation .....	97
2.6.1.7. Biofilm growth conditions induce efficient natural genetic transformation .....	99
<b>Objectives</b> .....	<b>103</b>
<b>Results</b> .....	<b>107</b>

<b>Chapter 1 : Construction of a protein-protein interaction network centered on competence and genetic transformation in <i>S. aureus</i></b> .....	109
1.1. The principle of yeast two-hybrid (Y2H).....	111
1.2. Construction of a <i>S. aureus</i> genomic DNA library in <i>E. coli</i> .....	111
1.3. Transfer of the libraries in the yeast strain PJ69-4 $\alpha$ .....	119
1.4. Identification of protein-protein interactions.....	127
1.5. Network of protein-to-protein interactions centred on competence development and genetic transformation in <i>S. aureus</i> .....	133
1.6. Discussion on our interesting interactions.....	137
1.7. Materials and Methods.....	149
1.8. Reference.....	169
<b>Chapter 2: Regulation of natural competence-mediated horizontal gene transfer in <i>S. aureus</i></b> .....	171
2.1. Summary.....	173
2.2. Introduction.....	175
2.3. Results.....	181
2.3.1. New protocols to naturally induce competence with high efficiency in CS2.....	181
2.3.2. Competence naturally develops at a specific cell density and growth phase in CS2....	181
2.3.3. Genetic transformation is improved.....	183
2.3.4. Is Quorum Sensing involved in competence development in <i>S. aureus</i> ?.....	185
2.3.5. <i>PcomG</i> , <i>Pssb</i> , <i>PcomC</i> and <i>PcomF</i> are not controlled by the same regulators.....	185
2.3.6. SigH and ComK1 control the expression of the natural transformation genes, not ComK2.....	187
2.3.7. SigH and ComK1 are essential for natural tranformation, not ComK2.....	187
2.3.8. SigH, ComK1 and ComK2 are all required for a full development of competence.....	189
2.3.9. Competence development is associated with virulence inhibition.....	189
2.3.10. In CS2, <i>PcomK1</i> is regulated by SigH.....	191
2.3.11. Oxygen concentration is an important factor to induce competence development .	193
2.4. Discussion.....	195
2.5. Materials and Methods.....	209
2.6. References.....	221
<b>Chapter 3: Natural development of competence in <i>S. aureus</i> under anaerobe conditions</b> .....	223
3.1. Introduction.....	225
3.2. Results.....	231
3.2.1. Anaerobic conditions can induce the expression of <i>comK1</i> , <i>comGA</i> , <i>comC</i> and <i>ssb</i> ....	231
3.2.2. ComK2 inhibits the development of competence under anaerobic conditions.....	231
3.2.3. SA2107 interacts with ComK2 and inhibits the development of competence under anaerobic conditions.....	233

3.2.4. NreC inhibits <i>comGA</i> expression under anaerobic conditions.....	235
3.2.5. NreC inhibits <i>comGA</i> expression through ComK2 under anaerobic conditions.....	235
3.3. Discussion .....	237
3.4. Materials and Methods .....	243
3.5. References .....	251
<b>Conclusions and Perspectives .....</b>	<b>253</b>
<b>References .....</b>	<b>269</b>



# Introduction

A

*Staphylococcus aureus*  
Domain: Bacteria  
Phylum: Firmicutes  
Class: Bacilli  
Order: Bacillales  
Family: Staphylococcaceae  
Genus: *Staphylococcus*  
Species: *S. aureus*

B



**Figure 1 Brief overview of *S. aureus*.**

A. Scientific classification. B. *S. aureus* N315 colonies on BHI agar.

# 1. *Staphylococcus aureus*

## 1.1. General presentation

In 1880, Scottish surgeon Sir Alexander Ogston first described staphylococci in pus from a surgical abscess in a knee joint <sup>1</sup>. He named this new species after the organization of the cells that looked like bunches of grapes under the microscope. In 1884, German physician Friedrich Julius Rosenbach differentiated two kinds of species by the color of their colonies: *Staphylococcus aureus* (from the Latin *aurum*, gold) and *Staphylococcus albus* (Latin for white). *S. albus* was later renamed *Staphylococcus epidermidis* because of its ubiquity on human skin <sup>1</sup>.

*S. aureus* belongs to the phylum of Firmicutes, the class of bacilli, the order of Bacilliales and the family of Staphylococcaceae (Figure. 1A). *S. aureus* cells, observed by microscopy, appear in the form of clusters, like grapes. The colonies are large, round, golden-yellow and when they are grown on blood agar plate (Figure. 1B), there is hemolysis <sup>2</sup>. The reproduction of *S. aureus* is asexual and occurs by binary fission. After fission, the two daughter cells remain attached, explaining why the bacteria are in clusters. *S. aureus* is a Gram-positive coccal bacterium, facultative anaerobic. The particularity of a facultative anaerobe is the capacity to convert ATP by aerobic respiration in the presence of oxygen and to switch for anaerobic respiration (fermentation) in the absence of oxygen <sup>2</sup>. *S. aureus* also produces the catalase enzyme which transforms hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) to oxygen and water. This catalase-activity is used to distinguish staphylococci from streptococci and enterococci. Previously, the coagulase test (conversion of fibrinogen to fibrin by the coagulase enzyme) was used to do this distinction, but it appears that all *S. aureus* strains are not coagulase-positive <sup>2</sup>.

*S. aureus* colonizes an important part of the human population (30% of persistent carrier and 30% of intermittent carrier). This pathogen has the potential to cause a large variety of complications, ranging from benign skin infections to severe diseases such as septicemia, osteomyelitis and endocarditis <sup>3</sup>. The mortality associated to *S. aureus* infections is correlated to the presence of virulence and antibiotic resistance genes. Over the past fifty years, *S. aureus* strains have acquired resistance to a broad spectrum of antibiotics to become a major cause of hospital-acquired infections but also of outbreaks among healthy individuals outside from the hospital.





## 1.2. Pathogenicity and virulence of *S. aureus*

In contrast to many other bacterial pathogens, which often rely on few toxins, *S. aureus* produces an astounding array of virulence factors to promote disease <sup>4</sup>. These include toxins, immune evasion factors and a vast array of protein and non-protein factors that enable the establishment and maintenance of an infection by colonization and immune evasion mechanisms.

### 1.2.1. *S. aureus* virulence mechanisms

#### 1.2.1.1. Origins of infection

*S. aureus* infections usually originate from asymptomatic colonization or, more rarely, in the hospital setting, from infected medical devices. The nares are traditionally regarded as the main colonization site, but *S. aureus* can also colonize many skin sites in addition to the intestine. Colonization of multiple body sites is usually observed and originate from frequent touching and nose picking <sup>5</sup>. Furthermore, it was reported that *S. aureus* nasal carrier with rhinitis can disperse high loads of *S. aureus* into the environment and may be the source of an outbreak of *S. aureus* infections <sup>6</sup>. In addition, food poisoning is a special case of acute *S. aureus* infection in which contaminated foods (containing toxins) are ingested <sup>7</sup>. Systemic infections originating from acute staphylococcal food poisoning are very rare. Indeed, whether sustained intestinal colonization by *S. aureus* can lead to gastro-intestinal and even systemic disease is not known yet. Finally, *S. aureus* can use, in an opportunistic fashion, a primary harm done by other pathogens or pre-existing conditions. This is for example the case during lung infections that have been initiated by a virus, such as the flu, during which *S. aureus* secondary infection is often the ultimate cause of death <sup>8</sup>.

It is also important to mention that systemic *S. aureus* infection is always dependent on bacterial breach through the epithelial protective layer. This is for example the case during skin infections that can develop from minor scratches and become more invasive <sup>9</sup>. Once the epithelial layer is breached, *S. aureus* can gain access to the circulatory system and seed a variety of vital organs, leading to more severe infections, such as septicemia, osteomyelitis and endocarditis <sup>10</sup>.

#### 1.2.1.2. Establishment of infection

After epithelial breach and systemic invasion, the success of a staphylococcal infection depends on the effective evasion of host defences. While in the bloodstream, *S. aureus* uses an important set of mech-



-anisms to avoid elimination by the host: toxins destroying phagocytes; mechanisms that trigger phagocytes apoptosis; inhibition of phagocytosis by aggregation, protective surface structures and biofilm formation; inhibition of neutrophil activation, chemotaxis and killing mechanisms.

In addition, invasion of organs and tissues by *S. aureus* not only requires immune evasion but also adhesion and further structural processes. Many of these steps are facilitated by surface-anchored proteins, tethered to the peptidoglycan, that interact with host cell receptors <sup>4</sup>. Once an organ is invaded, *S. aureus* forms encapsulated abscesses where bacterial proliferation inhibits additional infiltration of leukocytes. Interestingly, such high cellular densities are associated with nutrients rarefaction. To adapt with such conditions, *S. aureus* produces a series of cytolysins that can lyse cells and enzymes to digest the released nutrient macromolecules.

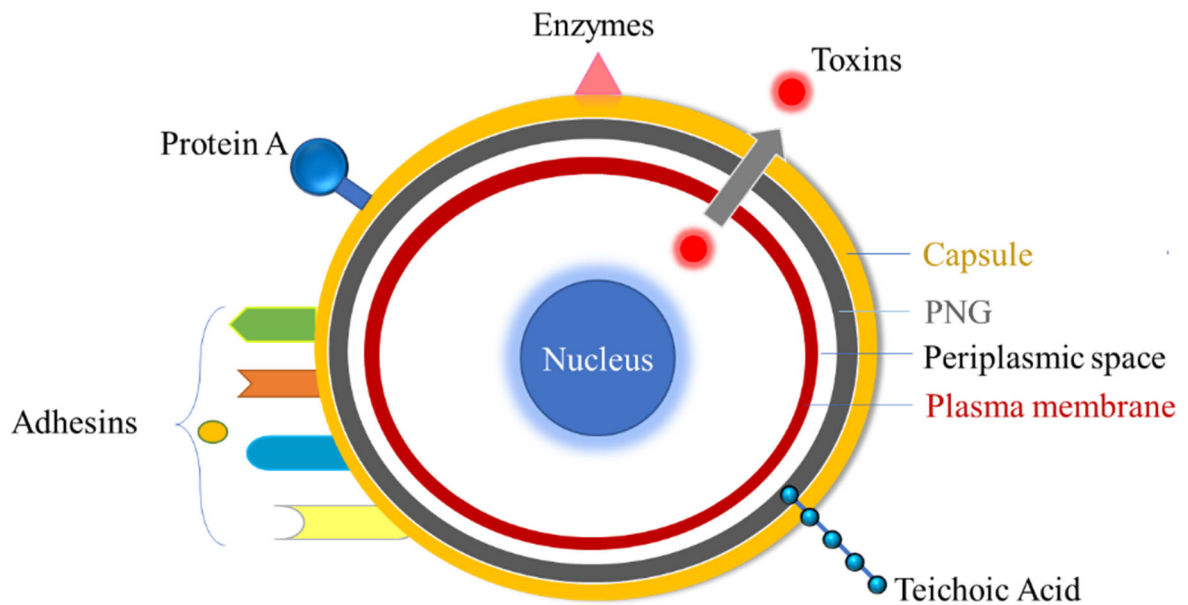
#### 1.2.1.3. Maintenance of infection

*S. aureus* utilizes different approaches to maintain an infection: matured abscesses, biofilm formation or persistence in phagocytic and non-phagocytic cells. All these strategies have in common to allow the bacteria to hide from attacks by the immune system by finding a shelter within an isolated structure. Rupture of the abscess, detachment from the biofilm and escape from the infected cell contribute to the spread of *S. aureus* and the chronicity of the infection.

#### The example of *S. aureus* biofilms.

An important way for *S. aureus* to maintain an infection is the formation of a microbially-derived sessile community, called biofilm. These can form on abiotic material of medical devices but also on tissue surfaces, such as on heart valves during endocarditis. Biofilm formation develops in three main stages: adhesion, proliferation and detachment. The multilayered biofilm generated by *S. aureus* is embedded within a glycocalyx or slime layer with heterogeneous protein expression throughout <sup>11</sup>. The main components of the biofilm include polysaccharide intercellular antigen (PIA) and/or proteinaceous cell-to-cell adhesions, as well as extracellular DNA. PIA is produced *in vitro* via products of the intercellular adhesion (*ica*) locus <sup>12</sup>. The staphylococcal respiratory response regulator, SrrAB, is essential for PIA induction under anaerobic environmental conditions by binding to a 100 bp DNA sequence upstream of the *icaADBC* operon <sup>13</sup>.

The main role of a biofilm during infection is to protect the bacteria from phagocytes attack. However, detachment of cells clusters from the biofilm can occur because of the extensive activity of degradative



**Figure 2. The virulence factors of *S. aureus*.**

Virulence factor	Functions
Capsule	Inhibition of phagocytosis and promotion of adherence
Peptidoglycan	Pyrogenicity, complement activation and inhibition of leukocyte migration
Teichoic Acid	Adherence to mucosal surfaces

**Table 1. Virulence factors associated with exopolysaccharides.**

exoenzymes (such as proteases) <sup>14</sup>. Therefore, the biofilms formed by *S. aureus* allows the infection to evolve from acute to persistent and chronic.

### 1.2.2. *S. aureus* virulence factors

The success of *S. aureus* as a pathogen and its ability to cause such a wide range of infections are the result of its extensive set of virulence factors <sup>15</sup>. Evidence of *S. aureus* virulence was first shown in 1941, when Skinner and Keefer reported that the mortality rate associated with *S. aureus* bacteremia in 122 patients at the Boston City Hospital was 82% <sup>16</sup>.

To date, a myriad of virulence factors has been discovered in *S. aureus*. Generally, virulence factors are the molecules produced by microorganisms that enable them to colonize the host, evade (or inhibit) host's immune response, shuttle between inside and outside of cells, and obtain nutrition from the host. For example, these include exopolysaccharides, surface-associated protein adhesins, proteins or enzymes related to immune evasion, and extracellular proteins including a plethora of toxins (Figure 2) <sup>17</sup>. Some factors are secreted in the extracellular environment (like toxins) while many others are part of the structural components of the bacterial cell, the majority of which are distributed on the surface of *S. aureus*.

#### 1.2.2.1. Exopolysaccharides

Exopolysaccharides are synthesized and secreted into the external environment or synthesized extracellularly by cell wall-anchored enzymes (Table 1). For example, layers of surface-associated covalently bound polysaccharides, structurally positioned as the outermost covering of bacterial cells are referred to as the capsule <sup>18</sup>. The capsule provides multiple functions to the bacterial cells such as adhesion receptors during colonization of tissues, absorption of water to prevent desiccation or enhancement of staphylococcal virulence by impeding phagocytosis. In *S. aureus*, capsular polysaccharides are important virulence factors which are currently under evaluation as potential vaccine antigens <sup>19</sup>.

Underneath the capsule, the cell wall layer is made of peptidoglycan, which is a complex mesh of sugars (glycan) crosslinked with pentapeptides. Glycan is composed of alternating units of N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG) molecules. The peptidoglycan provides a rigid support for the bacteria and maintains the characteristic shape of the cells.

<b>Abbreviation</b>	<b>Name and function</b>
Cna	Collagen-binding adhesin
Clfs	Clamp factors
Eap	Extracellular adherence protein
Efb	Extracellular fibrinogen-binding protein
Emp	Extracellular matrix protein
FnBPA,B	Fibronectin binding proteins A, B
Pls	Plasmin-sensitive protein
SpA	<i>S. aureus</i> protein A
vWbp	Von Willebrand factor binding protein

**Table 2. *S. aureus* virulence factors: adhesins**

Teichoic acid (TAs) are bacterial polymers of glycerol or ribitol phosphate and carbohydrates linked via phosphodiester bonds. There are two types of TAs: Lipo-Teichoic Acid (LTA), anchored in the lipid membrane and Wall Teichoic Acid (WTA), covalently bound to peptidoglycan. TAs contribute to staphylococcal adhesion, colonization, cell division and biofilm formation [20](#).

#### 1.2.2.2. Staphylococcal protein A

The staphylococcal protein A (SpA) is an important virulence factor that has multifaced roles. It has been demonstrated that SpA has an affinity to the Fc receptors of the immunoglobulin making *S. aureus* inaccessible to the opsonin (proteins of the innate and adaptive immune system that facilitate phagocytosis and cell lysis by “marking” antigens) thus protecting it from being engulfed and subsequently killed by the immune cells [21](#). SpA can also stimulate inflammation in the lung by binding to the receptor for tumor necrosis factor 1 (TNFR-1) [22](#). Finally, SpA facilitates the adherence of *S. aureus* to von Willebrand factor (vWF)-coated surfaces, an important step for adhesion [23](#).

#### 1.2.2.3. Adhesins

*S. aureus* has developed a wide variety of molecules to assist its adhesion to different substrate. Through adhesion, *S. aureus* realizes colonization and, eventually, infection of the host. This process is a required phase for pathogenesis of *S. aureus*.

These molecules are named adhesins (Table 2). They not only are anchored on the surface of the bacteria, but also can be secreted. The adhesins presented on the bacterial surface are called Microbial Surface Recognizing Adhesive Matrix Molecules (MSCRAMMs) [24](#), while the secreted molecules are called Secretable Expanded Repertoire Adhesive Molecules (SERAMs) [25](#). Most adhesins are proteins, except for Polysaccharide Intercellular Adhesin (PIA) and TAs.

Typical members of the MSCRAMMs family are SpA (presented above), Fibronectin binding proteins A and -B (FnBPA and -B), collagen-binding protein (CNA) and clumping factor (Clf) A and B proteins [24](#). FnBPA and -B tie fibrinogen to allow a bond between *S. aureus* and the host integrins, which have an important role in the colonization of the host [26](#). CNA participates in bacterium-host adherence and in immune evasion [27](#). ClfA is the major staphylococcal fibrinogen-binding protein and is responsible for the observed clumping of *S. aureus* in blood plasma [28](#).



<b>Abbreviation</b>	<b>Name and function</b>
Hla	$\alpha$ -toxin
Hlb	$\beta$ -toxin
Hld	$\delta$ -toxin
Hlg	$\gamma$ -toxin;
PVL	Panton-Valentine Leukocidin
LukAB/GH	Leukocidins AB /GH
LukD/E/	Leukocidins D/E

**Table 3. *S. aureus* virulence factors: toxins.**

<b>Abbreviation</b>	<b>Name and function</b>
Sak	Staphylokinase
OatA	O-acetyltransferase
ApsS/R/X	Antimicrobial peptide sensor

**Table 4. *S. aureus* virulence factor: enzymes and proteins.**

The SERAMs mainly include the coagulase (Coa), the extracellular fibrinogen binding protein (Efb), the extracellular matrix binding protein (Emp), and the extracellular adhesive protein (Eap). Importantly, secreted adhesins are often associated with endovascular diseases [25](#).

#### 1.2.2.4. Toxins

Toxins play an important role in *S. aureus*' infectious capacity and its success as a pathogen. They are secreted by the bacteria and interfere directly with the host [29](#). The majority of toxins-encoding genes are located on mobile genetic elements (MGEs, discussed below). The main *S. aureus* toxins can be divided into three major families: first, the Pore Forming Toxins (PFTs), which may work in a receptor-mediated or receptor-independent way; then, the Exfoliative Toxins (ETs) that interfere with the host receptors functions, but are not membrane-damaging; and finally, the superantigens, which are secreted enzymes that degrade host molecules or affect important host defense mechanisms (Table 3).

Among these toxins, PFT can be further divided into 4 types: Hemolysin- $\alpha$  (Hla), Hemolysin- $\beta$  (Hlb), phenol-soluble modulins (PSMs) and leukotoxins [30](#).

Hla is a secreted water-soluble monomer. It can bind to a receptor on the host cell membrane and, then, oligomerize into a heptameric structure, forming a transmembrane channel. This toxin has been demonstrated to affect epithelial and endothelial cells, T cells, monocytes and macrophages. Hlb is produced in large quantities in animal isolates [31](#). It is cytotoxic towards a wide range of human cell types. It can contribute to phagosomal escape and induction of biofilm formation [32](#).

PSMs have multiple functions during staphylococcal pathogenesis. *S. aureus* produces high amounts of strongly cytolytic PSMs, in particular the PSMa peptides encoded by the *psma* locus. Importantly, PSMs trigger an inflammatory response by interacting with innate immune response receptors [33](#).

Leukotoxins are structurally similar to Hla. Four leukotoxins are found in *S. aureus*: Pantone-Valentine Leukocidin (PVL), gamma-Hemolysin (Hlg), Leukotoxin ED (LukE, LuKD) and Leukotoxin AB/GH (LukAB/LukGH). They are made of two different protein components that assemble together to form  $\beta$ -barrel pores [34](#).

#### 1.2.2.5. Enzymes and other proteins

There are many other proteins or peptides functioning as virulence factors in *S. aureus*. Examples are shown in Table 4. Some of these virulence factors have the activity of enzymes. The O-acetyltransfer-



-ase (OatA) can modify the cell wall to make *S. aureus* resistant to lysozyme [35](#). The Staphylokinase (Sak) can bind to  $\alpha$ -Defensins of human neutrophils to make *S. aureus* successfully colonize the host [36](#). Finally, the antimicrobial peptide sensor (ApsS/R/X) can bind and impair antimicrobial peptides to make *S. aureus* resistant to these peptides [37](#).

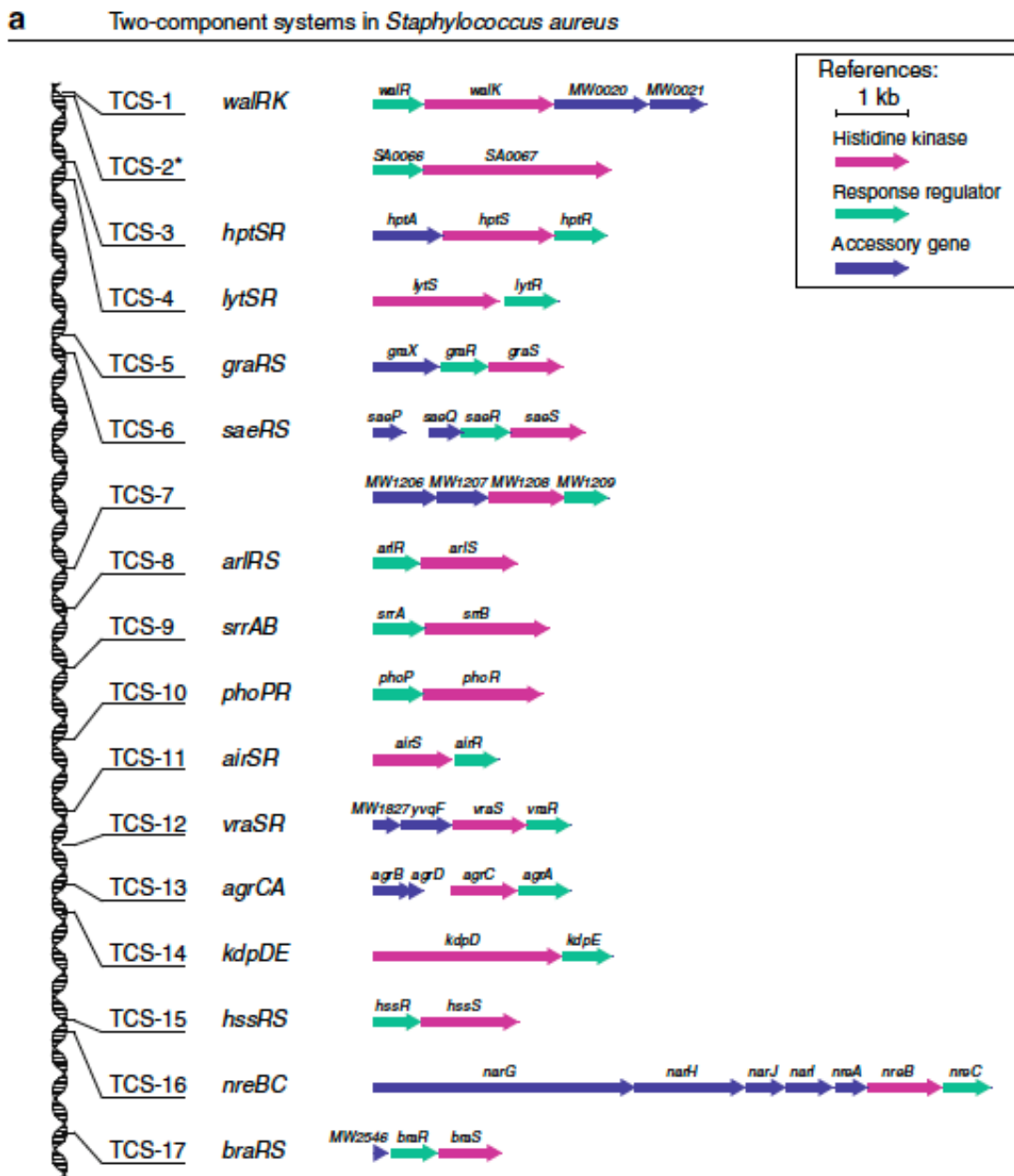


Figure 3. The 17 operons encoding TCSs in *S. aureus*' genome [38](#).

### 1.3. *S. aureus* environmental adaptation and regulatory networks

*S. aureus* has evolved a remarkably complex regulatory network to control virulence genes expression. One of the best studied mechanism relies on two-component systems (TCSs) that can sense various environmental signals and translate this information into a specific gene expression pattern. In addition, sigma factors and the newly discovered small RNAs also play an important role in controlling virulence gene expression. Sigma factors, which work with the RNA polymerase, control genes transcription. Small RNAs function through binding to proteins or by pairing with mRNAs to regulate their stability or translation. Thanks to these efficient strategies, *S. aureus* can respond and adapt to environmental signals, produce virulence factors and survive in the host. Therefore, understanding virulence genes regulation through these networks has important implications for drug and vaccine design.

#### 1.3.1. Two Component Systems (TCSs)

Two-component systems are the primary machinery to connect environmental stimuli and a bacterium's cellular physiology. A canonical TCS consist of a histidine kinase (HK) and a response regulator (RR). In response to extracellular stimuli, the transmembrane HK auto-phosphorylates on a conserved histidine residue. Then, the phosphorylated HK binds and transfer the phosphoryl group to a conserved aspartate residue located on the cytosolic RR. The phosphorylation of the RR activates an output domain, which can subsequently regulate gene expression, leading to changes in the cellular physiology. Most of *S. aureus* genomes encode a total of 16 conserved pairs of TCSs with an additional TCS present in the staphylococcal cassette chromosome *mec* of MRSA, which is linked to methicillin resistance <sup>38</sup> (Figure 3). Among these TCSs, AgrCA and SaeRS have been shown to be major regulators of virulence gene expression <sup>39</sup>. VraSR, GraXSR and BraRS are involved in antimicrobial peptides (AMPs) resistance and cell wall damage <sup>40</sup>. WalRK, ArlRS and LytSR are involved in cell wall metabolism, autolysis and cell death <sup>40</sup>. SrrAB, NreCBA and AirRS are involved in fermentation and nitrate metabolism in the absence of oxygen <sup>40</sup>. And finally, HssSR, KdpDE and PhoRP are involved in nutrient sensing and metabolism <sup>40</sup>.

#### The example of the *agr* system.

The accessory gene regulator (*agr*) system is one of the main global gene regulators controlling virulence expression and one of the best characterized systems in *S. aureus*. It was first identified as a

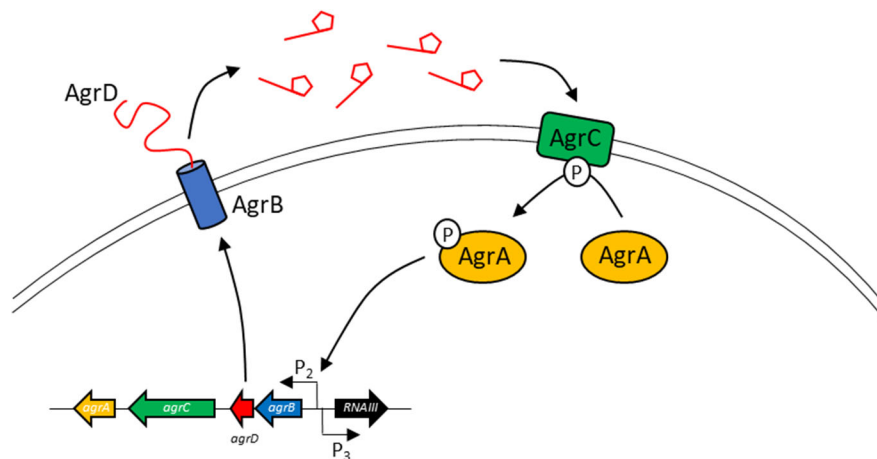


Figure 4. The Agr operon in *S. aureus*, an example of quorum sensing system.

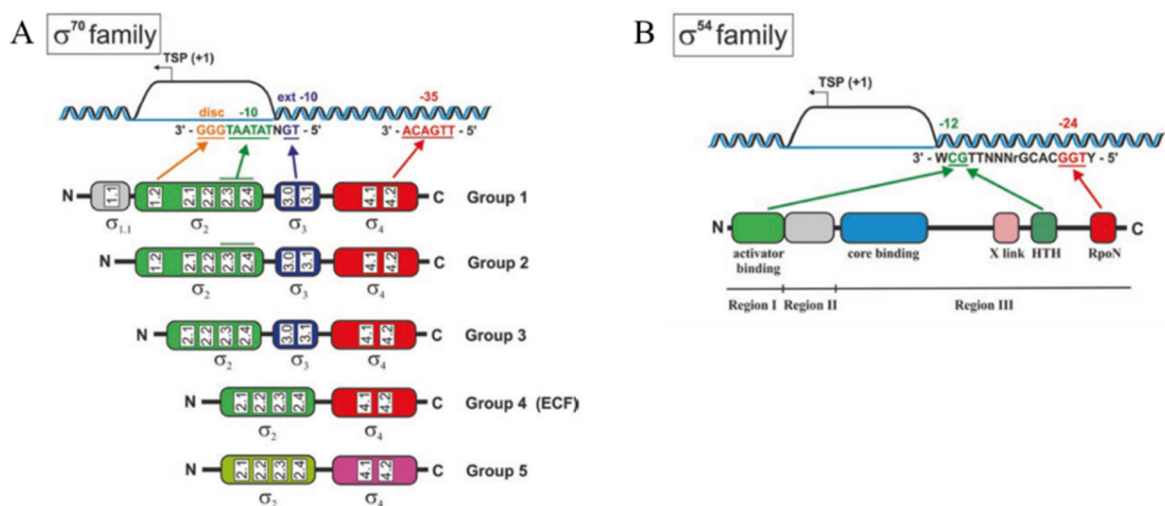


Figure 5. Structural domains of sigma factors and interactions with particular promoter elements [41](#).

A. Structural domains present in the 5 sub-groups in the  $\sigma^{70}$  sigma factor family. Structural data revealed that sigma factors of this family can be divided in 2 to 4 helical domains ( $\sigma_{1.1}$ ,  $\sigma_2$ ,  $\sigma_3$ ,  $\sigma_4$ ), connected by highly flexible linkers.

B. Structural domains present in  $\sigma^{54}$  sigma factor family. This family contains two highly conserved domains (Region I and III), allowing the DNA-binding activity directly recognizing the -24 and -12 promoter elements, separated by a flexible linker (Region II).

transposon-insertion mutant with a reduced capacity to produce multiple exotoxins [42](#). Importantly, *agr* is considered to be the best studied quorum sensing system in *S. aureus*. In this system, shown in Figure 4, an extracellular auto-inducing peptide (AIP) accumulates as the cell density increases. AIP is encoded by *agrD* and processed for export by AgrB. When the concentration of AIP reaches a threshold in the environment, the sensor kinase AgrC, transiently auto-phosphorylates when AIP binds to its extracellular sensor domain. Then, the phosphate is transferred to the response regulator AgrA. Phosphorylated AgrA binds to direct repeats within the P2 and P3 promoters of the *agr* locus. The P2 promoter drives the transcription of RNAII, that comprises the four genes of the quorum sensing module *agrBDCA*, forming an autoinduction cycle. In parallel, the P3 promoter drives the transcription of RNAIII (a stable regulatory RNA), which is the main effector of the quorum sensing response by either inducing or repressing expression of virulence genes.

### 1.3.2. Sigma factors

Sigma factors are subunits of all bacterial RNA polymerases [43](#). Together with the RNA polymerase, they are known as the RNA polymerase holoenzyme. Sigma factors provide promoter recognition specificity, contribute to DNA strand separation, and then dissociate from the core enzyme following transcription initiation [44](#). Functionally, structurally, and phylogenetically, sigma factors can be classified into two distinct families,  $\sigma^{70}$  and  $\sigma^{54}$ , based on their representatives in *Escherichia coli* [45](#).

According to the structural and functional similarities, the  $\sigma^{70}$  family is further sub-divided into four groups (Figure 5A, [41](#)). Group 1 contains essential primary sigma factors, closely related to  $\sigma^{70}$  of *Escherichia coli*, which directs expression of most housekeeping genes and contains four conserved domains ( $\sigma^{1.1}$ ,  $\sigma^2$ ,  $\sigma^3$ ,  $\sigma^4$ ). The sigma factors of group 2, 3 and 4 are termed alternative sigma factors. Alternative sigma factors enable bacteria to change the promoter specificity of the core RNA polymerase to enable the expression of genes that give them advantages in particular situations. Group 2 contains non-essential sequential homologues of primary sigma factors, lacking the  $\sigma^{1.1}$  domain. Several of them are involved in responses to a broad range of stresses. Group 3 contains only three domains ( $\sigma^2$ ,  $\sigma^3$  and  $\sigma^4$ ) and comprises structurally and functionally diverse alternative sigma factors, including flagellar sigma factors, heat shock sigma factors, general stress response sigma factors and sporulation sigma factors. Group 4 comprises distantly related extra-cytoplasmic function (ECF) sigma factors and contains only two domains ( $\sigma^2$  and  $\sigma^4$ ). The members of this group are largely involved in the sensing and responding to signals from outside of the cells or in the envelope stress response (Figure 5A).





The  $\sigma^{54}$  family is structurally and functionally distinct from the  $\sigma^{70}$  family (Figure 5B, [41](#)). In contrast to the transcription initiation mediated by sigma factors of the  $\sigma^{70}$  family,  $\sigma^{54}$ -dependent transcription absolutely requires hydrolysis of ATP, catalyzed by an activator protein bound to an upstream activator sequence (UAS).

#### 1.3.2.1. Primary sigma factor in *S. aureus*

In *S. aureus*, the primary sigma factor is named  $\sigma^{SA}$ , encoded by the *plaC* gene [46](#). Disruption of the *plaC* gene on the *S. aureus* chromosome is lethal. Analogous to other primary sigma factors, it is essential for *S. aureus* growth, and controls much of the house-keeping transcription, including *hla*, *sea* and *sec*. It was first identified by Rajendar Deora and Tapan K. Misra, based on its homology with  $\sigma^A$  from *B. subtilis* [46](#). In their work, the authors showed that purified  $\sigma^{SA}$  can cross-react with *B. subtilis*  $\sigma^A$  antibody, and that  $\sigma^{SA}$  associated with *E. coli* core RNAP, can initiate promoter-specific transcription from *S. aureus* and *E. coli* promoters. Investigated at the structural level,  $\sigma^{SA}$  is thermodynamically unstable, but is stabilized through binding with the core RNA polymerase [47](#).

#### 1.3.2.2. Alternative sigma factors in *S. aureus*

##### - Sigma B ( $\sigma^B$ )

The alternative sigma factor  $\sigma^B$  of *S. aureus* is by far the most widely studied.  $\sigma^B$  of *S. aureus* is encoded by the particularly conserved *sigB* operon, which encodes, in addition to  $\sigma^B$  itself, three *S. aureus* principal regulators: RsbU, RsbV and RsbW. Activation of  $\sigma^B$  depends upon RsbU in response to nutritional and physical stresses [48](#). The role of RsbV and RsbW in regulating  $\sigma^B$  activity in *S. aureus* has not yet been reported. Based on results obtained in *Bacillus cereus* [49](#), RsbV and RsbW could respectively act as anti-sigma factor antagonist and anti-sigma factor.

Identified by several proteomics and transcriptomics analyses in several strains of *S. aureus* under various conditions, more than 250 genes are positively or negatively regulated by  $\sigma^B$  [41](#). These genes are involved in many cellular processes, including cell envelope biosynthesis and turnover, intermediary metabolism, membrane transport, signaling pathways, oxidative stress resistance, pigmentation, protein secretion, biofilm formation [50](#), drug resistance, adaptation to stress and the progression of disease [51](#). Among these genes, many virulence factors are indirectly regulated by  $\sigma^B$  through its influence on the expression of major virulence regulators.

For example,  $\sigma^B$  regulates the level of RNAlII (presented above): while this key virulence regulator is repressed by  $\sigma^B$  in early exponential phase, it is activated upon entry into stationary phase (where  $\sigma^B$



is downregulated) which allows RNAIII to positively regulate the expression of a number of extracellular virulence factors (i.e. lipases, proteases, toxins) [52-54](#).

- Sigma S ( $\sigma^S$ )

In addition to  $\sigma^B$ , the ECF sigma factor  $\sigma^S$  is a novel alternative sigma factor in *S. aureus*, first described by Lindsey N. Shaw *et al.* [55](#). In this work, the authors used a biochemical approach to show that  $\sigma^S$  is able to associate with the core-RNAP, and initiate transcription from its own coding region. Moreover, they used a mutant strain to demonstrate that  $\sigma^S$  is important for *S. aureus* survival during starvation. Finally, they used a murine model of septic arthritis to establish that  $\sigma^S$  is an important component of *S. aureus* fitness and adaptation to stress.

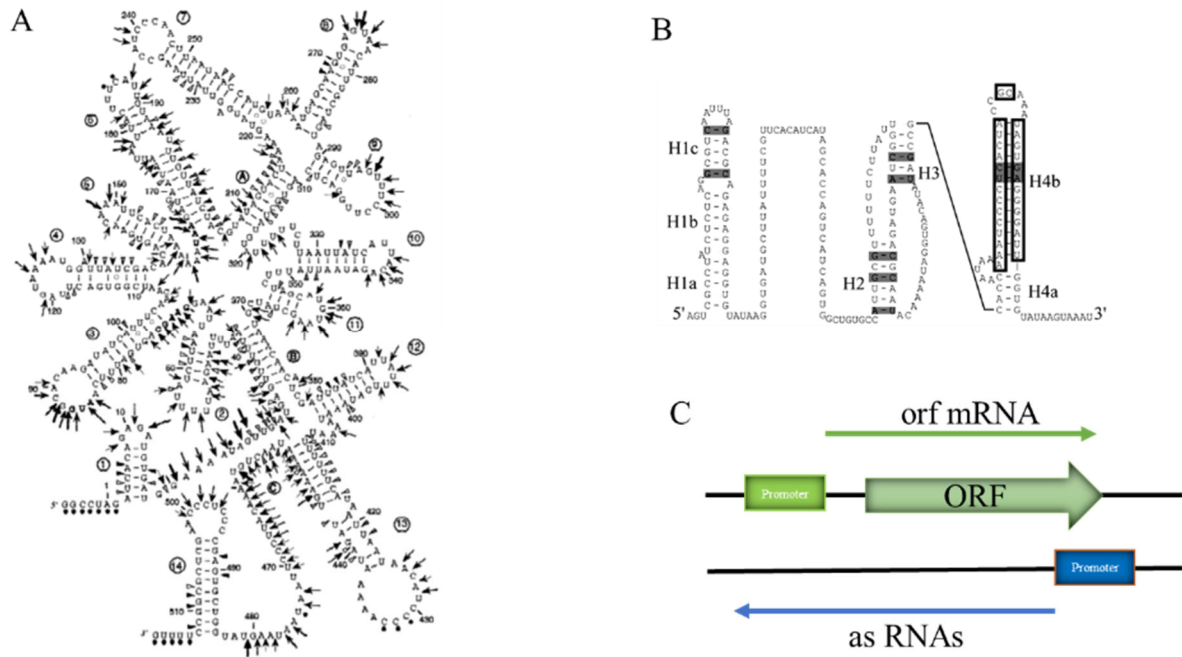
Further studies by the same laboratory showed that this sigma factor also plays an important role during the response to specific stresses, disease causation, the protection against DNA damage, cell wall disruption, and to allow the interaction with components of the innate immune system [56](#).

Finally, as shown in the work of Burda WN [57](#), the cysteine biosynthesis pathway transcriptional regulator, CymR, directly binds to the *sigS* promoter and negatively affects its transcription. Moreover, the authors showed that the lactose phosphotransferase repressor, LacR, and the membrane histidine kinase, KdpD, act as central regulators of *sigS* expression. In addition, it has been shown that LacR is an indirect regulator of *sigS* expression, while the response regulator, KdpE, directly binds to the promoter region of *sigS*. Taken together, these data suggest that a complex regulatory network exists in *S. aureus* in order to modulate the expression of  $\sigma^S$ .

- Sigma H (SigH or  $\sigma^H$ )

Importantly, a new staphylococcal alternative sigma factor,  $\sigma^H$ , was identified by Kazuya Morikawa and his colleagues [58](#). Based on the genetic environment of each *B. subtilis* alternative sigma factors, the authors found SigH, which controls, in *S. aureus*, the expression of the genes encoding proteins required for natural genetic transformation. The orthologue in *S. pneumoniae*, called ComX, is identified as a central competence regulator controlling the expression of all the genetic transformation genes. The orthologue in *B. subtilis*, encoded by *spo0H*, does not play the same role, but is involved in the sporulation process [59](#).

In addition, Liang Tao *et al.* [60](#) revealed that  $\sigma^H$  modulates prophage integration and excision in *S. aureus*. Using phage assays, the authors further showed that  $\sigma^H$  can affect the ability of spontaneous lysis and lysogenization in *S. aureus*, suggesting that  $\sigma^H$  plays a role in stabilizing the lysogenic state.



**Figure 6. Examples of rRNA in *S. aureus*.**

A. Secondary structure of RNAIII [61](#).

B. Secondary structure of the *sprA* sRNA, encoded from a pathogenicity island [62](#).

C. Genetic structure of an mRNA and its antisense RNA.

### 1.3.3. Regulatory RNAs

The universality of regulatory RNAs (rRNAs) playing a role in gene regulation in bacteria has been now well established [63](#). sRNA are members of the non-coding RNAs (ncRNA). The number of ncRNA identified in bacteria has considerably increased during this past decade [64](#). Usually, rRNAs are members of regulatory circuits involved in metabolism, stress adaptation, and virulence [65](#). Most of them exert regulatory functions by binding to proteins or by pairing with mRNAs [63](#). In *S. aureus*, rRNAs include the multifunctional regulatory RNA, called RNAIII [66](#) and *cis*-encoded antisense RNAs (asRNAs) [67](#).

#### - RNAIII

RNAIII, a stable 514-nt RNA molecule, is the first rRNA reported in *S. aureus* [68](#). RNAIII regulates the expression of many *S. aureus* genes mainly encoding exoproteins and cell-wall-associated proteins. The RNAIII secondary structure was confirmed by Benito *et al.* [61](#) (Figure 6A). Using enzymatic and chemical analyses, the authors showed that RNAIII secondary structure contains 14 hairpin (H) structures connected by unpaired nucleotides, and 3 helices formed by distant nucleotides that close off structural domains [61](#) (Figure 6A). The complex structure of RNAIII is a paradigm for RNA-mediated regulation of virulence genes [66](#).

#### - asRNAs

asRNAs are particularly abundant, but in most cases the function and mechanism of action for a particular asRNA remains elusive [67](#). The schematic diagram of asRNAs is shown in Figure 6C. In *S. aureus*, translation of the mRNA encoding the toxin SprA1 is inhibited by the binding of the asRNA SprA1<sub>AS</sub> (Figure 6B) to the Shine-Dalgarno (SD) sequence and the start codon in the *sprA1* mRNA [67](#). A deep-sequencing study of *S. aureus* reported a collection of short RNAs generated by RNase III-dependent digestion of overlapping mRNA–asRNA transcripts across the bacterial genome, suggesting that asRNAs have a general role in mediating RNA degradation in *S. aureus* and possibly other Gram-positive bacteria [69](#).

In addition, using comparative genomic approaches, 12 sRNA were found in *S. aureus* by Pichon and Felden [62](#). 7 sRNA are found in pathogenicity islands that carry a large number of bacterial virulence genes. The others are found in *S. aureus* genome. Some of these sRNAs have multiple copies and are located next to genes implicated in virulence and antibiotic resistance. Some of these sRNAs are expressed in different strains.



## 1.4. *S. aureus* resistance to antibiotics

In 1928, Alexander Fleming accidentally discovered penicillin in environmental mold, *Penicillium notatum*, contaminating a pure culture of *S. aureus* <sup>70</sup>. The pure compound of penicillin was successfully extracted in 1939 by Howard Walter Florey and Ernst Boris Chain <sup>71</sup>. After the first clinical use of penicillin in 1941, other types of penicillin were developed and used to treat several bacterial infections around the world. However, penicillin quickly became ineffective in the treatment of the infections of many bacteria including *S. aureus*. Interestingly, the report of penicillin-resistant *S. aureus* was published in 1940, a year before its first clinical use <sup>72</sup>.

To treat infections caused by the penicillin-resistant *S. aureus*, methicillin was introduced in 1959. However, only a year following the generalization of methicillin use in the clinic, the first Methicillin-Resistant Staphylococcus aureus (MRSA) isolate was reported in Britain ('MRSA history timeline', University of Chicago Medical Center) <sup>73</sup>. During the following years, a number of sporadic hospital outbreaks occurred throughout Western Europe. Rapidly, MRSA isolates were reported in Japan, Australia, and the United States.

In the late 1970s, infections of *S. aureus* became a challenging threat since MRSA strains were noticed with multiple additional drug resistances (MDR) <sup>74</sup>. To fight against *S. aureus* infections, different classes of novel antibiotics, such as vancomycin and linezolid, were developed. However, The first strain of *S. aureus* with reduced susceptibility to vancomycin; Vancomycin Intermediate S. aureus (VISA), was isolated in 1996 from a Japanese patient <sup>75</sup>. The first clinical isolate of Vancomycin Resistant S. aureus (VRSA) was reported in the United States few years later, in 2002 <sup>76</sup>. One option used to combat VRSA and VISA infections is linezolid. However, two clinical isolates of linezolid resistant MRSA were first reported in March 2011 from pus samples <sup>77</sup>.

### 1.4.1. Where are these antibiotic resistance genes located in *S. aureus* chromosome?

Mobile genetic elements (MGEs) are typically defined as fragments of DNA that encode a variety of virulence factors and antibiotic resistance genes as well as the enzymes that mediate their own transfer and integration into new host chromosomes <sup>78</sup>. MGEs play a central role in *S. aureus* adaptation process and are a means to transfer genetic information among and within bacterial species. MGEs may consist of chromosome cassettes, plasmids transposons, insertion sequences, bacteriophages, and pathogenicity islands, among which 5 members have been reported to mobilize antibiotic resistance genes.





### a/ The 'staphylococcal cassette chromosome *mec*' (SCC*mec*) element

Considering the threats caused by MRSA, numerous studies quickly determined that methicillin susceptible *S. aureus* (MSSA) evolved to MRSA due to acquisition of a 'staphylococcal cassette chromosome *mec*' (SCC*mec*) element. This highly diverse mobile genetic element carries a *mec* gene, encoding a novel specific penicillin-binding protein presenting a decreased affinity to  $\beta$ -lactams antibiotics <sup>79</sup>. SCC*mec* also encodes a set of *ccr* recombinases facilitating the genomic island excision and re-integration. Interestingly, besides methicillin and penicillin resistances, SCC*mec* often harbors a number of additional genes conferring resistance to a wide variety of antibiotics inserted in the *ccr* genes complex <sup>80</sup>. Several types and subtypes of SCC*mec* have been characterized based on the nature of *mec* and *ccr* genes <sup>80</sup>. As these different types and subtypes were verified to influence multi-drug resistance differently (modulating the antimicrobial minimal inhibitory concentration), the prevalence and occurrence of SCC*mec* is now highly monitored to identify, control, prevent and treat *Staphylococcus* mediated diseases <sup>80</sup>.

### b/ Plasmids

Staphylococci typically carry one or more plasmids per cell and these plasmids can be categorized into three groups. The first group is composed of small multi-copy plasmids that are cryptic and usually carry a single resistance determinant. The second group is composed of larger (15-30kb), low copy (4-6 per cell) plasmids that can carry several resistance genes. Finally, the third group is composed of conjugative multi-resistance plasmids, which widely contribute to the dissemination of antimicrobial resistances in *S. aureus* <sup>81</sup>.

Currently, at least 3 families of distinct staphylococcal conjugative plasmids were identified. The pGO1/pSK41 family was first documented <sup>82</sup>. These plasmids carry genes conferring a resistance to gentamicin, trimethoprim, penicillin, bleomycin, tetracycline, disinfectants, mupirocin, macrolides, lincosamide, streptogramin B and antiseptics <sup>83</sup>. Notably, this first family of plasmids has been recently reported to be linked to the resistance against the most important antimicrobials used in MRSA treatment: linezolid and vancomycin <sup>84, 85</sup>. The second family of conjugative plasmids, pWBG637, does not carry resistance genes <sup>86</sup>. Finally, the third conjugative plasmids family, pWBG4, was first identified in 1985 by Townsend et al <sup>87</sup>. Among this family, pWBG14 confers aminoglycoside, macrolide, lincosamide and spectinomycin resistance, while pWBG707 confers trimethoprim resistance <sup>88</sup>.

### c/Transposons



Transposons (Tn) are the predominant MGE encoding antibiotic resistance genes in *S. aureus*<sup>89</sup>. Among these Tn, the smaller ones are often present in multiple copies throughout *S. aureus* genome. They are inserted into the chromosome or other MGEs, such as SCC or plasmids. The best characterized examples in this group are Tn554 and Tn552, which encode resistance to MLSB antibiotics (macrolides: erythromycin, lincosamides: clindamycin, streptogramin B) and spectinomycin or penicillinase, respectively<sup>90,91</sup>. The larger Tn (> 18kpb), such as Tn4001 and T4003, are present as a single copy and encode resistances to antibiotics such as tetracycline<sup>92</sup>, trimethoprim, aminoglycosides<sup>93</sup>, or vancomycin<sup>85</sup>.

#### d/Insertion sequences

Insertion sequences (IS) also contribute to *S. aureus* resistance to antibiotics. IS are usually present as pairs constituting a composite transposon<sup>93</sup>. They can insert into various loci and their insertion may result in changes in the expression of genes within the vicinity. For example, IS256 and IS257 insertion activate the expression of the aminoglycoside resistance operon (*aacA-aphD*) and of the gene encoding the resistance for trimethoprim (*dfra*), by forming a hybrid promoter within the native promoter<sup>93,94</sup>.

#### e/Bacteriophages

Recently, bacteriophages and phage-related particles have been highlighted as MGEs that carry antibiotic resistance genes<sup>95</sup>. A report from American Academy of Microbiology<sup>96</sup> indicates that phage implication in the transfer of antibiotic resistance genes could be more important than previously thought, particularly in the environment. Several examples of phages carrying antibiotic resistance genes can be found in the literature. P1-like phages have been reported to carry  $\beta$ -lactamase genes in *Escherichia coli*<sup>97</sup>. P22 phages have also been found to carry  $\beta$ -lactamase genes in *Salmonella*<sup>98</sup>. Phages in *Streptococcus* contribute to the acquisition of multiple resistances to tetracycline, chloramphenicol, macrolide antibiotics, lincomycin and clindamycin or erythromycin<sup>99,100</sup>. The prophage W $\beta$  encodes fosfomycin resistance in *Bacillus anthracis*<sup>101</sup>. Finally, the extra-chromosomal phages (ExP $\Phi$ s) in *S. aureus* have been reported to mobilize antibiotic resistance and virulence genes<sup>102</sup>.

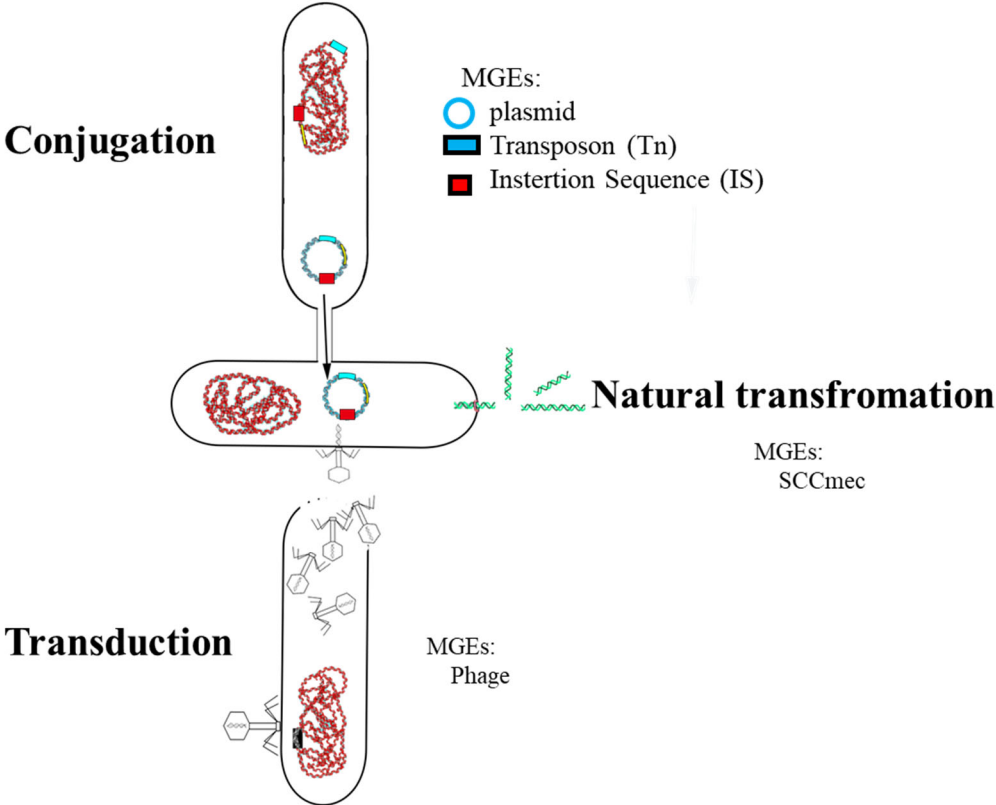


Figure 7. The three main Mechanisms of HGT in bacteria.

### 1.4.2. Transfer of antibiotic resistance genes in *S. aureus*

The rapid and broad dissemination of antibiotic resistance genes is mainly dependant on Horizontal Gene Transfer (HGT). HGT is the movement of genetic material between unicellular and/or multicellular organisms, other than by the vertical transmission of DNA from parent to offspring. It can be performed through three main mechanisms in bacteria: conjugation, transduction and natural genetic transformation (Figure 7).

#### - Conjugation

Bacterial conjugation was discovered and first described by Lederberg and Tatum in 1946 as a process in which transfer of genetic material from one bacterial strain to another is dependent upon cellular contact between members of the two strains [103](#). Conjugation allows the transfer of conjugative plasmids and other conjugative elements. Conjugation occurs through several successive events: first, recognition of the recipient cell by the donor cell thanks to a conjugation pilus (member of the type IV secretion system superfamily, T4SS); second, pilus retraction and contact between the two cells; third, transfer of DNA from the donor to the recipient cell; forth, replication of the conjugative DNA in both cells; and fifth, dissociation of the two cells [104](#). During this process, the conjugative pilus (an extracellular filaments composed of pilins), is implicated in the recipient cell recognition and the contact between the donor and recipient cells upon retractation [104](#).

#### - Transduction

The transfer of DNA sequences between non-contiguous cells by phages is known as transduction, which also serves as an assistant to transfer *S. aureus* pathogenicity islands (SaPIs) [105](#). Practically, DNA of viral origin accounts for 20% of the bacterial genome, including fully functional prophages or remnant phages [95](#). Bacteriophages can be transferred by means of specialized or generalized transduction. Specialized transduction only mobilizes a few specific genes, while generalized transduction transfers any fragment of the bacterial genome [106](#). Normally, specialized transduction depends on “*cos*” mechanisms, like the T7 phage. Generalized transduction relies on the *pac* site homologues in the bacterial chromosome, like P1 and T4 phages.

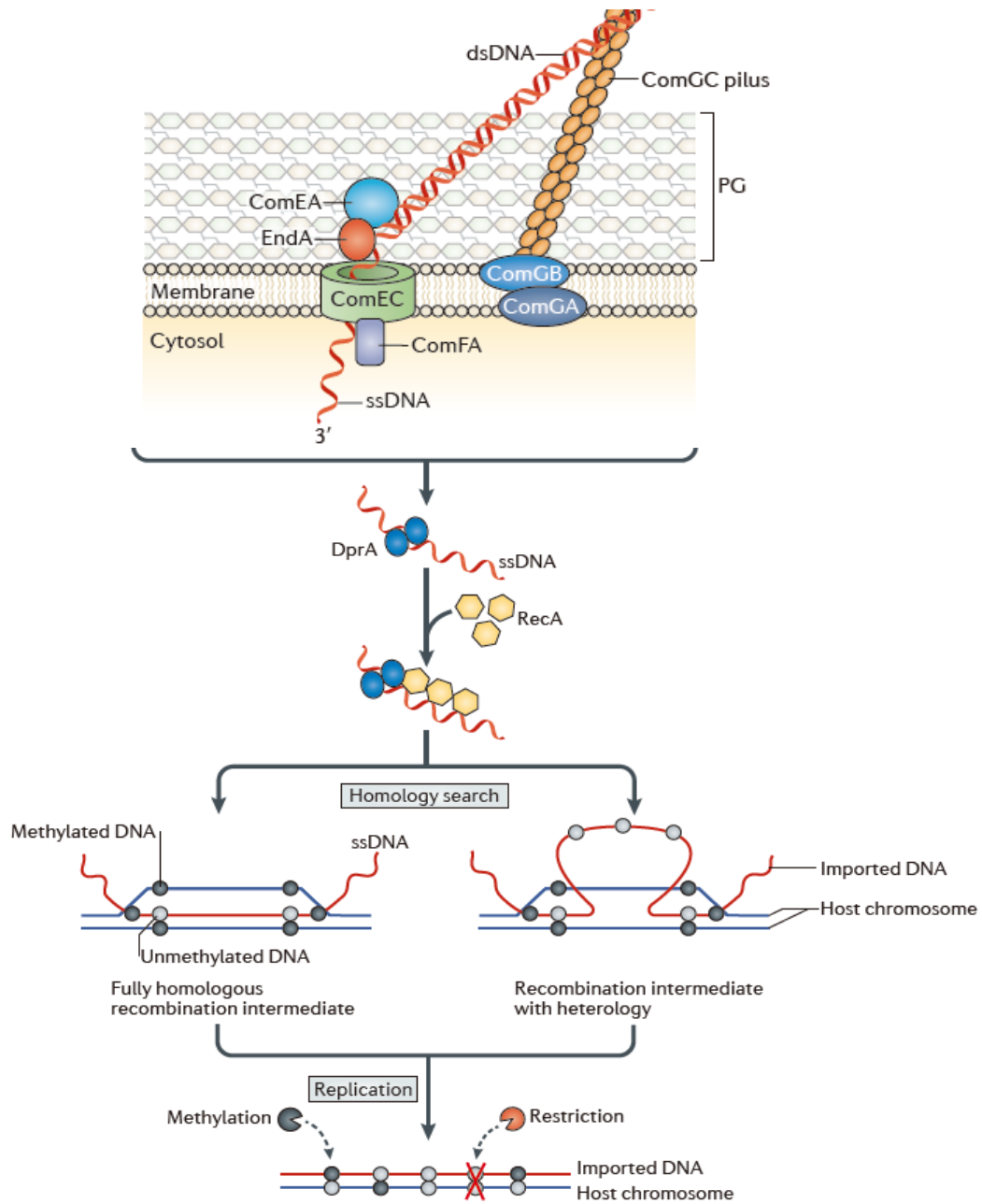
#### - Natural genetic transformation

Recently, the demonstration of the ability of *S. aureus* to become naturally competent for genetic transformation (the third main mechanism of HGT in bacteria) [58](#) changed our way of looking at anti-



-biotic resistance genes acquisition and accumulation in this human pathogen. Through natural genetic transformation, exogenous DNA present in the environment is internalized and integrated into the recipient cell genome by homologous recombination. Importantly, in the original publication, the full length SCCmecII element (45kb) was successfully transferred through natural transformation to a methicillin-sensitive strain, conferring methicillin resistance to the resulting *S. aureus* transformants [58](#).





**Figure 8. An overview of the genetic transformation process [107](#).**

Genetic transformation is composed of three main steps: (i) the binding of exogenous DNA at the competent cells surface, (ii) the import of this DNA across the cell wall and membrane, and (iii) its recombination in the host chromosome if there is a homology.

## 2. Genetic transformation and competence development

### 2.1. Genetic transformation

Up to now, a total of 82 species has been shown to be naturally transformable or to have all the genes required [107](#). Generally, all transformable bacteria share common mechanisms in the course of genetic transformation: exogenous DNA binding, uptake and processing leading to its homologous recombination in the host chromosome. Importantly, these different steps rely on highly conserved proteins present in all Gram-positive model organisms (shown in Table 5).

Process	Protein	Function
DNA up-take	ComEA	Membrane associated protein helix loop helix DNA binding motif (dsDNA receptor)
	ComEB	Late competence protein induced during competence
	ComEC	Membrane channel
	ComFA	ATP translocase
	ComFC	amidophosphoribosyltransferases
	ComGA	traffic ATPase, binds and transports transforming DNA
	ComGB	Polytopic membrane protein, DNA transport machinery
	ComGC	DNA uptake, major component of the transformation pilus
	ComGD	DNA transport machinery, minor pilin
	ComGE	DNA transport machinery, minor pilin
	ComGF	DNA transport machinery, minor pilin
Recombination	ComC	Leader peptidase (prepilin peptidase)
	DprA	DNA processing protein A mediates RecA loading onto internalized ssDNA
	RecA	Recombinase A binds to ssDNA and mediates homologous recombination
	SsbB	Single strand binding protein B protects and stabilizes internalized ssDNA
	CoiA	Recombinase protein

**Table 5. The conserved proteins involved during natural genetic transformation.**

#### - DNA binding

The movement of exogenous double-stranded DNA (dsDNA) from outside a bacterial cell to the cytoplasm requires a complex transformation apparatus. This apparatus is devoted to the assembly of a competence pseudopilus, a structure similar to Type IV pili (T4P), which participates to the first steps of genetic transformation. Numerous studies have characterized the structural components of this pseudopilus in Gram-positive bacteria, such as *Bacillus subtilis* (*B. subtilis*) and *Streptococcus pneumoniae* (*S. pneumoniae*).

In Gram-positive bacteria, the competence pseudopilus spans throughout the thick cell wall (as shown in Figure 8). It is composed of the ComG proteins [108](#): a traffic NTPase, ComGA (providing the energy for the pseudopilus assembly), a polytopic membrane protein, ComGB, four pseudopilins (one major



pilin, ComGC, and 3 minor pilins, ComGD, ComGE and ComGG), as well as a protein with no known orthologues in other systems, ComGF. Other proteins are involved in the construction of this pseudopilus: a prepilin peptidase, called ComC in *B. subtilis* (also known as CilC in *S. pneumoniae*) cleaves the N-terminal leader peptide from pseudopilins, while BdbD (a protein oxidoreductase) introduces an intramolecular disulphide bond into the major pilin, ComGC, which is essential for its stability [109](#).

Interestingly, the competence pilus has been shown to be involved in the initial binding of exogenous DNA in *S. pneumoniae* [110](#). However, in the *comG* operon, only *comGA* was found essential for DNA binding in *B. subtilis*, suggesting that the pseudopilus itself was not involved in this first step. Indeed, few years later, it was shown that wall teichoic acids, specifically produced and modified during competence were responsible of this initial binding of exogenous DNA at the surface of *B. subtilis* competent cells [111](#).

- DNA uptake

The translocation of the exogenous DNA across the cell wall and the cytoplasmic membrane requires four components: the pseudopilus, the DNA receptor ComEA, the permease/channel protein ComEC and the ATP-binding protein ComFA.

The pseudopilus, even though not involved in the initial DNA binding in *B. subtilis*, has to be present to allow its transport across the cell wall [112](#) and direct the transforming DNA to the next actor, ComEA. ComEA is bound to the membrane through its N-terminal region. The C-terminal extremity of ComEA contains the DNA-binding domain, which interacts with dsDNA without any sequence specificity. Then, ComEA presumably delivers dsDNA to a protein that generates single-stranded DNA (ssDNA) for internalization. In *S. pneumoniae*, this protein is a membrane-bound nuclease, called EndA [113](#). However, an EndA orthologue has not yet been identified in other species [114](#).

ssDNA is then delivered through the cytoplasmic membrane thanks to the channel formed by a polytopic membrane protein called ComEC. ComEC is present in the membrane as a homodimer, with five transmembrane domains present per subunit, and contains an ABC transporter permease motif [115](#). The energy to transport the transforming ssDNA across the membrane is finally provided by ComFA [114](#), an ATP-driven DNA translocase that has been proposed to pump the ssDNA in the recipient cell cytoplasm.

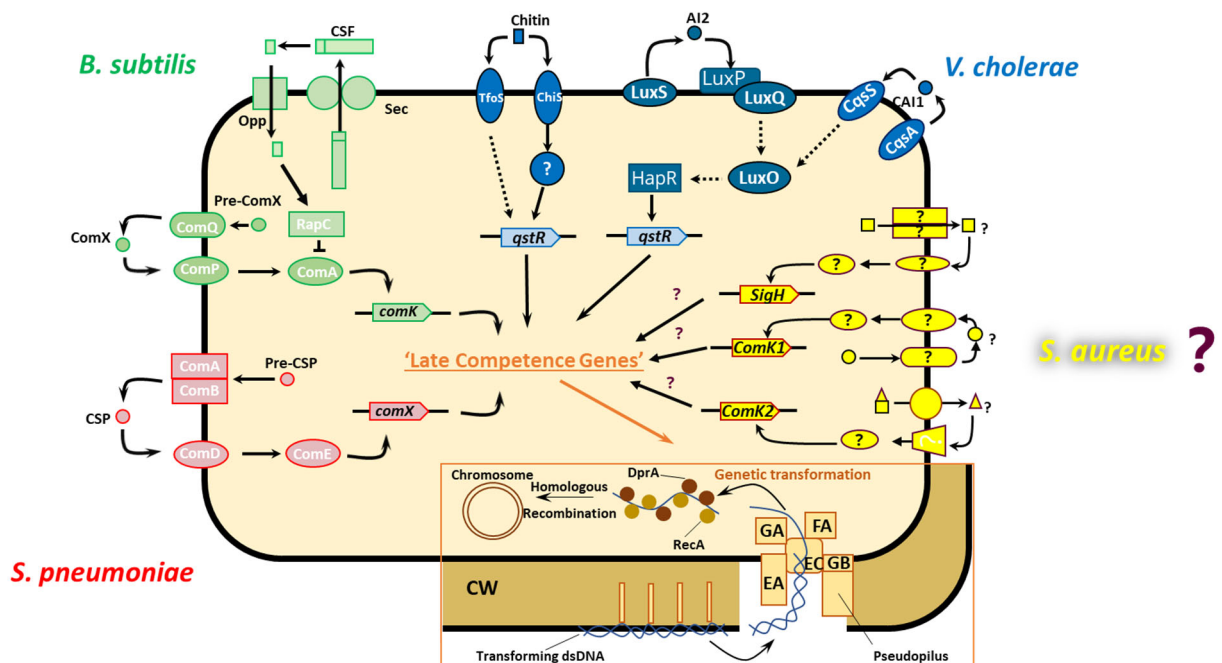
- DNA processing

Once present in the cytoplasm, the transforming ssDNA is bound by ssDNA-binding proteins, Ssb and DprA. The Ssb protein was documented to be a reservoir of internalized ssDNA during genetic



transformation [116](#). DprA would interact with not only naked ssDNA, but also Ssb-coated ssDNA. Once associated with ssDNA, it has been proposed that DprA would facilitate the loading of the recombinase, RecA, leading to its homologous recombination.

RecA once loaded on the transforming ssDNA, promotes a search for homology along the recipient cell chromosome, followed by strand exchange. Dependent on the position of the heterologous sequence in the imported DNA, two types of transformation heteroduplexes could be generated: a fully homologous recombination intermediate and a recombination intermediate with a heterologous stranded loop, as shown in Figure 8. Upon replication, the methylation and restriction activities of the restriction-modification systems (R-M) compete for access to the imported DNA, ultimately determining the fate of the newly incorporated exogenous DNA [107](#).



**Figure 9. Competence regulation and genetic transformation in *S. pneumoniae*, *B. subtilis* and *S. aureus*.**

Competence regulation is presented in three important model organisms: *B. subtilis* (green), *V. cholerae* (blue) and *S. pneumoniae* (red). Apart from the three regulators identified, all the regulation pathways are only hypothetical in *S. aureus* (yellow). The main conserved steps and actors of genetic transformation are presented in the right down corner.

Figure inspired from [107](#).

## 2.2. Development of natural competence in important model organisms

To perform genetic transformation, bacteria need to enter a differentiated state called competence (or natural competence). Competence is a transient opportunity for DNA internalization and thereby enables subsequent transformation [107](#). Competence development is usually controlled by one or two central competence regulators (Figure 9). In *B. subtilis*, this central competence regulator is the transcriptional factor, ComK [117](#), while in *Streptococcus pneumoniae* (*S. pneumoniae*), this role is held by the alternative sigma factor, ComX [118](#). Finally, *Vibrio cholerae* (*V. cholerae*) displays a different regulatory organization as two central competence regulators called, TfoX and QstR, have been shown to be essential.

These central regulators are activated by species-specific regulatory cascades, responding to different environmental signals, such as antibiotics, nutritional status and more generally, stressful conditions. Activation of these central regulators leads to the expression of the late competence genes, among which are found all the genes necessary for genetic transformation. Most of the knowledge accumulated on natural competence and genetic transformation comes from *B. subtilis*, *S. pneumoniae* and *V. cholerae*. However, *S. aureus* is one of the latest bacteria for which natural competence for genetic transformation has been investigated [58](#). In this new model organism, very little is known. Up to now, 3 potential central competence regulators have been identified in *S. aureus*: an alternative sigma factor called SigH, homologous to ComX in *S. pneumoniae*, and two transcriptional factors called ComK1 and ComK2, homologous to ComK in *B. subtilis*. Importantly, it remains unclear why there are three potential central regulators in *S. aureus* and what their exact role is.



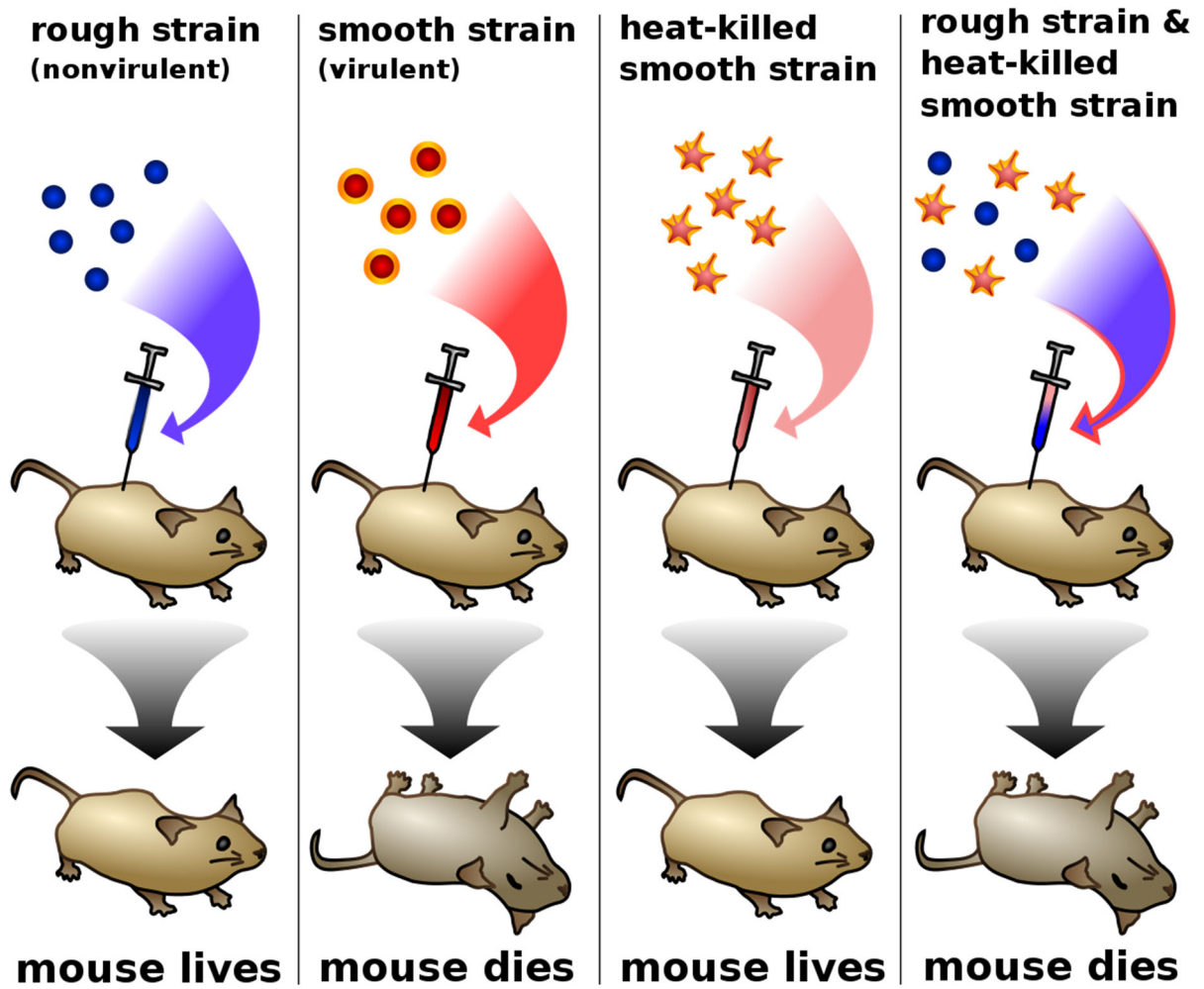


Figure 10. Griffith experiment. Link: [https://en.wikipedia.org/wiki/Griffith%27s\\_experiment](https://en.wikipedia.org/wiki/Griffith%27s_experiment)

## 2.3. Competence development in *S. pneumoniae*

### 2.3.1. Presentation of the bacterium

*Streptococcus pneumoniae* (*S. pneumoniae*) asymptotically colonizes the oral cavity and upper respiratory tract of humans. However, it is potentially pathogenic in case of migration to other regions, leading to a board variety of diseases, such as otitis media, pneumonia, meningitis, and septicaemia [119](#). To this day, it continues to be a major public health problem. To fight pneumococcal diseases, antimicrobial therapy was introduced after the Second World War, and subsequently, the first effective vaccine was produced more than half a century ago. However, these clinical treatments generated a selection pressure that accelerated the evolution of strains that were highly resistant against antibiotics and vaccines [120](#). The ability of *S. pneumoniae* to escape antimicrobials action and host immune response is due to its remarkable genomic plasticity, mainly obtained from rapid HGT *via* natural genetic transformation [121](#). Gene clustering and phylogenetic analyses indicated that over 60% of the pan-genomes of *Streptococcus* species were acquired through HGT events [122](#), thus explaining *S. pneumoniae*'s genomic plasticity. During evolution of *S. pneumoniae*, natural genetic transformation is recognized to be a primer driver by promoting genetic variability through HGT [123](#).

### 2.3.2. Griffith's historical experiment

Natural genetic transformation was first discovered *in vivo* by Frederick Griffith in 1928 while observing the transfer of morphological traits between different strains of *S. pneumoniae* [124](#). In this historical experiment, shown in Figure 10, Griffith used two types of *S. pneumoniae* strains that differed by the appearance of their colonies on petri plates containing a nutrient agar medium.

The first strain had a smooth appearance due to the presence of a polysaccharide capsule. The capsule is an important virulence factor in *S. pneumoniae* which made the 'smooth' strain lethal when injected to a mouse. In contrast, the second strain had a rough appearance due to the absence of this capsule. As a consequence, a mouse would survive to the injection of this second 'rough' strain. Finally, Griffith observed that the mice that were injected with a mixture of the living rough non-virulent strain with the killed (heat inactivated) smooth strain would not survive. In addition, the colonies that were isolated from the dead mice all displayed a smooth aspect on plates. For the first time, Griffith demonstrated the transfer of DNA (in this case, the genes encoding the capsule synthesis and conferring the virulence traits) by natural transformation between two *S. pneumoniae* strains.

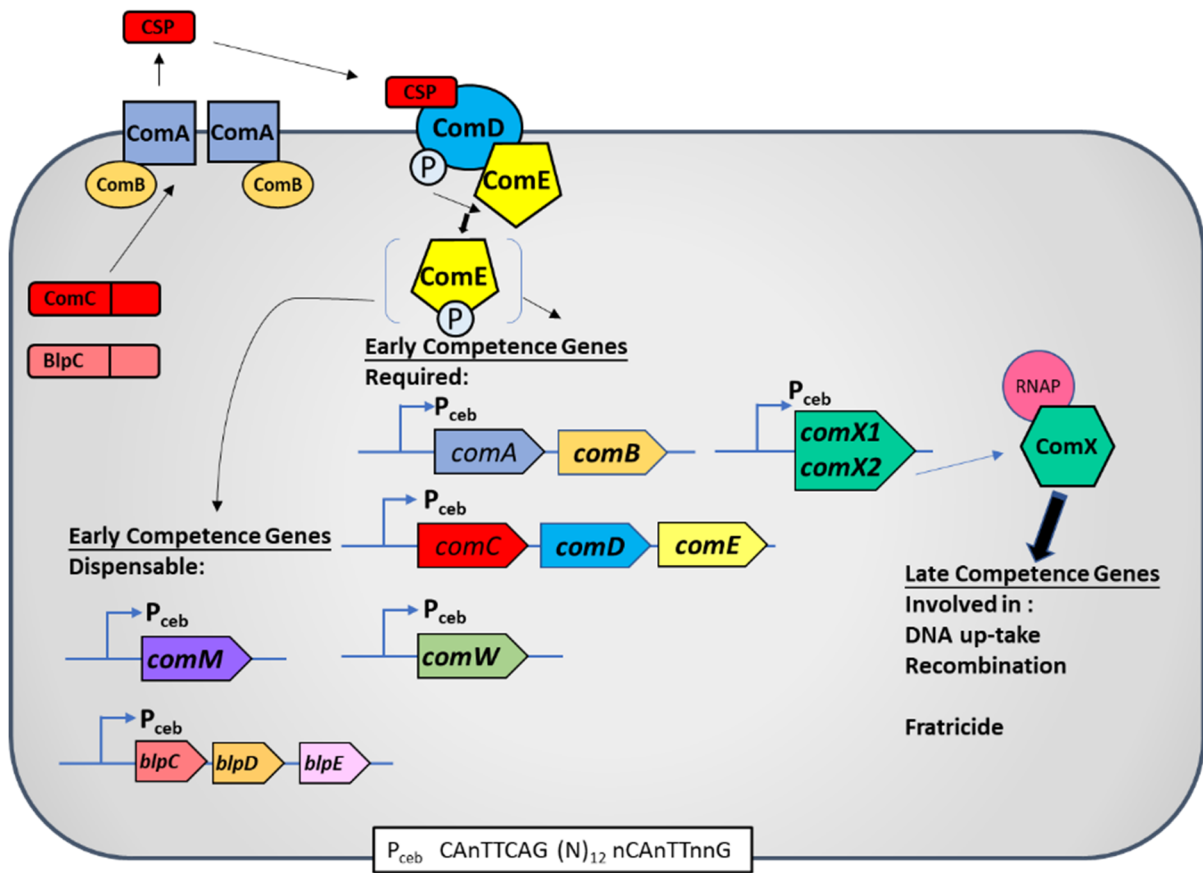


Figure 11. The ComABCDE quorum sensing pathway of *S. pneumoniae*.

### 2.3.3. Competence regulatory pathways in *S. pneumoniae*

#### 2.3.3.1. Inducer of competence (stimuli)

Peptides, or autoinducers, as an induction cue are a recurring theme in transformable species [125](#). In *S. pneumoniae*, a peptide named competence-stimulating-peptide (CSP) acts as the inducing molecule for competence. It also functions as an alarmone or a timing device to promote competence in neighbouring pneumococci [107](#). Its synthesis and export coordinate the response to cellular triggers and environmental stimuli. Over the years, a large number of CSP have been characterized in streptococcal species [121](#).

In *S. pneumoniae*, and most of transformable species, cell density is the major cellular trigger for competence induction. Quorum sensing (QS) is a bacterial intercellular communication mechanism for controlling gene expression in response to population density. QS allows bacterial cells to communicate and behave collectively, probably providing significant benefits through cellular differentiation and adaptation. In *S. pneumoniae*, the CSP accumulates extracellularly, providing an estimation of the cell density, and, as a result, leads to the induction of competence. Importantly, induction of competence in *S. pneumoniae* has been described as one of the first examples of quorum-sensing systems [126](#).

Environmental stresses that threaten cell survival can also induce competence. Indeed, sub-lethal concentration of antibiotics, such as aminoglycoside and fluoroquinolone, and DNA-damaging agents, such as mitomycin C were shown to induce competence in *S. pneumoniae* [127](#). A recent study suggested that these compounds increase the rate of decoding errors during translation [107, 128](#). In addition, a cell surface protease/chaperone, HtrA, participates in competence induction based on translation fidelity. HtrA degrades the CSP to repress competence when translation fidelity is high, whereas CSP accumulation is allowed when HtrA is saturated by misfolded proteins [107, 128, 129](#). Finally, it has been shown that an increase in the mutational load, induced by the deletion of *mutX*, *hexA* or *hexB*, can also stimulate the development of competence in *S. pneumoniae* [130](#). This result highlights the potential connexion existing between the replication stress and the induction of competence. Indeed, all the antibiotics associated with a DNA replication stress, that would increase the copy number of genes located near the origin of replication (*oriC*), such as the *comAB* and *comCDE* operons, increase the CSP secretion, induce the early auto-catalytic loop and consequently the development of competence [131](#).

#### 2.3.3.2. Regulation of competence



### 2.3.3.2.1. Early regulatory pathways

The main early regulatory pathway involved in competence development in *S. pneumoniae* requires a quorum-sensing system, composed by five proteins encoded by the *comABCDE* genes [132](#). Among these genes, *comAB* encodes an ABC transporter [133](#), *comC* encodes the precursor of a secreted signal molecule, called the pre-CSP and *comDE*, located downstream of *comC*, respectively encode a transmembrane histidine kinase (ComD) and a response regulator (ComE), forming a two-component signal transduction system (TCS) [134](#).

First, ComAB cleaves the leader peptide of the pre-CSP and simultaneously translocates the matured CSP peptide across the cytoplasmic membrane [133](#). As a result, the CSP accumulates outside of the cells. Then, the extracellular CSP is sensed by the histidine kinase ComD. Activation of ComD results in its autophosphorylation and the subsequent transfer of a phosphoryl group to its cognate response regulator ComE. Phosphorylated ComE (ComE-P) then binds to a conserved promoter sequence, called ComE-binding site, or Ceb, found in the promoter region of *comAB* [135](#), *comCDE* [136](#), and the alternative sigma factor gene *comX* [137](#). Importantly, as ComE-P activates the expression of the early competence regulatory genes (i.e. *comABCDE*), it forms an autocatalytic loop that amplifies the signal and finally drives the entire population of *S. pneumoniae* into the competent state [121](#), [125](#). When the ComE-P concentration reaches a threshold, it activates the expression of *comX*. ComX is considered as the central competence regulator as it controls the expression of the late competence genes among which are found all the genes involved in natural transformation (see paragraph 2.3.3.2.2).

In addition to *comABCDE* and *comX*, *comW* is another important early competence gene, as demonstrated by DNA microarrays [138](#). A Ceb site is also present in its promoter. *comW* encodes a globular protein which can interact with DNA, independently from the sequence [139](#). It functions during assembly of the RNA polymerase- $\sigma^x$  holoenzyme to help promote transcription from  $\sigma^x$ -targeted late competence genes promoters. However, it remains unclear how ComW facilitates the holoenzyme assembly.

Additionally, *blpABC* and *comM* are new members of the early competence genes, given that a Ceb site is also found in their promoter and that their expression is upregulated by CSP [140](#), [141](#). They are both thought to be dispensable for competence [132](#), but their activity intertwines with competence development [142](#). In *S. pneumoniae*, the *blp* locus controls the pneumococcal bacteriocin (pneumocin) production. *blpAB* encodes an ABC transporter and *blpC* encodes a peptide pheromone. Like *comABCDE*, the *blp* locus is stimulated by the accumulation of BlpC [143](#). Pre-BlpC is processed at a double glycine motif and exported via its cognate transporter BlpAB [141](#). BlpC binds to and activates its histidine kinase receptor, BlpH, resulting in the phosphorylation of the response regulator, BlpR [144](#).



Indeed, it has been shown that BlpAB has a high degree of similarity with ComAB and BlpC is homologous to CSP [143](#). Finally, *comM* encodes an immunity protein by which competent cells can protect themselves against the lethal activity of CbpD that is encoded by a late competence gene [140](#).

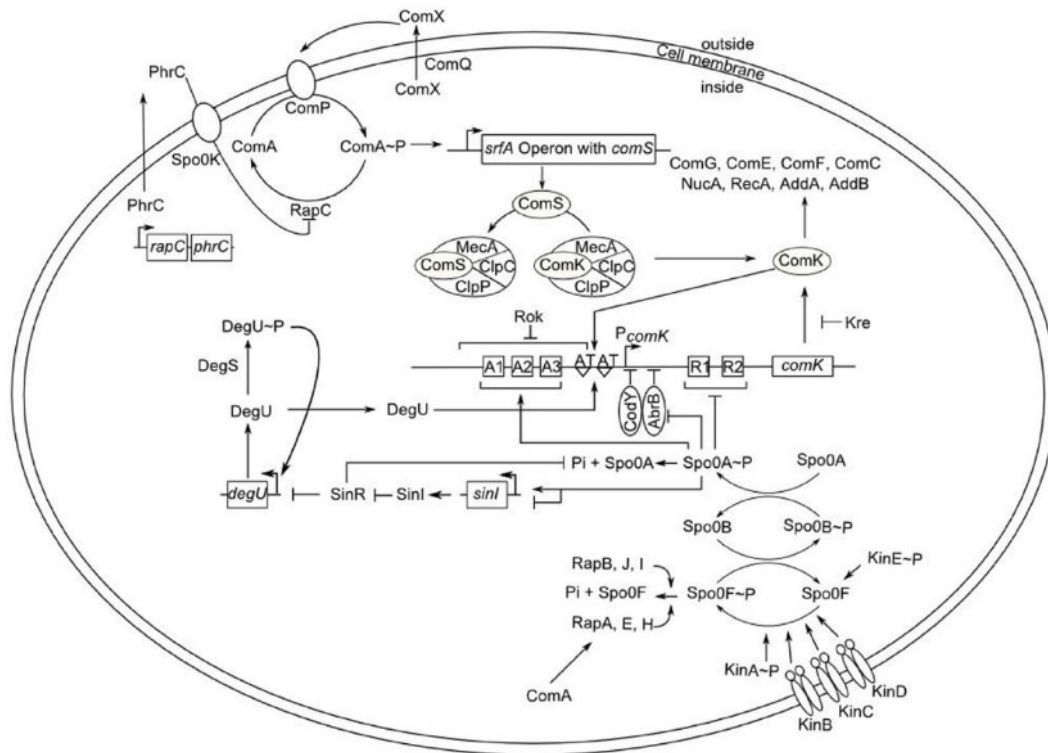
#### 2.3.3.2.2. Late competence genes

The late competence genes expression is controlled by the alternative sigma factor, ComX, which is encoded by two identical genomic copies of *comX* within *S. pneumoniae* genome [145](#). ComX can form a complex with the RNA polymerase to recognize an 8bp sequence (TACGAATA, called the “cinbox”) in the promoter of approximately 80 late competence genes [107](#). All the genes involved in genetic transformation are controlled by ComX (Table 5 & Figure 8), with the exception of *endA*. Importantly, some of the late competence genes do not directly participate in genetic transformation, and are rather involved in stress response or fratricide.

#### 2.3.3.2.3. Shut-off of competence

Like the induction of competence in *S. pneumoniae*, its shut-off is abrupt. In the laboratory, when the CSP is added to the culture medium during early exponential phase, the pic of competence is reached after 15 to 20 min and then gradually declines. During this shut-off phase, cells are unresponsive to the presence of high concentrations of CSP in the medium (about 60-80 min of refractory period). The mechanism for this shut-off of competence is directly dependent on DprA, the transformation-dedicated RecA loader [146](#). Indeed, inactivation of DprA up-regulates *comX* expression. Moreover, its engineered expression at the early stage of competence reduces *comX* and *comC* expression, independently of any other late *com* gene product, suggesting that DprA plays a role in the shut-off of all the early *com* promoters. Further studies showed that the shut-off of the *comE* promoter involves a direct interaction between ComE-P and DprA and two types of mechanisms were proposed [146](#). DprA could either act as anti-activator, antagonizing the interaction of ComE-P with its target promoters, or affect ComE phosphorylation state.





**Figure 12. Overview of competence regulation in *B. subtilis* [147](#).**

The main pathways leading to competence development in *B. subtilis* are composed of two quorum sensing systems, characterized by two exported peptides, ComX and PhrC, both modulating the activity of the transcriptional activator ComA. Once activated, ComA promotes the expression of the *srf* operon in which is embedded the *comS* gene. The anti-adaptor protein ComS binds to MecA, preventing the degradation of the central competence regulator ComK by the ClpCP protease system. *comK* transcription is also highly controlled by various regulators, that can be either inhibitors, like Rok, AbrB, Spo0A~P, and CodY, or activators, like DegU, Spo0A~P and ComK itself. The stability of the *comK* mRNA is also influenced by Kre.

## 2.4. Competence development in *Bacillus subtilis*

### 2.4.1. Presentation of the bacterium

*Bacillus subtilis* (*B. subtilis*), also known as the hay bacillus or grass bacillus, is a Gram-positive, catalase-positive bacterium, found in the soil and the gastrointestinal tract of ruminants and humans. This bacterium is also an historical model organism for the study of natural competence and genetic transformation. The first report on transformation of *B. subtilis* was published by Spizizen in 1958 [148](#). Since then, the mechanisms regulating natural competence of *B. subtilis* have been intensively studied. Until now, a highly interconnected regulatory network has been revealed. The central regulator in this regulatory network is ComK, a transcription activator which is specifically synthesized during the competence state [149](#). Once activated, ComK drives the expression of more than 100 late competence genes, among which are found all the genes involved in genetic transformation [150](#).

### 2.4.2. Competence regulatory pathways in *B. subtilis*

#### 2.4.2.1. The stress and stimuli inducing competence in *B. subtilis*

Nutritional status is an important signal for the induction of competence in *B. subtilis*. When nutrients are exhausted in the competence-stimulating medium (poor, chemically-controlled medium), *B. subtilis* spontaneously enters into the competent state at the transition between logarithmic and stationary growth phase [125](#). Indeed, competence in *B. subtilis* can develop in amino-acid-supplemented, glucose-minimal salts-based medium, but does not develop in meat-broth-based complex medium [151](#). However, only 10%-20% of *B. subtilis* cells in a given culture can enter the competent state. Because of the presence of two clearly distinguishable cell types, competent and non-competent cells, the biological system that switches competence ON is defined as bistable [117](#). This bistability is attributed to the stochastic fluctuations in the cellular levels of *comK* mRNA.

In addition to medium-specific modes, competence induction is also dependent on cell density which is regulated by quorum sensing. Two signalling peptides, participating in the development of competence have been well described in *B. subtilis*. The main peptide is a 9 to 10 amino acid oligopeptide, called ComX [152](#) (different from the alternative sigma factor ComX in *S. pneumoniae*). The second is a pentapeptide, named CSF for competence-stimulating factor [153](#). Ultimately, these two signalling molecules lead to *comK* expression and ComK accumulation, resulting in the induction of competence (Figure 12).



### 2.4.2.2. Regulation of competence

#### 2.4.2.2.1. Inhibition of competence development (exponential phase)

##### i. Inhibition of *comK* expression

When growth conditions are not favourable for competence development (plenty of nutrients, low cell density), *comK* expression is highly repressed. Several transcriptional repressors have been shown to directly bind to the *comK* promoter [117](#). These repressors include AbrB, CodY and RoK (Figure 12). AbrB is regulated by the global transcription factor Spo0A. CodY, the GTP-binding transcription factor, can sense the intracellular GTP concentration as an indication of the nutritional status, and regulate many stationary growth phase genes such as the dipeptide transport operon *dpp*, but also *spo0A* [154](#). Finally, the third repressor of *comK* expression, RoK, is the most important. Hoa *et al.* particularly showed that Rok binds to the *comK* promoter to repress its expression [155](#). In addition, Rok and ComK are able to bind to the *rok* promoter and also repress its expression. This mechanism establishes feedback loops that somehow decide which cells enter into competence [117](#). Cells with higher concentrations of RoK would inhibit the expression of *comK*, and therefore remain in the non-competence state. At the opposite, cells with higher ComK concentrations would inhibit *rok* expression, leading to more ComK production and therefore the induction of competence.

##### ii. ComK protein degradation

Even though the expression of *comK* is tightly repressed during exponential growth, few transcripts can be made (due to the intrinsic leakiness of the promoter) and few ComK proteins synthesized. In order to prevent any ComK transcriptional-related activity, *B. subtilis* developed an additional pathway leading to the degradation of all the ComK proteins produced during exponential growth. This pathway involves an important competence negative regulator, the adaptor protein MecA. Revealed by biochemical analyses [156](#), MecA can form a ternary complex with ComK and ClpC in the presence of ATP (Figure 12). Simultaneously, ClpC recruits the ClpP protease, targeting ComK for degradation [157](#). Therefore, the positive feedback loop (presented below) leading to ComK synthesis is interrupted.

#### 2.4.2.2.2. Activation of competence

##### i. Transcriptional activation of *comK*

At the end of the exponential phase, the expression of *comK* is enabled. First, ComK can bind to its own promoter and induce its own transcription, forming a positive feedback loop. This loop is a critical



step in establishing competence. Only the cells where the amount of ComK protein exceeds a threshold, can enter into the competence ON state. The proportion of competent cells increases to reach a maximum during the approach to stationary phase (10 to 20% of the population, [158](#)), while most cells remain non-competent.

In addition, DegU also promotes the expression of *comK*. Briefly, DegU is phosphorylated by DegS in response to unclear intracellular stimuli. Phosphorylated DegU binds to the *comK* promoter, acting as a transcriptional activator (Figure 12). Moreover, revealed by Hamoen *et al.* [159](#), DegU can enhance the affinity of ComK for its own promoter and *degU* deletion would completely abolish *comK* competence-specific activation.

Finally, the response regulator Spo0A also promotes *comK* expression. Revealed by Nicolas Mirouze *et al.* [158](#), as the concentration of Spo0A~P increases during the exponential phase, it first induces *comK* expression. To do so, Spo0A~P binds to 3 activation sites (over-lapping the Rok binding site) present in the *comK* promoter, amplifying the basal expression from the *comK* promoter. The authors have hypothesized that Spo0A antagonizes the repressing effect of Rok during this period.

## ii. ComS activation for ComK stability

As cell density increases during growth and the cells approach stationary phase, a quorum-sensing system is initiated, leading to the synthesis of the small protein called ComS [160](#). This quorum-sensing system comprises the ComP-ComA two-component system and ComX (Figure 12)[117](#). ComX acts as a signalling molecule. Initially, ComX is synthesized as a 55-residue precursor peptide which is then post-translationally modified by ComQ [161](#). In *B. subtilis* genome, the *comQ* gene is located upstream of *comX*, while *comA* and *comP* are located downstream of *comX*.

ComQ exports the matured ComX peptide, leading to the extracellular accumulation of ComX. ComP, a membrane-spanning protein histidine kinase, senses the level of extracellular ComX. Once the level of ComX achieves a threshold, ComP is activated and subsequently phosphorylates its cognate transcription factor ComA. Upon activation of the two-component system, phosphorylated ComA binds to the *srfA* promoter and thus induces the 30kb-spanning *srfA* operon transcription [162](#). Within this operon, the second open reading frame encodes a 46-amino-acid peptide called ComS. ComS plays a key role in ComK stability. Indeed, *in vitro* experiments indicated that ComS acts as an anti-adaptor protein able to bind the adaptor protein MecA (Figure 12), preventing the targeting of ComK for degradation [117](#).



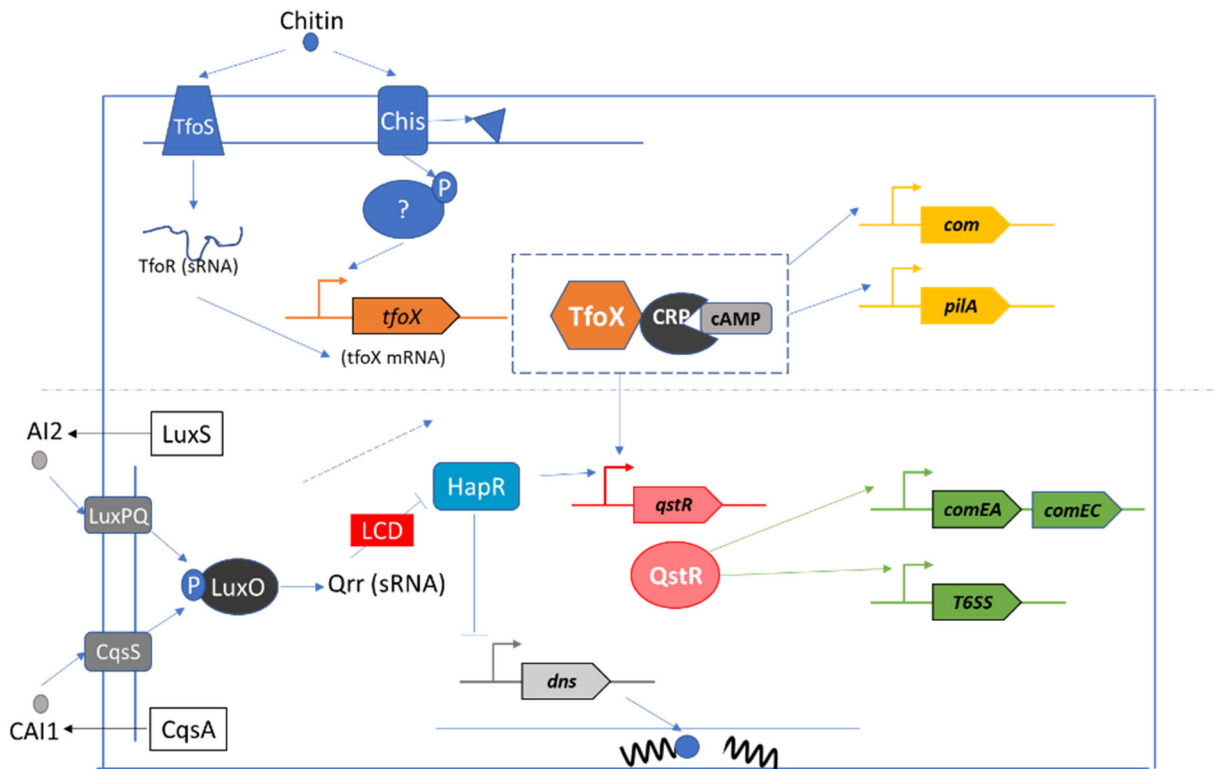
In addition, the CSF-mediated pathway is indirectly involved in the regulation of ComK stability. In this regulatory cascade, PhrC, the CSF precursor, is exported by the Sec system [163](#), and processed into matured CSF by the Subtilisin, Epr or Vpr proteases [164](#). These proteases are regulated by the alternative sigma factor SigH in *B. subtilis*. Matured CSF is transported into the cell by the oligopeptide permease Spo0K (opp) [165](#). Intracellular CSF is able to antagonize RapC acting on the ComA response regulator for *srfA/comS* expression (Figure 12). In *B. subtilis* genome, *rapC* is located upstream of *phrC* and co-transcribed with *phrC*. RapC can interact with ComA, preventing phosphorylated ComA from binding to its target DNA promoter [166](#). Thus, when PhrC levels rise sufficiently, RapC activity is repressed, ComA~P levels accumulate, and *srfA/comS* expression increases.

#### 2.4.2.2.3. Shut-off of competence

During competence development in *B. subtilis*, there is a window of opportunity allowing cells to enter into competent state. This window starts at the end of exponential phase, reaches a maximum during the approach to stationary phase and then decreases. During this window, only 10%-20% of the population enters competence, and most cells remain non-competent. Why don't most cells enter into competent state? One explanation is that Spo0A decreases the *comK* expression by binding to inhibition sites (located over the -35 -10 boxes) present in the *comK* promoter [158](#). In addition, Rok can also bind to the promoter of *comK*, leading to repression of the *comK* expression [158](#). As a result, the concentration of ComK goes below the threshold necessary for competence, leading to competence shut-off.

However, the mechanisms leading a competent cell to escape from competence remains unclear. As demonstrated in [167](#), cell division and cell elongation are inhibited during the escape from competence. It was shown that competent cells were delayed in cell elongation, because of ComGA, and cell division, because of Z-ring formation inhibition by the Maf protein. Based on this, the authors proposed that either ComGA degradation and/or Maf inactivation would allow the competent cells to resume growth and division, leading to diluting the ComK protein and, ultimately, shutting off competence.





**Figure 13. A summarized model for competence development in *V. cholerae*.**

In this bacterium, there are two central competence regulators, called TfoX and QstR. *tfoX* expression is activated by the presence of chitin in the environment. Meanwhile, two quorum sensing systems, characterized by two exported peptides (AI2 and CAI1) promote the expression of *qstR*. In addition, TfoX has been shown to activate *qstR* expression. Finally, once TfoX and QstR are activated, they both control different set of late competence genes.

## 2.5. Competence development in *Vibrio Cholerae*

### 2.5.1. Presentation of the bacterium

*Vibrio cholerae* (*V. cholerae*), is a gram-negative, comma-shaped bacterium. It is a facultative anaerobe, and can undergo respiratory and fermentative metabolism. This bacterium inhabits aquatic environments and is able to attach to the chitinous surface of zooplankton, crabs, shrimps, and other shellfish [168](#). However, some strains (like O1 and O139) of *V. cholerae* can enter human being's digestion system *via* undercooked or raw marine life species, causing the cholera disease. The most compelling cholera event happened in 1992. A vast cholera epidemic broke out in Bangladesh and India, and subsequently spread throughout Asia. As revealed by genetic analysis, this outbreak began due to the acquisition of a gene cluster that converted the ancestral *V. cholerae* O1 E1 Tor serogroup to a completely new serogroup, named O139 Bengal. Importantly, this serogroup conversion, from O1 to O139, was observed through genetic transformation when the two strains were co-cultured on chitin surface by Melanie Blokesch et al. [169](#). In addition, the authors also found that this serogroup conversion could occur as a single-step exchange of large fragments of DNA [169](#). These data indicate that natural transformation might be a common mechanism for *V. cholerae* to adapt and survive in stressful environmental niches or become more pathogenic for humans.

### 2.5.2. Competence regulatory pathways in *V. cholerae*

#### 2.5.2.1. Induction of competence

The chitin is an essential inducer of competence in *V. cholerae* [169](#). Interestingly, and as indicated by microarray expression profiling, the production of the central competence regulator TfoX is induced in the presence of chitin oligosaccharides but not in the presence of the chitin monomer N-acetylglucosamine (GlcNAc) or deacetylated chitosan dimers [170](#).

Apart from chitin-induced competence, two auto-inducers are involved in the induction of competence in *V. cholerae*: the species-specific cholera auto-inducer (CAI-1) and a universal auto-inducer (AI-2). CAI-1 is the main driver of competence. These two auto-inducers are also linked to quorum sensing (QS) (Figure 13).

#### 2.5.2.2. Regulation of competence

Competence regulation in *V. cholerae* has been shown to be bipartite [171](#). This specific bipartite regulation consists of two pathways: the TfoX-dependent and the QstR-dependent regulatory pathways.



During the first steps of competence induction, chitin is sensed by the histidine kinase Chis (Figure 13). In the absence of chitin, Chis is deactivated by a periplasmic binding protein, which is a part of a Chitin ABC transporter. In the presence of Chitin, Chis dissociates from the periplasmic binding protein, leading to its activation. Once activated, Chis potentially phosphorylates an unknown response regulator, ultimately leading to the expression of the downstream genes, among which is found *tfoX*. In addition, chitin oligosaccharides might be also sensed by the periplasmic domain of a transmembrane transcriptional regulator, TfoS (Figure 13)<sup>172</sup>. In this case, TfoS activation would lead to the production of a small RNA, TfoR (Figure 13), which ultimately initiates *tfoX* translation upon base-pairing with the *tfoX* mRNA.

TfoX acts in conjunction with the cAMP receptor protein (CRP) complexed with cAMP (Figure 13). TfoX assists CRP-cAMP in binding to the promoter of certain genes, including *tfoX*. Most of TfoX-regulated genes are involved in the biosynthesis or as structural components of DNA up-take machinery, transformation and type VI secretion system (T6SS). The DNA-uptake machinery of *V. cholerae* is composed of a central type IV pilus structure, the periplasmic DNA-binding protein ComEA, and the inner membrane translocator ComEC. T6SS are a wide-spread molecular killing device, which mediates the killing of neighbouring bacteria. Thanks to the T6SS activity, DNA is released from dead bacteria, absorbed by competent bacteria and used as transforming material.

Notably, the expression of *comEA*, *comEC* and T6SS gene cluster not only require TfoX, but also require QstR (Figure 13)<sup>173</sup>. Indeed, *qstR* expression requires TfoX. As mentioned above, *qstR* expression is also dependent on QS. Extracellular CAI-1 and AI-2 are respectively produced by the synthases CqsA<sup>174</sup> and LuxS<sup>175</sup>, and then sensed by the receptor CqsS<sup>176</sup> and the receptor complex LuxPQ (Figure 13)<sup>177</sup>. These two systems are membrane-bound two-component receptors. At low cell density, in the absence of auto-inducers, both receptors function as kinases that phosphorylate their shared response regulator LuxO<sup>178</sup>. Phosphorylated LuxO activates the expression of genes encoding quorum-regulatory sRNAs, called Qrr (Figure 13). The Qrr sRNAs repress *hapR* expression by base-pairing with the *hapR* mRNA (Figure 13). HapR is a *qstR* transcriptional activator. Under these conditions, expression of *qstR* is inhibited. At high cell density, the auto-inducers accumulate extracellularly. They bind to CqsS and LuxPQ, respectively, an event which allows the receptors to switch into their phosphatase activity<sup>179</sup>. LuxPQ and CqsS de-phosphorylate LuxO, leading to the cessation of *qrr* expression, and therefore activating the *hapR* mRNA for translation. Thereby, HapR accumulates and drives the transcriptional expression of *qstR*, resulting in the expression of the specific late competence genes. In addition, HapR represses the expression of the nuclease *dns* by binding to the upstream region of *dns* promoter<sup>180</sup>, thereby preventing free DNA degradation to make it available for up-take and transformation.

	<i>B. subtilis</i>	<i>S. pneumoniae</i>	<i>S. aureus</i>
prepilin peptidase	<i>comC</i>	SP1808	SA1486
DNA receptor, transport (see text for details)	<i>comGA</i>	SP2053( <i>cglA</i> )	SA1374
	<i>comGB</i>	SP2052( <i>cglB</i> )	SA1373
	<i>comGC</i>	SP2051( <i>cglC</i> )	SA1372
	<i>comGD</i>	SP2050( <i>cglD</i> )	SA1371
	<i>comGE</i>	SP2049( <i>cglE</i> )	SA1370
	<i>comGF</i>	SP2048( <i>cglF</i> )	SA1369
	<i>comGG</i>	SP2047( <i>cglG</i> )	×
	<i>comEA</i>	SP0954( <i>celA</i> )	SA1418
	<i>comEB</i>	×	SA1417
	<i>comEC</i>	SP0955( <i>celB</i> )	SA1416
	<i>comFA</i>	SP2208( <i>cflA</i> )	SA0705
	<i>comFB</i>	×	×
	<i>comFC</i>	SP2207( <i>cflB</i> )	SA0706
single strand DNA binding protein	<i>ssb</i>	SP1540	SA1792
DNA processing	<i>smf</i>	SP1266( <i>dprA</i> )	SA1092
competence protein	<i>yjbF</i>	SP0978( <i>coiA</i> )	SA0858
competence damage inducible protein	<i>cinA</i>	<i>cinA</i> (SP1941)	<i>cinA</i>
recombination	<i>recA</i>	<i>recA</i> (SP1940)	<i>recA</i>
mismatch repair	<i>mutS</i>	<i>mutS</i>	<i>mutS</i> 1137
nuclease	<i>nucA</i>	SP1964( <i>endA</i> )	?
alanine racemase	<i>dal</i>	SP1698	<i>alr</i>
UDP-N-acetylglucosamine 1-carboxyvinyl transferase	<i>murA</i>	SP1081	<i>murA</i> (SA1902)

**Table 6. Most of the late competence genes involved in genetic transformation are present in *S. aureus* N315 genome [181](#).**

## 2.6. Competence development in *S. aureus*

### 2.6.1. Current state of the art

To this day, there are only 3 published papers and one publication deposited in bioRxiv discussing the development of natural competence in *S. aureus*. The first paper was published in 2003 by Morikawa *et al.* [181](#). In this work, the authors identified for the first time a new staphylococcal alternative sigma factor (SigH) that controls the expression of some of the late competence genes. In the second paper, also published by Morikawa and his colleagues in 2012 [58](#), the authors showed that competence in *S. aureus* could be naturally induced in various chemically-defined media and conditions but also that SigH plays an important role in the development of competence and genetic transformation. The third paper was published in 2014 by Fagerlund *et al.* [182](#). In this publication, the authors found two homologues of ComK from *B. subtilis* present in *S. aureus*' genome. They also characterized SigH, ComK1 and ComK2 regulons by over-expressing individual or combination of regulators. Finally, Morikawa's laboratory deposited in bioRxiv, in December 2020, a last paper where they tested the potential role, in the development of competence, of the 17 two-component systems present in *S. aureus*' genome. They identified three interesting TCS, among which two activates and one inhibits competence. They also identified a new protocol, involving biofilm formation, allowing improved natural transformation.

#### 2.6.1.1. *S. aureus* has all the genes required for competence development and natural transformation

The strain N315 of *S. aureus* is mostly used to study the competence development in this species. N315 is a MRSA strain insolated in 1982. Whole genome sequences of this strain were analyzed by Makota Kuroda *et al.* in 2001 [183](#). This work allowed us to compare the gene organization between *S. aureus* and other historical transformable bacteria. For example, by comparison with *B. subtilis* and *S. pneumoniae*, most of the genes involved in competence development and genetic transformation can be found in *S. aureus*' genome, as shown in Table 6.

#### 2.6.1.2. *S. aureus* has three potential central competence regulators

##### 2.6.1.2.1. SigH

In Kuroda *et al.* [183](#), the authors wanted to identify genes encoding possible alternative sigma factors

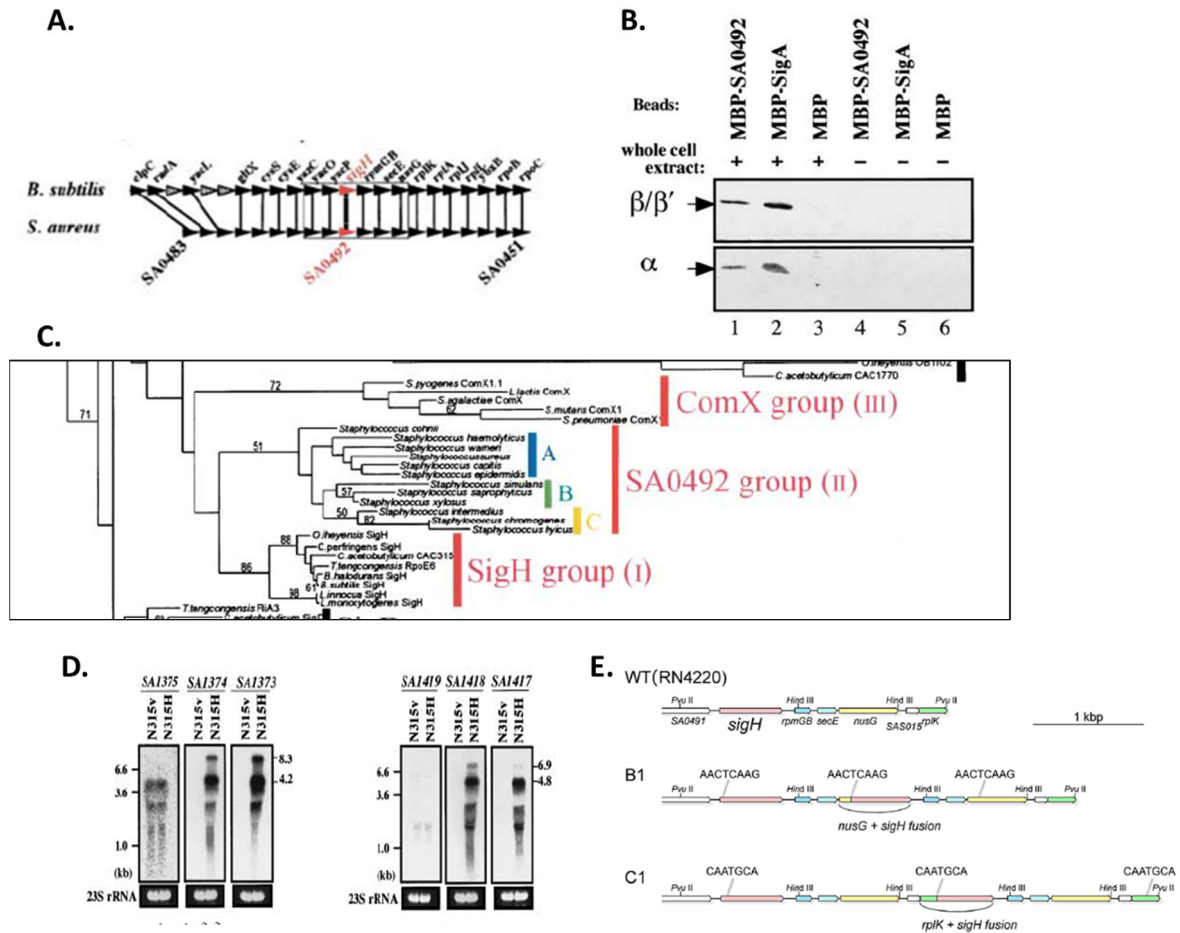


Figure 14. Identification of SigH (SA0492) in *S. aureus* <sup>181</sup>.

**A. Genetic organization around the genes encoding sigma factors in *B. subtilis* and the corresponding regions in *S. aureus*' genome.**

**B. Interaction between SA0492 (SigH) and the RNA polymerase demonstrated by a MBP pull-down assay.** MBP-SA0492 (lanes 1 and 4), MBP-SigA (lanes 2 and 5) and MBP alone (lanes 3 and 6) were mixed with *S. aureus* whole cell extracts (lanes 1, 2 and 3) or with a buffer (lanes 4, 5 and 6). The tightly bound proteins were eluted with a maltose solution. The eluates were fractionated by SDS-PAGE, followed by Western blotting using an anti-*B. subtilis* core RNA polymerase antibody. The positions of the signals for the α and β/β' subunits are marked by arrows.

**C. Unrooted phylogenetic tree (part) of the σ70-family sigma factors in 18 Firmicutes <sup>181</sup>.** The 93 amino acid residues within the conserved region (corresponding to the SA0492 amino acid residues, I(35)-G(42), N(44)-T(65), D(67)-K(92), A(129)-K(145), H(165)-R(184)) were used for the analysis. The numbers above or below the branches represent the bootstrap percentage values in 1000 replications. The values inferior to 50% are not shown. Scale bar shows the expected number of amino acid replacements per site. The names of (putative) sigma factors that have no conventional names were indicated as the nomenclature in KEGG (e.g. BH3216, CAC1766, lin0443).

**D. Two distinct operons can be regulated by SA0492 (SigH).** Northern blot analyses of the mRNAs derived from the *comG* and *comE* operons demonstrate their transcriptional induction by SA0492. The positions of the probes used are illustrated by bold lines in the gene maps. The positions of the RNA markers are indicated on the left of the panels. Ethidium bromide staining of 23S rRNA shows an equal loading of RNA.

**E. Activation of *sigH* expression involves a chromosomal rearrangement with a tandem gene duplication/fusion.** Schematic representations of the *sigH* loci of RN4220 and SigH active derivatives B1 and C1. Arrows represent coding sequences, and are colored by gene, e.g. red represents *sigH*, yellow represents *nusG* etc. The direct repeat sequences encompassing duplication units are indicated above the gene maps. These direct repeats originally exist in the RN4220 sequence, and the duplication event increases their number from two (in RN4220) to three (in B1/ C1). Gene maps are illustrated to scale and the bar represents 1 kbp.

present in N315 by comparing the genetic environment of genes encoding alternative sigma factors in a close Gram-positive relative, *B. subtilis*. Indeed, the authors found that *S. aureus* gene *sa0492* has the same genetic environment as *sigH* of *B. subtilis*, as shown in Figure 14A. Furthermore, the authors analyzed the potential counterpart of the *sa0492* gene in 12 *Staphylococcus* species. The sequence analysis showed that *sa0492* positional counterparts were found in all *strains*, the Shine-Dalgarno (SD) sequences were found in all cases and no frame-shifting mutation could be identified. Finally, analysis of SA0492 amino acid sequence showed that the most conserved regions were regions 2 and 4 of the  $\sigma^{70}$ -family sigma factors. To further demonstrate that SA0492 is a sigma factor, they used pull-down assays to show that SA0492 could bind to the core RNA polymerase (Figure 14B).

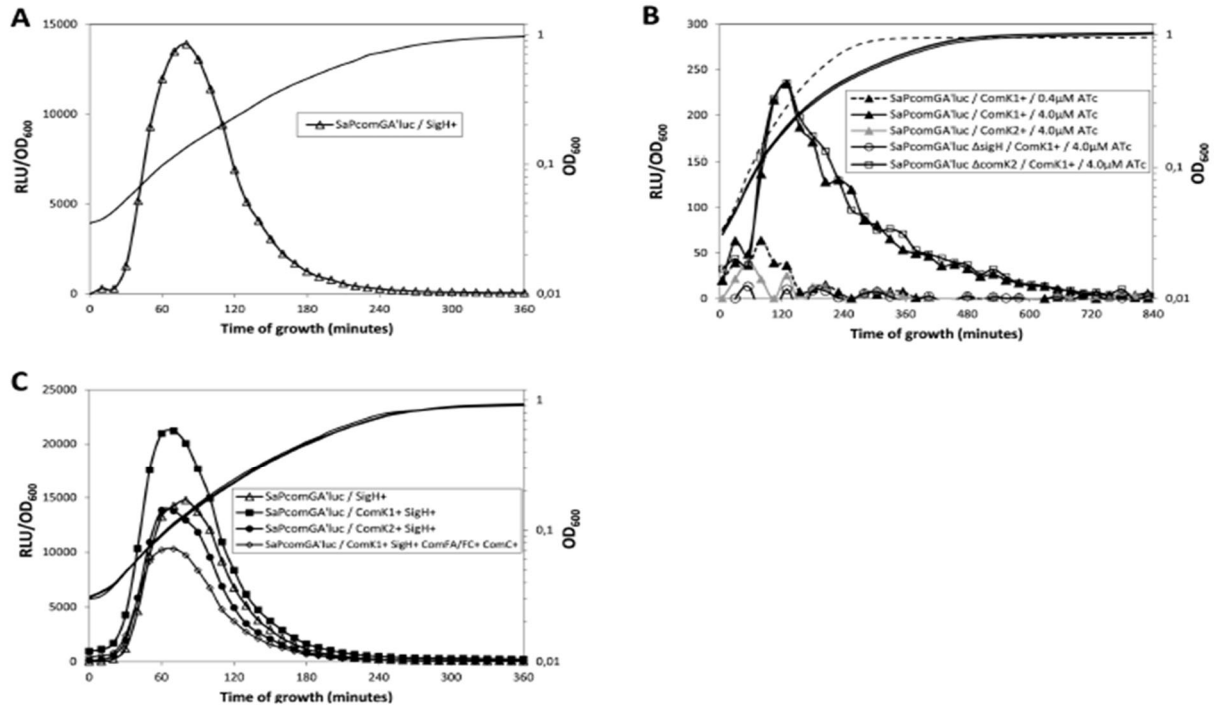
In the same study, the authors then conducted phylogenetic analysis of *sa0492* and its positional counterparts in all known  $\sigma^{70}$ -family sigma factors in *Firmicutes* for which genomic sequences were available. The phylogenetic trees showed that SA0492 and its positional counterparts in staphylococci gathered into one group and that the *sa0492* group had a close relationship with the *sigH* (including *B. subtilis*) and *comX* (including *S. pneumoniae*) groups. This phylogenetic proximity with a central competence regulator (*S. aureus* does not sporulate) suggested that *sa0492* could encode an alternative sigma factor controlling the development of competence in *S. aureus* (Figure 14C). Indeed, the author finally demonstrated that *sa0492* over-expression could lead to a drastic induction of the *comG* and *comE* operons expression (Figure 14D).

Further studies were conducted by the same team in 2012 [58](#). The authors constructed a double-reporter plasmid using the *comG* and *comE* promoters to monitor SigH activity. *comG* and *comE* expression were used to isolate SigH active cells. Under specific culture conditions, the authors showed that SigH was expressed stochastically at a frequency of roughly  $10^{-2}$ , through post transcriptional regulatory mechanisms. Inverted and repeated (IR) sequences, located upstream of the translation initiation site of *sigH* mRNA, likely form a secondary structure trapping the ribosome binding site, or serving as a post-transcriptional regulatory target, restricting SigH activation to a minor fraction of the cell population (roughly 1.6 %).

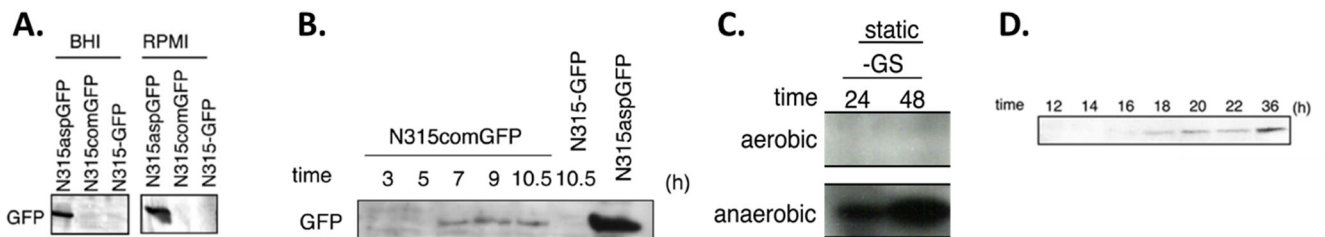
Interestingly, by further analyzing the *sigH* locus in some clones, the authors also found that the *sigH* gene could also be duplicated to form a translational fusion with the downstream *nusG* or *rplK* genes (Figure 14E), with a frequency of  $10^{-5}$  to  $10^{-9}$ . Interestingly, this duplication can be cured at a high frequency of  $10^{-2}$ . The authors proposed a theory involving short junction (SJ) tandem duplication occurring through chromosomal rearrangements, in order to explain the *sigH* gene duplication. As a consequence, SJ tandem duplication could effectively relieve the expression of the chimeric gene from the post transcriptional control of the *sigH* 5' UTR IR element.



## Introduction



**Figure 15. Luminescence assay showing expression from the *comGA* promoter upon overexpression of SigH, ComK1 and ComK2** [182](#). A. Expression of SigH in the SaPcomGA'luc strain ( $\Delta$ ), induced by addition of 0.4  $\mu$ M ATc. B. Expression of ComK1 (black lines) or ComK2 (grey lines) in strains containing the PcomGA'luc fusion ( $\blacktriangle$ ), and expression of ComK1 in SaPcomGA'luc strains additionally containing deletions in sigH ( $\circ$ ) and comK2 ( $\square$ ). Induction with ATc was with 0.4  $\mu$ M (dotted lines) or 4.0  $\mu$ M (solid lines). C. Expression of SigH alone ( $\Delta$ ), ComK1 plus SigH ( $\blacksquare$ ), ComK2 plus SigH ( $\bullet$ ), and ComK1, SigH, ComFA, ComFC and ComC ( $\diamond$ ) induced by addition of 0.4  $\mu$ M ATc, in the SaPcomGA'luc strain. Values shown are means of triplicate wells and the data are representative of at least three independent experiments.



**Figure 16. Western blot analyses directed against GFP showing that SigH is not active in standard culture conditions but is in CS2 and GS synthetic media** [58](#).

A. Western blot against GFP showing that SigH is not active in standard cultures. Strains N315aspGFP (*asp23* promoter), N315comGFP (*comG* promoter) and N315-GFP (promoter-less) were grown aerobically in BHI or RPMI 1640 medium.

B. Western blot against GFP showing that SigH activity is induced in CS2 medium. Incubation times (hours) are indicated above the panel. Positive control cells of N315aspGFP were grown in TSB for 10.5 h.

C. Western blot analysis of SigH-dependent GFP production in strain N315comGFP. Cells were grown at 37°C in -GS or CS1 medium for 24 or 48 hours with or without shaking under aerobic (upper panel) or anaerobic (lower panel) conditions. Results with the highest induction efficiency are shown, with a tendency indicating that static growth rather than shaking could enhance SigH-dependent GFP production. The removal of glycine and serine from CS1 medium improved SigH activation (-GS medium).

D. Time course of GFP induction in N315comGFP cells during anaerobic static growth in -GS medium. Cells were grown in separate tubes and harvested at the indicated times, suggesting that GFP induction seems to be a specific cellular response rather than a spontaneous event.

#### 2.6.1.2.2. ComK1 and ComK2

The other two potential central competence regulators present in *S. aureus* are ComK1 and ComK2, first described by Annette Fagerlund *et al.* [182](#). In this study, the authors analyzed the genome of *S. aureus* N315 and found two genes (*comK1* and *comK2*) encoding putative homologues of ComK from *B. subtilis*. To analyze their potential role, they used a reporter strain to monitor the expression of competence genes in *S. aureus*. This reporter strain contains a transcriptional fusion between the *comGA* promoter and the gene encoding the luciferase (*luc*), integrated in the chromosome at an ectopic location in *S. aureus* ATCC51811, creating a strain named *SaP<sub>comGA</sub>'luc*. When *sigH* was overexpressed in the *SaP<sub>comGA</sub>'luc* strain, activation of the *comGA* promoter was expectedly detected. To determine if ComK1 and/or ComK2 were involved in the transcriptional regulation of the *comG* operon, *comK1* and *comK2* were overexpressed in *SaP<sub>comGA</sub>'luc*. The results showed that ComK1 was able to activate the transcription from the *comGA* promoter at a low level, but not ComK2. When *sigH* deletion was introduced into the *SaP<sub>comGA</sub>'luc* strain, they found that the ComK1-dependant *comGA* promoter activity required the presence of SigH. However, when *comK1* and/or *comK2* deletions were introduced into the *SaP<sub>comGA</sub>'luc* strain overexpressing *sigH*, no significant difference in *comGA* expression was observed. This result indicated that neither ComK1 nor ComK2 was required for SigH-dependent transcription from the *comGA* promoter. Furthermore, when simultaneous overexpression of *sigH* and either *comK1* or *comK2* was performed, the authors found that only SigH and ComK1 displayed a synergistic activation of the *comGA* promoter (Figure 15).

#### 2.6.1.3. Competence in *S. aureus* is induced in CS2 and GS medium

In Morikawa *et al.* [58](#), to identify specific growth conditions that could lead to natural competence, the authors constructed a competence-reporter plasmid expressing the *gfp* gene under the control of the *comG* promoter (i.e. pRIT-com-GFP). Then, they introduced this plasmid into the *S. aureus* N315 strain and examined the competence-associated GFP expression by Western blot under different growth conditions. As shown in Figure 16, BHI and RPMI (two rich media) could not lead to the induction of the *gfp* expression from the *comG* promoter. However, two complete synthetic media, called CS2 (under aerobic conditions) and GS (under anaerobic conditions), could clearly induce the expression of *gfp* and therefore the development of competence in *S. aureus*.

In addition, the authors also tested other conditions, such as hypertonic shock (1 M KCl or 2 M NaCl, for 20 min), heat shock (46°C for 20 min), freeze-thaw cycles, treatment with detergents (CHAPS, NP40, Tween 20, Triton X-100), or antibiotics (bacitracin, cefazolin, gentamycin, oxacillin, penicillin G, tetra-

Locus tag MW2		Gene name	Protein function	SigH	ComK1	ComK1 plus SigH	ComK2	ComK2 plus SigH
Competence genes								
MW1496	SA1374	<i>comGA</i>	Late competence protein, traffic ATPase, binds and transports transforming DNA	392	8	1613	419	
MW1495	SA1373	<i>comGB</i>	Polytopic membrane protein, DNA transport machinery	394	8	1883	438	
MW1494	SA1372	<i>comGC</i>	DNA uptake, major component of the transformation pilus	275	5	1353	288	
MW1493	SA1371	<i>comGD</i>	DNA transport machinery, minor pilin	249	5	1318	261	
MW1492	SA1370	<i>comGE</i>	DNA transport machinery, minor pilin	174	5	853	181	
MW1491	SA1369	<i>comGF</i>	DNA transport machinery, minor pilin	205	5	1063	211	
MW1541	SA1418	<i>comEA</i>	Competence protein, membrane DNA receptor	60	3	152	58	
MW1540	SA1417	<i>comEB</i>	Late competence protein induced during competence	57	3	57	23	
MW1539	SA1416	<i>comEC</i>	Channel protein for DNA binding and uptake	24	107	107	27	
MW1132	SA1092	<i>dprA</i>	DNA recombination-mediator protein A	12	24	24	11	
MW0881	SA0858	<i>coiA</i>	Protein of unknown function, induced during competence	47	174	174	46	
MW0882	SA0859	<i>oppF</i>	Oligoendopeptidase F	2	4	4	2	
MW2021	SA1899	<i>ssb</i>	Single-strand DNA-binding protein		66	208		
Stress response genes								
MW0730	SA0723	<i>clpP</i>	ATP-dependent Clp protease proteolytic subunit			2		
MW0857	SA0835	<i>clpB</i>	ATP-dependent chaperone			2		
MW1953	SA1836	<i>groEL</i>	60 kDa chaperonin			3		
MW1954	SA1837	<i>groES</i>	10 kDa chaperonin			4		
MW1534	SA1411	<i>hrcA</i>	Heat-inducible transcriptional repressor			3		
MW1533	SA1410	<i>grpE</i>	Heat shock protein			3		
MW1532	SA1409	<i>dnaK</i>	Molecular chaperone			3		
MW1531	SA1408	<i>dnaJ</i>	Chaperone protein			2		
Genes with other functions								
MW0504	SA0507		M20 Peptidase Aminoacylase 1 family protein			4		
MW0517	SA0520	<i>sdrD</i>	Serine-aspartate repeat-containing protein D			2		
MW1251	SA1196		DNA-damage repair protein, UmuC-family protein			2		
MW1485	SA1363		Lipoate-protein ligase A			2		
MW1487	SA1365	<i>gcvPB</i>	Glycine dehydrogenase subunit 2	5	19	19	5	
MW1488	SA1366	<i>gcvPA</i>	Glycine dehydrogenase subunit 1	5	22	22	5	
MW1489	SA1367	<i>gcvT</i>	Glycine cleavage system aminomethyltransferase T	6	25	25	6	
MW1490	SA1368	<i>aroK</i>	Shikimate kinase	15	76	76	16	
MW1530	SA1407	<i>prmA</i>	Ribosomal protein L11 methyltransferase			2		
MW1535	SA1412	<i>hemN</i>	coproporphyrinogen III oxidase			5		
MW1536	SA1413	<i>lepA</i>	Elongation factor 4			6		
MW1538	SA1415	<i>hoiA</i>	DNA polymerase III subunit delta	7	33	33	2	
MW1604	SA1485	<i>radC</i>	Protein of unknown function, induced during competence (Attaiech et al., 2008)			5		
MW2020	SA1898	<i>radC</i>	Transglycosylase, ScaD precursor			5		
MW2205	SA2081	<i>sceD</i>	Urea transporter			2		
Hypothetical genes								
MW0405	SA0409		Hypothetical 64 aa protein			3		
MW0582	SA0575		Hypothetical 74 aa protein (DUF2922)				32	
MW0583	SA0576		Hypothetical 67 aa protein			5	6	49
MW1210	SAS041		Hypothetical 62 aa protein				23	
The overexpressed genes								
MW0489	SA0492	<i>sigH</i>	RNA polymerase sigma factor H	473	126	314	213	
MW0906	SA0882	<i>comK1</i>	Competence transcription factor ComK homologue 1			183		
MW1718	SA1596	<i>comK2</i>	Competence transcription factor ComK homologue 2			484		625

a. All genes with fold change  $\geq 2$  and FDR-corrected *P*-values  $< 0.05$  are listed.

Table 7. Upregulated late competence genes identified by RNA-seq [182](#).

-cycline, vancomycin at the Minimal Inhibitory Concentrations for 20 min) as potential environmental signals for the induction of competence. Unfortunately, no effect could be detected.

#### 2.6.1.4. Global transcriptional analysis of SigH, ComK1 and ComK2 regulons

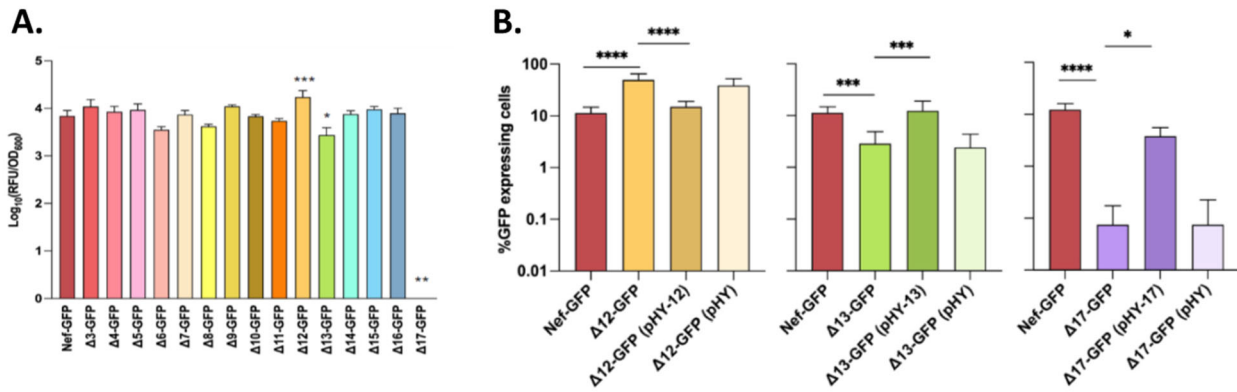
To determine the regulons controlled by SigH, ComK1 and ComK2, Fagerlund and co-workers performed global transcriptomic analyses using RNA-sequencing (RNA-seq) [182](#). To do so, the authors sequenced the total mRNA content of strains overexpressing *sigH*, *comK1*, *comK2* or combination of *comK1* plus *sigH* and *comK2* plus *sigH* (Table 7). Interestingly, a total of 40 genes was significantly upregulated in the different strains studied. Importantly, the largest class of genes had predicted functions related to natural genetic transformation.

The genes upregulated in response to *sigH* overexpression alone were the 6 genes of the *comG* operon (*comGA-F*), the 3 genes of the *comE* operon (*comEA-C*), *dprA*, and the *coiA-oppF* operon. Upon overexpression of both *comK1* and *sigH*, the increase in expression of the *comG* operon was 5-fold higher than when *sigH* was overexpressed alone, confirming their potential synergy.

Overexpression of *comK1* alone induced the expression of 7 genes. Among these were found 5 of the six genes in the *comG* operon and *comEA*, even though this induction was much smaller than when *sigH* was overexpressed. Interestingly, the most induced gene associated to *comK1* overexpression was *ssb*. Oddly, *ssb* was the only member of the ComK1 regulon that was not transcriptionally upregulated upon overexpression of *sigH* alone. However, simultaneous over-expression of *comK1* and *sigH* had a synergistic effect on the *ssb* gene expression, which was 3 times higher compared to when *comK1* was over-expressed alone. All these differences indicated that the transcriptional regulation of the *ssb* promoter presumably differs from that of the *comG* and *comE* operons.

Finally, overexpression of *comK2* alone didn't induce the expression of any transformation genes, even when overexpressed with *sigH*. Therefore, ComK2 did not appear to be involved in competence gene regulation.

Furthermore, the transcriptome analysis showed that the transcriptional activity of genes encoding essential genetic transformation proteins such as ComC, ComFA, ComFC and RecA didn't increase upon overexpression of *sigH*, *comK1* or *comK2*. Based on this result, Fagerlund proposed that these genes could be separately regulated by unknown activator(s) or repressor(s).



**Figure 17. TCS-VraSR, TCS-AgrCA and TCS-BraSR are involved in *comG* promoter activity <sup>184</sup>.**

A. The Y axis shows the increase in RFU/OD<sub>600</sub> values, which was calculated by subtracting the minimum GFP intensities from the maximum GFP intensities during 12-24 h of growth. The mean of n=3-8 independent experiments are shown.

B. The percentage expressing-GFP cells was determined after 12-24 h of growth by fluorescence microscopy. The mean of n=4-9 independent experiments are shown. Error bars represent s.d. Statistical significance was determined by one-way ANOVA with Tukey's multiple comparison test. \*P<0.05, \*\*P<0.01, \*\*\*P<0.001, \*\*\*\*P<0.0001.

	Recipient strain					
	N315h	N315Δ <i>comE</i> h	N315Δ <i>comG</i> h	N315	N315ex w/oφ pRIT-sigH	N315ex w/oφ pRIT5H
<b>Donor</b>						
Plasmid DNA (10 μg pT181)	4.0×10 <sup>-9</sup> ±3×10 <sup>-9</sup> (n=11)	ND	ND	ND		
	ND (n=1)	ND (n=3)	ND (n=3)	ND (n=3)		
	9.0×10 <sup>-7</sup> ±6.6×10 <sup>-9</sup> (n=12)			2×10 <sup>-10</sup> (n=1)		
COL/pT181		ND (n=5)	ND (n=4)	ND (n=9)		
Plasmid DNA (10 μg pHY300)					1.5×10 <sup>-6</sup> ±6.3×10 <sup>-7</sup> (n=3)	ND (n=2)

Frequencies = number of transformants/cfu of recipient strain; mean ± SD.  
ND: none detected.

**Table 8. Transformation frequencies of *S. aureus* strains <sup>58</sup>.**

N315, wild type strain. N315h carries the pRIT-sigH plasmid (used for *sigH* overexpression). N315Δ*comE*h has the same genetic background as N315h except for *comE* deletion. N315hΔ*comG*h has the same genetic background as N315h except for *comG* deletion. N315ex w/oφ is a strain that lost SCCmec and the φ prophage. N315ex w/oφ is shown carrying the pRIT-sigH or pRIT5H (empty) plasmids.

#### 2.6.1.5. TCS VraSR, AgrCA and BraSR are involved in the expression of the *comG* promoter

Very recently, Morikawa and his colleagues deposited a new publication in BioRxiv [184](#). In this paper, the authors inactivated the 15 two-component systems (TCS) present in the strain N315ex w/o  $\Phi^{13}$  (named Nef, Figure 17) to delineate the environmental conditions and early regulation pathways leading to the development of competence for natural transformation in *S. aureus*.

To do so, they monitored the *comG* operon expression under the microscope (using a *PcomG-gfp* transcriptional fusion) in a wild type strain (Nef), after 12 or 24 h of growth in CS2, compared to individual TCS mutants. In brief, compared to the Nef strain, both *agrCA* ( $\Delta 13$ ) and *braSR* ( $\Delta 17$ ) mutants were less competent while a *vraSR* mutant ( $\Delta 12$ ) appeared more competent. In addition, complementation of these individual mutants in trans restored a wild type development of competence. Based on these results, the authors concluded that among the 17 TCS present in *S. aureus* genome, only VraSR, AgrCA and BraSR seem to be involved in the regulation of natural competence.

The idea that VraSR, mainly associated to vancomycin resistance [185](#) and BraSR to bacitracin and nisin resistance [186](#) could control the development of competence in *S. aureus* is very interesting. This model would clearly expose the importance of natural genetic transformation for the emergence of *S. aureus* multi resistant strains. In addition, the potential involvement of the AgrCA system is also very exciting as Quorum Sensing (QS) systems have been shown to be involved in competence development in all the historical model organisms.

#### 2.6.1.6. *S. aureus* is able to accomplish genetic transformation

Genetic transformation in *S. aureus* was detected for the first time by Morikawa and co-workers in 2012 using plasmids as exogenous DNA ([58](#), Table 8). However, transformation efficiencies with a wild-type strain grown for 8h in CS2 medium were barely above the detection limit, around  $10^{-10}$  (to be precise, one experiment gave transformants out of 13 attempts). This result could be drastically improved by using a strain overexpressing *sigH* (from the pRIT-*sigH* plasmid), which could reach transformation efficiencies between  $10^{-6}$  and  $10^{-9}$ , depending on the experiment and the genetic background of the recipient strain.

In addition, the authors wanted to clearly establish that they were studying genetic transformation and not any artifact. To do so, they tried to transform strains lacking the *comE* or *comG* operons. As no transformation events could be detected in the absence of genes encoding essential proteins for this HGT mechanism, the authors concluded that the colonies obtained were surely the result of gene-





-tic transformation events.

Interestingly, in the same study, it was also shown that genetic transformation in *S. aureus* could be achieved through the transfer of chromosomal DNA. Crucially, the authors used as a chromosomal marker of genetic transformation the full length *SCCmec* element (52 kbp). In such experiment, they used a strain (called N315ex-h) that lost the *SCCmec* element and carries the *sigH* overexpression plasmid (pRIT-sigH) as the recipient strain. During transformation, the recipient competent cells were incubated with purified genomic DNA of a strain containing *SCCmec*. Potential transformants were selected through their ability to grow in the presence of kanamycin (the *aadD* gene is present in *SCCmec*), verified by PCR for the presence of the *mecA* gene but also the entire *SCCmec*, revealing transformation efficiencies using genomic DNA around  $3.1 \times 10^{-8}$ .

Importantly, the *sigH* overexpression used throughout the paper is not natural, as it bypasses all the stimuli and early regulatory pathways that naturally lead to the central regulators' activation. The important difference observed between natural transformation ( $10^{-10}$ ) and "forced" transformation (up to  $10^{-6}$  through *sigH* overexpression), associated with a very low number of cells naturally inducing competence (i.e. 1,6 % of the population) clearly indicated that ideal environmental conditions to induce this HGT mechanism have not been found yet.

#### 2.6.1.7. Biofilm growth conditions induce efficient natural genetic transformation

In their last publication [184](#), Morikawa and his colleagues also showed how biofilm-forming conditions can induce *comGA* expression and more generally natural transformation. In these experiments, the authors cultured a wild type strain (Nef-GFP) statically in CS2 medium to form a biofilm. Using a *PcomG-gfp* fusion, they observed that the percentage of GFP-expressing cells increased during three days to reach a maximum of 2%. Interestingly, under such conditions, a wild-type strain would display a very high transformation efficiency (compared to planktonic conditions) around  $10^{-7}$  to  $10^{-6}$ . However, unexpectedly, they found that in biofilm-forming conditions a *comG* mutant was still transformable (even though the transformation efficiency was lower, around  $10^{-8}$ ) while a *comE* mutant was not transformable. It is important to mention that the detection limit is quite low in these experiments (between  $10^{-7}$  to  $10^{-8}$ ), so the real difference between *comG* and *comE* mutants might not be that important.

Moreover, *agrCA* and *braSR* mutants were also tested in biofilm-forming conditions. Interestingly, while the *agr* mutant strain was poorly affected, with a transformation efficiency decreased by less





than a 10-fold factor, the *braSR* mutant strain displayed a very poor transformation efficiency for three days, below the detection limit ( $10^{-9}$ ) and reached almost  $10^{-8}$  on day 4. Interestingly, when bacitracin (an antibiotic for which *S. aureus* resistance is controlled by *braSR*) was added to the biofilm at a low concentration ( $0,5\mu\text{g/mL}$ ), the transformation efficiency was reduced by a 10-fold factor and completely abolished with  $5\mu\text{g/mL}$ . Therefore, the author proposed that growth in biofilm conditions increased transformation efficiency through the Agr and Bra systems.

Finally, the authors showed how the full length SCCmec element could be transferred in biofilms. To do so, they used a MSSA clinical isolate as a recipient strain and heat-killed MRSA or methicillin-resistant coagulase-negative staphylococci strains (MR-CoNS) as donors. First, they selected the *mecA* locus (located in SCCmec) through growth on cefmetazole plates. Then, they used these transformants as new donors to test their transformability in biofilm and found that SCCmec elements can be transferred to MSSA strains by natural genetic transformation in biofilm, with efficiencies ranging from  $10^{-7}$  to  $10^{-8}$ .

Biofilm formation is intimately associated to *S. aureus* infections, providing a mechanism facilitating colonization, dissemination and protection [11](#). The results presented in this publication, using a new protocol and biofilm-forming conditions, are quite interesting as natural competence for genetic transformation, at relatively high levels, is demonstrated in wild-type strains. Indeed, in previous studies [58](#), *sigH* overexpression was required to reach such efficiencies. Development of natural competence for genetic transformation in planktonic conditions still need to be investigated to understand if and how *S. aureus* can use this HGT mechanism outside from biofilms.



# Objectives

Objectives

The ability of *S. aureus* to induce natural competence for genetic transformation has only been described recently. As a result, both processes remain poorly understood. However, the limited information available already stresses the importance of studying this physiological differentiation in this important human pathogen. Indeed, the competence-specific machinery combines early regulators originally found in the two main model organisms: a cryptic secondary sigma factor, SigH (comparable to ComX in *S. pneumoniae*), and two genes encoding putative homologs of the transcription factor ComK (present in *B. subtilis*). Interestingly, the use of multiple regulators might have evolved to allow *S. aureus* to be able to use a wider range of cues to decide whether or not to become competent for genetic transformation.

In addition, even though the original publications can be considered as cornerstones in the field, the experimental conditions described to naturally induce competence clearly weren't optimum and forced the authors to over-express central regulators. By bypassing natural environmental signals and signal transduction pathways, this strategy clearly limited the amount of information available for investigation.

This is why, during my PhD, I took advantage of optimized protocols, developed in the host laboratory, to efficiently induce this environmental adaptation (up to 80% of competent cells in a given planktonic culture) in order to investigate the NATURAL development of competence in the human pathogen *S. aureus*. To achieve such objective, I dedicated my PhD work to investigate three main questions:

- 1/ What are the environmental signals leading to the NATURAL development of competence in *S. aureus*?
- 2/ What are the steps and actors integrating these signals and transmitting the information to the central competence regulators?
- 3/ What is the exact role of each potential central competence regulator?

To answer these questions, my PhD was divided in two main parts: first, using the Yeast two-hybrid technique I constructed a protein-protein interaction network centered on competence development and natural genetic transformation. Then, in the second half of my PhD, I performed genetic studies, investigating the impact of candidate genes deletion on the development of competence.

During the first half of my PhD, I personally constructed a *S. aureus* genomic library in yeast. Then, I screened this library, using the yeast two-hybrid technique, to identify proteins or peptides able to interact with the three central competence regulators and proteins of interest involved during genetic transformation. The potential partners identified in this first part of my PhD were essential to propose

## Objectives

hypotheses to answer the three questions cited above.

During the second half of my PHD, I conducted genetic studies to investigate at the transcriptional level the regulation of natural competence. To do so, I constructed a series of transcriptional fusions between promoters of interest and the gene encoding the Green Fluorescent Protein (GFP). These transcriptional fusions were used to investigate the transcriptional regulation of the three central competence regulators and several late competence genes in different genetic contexts. This is why I also constructed a series of gene deletions to understand the impact of candidate regulators on the development of competence.

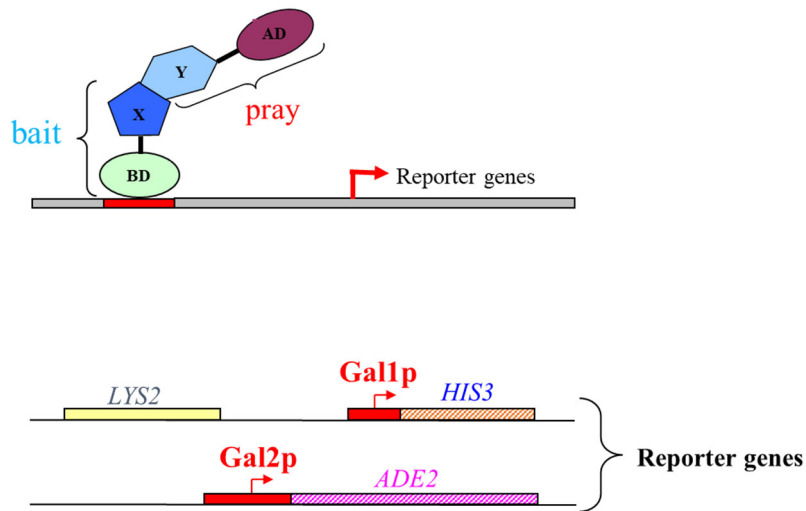
Using this approach, I investigated two main projects: first, I showed how in CS2, all the central competence regulators are involved in a complete and natural development of competence while only SigH and ComK1 activate the expression of the late competence genes involved in genetic transformation. Then, I showed how in GS medium (anaerobic conditions), ComK2 can inhibit the expression of the natural genetic transformation genes, probably through an interaction with a protein of unknown function identified in the Y2H study. Finally, I showed how the concentration of oxygen can modulate the natural development of competence. I particularly demonstrated how different oxygen conditions can be integrated by different oxygen-sensing systems to activate or inhibit competence development.

# Results



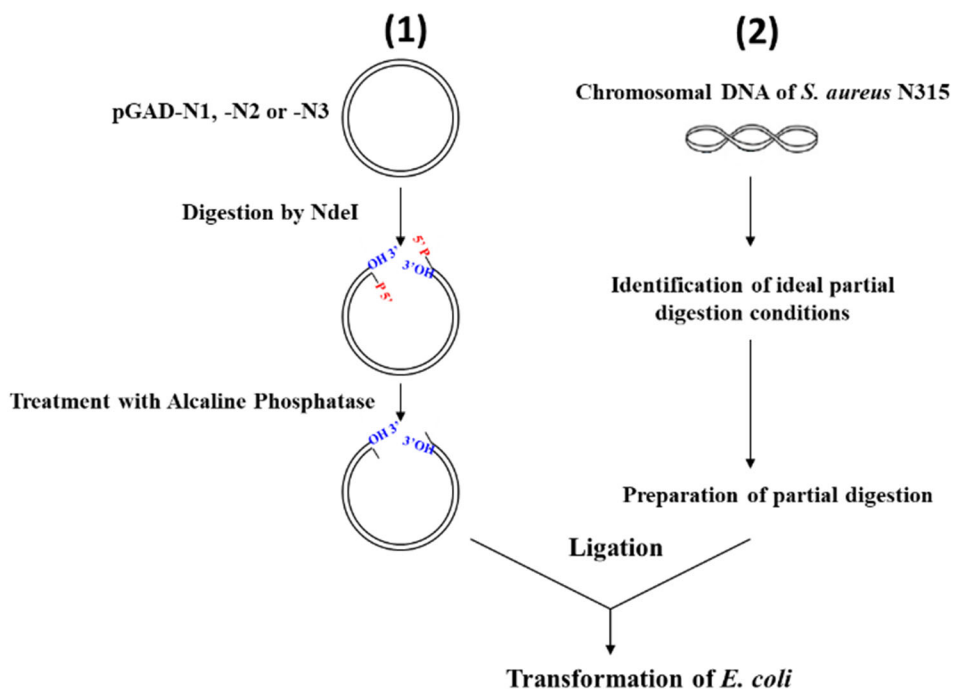


**Chapter 1 : Construction of a protein-protein interaction network  
centered on competence and genetic transformation in *S. aureus***



**Figure 1. Principle of the yeast two-hybrid technique.**

BD, DNA binding domain; AD, transcriptional activating domain; X and Y, protein of interest; *Gal1p* and *Gal2p*; promoters recognised by the reconstituted GAL4 transcriptional activator subsequent to the interaction between X and Y; *HIS*, gene encoding histidine; *ADE2*; gene encoding adenine.



**Figure 2. Strategy for our genomic DNA library construction in *E. coli*.**

## 1.1. The principle of yeast two-hybrid (Y2H)

The Y2H technique was pioneered by Fields and Song in 1989 <sup>1</sup>. It was originally designed to detect protein-protein interactions using the GLA4 transcriptional activator of the yeast *Saccharomyces cerevisiae*. GLA4 is artificially split into two separate domains: the DNA binding domain (BD) and the transcriptional activation domain (AD). The separate domains lack function alone. In the conventional Y2H approach used in this project, AD and BD domains are introduced into recombinant plasmids respectively called pGAD and pGBDU. These two domains could be synthesized as chimeric proteins fused to proteins of interest (X or Y in Figure 1). Each plasmid is introduced in a different haploid yeast strain ( $\alpha$  or  $a$ ). The sexual reproduction of yeast can then generate diploid cells where the two plasmids, pGAD-Y and pGBDU-X, coexist. The interaction between proteins X and Y, inside a diploid cell, brings the AD domain into close proximity to the BD domain, thereby reconstituting a functional GAL4 transcription activator able to drive the expression of selected reporter genes (Figure 1). The reporter genes *HIS3* (allowing growth in the absence of histidine) and *ADE2* (allowing growth in the absence of adenine) are used to select positive interactants following growth on media lacking the corresponding amino acid.

## 1.2. Construction of a *S. aureus* genomic DNA library in *E. coli*

### 1.2.1. Our strategy (Figure 2)

We decided to generate our genomic library through digestion of the *S. aureus* chromosome by restriction enzymes. The digested genome would be then cloned into 3 plasmids called pGAD-N1, -N2 and -N3. Each plasmid only differs by one nucleotide allowing to express any given fragment in the three phases.

The pGAD plasmids were digested by the restriction enzyme NdeI, generating AT extremities, and, subsequently, dephosphorylated to prevent the linear plasmids from ligating with themselves. The genomic library was prepared by partial digestion of genomic DNA of the *S. aureus* N315 strain using restriction enzymes generating extremities compatible with NdeI (i.e. MseI, Csp6I and BfaI). Once optimal partial digestion conditions were found, genomic fragments generated by each enzyme were cloned in the three pGAD plasmids. We therefore generated 9 intermediate libraries corresponding to the 9 pGAD/enzyme couples. All the intermediate libraries cloned in the same plasmid (for example pGAD-N1/MseI, pGAD-N1/Csp6I and pGAD-N1/BfaI) were then mixed to prepare three final libraries cloned in the three lecture phases. The 3 final libraries were finally ready to be transferred in yeast.



### 1.2.2. Elected restriction enzymes & estimated coverage of the genomic library

Given that the pGADs digestion by NdeI generates AT extremities, we had to choose enzymes generating compatible ends. Based on their high number of restriction sites present in *S. aureus*' genome, we decided to use the three enzymes: MseI, Csp6I and BfaI (Table1).

	Generated ends (enzyme used for cloning)	Number of cleavage sites in the genome	Average size of generated fragments (bp)
<b>Mse I</b>	<b>AT (Nde I)</b>	<b>36889</b>	<b>76.3</b>
<b>Csp 6I</b>	<b>AT (Nde I)</b>	<b>7307</b>	<b>385.2</b>
<b>Bfa I</b>	<b>AT (Nde I)</b>	<b>4351</b>	<b>646.9</b>

**Table 1. Elected restriction enzymes.**

We then performed a statistical analysis to evaluate, for each enzyme, the number of individual transformants needed in *E. coli* to ensure that every fragment generated would be present in our library. Using the equation presented below, we have estimated that if we fixed the probability to miss a given fragment at 1%, we would need a minimum of 1,6 million (MseI), 0,34 million (Csp6I) and 0,2 million (BfaI), respectively (Table 2). Of course, this minimum number of clones is larger for enzymes with a high number of cleavage sites present in the genome.

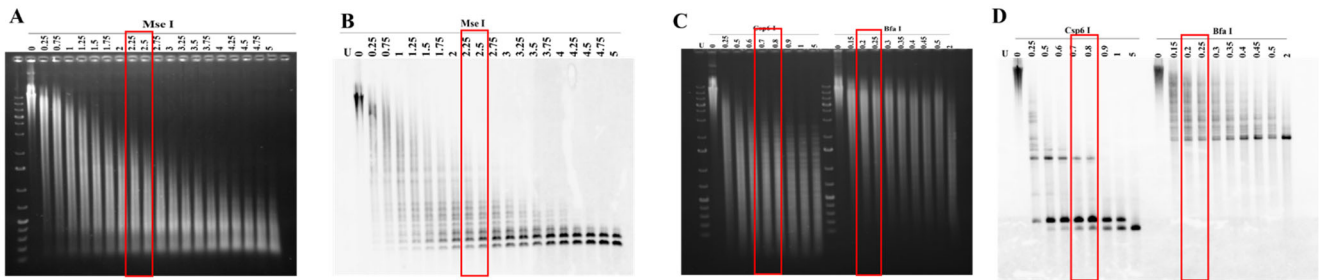
$$N = \frac{\ln F1}{\ln (1-P)} \times 10 \text{ (Finkel } et al., 1994)$$

N= Number of necessary transformants; F1 = probability accepted of missing a given insert;

P= l/G; l= average size of fragments generated by an enzyme; G= size of the genome=2,814,816 bp

	Number of the estimated transformants	
	F1=0.05	F1=0.01
<b>Mse I</b>	1,105,080	1,698,778
<b>Csp6 I</b>	218,895	336,495
<b>Bfa I</b>	130,336	200,358

**Table 2. The minimal number of transformants needed in *E. coli*.**



**Figure 3. Conditions for gDNA partial digestions.**

*S. aureus* gDNA was digested by a range of restriction enzymes units for a fixed amount of time.

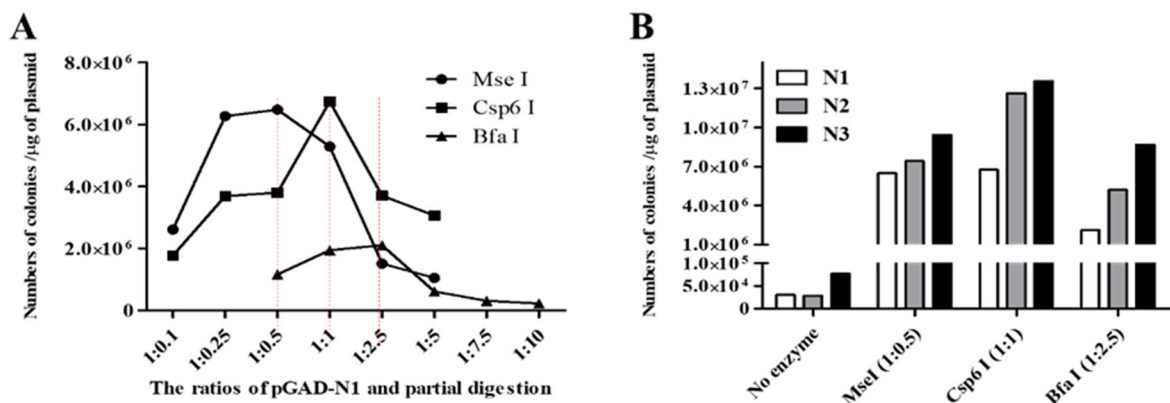
A. Agarose gel showing MseI gDNA partial digestion.

B. Corresponding Southern bolt for MseI gDNA partial digestion.

C. Agarose gel showing Csp6I and BfaI gDNA partial digestion.

D. Corresponding Southern bolt for Csp6I and BfaI gDNA partial digestion.

DNA ladder is Generuler 1kb Puls (purchased from Thermo Fisher). The red boxes indicate the “ideal” partial digestions.



**Figure 4. Optimal ratio of ligation between vector and partial digestions.**

A. Number of *E. coli* transformants obtained from ligations between linearized pGAD-N1 and each enzyme partial digestion made at different ratios (w/w).

B. Actual number of *E. coli* transformants obtained at the ideal ratios.

### 1.2.3. *S. aureus* genomic DNA partial digestions

To prepare our libraries, we decided to use partial genomic digestions (rather than total digestion). This strategy allowed us to complexify our libraries by generating several potential fragments containing the same sequence. In addition, partial digestions should allow us to express a given fragment in more than one phase.

Therefore, we first decided to identify, for each enzyme, ideal partial digestion conditions. To do so, a range of units of each enzyme was used to digest a small fixed amount of gDNA (2 µg) for a fixed time. Then, digestion products were analysed by agarose electrophoresis and southern blot, using a probe complementary to a 500bp genomic fragment with no MseI, Csp6I and BfaI cleavage site. With such probe, a total digestion would therefore result in a single band, that would be at a different size with each enzyme because they recognize different cleavage sites (for example, see the difference between Csp6I and BfaI in Figure 3D). Fragments generated by partial digestion would then spread above the total digestion band. “Ideal” partial digestion conditions were then chosen based on the host laboratory experience through the construction of previous genomic libraries. The presence of a faint total digestion band and a clear spread of partial digestion above were important characteristics to consider. Based on previous libraries construction, such conditions should generate fragments with an average size of 100 to 200 bp. The ideal partial digestions conditions for each enzyme are indicated by red boxes. Once optimal conditions were obtaining for each enzyme, we upscaled the experiment and partially digested bigger amounts (20µg) of genomic DNA to construct our libraries. The partial digestion profile was again checked by southern blot (not shown).

### 1.2.4. Cloning strategy optimization

Once genomic DNA was digested, we decided to determine the optimal ratio of ligation between each vector and partial digestions. To do so, a range of ratios of linearized pGAD-N1 (digested by NdeI and dephosphorylated) to each partial digestion was tested. As shown in Figure 4A, this optimal ratio was 1:0,5 for Mse I partial digestions, 1:1 for Csp6 I partial digestions and 1:2,5 for Bfa I partial digestions. Importantly, the actual number of transformants obtained at the ideal ratios meets the minimal number of transformants estimated in Table 2. These conditions were also tested and confirmed with the two other vectors (i.e. pGAD-N2 and -N3), as shown in Figure 4B.



Number (millions) of transformants in <i>E. coli</i>				Actual probability (F1) of missing a given insert			
	Mse I	Csp6 I	Bfa I		MseI	Csp6 I	Bfa I
N1	5.56	10.68	1.75		3.E-07	2.E-107	2.E-11
N2	41	22.84	6.3		5.E-49	7.E-229	5.E-39
N3	18.9	23.2	8.3		6.E-23	2.E-232	3.E-51

$$N = \frac{\ln F1}{\ln (1-P)} \times 10$$

Figure 5. Actual coverage of our intermediate libraries in *E. coli*.

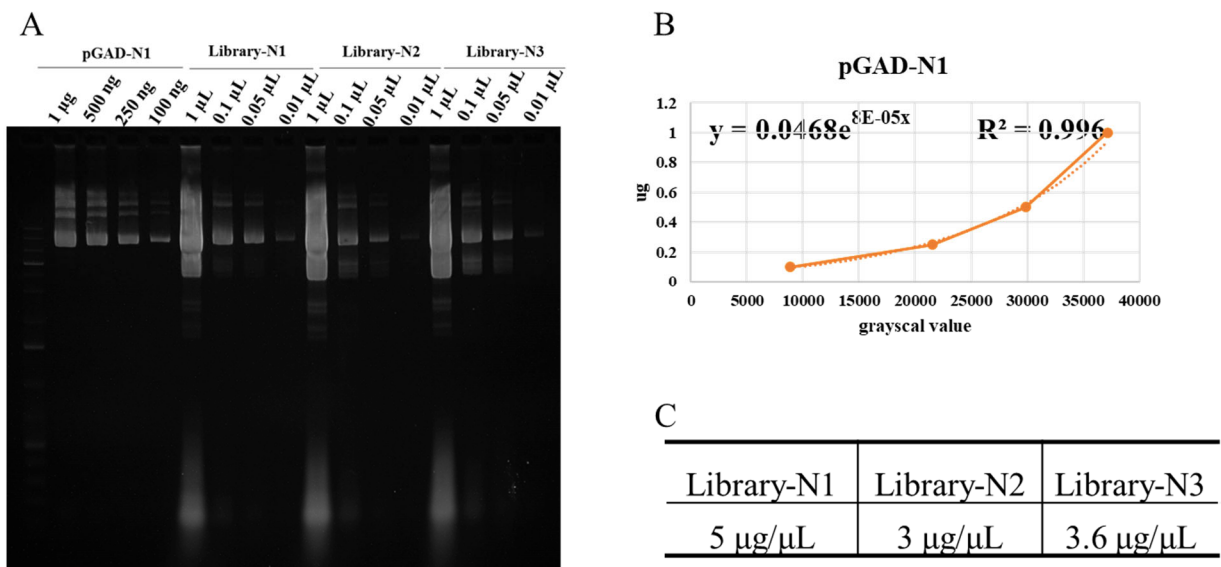


Figure 6. Libraries purified from *E. coli*.

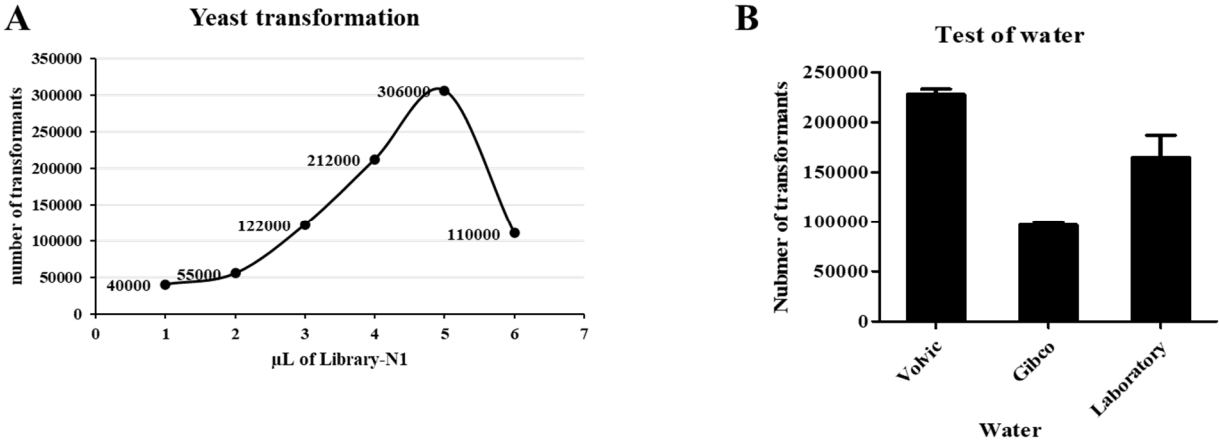
- A. Comparison of library-N1, N2 and -N3 with the empty pGAD-N1 vector on an agarose gel.
- B. Curve and fitted equation for the correlation between the fluorescence intensity detected in each lane of empty pGAD-N1 in panel A and the actual weight of pGAD-N1 plasmid deposited in each lane.
- C. Concentration of each library estimated thanks to the equation in panel B.

### 1.2.5. Libraries transformation in *E. coli*.

Using the optimal conditions described above, the experiment was once again upscaled and the ligation of each couple “pGAD/partial digestion generated by one of the three enzymes” was performed in a total of 300  $\mu\text{L}$  (1.5  $\mu\text{g}$  of plasmid). After over-night ligations, the product was purified by ethanol precipitation (described in Materials and Methods) and then transformed into *E. coli*. A total of 15 electroporations were required for each pGAD/partial digestion couple. After combining the 15 electroporations for each couple, 100  $\mu\text{L}$  of cells were plated on LB+Amp plates to estimate the total number of transformants. The results are shown in Figure 5. The actual number of the transformants obtained for each couple is much higher than the minimal number of transformants presented in the Table 2. Moreover, if we insert the actual number of transformants obtained back in the equation, the actual probability of missing a given insert is now very low (at least  $10^{-11}$ ). This result indicated that the actual coverage of our libraries is extremely good.

### 1.2.6. Libraries extraction from *E. coli*.

After electroporation into *E. coli* (for each pGAD/partial digestion couples), the electroporated cells were used to inoculate 1 Liter of LB and incubated at 37 °C overnight. This step is used to amplify our libraries and generate more plasmids to transfer in yeast. The next day, the cells were centrifuged and each pellet saved at -80 °C. Once all the partial digestions generated by the three enzymes were cloned in one of the three plasmids (for example pGAD-N1/MseI, pGAD-N1/Csp6I and pGAD-N1/BafI) and transformed in *E. coli*, the electroporated cells were all mixed together and the recombinant plasmids extracted. We therefore obtained three genomic libraries each cloned in one of the three plasmids. To estimate the concentration of each library, we first loaded different fixed amounts of empty pGAD and compared the band intensity with different volumes of each library (Figure 6A). We calculated the fluorescence of each lane of the empty pGAD-N1, generating a reference curve for which we generated a fitted equation (Figure 6B). Finally, we integrated the fluorescence of each lane for each library and entered the value in the equation to obtain each library DNA quantity. Knowing this result and the volume deposited in the lane, the concentration of each library could be calculated and is shown in Figure 6C.



**Figure 7. Yeast transformation optimization.**

A. Number of transformants obtained with different amount ( $\mu\text{g}$ ) of Library-N1.

B. Number of transformants obtained using different water to prepare the growth medium and a fixed amount of library.

### 1.3. Transfer of the libraries in the yeast strain PJ69-4 $\alpha$

#### 1.3.1. Yeast transformation optimization

Before transferring the libraries, we decided to find optimal conditions for yeast electroporation. First, we determined the ideal amount of DNA required for a fixed volume of yeast competent cells (50 $\mu$ l). As shown in Figure 7A, we tested a range of Library-N1 volumes. We found that 5  $\mu$ L of Library-N1 (corresponding to 25  $\mu$ g of DNA) was the best condition. In parallel, we tested the water that would be used to prepare the medium in which yeast competent cells are generated. As shown in Figure 7B, the Volvic<sup>®</sup> water (purchased from the supermarket) was better than the water sold by Gibco<sup>®</sup> or water directly from the laboratory sink. Finally, optimal conditions found with Library-N1 (25  $\mu$ g of DNA and Volvic<sup>®</sup> water) were tested on Library-N2 and -N3. As shown in Table 3, high numbers of transformants were obtained for both libraries, and comparable to that found with Library-N1 (Figure 7A).

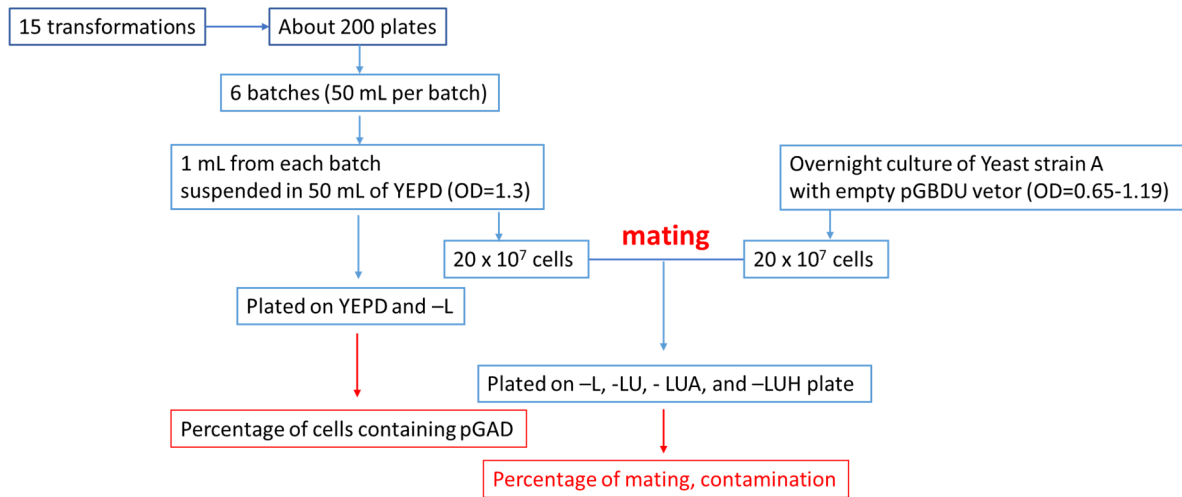
Using these numbers, we estimated the expected number of transformants that would be obtained after magnifying these transformations 10 times (i.e. 500 $\mu$ l of yeast competent cells and 250  $\mu$ g of library-N1) and repeating the experiment 15 times. As shown in Table 3, the estimated number of transformants is 10 times higher than the minimum number of transformants that was required in *E. coli*. This indicated that we had ideal conditions to transfer all the libraries complexity into yeast.

	N1	N2	N3
<b>Actual transformants</b>	306000	157000	55200
<b>Estimation</b>	45.9 million	23.5 million	8.3 million

**Table 3. The number of transformants obtained and estimated**

#### 1.3.2. Transfer of the libraries in the yeast strain PJ69-4 $\alpha$

Using the optimal conditions described above, we transformed the three libraries into the haploid yeast strain PJ69-4 $\alpha$ . The transformed cells were then plated on 200 plates containing a minimum medium lacking leucine (-L), which allowed us to select the cells that incorporated the pGAD plasmid, and incubated at 28 $^{\circ}$ C for at least 5 days. In parallel, an aliquote of each electroporation was used to estimate the actual total number of transformants obtained for each library. As shown in Table 4, the total number of transformants for each library was higher than the previous estimation (Table 3) and ensured a very good coverage of our libraries.



**Figure 8. Procedure to harvest and test the libraries in yeast.**

YEDP, rich medium; -L, minimum medium lacking leucine; -LU, minimum medium lacking leucine and uracil; -LUA, minimum medium lacking leucine, uracil and adenine; -LUH, minimum medium lacking leucine, uracil and histidine.

Library-N1	Library-N2	Library-N3
91.84 million	36.23 million	62.56 million

**Table 4. Actual number of transformants obtained in yeast for each library.**

### 1.3.3. Harvest and test of the final libraries in yeast

After 3 to 4 days of incubation at 37°C, we divided the 200 plates of each library into 6 batches (each made from 30 to 40 plates), as shown in Figure 8. The aim of dividing in batches is to avoid the contamination of the entire library. For each batch, the cells were harvested, and finally resuspended in 50 mL of 20% glycerol. 1mL of each batch was saved separately to evaluate the percentage of cells containing a pGAD plasmid (library), the mating efficiency of our libraries and potential contaminations.

#### 1.3.3.1. The percentage of cells containing a pGAD plasmid in each library

Therefore, 1mL of each batch was used to inoculate 50 mL of YEPD (rich medium). First, part of this culture was plated on YEPD and -L medium. The YEPD plates allowed us to evaluate the total number of cells (with no selection) while the -L plates allowed us to calculate the actual number of cells containing a pGAD plasmid (library). Colonies containing a pGAD and able to grow on -L plates are called Leu<sup>+</sup>. As shown in Table 5, between 60 to 85% of the cells present in each batch of our three libraries are Leu<sup>+</sup>, a result comparable to other libraries constructed in the past.

In addition, we evaluated that we obtained 60 million of Leu<sup>+</sup>/mL in library N1, 70 million of Leu<sup>+</sup>/mL in library N2 and 106 million of Leu<sup>+</sup>/mL in library N3.

N1		Bach 1		Bach 2		Bach 3		Bach 4		Bach 5		Bach 6	
	Dilution time	YEPD	-L	YEPD	-L	YEPD	-L	YEPD	-L	YEPD	-L	YEPD	-L
	-4	52	34	44	24	43	32	54	36	35	26	43	30
The number of cells containing pGAD is $6 \times 10^7$ Leu-/mL													
N2		Bach 1		Bach 2		Bach 3		Bach 4		Bach 5		Bach 6	
	Dilution time	YEPD	-L	YEPD	-L	YEPD	-L	YEPD	-L	YEPD	-L	YEPD	-L
	-3	394	335	449	311	376	221	279	247	418	299	339	252
The number of cells containing pGAD is $7 \times 10^7$ Leu-/mL													
N3		Bach 1		Bach 2		Bach 3		Bach 4		Bach 5		Bach 6	
	Dilution time	YEPD	-L	YEPD	-L	YEPD	-L	YEPD	-L	YEPD	-L	YEPD	-L
	-3	761	559	596	517	718	432	883	524	831	578	909	512
The number of cells containing pGAD is $1.06 \times 10^8$ Leu-/mL													

**Table 5. Number of total yeast cells (YEPD) and pGAD-containing cells (-L) in each batch for our three libraries.**

	Bach 1		Bach 2		Bach 3		Bach 4		Bach 5		Bach 6	
Dilution time	-L	-LU	-L	-LU	-L	-LU	-L	-LU	-L	-LU	-L	-LU
-4	334	217	186	139	260	190	251	182	185	114	246	183
	-LUA	-LUH	-LUA	-LUH	-LUA	-LUH	-LUA	-LUH	-LUA	-LUH	-LUA	-LUH
No dilution	None	None	None	None	None	None	None	None	None	None	None	None
No dilution	None	None	None	None	None	None	None	None	None	None	None	None
No dilution	None	None	None	None	None	None	None	None	None	None	None	None

Table 6. Library N1 test through mating with the PJ69-4a yeast strain (empty pGBDU).

	Bach 1		Bach 2		Bach 3		Bach 4		Bach 5		Bach 6	
Dilution time	-L	-LU	-L	-LU	-L	-LU	-L	-LU	-L	-LU	-L	-LU
-4	200	112	241	95	177	55	197	93	246	85	128	79
	-LUA	-LUH	-LUA	-LUH	-LUA	-LUH	-LUA	-LUH	-LUA	-LUH	-LUA	-LUH
No dilution	None	None	None	None	8	9	None	None	None	1	118	hundred
No dilution	None	None	None	None	8	3	None	None	None	None	hundred	hundred
No dilution	None	None	None	None	7	6	None	None	None	None	hundred	hundred

Table 7. Library N2 test through mating with the PJ69-4a yeast strain (empty pGBDU).

	Bach 1		Bach 2		Bach 3		Bach 4		Bach 5		Bach 6	
Dilution time	-L	-LU	-L	-LU	-L	-LU	-L	-LU	-L	-LU	-L	-LU
-4	352	142	352	150	282	119	279	127	408	137	216	117
	-LUA	-LUH	-LUA	-LUH	-LUA	-LUH	-LUA	-LUH	-LUA	-LUH	-LUA	-LUH
No dilution	None	None	None	None	None	None	None	None	None	None	None	None
No dilution	None	None	None	None	None	None	None	None	None	None	None	None
No dilution	None	None	None	None	None	None	None	None	None	None	None	None

Table 8. Library N3 test through mating with the PJ69-4a yeast strain (empty pGBDU).

	Library-N1	Library-N2	Library-N3
Number of aliquots	172	120	176

Table 9. Number of aliquots obtained for each library.

### 1.3.3.2. The mating efficiency and contamination of each library

Meanwhile, we prepared an overnight culture of the yeast strain PJ69-4a, containing an empty pGBDU vector. When the two cultures (PJ69-4a from the 1mL aliquot of each batch and PJ69-4a + empty pGBDU) were ready, we perform a mating experiment (where the two cultures are mixed), which actually correspond to a real screening experiment. The mated cells were finally plated on -L, -LU, -LUA and -LUH media. The -L medium allowed us to calculate the number of cells containing the pGAD (library). The -LU medium allowed us to calculate the number of cells containing both the pGAD and the pGBDU (diploid cells). Finally, the -LUH and -LUA media allowed the selection of potential auto-interactors (peptides expressed from the library able to directly interact with the “binding domain”) and estimate potential contaminants.

Division of the number of colonies detected on -LU by the number of colonies detected on -L gave us the mating efficiency of our libraries (ability to generate diploid cells). While a good mating efficiency starts at 30%, the average mating efficiency ranged from 40% for Library-N2 and -N3 and up to 70% for Library-N1, reflecting our ability to perform Y2H screening experiments.

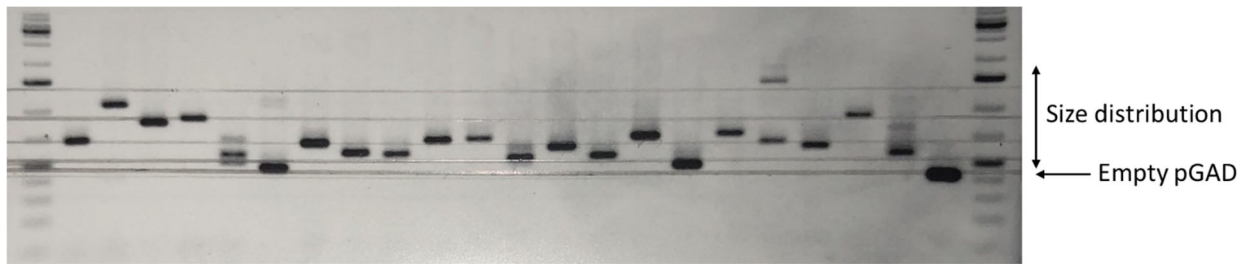
Finally, we should not have obtained any colonies on -LUH or -LUA plates. Indeed, the pGBDU vector is empty and the peptides encoded by the fragments cloned in the pGAD should not have any partner to interact with. Any colony should therefore be considered as a false positive (a library peptide able to directly interact with the BD domain) or a contaminant. As presented in Tables 6, 7 and 8, only 2 batches from Library-N2 showed colonies on -LUH and -LUA plates and based on the aspect of the colonies, we believe that they were due to contamination. The contaminated batches were removed and all the batches for a given library mixed together.

### 1.3.3.3. Final preparation of our genomic library in yeast

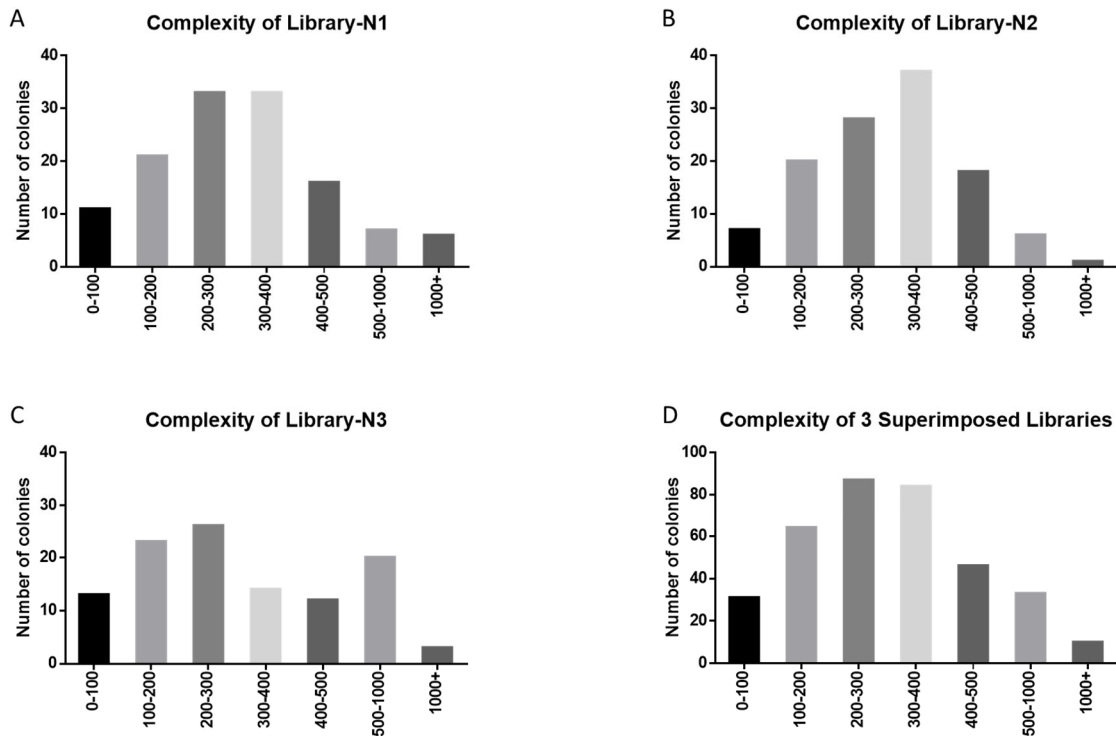
Based on these results, all the good batches for a given library were mixed together and aliquoted. Each aliquot would then be used to perform a Y2H screening experiment. Therefore, we had to be sure that each aliquot would contain a sufficient number of cells containing the pGAD (Leu<sup>+</sup>) to fully cover our libraries in each experiment.

We already estimated the number of Leu<sup>+</sup> cells/mL in Table 5. In addition, we know from experience that for libraries of this size, the minimum number of Leu<sup>+</sup> cells per aliquot should to be around 5 million. Given the number of Leu<sup>+</sup> cells/mL in each library (Table 5), we could estimate the volume per aliquot as well as the total number of aliquots available for each library (Table 9).





**Figure 9.** Evaluation of the size of genomic fragments cloned in library N1 on agarose gel. The empty pGAD is shown in the last lane and used as a size reference. The size of each fragment is then calculated thanks to the 1kb plus DNA ladder.



**Figure 10.** Size distribution of the genomic fragments cloned in the three libraries.

These graphs show the number of genomic fragments identified (ordinate) in each size range (bp, abscissa) for our three libraries (A, B and C) as well as for the overall complexity (D).

#### 1.3.3.4. Size distribution of genomic fragments cloned in each library

Finally, to test the true quality of our libraries, we decided to evaluate the size distribution of the genomic fragments cloned in our 3 libraries and transferred in yeast. To do so, we plated each library in order to isolate colonies. We then performed a PCR, directly on colonies, using primers hybridizing upstream and downstream from the multi-cloning site present in pGAD and where the gDNA fragments are cloned. The PCR products are finally loaded in an agarose gel and the size of each fragment estimated in comparison to the empty pGAD and the 1kb plus DNA ladder (Figure 9). At least 110 individual clones were analyzed for each library and the results are presented in Figure 10 and Table 10.

For each library, the genomic fragments are predominantly found between 100 and 400 bp, representing 66.2% of all the fragments analyzed. Smaller fragments (less than 100 bp) were also found and represent 8.73 % of our libraries. Finally, genomic fragments with a larger size (between 400 and 1000 bp), represent 22.25%. Interestingly, even larger fragments were found in our libraries (more than 1 kb) but they only represent 2.82%.

Considering the shape of the curves representing the size distribution of fragments in each library, it seems that Library-N1 and -N2 display the same profile, while Library-N3 (Figure 10C), seems to be composed of larger fragments, even though the predominant fragments are comparable to that of library N1 and N2.

Finally, we superimposed the data obtained in our 3 libraries in one graph. As shown in Figure 10D, the fragments are mainly found between 200 bp and 400 bp (schematized by a peak), which is slightly different from the distribution found in *S. pneumoniae* genomic libraries (previously made by Nicolas Mirouze) where predominant fragments displayed a size between 100 and 200 bp.

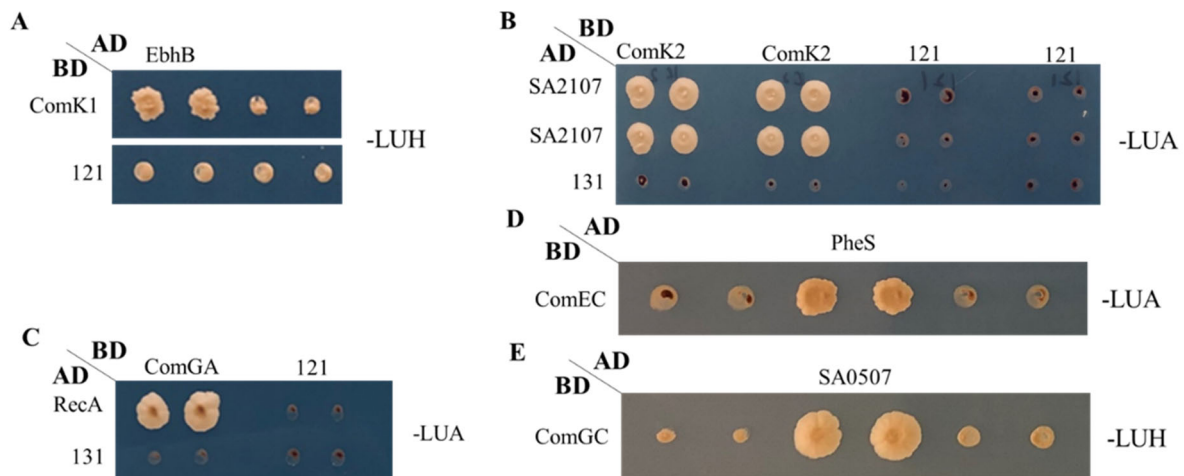
Size	N1	N2	N3	Total	% of inserts	
<b>0-100</b>	11	7	13	31	8.73%	74.93%
<b>100-200</b>	21	20	23	64	66.20%	
<b>200-300</b>	33	28	26	87		
<b>300-400</b>	33	37	14	84	22.25%	25.07%
<b>400-500</b>	16	18	12	46		
<b>500-1000</b>	7	6	20	33	2.82%	
<b>1000+</b>	6	1	3	10		
	127	117	111	355		

**Table 10.** Size distribution of the genomic fragments cloned in our three libraries transferred in yeast.

## Chapter 1

AD	N.O. colonies	Gene name	Function	Identities (aa) Start-end/Full size
SigH	3	gsaB	glutamate-1-semialdehyde aminotransferase	300-317(18)/429
	2	SA0019	conserved hypothetical protein/Two-component signal transduction system YycFG/regulatory protein YycH [Signal transduction mechanisms]	77-147(60)/453
	2	SA0828	hypothetical protein, similar to ATP-dependent nuclease subunit A	1050-1177(128)/1217
	2	SA2181	hypothetical protein/GAF domain; pfam13185/two-component sensor histidine kinase	1-119(119)/155
ComK1	2	lig	polydeoxyribonucleotide syntase	1-17(17)/667
	3	SA1885	hypothetical protein, similar to ATP-dependent RNA helicase	249-330(82)/506
	2	SA0620	secretory antigen SsaA homologue	140-166(27)/265
	4	ebhB	hypothetical protein, similar to streptococcal adhesin emb	2961-3295(335)/3890
	1	ebhA	hypothetical protein, similar to streptococcal adhesin emb	2092-2108(17)/6713
ComK2	19	SA2107	hypothetical protein	1-109/109
DprA	1	SA1448	conserved hypothetical protein/translocase of outer membrane tom70	1-43(43)/222
	1	araB	probable L-ribulokinase	419-456(38)/545
RecA	4	SA0828	hypothetical protein, similar to ATP-dependent nuclease subunit A	1050-1177(128)/1217
	2	SA1363	conserved hypothetical protein	27-65(39)/276
	2	opp-1C	oligopeptide transporter putative membrane permease domain	88-156(69)/289
ComEA	2	opp-1C	oligopeptide transporter putative membrane permease domain	88-156(69)/289
	5	SA1085	conserved hypothetical protein/Uncharacterized membrane protein YfhO [Function unknown]	1-167(167)/868
	2	tagB	teichoic acid biosynthesis protein B	34-57(24)/156
	4	scrB	sucrose-6-phosphate hydrolase	71-122(52)/494
	2	ebhA	hypothetical protein, similar to streptococcal adhesin emb	3615-3728(114)/6713
	1	araB	probable L-ribulokinase	419-456(38)/545
ComEB	2	SA2395	L-lactate dehydrogenase	129-158(30)/319
	117	ComEB	Late competence gene	18-153/153
ComFA	4	sdrC	Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding protein	682-904/953
	7	sdrD	Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding protein	1088-1336/1385
	3	sdrE	Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding protein	896-1089/1141

**Table 11. Interesting candidates for interactions obtained from library screenings.**



**Figure 11. Interactions verified by “rescue” using Y2H approach.**

- Interaction between ComK1 and EbhB.
- Interaction between ComK2 and SA2107.
- Interaction between ComGA with RecA.
- Interaction between ComEC and PheS.
- Interaction between ComGC and SA0507.

121, yeast strain a containing empty pGBDU. 131, yeast strain  $\alpha$  containing empty pGAD.

-LUH and -LUA, minimum medium respectively lacking Leucine, Uracil and Histidine or Adenine.

## 1.4. Identification of protein-protein interactions

### 1.4.1. The choice of baits

In order to construct the protein-to-protein network, we have screened our libraries with several proteins involved at different levels during competence development and genetic transformation in *S. aureus*:

- 3 potential central competence regulators: SigH, ComK1 and ComK2;
- Proteins involved in genetic transformation:
  - o allowing the binding and the transport of exogenous DNA:  
ComGA, ComGB, ComC, ComEA, ComEB, ComEC, ComFA, ComFC and ComC;
  - o contributing to the homologous recombination in host chromosome:  
DprA and RecA.

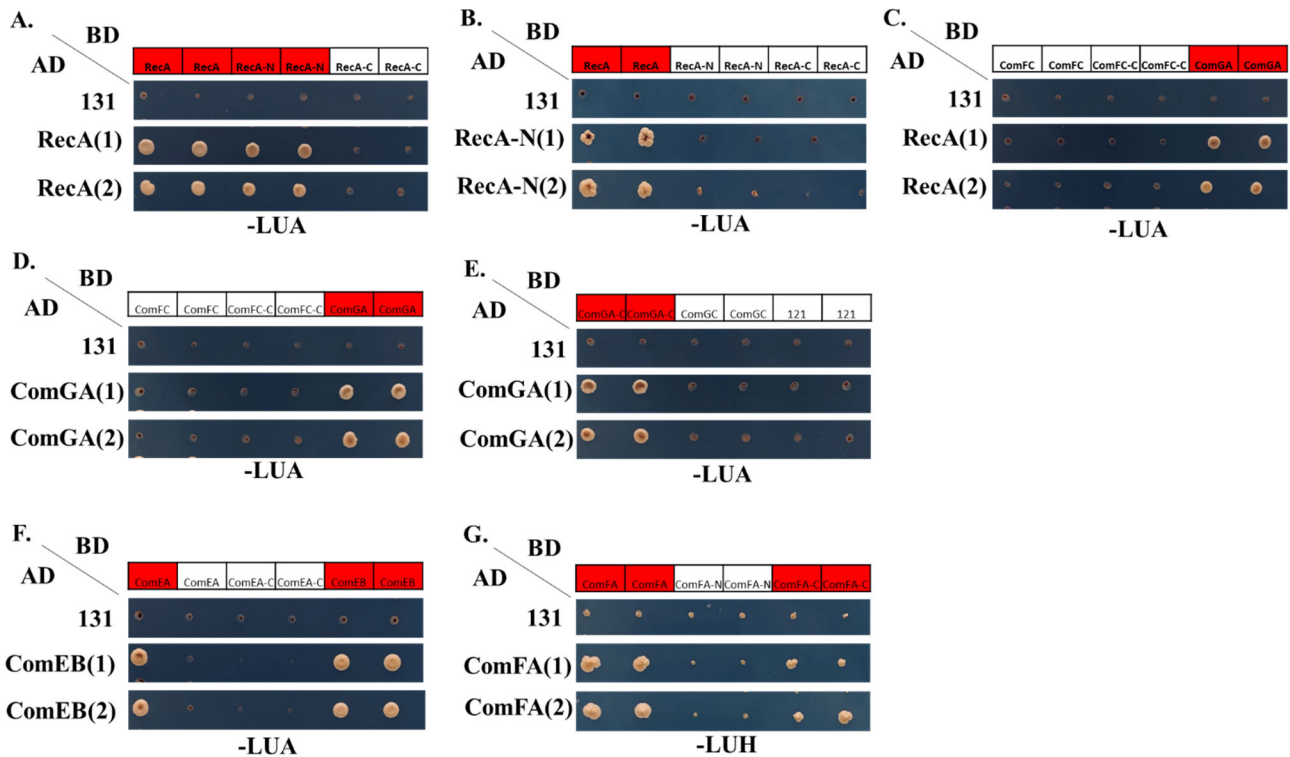
According to the principle of Y2H (Figure 1), these 12 chosen proteins would be expressed as baits by cloning their encoding gene fused to the BD domain of the GAL4 transcriptional activator into the pGBDU and transforming such plasmids into the yeast strain PJ69-4a. By screening our genomic libraries (-N1, -N2 and -3) with these baits, their potential partners, present among a huge number of potential peptides, could be discovered.

In addition to the screenings of our gDNA libraries, we also tested the ability of our selected proteins to directly interact with each other (an approach that we call “matrix”). Thus, the 12 whole genes as well as independent domains present in N- or C-terminal of the same genes (for a total of 23 independent constructs) were cloned as baits in the pGBDU (the full-length genes were also used in the screenings) and as preys in the pGAD. Such strategy, allowed us to test the simple hypothesis that proteins involved in the same pathways have a high probability to directly interact.

### 1.4.2. Overall results from our library screenings

The genomic library was screened using a mating strategy presented in Materials and Methods. After obtaining colonies able to grow on the selective medium lacking histidine, the genomic fragment cloned into the pGAD was amplified by PCR using specific primers, and then sequenced and finally compared to *S. aureus*' genome.

In total, 12 screenings were carried out. Out of all these screening, 896 candidates for interactions were isolated and identified by sequencing the PCR products. Among all the candidates, 56% comes from library-N1, 19% comes from library-N2, and 25% comes from library-N3. According to the sequence of the DNA inserts, 30.36% of these clones (i.e. 272) expressed the fusion protein in phase.



**Figure 12. Overall results from the “matrix”.**

A. Full-size RecA interacts full-size and N-terminal RecA.

B. N-terminal RecA interacts with full size RecA.

C. Interaction between RecA and ComGA.

D. Full size ComGA-BD fusion interacts with Full size ComGA-AD fusion.

E. Full size ComGA-BD fusion interacts with C-terminal ComGA-AD fusion.

F. Interaction between ComEB and ComEA, and Self-interaction of ComEB.

G. Self-interaction of ComFA.

131, yeast strain  $\alpha$  containing empty pGAD.

121, yeast strain  $\alpha$  containing empty pGBDU, is not shown everywhere, but has been tested every time.

-LUA, minimum medium lacking Leucine, Uracil and Adenine. -LUH, minimum medium lacking Leucine, Uracil and Histidine

Among these in-phase clones, 117 clones corresponding to ComEB (in the ComEB screenings), and 19 clones to SA2107 (in the Comk2 screenings). In fact, we obtained a total of 109 potential interacting proteins. They have all been presented in the network (Figure 13).

Importantly, more than half of the potential partners are “false positive”. It is inevitable to obtain false-positive interactions in Y2H, and this is criticized as the main factor affecting the quality of interaction networks generated by this approach. False positives can be explained by: (i) peptides that are not expressed in the right phase; (ii) peptides that are expressed in the right phase but that do not really interact with the prey but rather with the BD domain; (iii) a direct interaction between the bait protein with the AD domain; (iv) a chromosomal mutation in yeast affecting the general transcriptional content of the cell; (iv) or a mutation in the sequence encoding the bait or pray.

Importantly, valid interactions can be divided into three classes: (i) Class 1: one peptide can be detected in only one clone in one library; (ii) Class 2: the same peptide can be found several times in independent clones from the same library; (iii) Class 3: different peptides corresponding to the same protein can be detected in one or several libraries. Among these interactions, some have been verified by “rescue” (5 interactions), in which the genomic fragment encoding a peptide able to interact with a protein of interest is cloned back into an empty pGAD, transformed in the PJ69-4 $\alpha$  haploid cell and mated with a strain harbouring the original gene of interest cloned into the pGBDU (see Materials and methods).

### 1.4.3. Overall results from the “matrix”

A total of 23 constructs involving the whole proteins or protein domains have been tested for their ability to interacting with each other. It provided a great information and also support for our network. Indeed, some interactions identified from our library were also detected in the matrix, like the ComEB self-interaction (Table 11&Figure 12F). In total, 6 interactions have been identified, as shown in Figure 12.

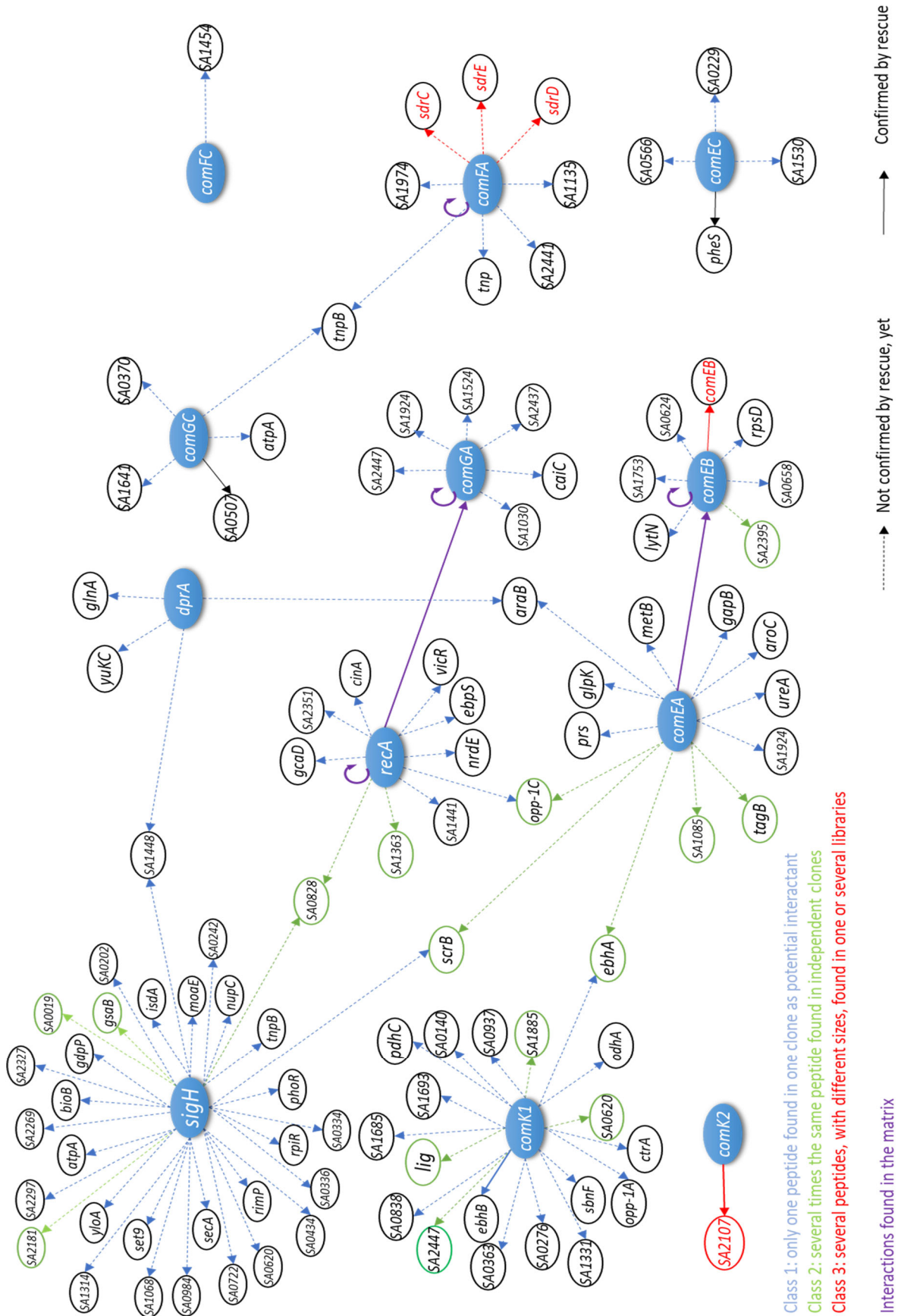
Unfortunately, some expected interactions, such as the interaction of RecA with DprA, were not detected in our matrix<sup>2</sup>. In fact, these interactions have been well characterized in other organisms by Y2H. The causes could be various. First, these interactions may not exist in *S. aureus*. Then, the three-dimensional conformation of the proteins fused to the AD or BD domains does not allow the detection of the interaction. Finally, there is a mutation in the sequence of the fusion proteins or even in the genome of the recipient cells.

In addition, except for RecA self-interaction, the other interactions cannot be detected bi-directionally. It means that we cannot reproduce these interactions if we changed their fused GAL4 domain to



another. However, this is in concordance with the observations made in previous studies with *S. pneumoniae* genomic libraries (made by Nicolas Mirouze). Generally, it has been shown that only 5% of all the interactions detected by Y2H were found bi-directional (P. Noirot, personal communication). A likely explanation is that these interactions rely on the difference in structural conformation adopted by a given protein or peptide fused to one or the other GAL4 domain, affecting the accessibility or folding of a potential interaction surface. This can also explain why some interactions could be detected in a diploid cell expressing a domain or part of a protein but not with the whole protein, or vice versa.





**Figure 13. Our global network of protein-to-protein interactions centered on natural competence and genetic transformation in *S. aureus***

## 1.5. Network of protein-to-protein interactions centred on competence development and genetic transformation in *S. aureus*

All the interactions detected, in both the screenings and the matrix, are compiled in our general protein-protein interaction network (Figure 13). Within this network, a total of 114 interactions have been presented (materialized by arrows), involving our 12 proteins of interest (in blue). The confirmed (by rescue) and unconfirmed interactions are distinguished by bold and dashed arrows, respectively. The three classes of interactions are represented by different colours: blue for Class 1, green for Class 2, and red for Class 3.

As shown in the network, the majority of the connections are unconfirmed (dashed arrows). These interactions all belong to Class 1 or Class 2, with two exceptions: the ComK1/EbhB interaction (Class 1, Table 11 & Figure 11A) and the ComEB self-interaction (Class 2, Table 11 & Figure 12F). As expected, the Class 3 interactions have all been confirmed by “rescue” (described in Materials and Methods). However, considering the fact that Class 2 interactions have been detected in more than one clone, we were surprised that very few of these could be confirmed. This result could have several explanations: these interactions are weak and difficult to reproduce in yeast (maybe requiring us to optimize the rescue protocol) but worth of *in vivo* studies; these interactions are real false-positives; or just a problem of selection on LUH medium.

### 1.5.1. Central competence regulators

First, we focused on the three potential central competence regulators: SigH, ComK1 and ComK2. Unfortunately, very few partners have been clearly identified. As presented in the network, SigH has no confirmed partners. For ComK1, only one weak interaction was detected (grown on -LUH, the less stringent medium) with an adhesin named EbhB (Figure 11A). Finally, ComK2 only interacts with SA2107, a protein of unknown function (Figure 11B). Interestingly, in Chapter 3, we investigated and proposed a role for SA2107, alongside ComK2, during competence development in *S. aureus* growing anaerobically in GS medium.

In addition to these confirmed interactions, some unconfirmed ones are also quite interesting, and therefore deserve to be considered as potentially relevant *in vivo*: we can cite for example, the SigH/YycH and SigH/NreA interactions or the ComK1/SrmB interaction. All these are classified as Class 2 interactions (Table 11 & Figure 13).



### 1.5.2. Genetic transformation proteins

Then, considering the proteins involved in the genetic transformation, we luckily discovered more interesting interactions. A brief overview of these interesting interactions has been summarized below. Among the confirmed interactions, we detected an interaction between ComGA and RecA. This result was confirmed by “rescue” (Figure 11C) and also detected in the matrix (Figure 11C). Furthermore, we found in the matrix that both ComGA and RecA can interact with themselves (Figure 12A,B,D&E).

In addition, the ComEB/ComEA interaction (Figure 12F), ComEB self-interaction (Figure 12F) and ComFA self-interaction (Figure 12G) were also detected in our matrix experiment. Interestingly, the ComEB self-interaction was also identified in our library screening.

Meanwhile, during the library screenings, we found that ComEC can interact with PheS, which is Phe-tRNA synthetase alpha chain (Figure 11D), and that ComGC can interact with a YhaA-like protein, which is similar to an N-acyl-L-amino acid amidohydrolase (Figure 11E). These two interactions were both confirmed by rescue.

Finally, some unconfirmed interactions involving genetic transformation proteins seem also quite interesting and worthy of further consideration. Among these, we found that RecA may interact with AddA and LplA (Table 11), and ComEA may interact with YfhO (Table 11). All these interactions belong to Class 2. Finally, ComFA putatively interacts with several Sdr proteins (Table 11) (Class 3).



## 1.6. Discussion on our interesting interactions

### 1.6.1. Confirmed interactions

These interactions were found during the library screenings and then verified by rescue, or directly shown during our matrix experiment.

#### 1.6.1.1. ComK1/EbhB interaction

As mentioned above, the ComK1/EbhB interaction (Figure 11A) is weak, but evident. However, when it comes to what is known about the roles of the individual proteins, this interaction seems difficult to apprehend. ComK1 in *S. aureus*, which is homologous to ComK in *B. subtilis*, is a cytoplasmic competence transcriptional regulator. EbhB is a hypothetical membrane protein, similar to the streptococcal adhesin Emb. How is it possible that such different proteins encounter each other *in vivo*? The first explanation could be that the maturation or export of the precursor of EbhB might require ComK1. Then, considering the exact role of ComK1 during competence development in *S. aureus*, is it possible for EbhB to participate in this physiological activity? However, it has never been documented that EbhB could have a role during the development of competence or genetic transformation in any other organism.

Nonetheless, it would be interesting to figure out if EbhB is involved in the competence development in *S. aureus*, or if ComK1 is essential for the maturation and export of EbhB. To answer the former question, we could analyse the expression of the late competence genes, using our transcriptional fusions, in the absence of *ehbB*. To figure out the latter question, we could compare the localisation of the EbhB protein in a wild type strain and *comK1* deletion mutant by microscopy.

#### 1.6.1.2. ComK2/SA2107 interaction

Unexpectedly, SA2107 (Figure 11B) is the only partner associated to ComK2 during our library screening. SA2107 has been detected many times in different libraries (Class 3 interaction). Therefore, it was not surprising that this interaction was verified by rescue, as shown in Figure 11B.

Unfortunately, useful information on these two proteins is limited. Few studies reported ComK2 in *S. aureus* <sup>3</sup>, and no study is available for SA2107. We know that ComK2 (i.e. SA1596) in the *S. aureus* N315 strain is the second putative homologue of ComK from *B. subtilis* and shows 18% identity and 32% similarity with ComK2 of *Bacillus cereus* <sup>3</sup>. On the other hand, SA2107 is a small protein with a



full length of 109 aa and no functional domain predicted.

As explained in the introduction, ComK2 might have a role during the development of competence in *S. aureus*. Therefore, SA2107 might also be involved in this process. Indeed, in Chapter 2, we confirmed that ComK2 is required for the full development of competence in *S. aureus* growing in CS2 medium. In addition, in Chapter 3, we investigated and proposed a role for SA2107, alongside ComK2, during competence development in *S. aureus* growing anaerobically in GS medium. This interaction, detected by Y2H, is the first confirmed *in vivo* in our network. Further investigations will be required to understand, at the molecular level, the purpose of this interaction and how it allows the control of the expression of important late competence genes involved in genetic transformation.

#### 1.6.1.3. RecA/ComGA interaction

The interactions of RecA with ComGA (Figure 11C&12C) and itself (Figure 12A&B) have been reported previously <sup>4,5</sup>. These two studies employed Foerster Resonance Energy Transfer (FRET) to demonstrate these interactions. FRET is a non-radiative process whereby energy from an excited fluorescent molecule (Donor) is transferred to a second, non-excited fluorophore (Acceptor) in its direct vicinity. FRET is therefore used as a molecular ruler due to its sensitivity in the range of 2-10 nm.

FRET experiments have the advantage to test the vicinity of proteins *in vivo*. However, they don't prove an interaction. Here, thanks to the Y2H technique we demonstrated the direct interaction between the two proteins in yeast. Therefore, we can conclude that ComGA and RecA truly interact in the competent cells.

Could the ComGA/RecA interaction be important/essential during genetic transformation? As mentioned in the introduction, ComGA is essential for the binding and transport of exogenous DNA and plays a role in the construction of the pseudopilus. Meanwhile, RecA controls the homologous recombination of the internalized ssDNA with the host chromosome. Therefore, ComGA participates in the first step of genetic transformation, while RecA is dedicated to the last step. The contact between ComGA and RecA might be important to coordinate these two sequences of events. The incoming transforming ssDNA could therefore be transferred to the "procesome" (composed of DprA, Ssb and RecA) as soon as it enters the cytoplasm. Therefore, a highly integrated multiprotein complex machine could be generated to fully coordinate DNA binding, uptake, protection and recombination during genetic transformation.





#### 1.6.1.4. ComEB/ComEA interaction

In the context of competence, it is reasonable to postulate that ComEA and ComEB, two proteins encoded by genes present in the same operon, could interact. Indeed, ComEA has already been demonstrated to be associated with ComEC, encoded by the third gene of the *comE* operon, in pull-down experiments<sup>5</sup>

*comEB* encodes a deoxycytidylate monophosphate (dCMP) deaminase<sup>6</sup>, which forms a hexamer (result confirmed by the ComEB self-interaction detected in this study). Importantly, ComEB was shown as not required during genetic transformation in *B. subtilis*<sup>6,7</sup>. The fact that ComEB may interact with ComEA, an essential protein for DNA transport during genetic competence raises different questions:

- Is the transformation apparatus in *S. aureus* different from that in *B. subtilis*? The interaction between ComEB and ComEA could therefore underline a new role for ComEB. For example, we could imagine that ComEB would interact with ComEA to promote its localisation nearby the transformation apparatus. Indeed, ComEA has been shown to localize throughout the membrane while the other transformation proteins tend to concentrate at the cell pole in *B. subtilis* [184]. According to this hypothesis, ComEB would be important/essential for genetic transformation in *S. aureus*, revealing an important difference with other model organisms.
- Is the interaction between ComEB and ComEA related to other functions, outside from genetic transformation? No other phenotype have been associated to ComEA to this day. Further studies would be required to test this hypothesis.

#### 1.6.1.5. ComFA self-interaction

This self-interaction (Figure 12G), together with the interaction of ComFA with ComFC and DprA, has already been documented<sup>8</sup>. As mentioned above, they are also devoted to the assembly of the multiprotein complex machinery dedicated to the binding and uptake of ssDNA. The di- or multimerization of ComFA might be important for its proposed function as a DNA-helicase that would “pump” ssDNA in the cytoplasm during genetic transformation.

#### 1.6.1.6. ComGC/YhaA interaction

YhaA, encoded by the *sa0507* gene in the *S. aureus* strain N315, belongs to the peptidase M20 family. These proteins are zinc binding homodimeric enzymes involved in the hydrolysis of N-acetylated



proteins. Proteins with serine and alanine in their terminus are the most frequently acetylated. Coincidentally, the N-terminal part of ComGC, the major pilin involved in the pseudopilus composition, includes this kind of residues. We can therefore hypothesize that YhaA might be involved in ComGC maturation (in addition of the cleavage performed by ComC<sup>9</sup>, ) or degradation.

### 1.6.2. Interactions that could not be confirmed yet

As shown in our network (Figure 13), a lot of interactions remain to be confirmed. Among them, some interaction that could be interesting have been summarized below.

#### 1.6.2.1. SigH/YycH interaction

This first putative interaction, detected in two independent clones during a SigH screening (Table 11), is quite interesting. Indeed, *yycH*, is the third gene of the *yycFGHIJ* operon. The two first genes of this operon encode an important TCS in *S. aureus* also named WalRK<sup>10</sup>. Importantly, the WalKR TCS is essential for cell viability and regulates cell wall homeostasis, cell membrane integrity and cell division<sup>11</sup>. Even though YycH (also known as WalH) has been shown to interact with WalK, no phenotype has been associated to *yycH* deletion.

However, the possibility of SigH interacting with WalH opens a new line of investigation. Indeed, we are interested in the stimuli sensed by the cells and the signal transduction pathways transmitting this information to the central competence regulators. This interaction could link an important central competence regulator, SigH, with the only essential TCS in *S. aureus*, WalKR. Further investigations will be required to determine if WalH is important for *sigH* expression and if this regulation requires WalKR. As WalKR is linked to cell division, a cellular process normally inhibited in competent cells<sup>12</sup>, we could also propose that this interaction could be involved in the inhibition of cell growth during competence development in *S. aureus*. As this TCS is essential, proving these hypotheses might be difficult.

#### 1.6.2.2. SigH/NreA interaction

Interestingly, this interaction was also found in two independent clones during the SigH screening (Table 11). Similarly to what we discussed in the paragraph above, *nreA* is the first gene of the *nreABC* operon. In this operon, NreB and NreC form another TCS involved in the direct sensing of nitrate/O<sub>2</sub>



able to upregulate the fermentation pathway to survive in the absence of oxygen. In this system, *nreA* encodes a nitrate sensor which dissociates from NreB in the presence of nitrate and relieves its inhibition on NreB <sup>13</sup>.

The SigH/NreA interaction could therefore reflect a new regulatory pathway linking the development of competence (through the SigH activation) and the presence or absence of nitrate. Importantly, in the third chapter of this manuscript, this same TCS (i.e. NreBC) was shown to be involved in the development of competence, through the ComK2 central competence regulator, but this time under anaerobic conditions. Therefore, the link existing between NreABC and the central competence regulators, SigH and ComK2, might allow *S. aureus* to integrate different stimuli to modulate the development of competence.

#### 1.6.2.3. RecA/AddA interaction

The interaction between RecA and AddA has been detected in four independent clones during our RecA screening (Table 11). AddA, encoded by the *sa0828* gene in the *S. aureus* N315 strain, encodes a helicase-exonuclease found in the firmicutes and the alpha-proteobacteria. Interestingly, AddAB in gram-positive bacteria, as well as RecBCD in gram-negative bacteria, is responsible for DNA repair of double-strand breaks <sup>14</sup>. Once bound to the extremities of the broken DNA, AddAB degrades one strand generating a 3'-overhang ends. Then, RecA catalyses the reparation through homologous recombination <sup>15</sup>. In the context of genetic transformation, AddAB does not recruit RecA to perform exogenous DNA integration in the chromosome by homologous recombination. This part is ensured by DprA <sup>16</sup>. Finally, it was shown in *Helicobacter pylori*, that deletion of *addA* would induce the expression of the competence genes and promote genetic transformation <sup>17</sup>.

The potential link existing between RecA and AddA is therefore quite interesting. Further investigation will be required to understand if this interaction is real *in vivo*, if it is involved during genetic transformation or during double-strand breaks reparation, and how it allows the two proteins to ensure their respective functions.

#### 1.6.2.4. ComEA/YfhO interaction

The interaction between ComEA and YfhO has been found in five independent clones during ComEA screenings (Table 11). YfhO, encoded by the *sa1085* gene in the *S. aureus* N315 strain, is a lipoteichoic-acid (LTA)-specific glycosyltransferase that adds  $\alpha$ -GlcNAc moieties to LTAs in *S. aureus* <sup>18</sup> but has also



been proposed to be involved in ribitol phosphate Wall Teichoic Acids (WTA) glycosylation with GlcNAc moieties in *Listeria monocytogenes* [19](#). Interestingly, a recent study has proposed that the initial DNA binding at the surface of competent *B. subtilis* cells was performed by WTAs, specifically produced and modified during competence [20](#). This interaction between ComEA, a DNA binding protein involved during exogenous DNA transport and a protein able to modify WTA/LTA is therefore quite interesting. This could mean that, as in *B. subtilis*, WTA (or LTA) could be involved in exogenous DNA binding at the surface of *S. aureus* competent cells. In addition, this interaction could be important to promote a step that is not yet well understood: the transfer of exogenous DNA from its initial DNA binding site at the cell surface to the multi-protein transformation apparatus, found in the cell wall and anchored in the membrane.

#### 1.6.2.5. ComFA/SdrC/SdrD/SdrE interactions

The interaction between ComFA and the three Sdr proteins have been observed in 14 independent clones, in different libraries (class 3 interaction) during ComFA screenings (Table 11). SdrC/D/E belong to the Microbial Surface Components Recognizing Adhesive Matrix Molecules (MSCRAMM) family. The MSCRAMM are important virulence factors in *S. aureus*, responsible for the initial contact with the host cells.

Surprisingly, ComFA is a cytoplasmic protein associated to the membrane inner-face that has been proposed to be an helicase, able to “pump” exogenous ssDNA, coming out of the channel ComEC, into the cytoplasm. Therefore, these proteins (ComFA and Sdr) should not be active in the same cellular compartment. We could propose some hypotheses to explain this result:

- These interactions are not true *in vivo* and ComFA tends to bind/stick non-specifically to structural domains, found in the Sdr proteins
- These interactions have a real *in vivo* purpose during, or outside, from genetic transformation. The actual state of our knowledge does not provide any hints about the function of such interactions and further experiments will be required.





## 1.7. Materials and Methods

### 1.7.1. Bacteria, Plasmid, and Growth conditions

*S. aureus* strain N315 was generously donated by Prof Morikawa from the University of Tsukuba - Japan. N315 was routinely grown at 37 °C on BHI agar or in BHI broth with aeration. *E. coli* Strain DH10 $\beta$  was grown in LB broth or LB agar. Yeast strain PJ69-4a and PJ69-4 $\alpha$  (Table S1) were grown in YEPD broth or YEPD agar. When appropriate, the following antibiotics were used: erythromycin (5  $\mu$ g/mL) or chloramphenicol (10  $\mu$ g/mL) for *S. aureus*, ampicillin (100  $\mu$ g/mL) for *E. coli*, and tetracycline (100 $\mu$ g/mL) for yeast.

Plasmids pGAD-N1, pGAD-N2 and pGAD-N3 were used to construct genomic library. Plasmid pGBDU-C1 was used to construct bait. *E. coli* containing pGADs or pGBDU plasmids were selected by ampicillin. Yeast containing pGADs or pGBDU plasmids were selected on medium lacking leucine or uracil, respectively.

### 1.7.2. Extraction of genomic DNA of *S. aureus* N315

*S. aureus* N315 was cultured in 10 mL of BHI medium with 5  $\mu$ g/mL Erythromycin and incubated at 37 °C shaking at 180rpm overnight. The cells were harvested by centrifugation at 11000 g for 1 min. The genomic DNA (gDNA) was extracted using a commercial kit (NucleoSpin Microbial DNA Kit, REF#:740235.50, purchased from MN).

### 1.7.3. Southern blot for the determination of optimal partial digestion

#### 1.7.3.1. Digestion of gDNA using each of three elected restriction enzymes (*Mse*I, *Csp*6I and *Bfa*I)

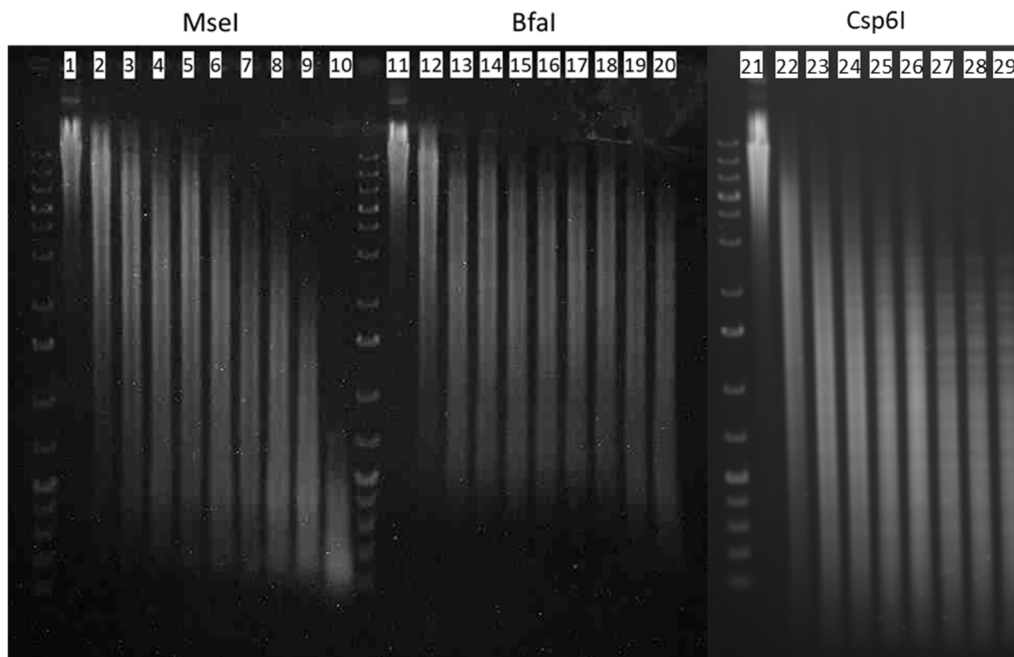
As shown in the table below, a range of *Mse*I, *Bfa*I and *Csp*6I was used to digest 2  $\mu$ g of gDNA in 20  $\mu$ L of reaction volume. *Mse*I was activated at 65 °C for 30 min and inactivated by adding 0.8  $\mu$ L of 0.5 M EDTA into each sample. Both *Bfa*I and *Csp*6I were activated at 37 °C for 30 min and inactivated at 65 °C for 20 min. Digested samples were stored at -20 °C.

Enzyme	<i>Mse</i> I									
NO. Sample	1	2	3	4	5	6	7	8	9	10
Enzyme (U)	0	0.25	0.5	0.75	1	1.25	1.5	1.75	2	5
Enzyme	<i>Bfa</i> I									
NO. Sample	11	12	13	14	15	16	17	18	19	20
Enzyme (U)	0	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	2

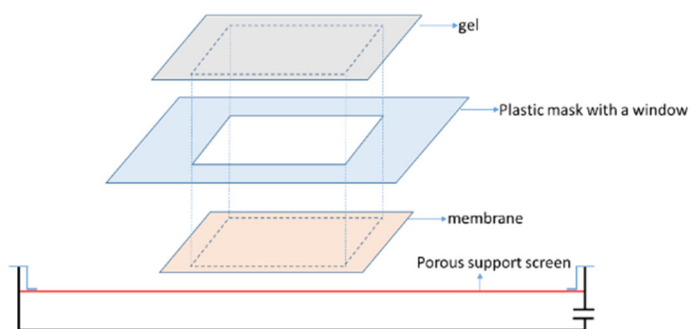
Enzyme	Csp6I									
NO. Sample	21	22	23	24	25	26	27	28	29	
Enzyme (U)	0	0.25	0.5	0.6	0.7	0.8	0.9	1	5	

1.7.3.2. Electrophoresis

10  $\mu$ L of each digested sample was loaded in 1 % agarose gel and submitted for electrophoresis at 100 V for 4 hr. The example is shown below:



1.7.3.3. DNA transfer from gel to membrane by vacuum blotting



The device used is shown in the picture above and the protocol is described as follow:

- (1) Prepare fresh solutions as shown below.

Solution	Reagent	Final concentration or volume
Depurination solution	<i>HCl</i>	0.25 M
Denaturation solution	<i>NaCl</i>	1.5 M
	<i>NaOH</i>	0.5M
Neutralizing solution	<i>NaCl</i>	1.5 M
	<i>Tris</i>	0.5M
20X SSC	<i>Na<sub>3</sub>citrate</i>	0.3 M
	<i>NaCl</i>	3 M
	<i>HCl</i>	Q.s. pH 7.0
	<i>H<sub>2</sub>O</i>	Q.s. 1 L

- (2) Cut the Hybond N<sup>+</sup> membrane (its size is bigger than the window) and pre-wet it in distilled water.
- (3) Prewet the porous support screen in distilled water and put it into the device with the smooth side up.
- (4) Mark the position of the window of the plastic mask on the support.
- (5) Put the membrane on the support and make sure the position of the window is covered. Make sure no bubbles.
- (6) Put the plastic mask on the top of the membrane. No bubbles.
- (7) Put the gel on the top of plastic mask and make sure the DNA scale is in the middle of the window. No bubbles.
- (8) Turn on the Vacuum pump, set the parameter of pressure between 40 and 60 and check whether the device is leaking.
- (9) Pour the Depurination Solution in the device and keep the gel covered for 7 min.
- (10) Remove the above solution completely.
- (11) Pour the Denaturation Solution in the device and run for 7 min.
- (12) Change the solution to Neutralising Solution as done above and run for 7 min.
- (13) Remove neutralising solution. Pour 20X SSC in the device until the gel is covered completely. Pump for 30 min and make sure the gel remains to be covered all the time.
- (14) Recycle the remaining 20 X SSC in the device. Stop pumping. Take out of the membrane and wrap it with filter paper.
- (15) Dry membrane at 37 °C for 30 min and then at 80 °C for 2hr.
- (16) Store it at room temperature.

#### 1.7.3.4. Hybridization

## Chapter 1

(17) Prepare hybridization buffer (from kit, REF#:RPN3000OL1, purchased from GE Healthcare) as follow.

<i>Hybridization buffer</i>	<i>100 mL</i>
<i>NaCl</i>	<i>2.922 g</i>
<i>Blocking agent</i>	<i>5 g</i>

Store at -20 °C.

(18) Prepare the probe by PCR (with primers of Sbprobe-1/2 shown in Table 3) and purify the PCR product by kit.

(19) Preheat the required volume (0.125 ml/cm<sup>2</sup>) of complete hybridization buffer at 42°C.

(20) Place the membrane in the above buffer to prehybridize for at least 20 min.

(21) Label probe by kit as follows:

- Dilute the probe to 10 ng/μL. Calculate the volume of probe according to the need of 1.25-2.5 ng of probe per cm<sup>2</sup> of membrane.
- Put the probe in boiling water for 5 min.
- Immediately cool it down on ice for 5 min.
- Gently mix the same volume of Labelling reagent as the probe.
- Add the same volume of glutaraldehyde as the probe, mix gently and spin briefly.
- Incubate at 37 °C for 10 min.

(22) Add the labelled probe into hybridization buffer and incubate at 42°C Cover night.

### 1.7.3.5. Detection

(23) Prepare and preheat wash buffer at 42 °C as follow.

<i>Urea</i>	<i>360 g</i>
<i>SDS</i>	<i>4 g</i>
<i>20X SSC</i>	<i>25 mL</i>
<i>H<sub>2</sub>O up to 1 L</i>	

(24) Take out of membrane from hybridization buffer and put it in 100 mL wash buffer.

(25) Shake gently at 42 °C for 20 min.

(26) Repeat the above two steps once.

(27) Prepare about 500 mL of 2X SSC.

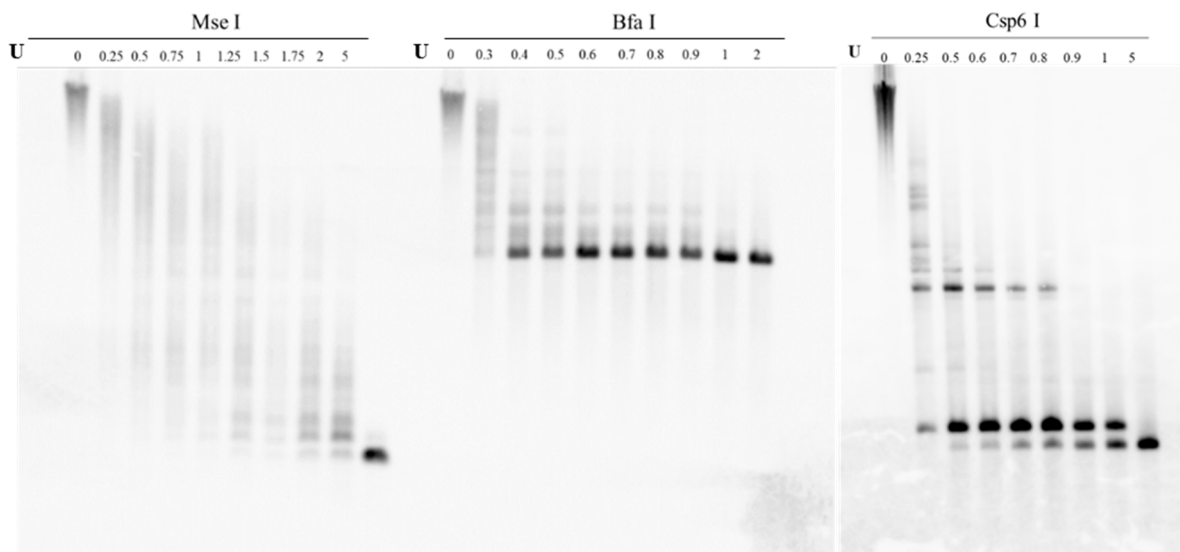
(28) Wash membrane with 2X SSC twice and shake gently for 5 min at RT.

(29) Mix 15 mL of buffer A and 15 mL buffer B (in kit).

(30) Take out of membrane and put DNA side down in the mixture of detection buffer.

(31) Wrap membrane in a plastic film and put it in the machine to take photo.

(32) Exposure for 7 min. The example is shown below.



#### 1.7.4. Preparation of *E. coli* electrocompetent cells

- (1) Streak the strain of DH10 $\beta$  on LB plate and incubate at 37 °C over night.
- (2) Inoculate 20 mL of SLB with 3 colonies. Dilute cell suspension by 100 or 1000 times in 20 mL of SLB. Incubate at 37 °C shaking at 200 rpm over night.
- (3) Preheat 1000 mL SLB in 37 °C.
- (4) Measure OD of over-night culture. Choose a culture with OD between 4 and 5.
- (5) Inoculate 1000 mL of preheated SLB with 10 mL of over-night culture. Incubate culture at 37 °C shaking at 200 rpm until OD approximately reaches 0.8-0.9.
- (6) Chill the culture on ice water immediately for 30 min. For all subsequent steps, keep the cells as close to 0 °C as possible and chill all container in ice before operation. To harvest, transfer the cells to a cold bottle and spin at 4000xg for 15 min.
- (7) Carefully pour off and discard the supernatant. It is better to sacrifice the yield by pouring off a few cells than to leave any supernatant behind.
- (8) Gently resuspend pellet in 1 L of cold water for each bottle. Centrifuge at 4000xg for 15 min, pour off and discard the supernatant.
- (9) Resuspend the pellet in 500 mL of cold water. Centrifuge at 4000xg for 15 min, pour off and discard the supernatant.
- (10) Resuspend the pellet in 20 mL of cold 10 % glycerol. Transfer to a 50 mL FALCON tube. Centrifuge at 4000xg for 15 min, pour off and discard the supernatant.
- (11) Resuspend the pellet in 3 mL of cold 10% glycerol. Aliquot the cell suspension by 100  $\mu$ L per tube, stored at -80 °C.

### 1.7.5. Preparation of pGADs plasmids for library construction

#### 1.7.5.1. Nde I digestion & dephosphorylation

Taking pGAD-N1 as an example, 60 µg of the plasmid were digested by NdeI and dephosphorylated at 37 °C for 4 hr. Enzymes were inactivated at 65 °C for 20 min.

Plasmids ( 60 µg )	64.2 µL
10x buffer	60 µL
NdeI	60 µL
Alhaline phosphatase	10 µL
H2O	355.8 µL
<b>Total volume</b>	<b>600 µL</b>

#### 1.7.5.2. DNA purification

##### (1) Phenol/Chloroform extraction

First, the mix of phenol, chloroform and Isopropanol was prepared by the ratio of 25:24:1. Secondly, 600 µL of the mixture was mixed with 600 µL of the digested sample by inverting and vertexing for 2 min. Finally, after centrifuge at room temperature at 16000 g for 5 min, the upper layer of supernatant was collected in a tube.

##### (2) Ethanol precipitation

About 300 µL of the supernatant was collected from the above process. Glycogen, 3M NH<sub>4</sub>Ac and 100% ethanol were added into the supernatant by the ratio shown in the table below.

Supernatant	About 300 µL	1 Vol
Glycogen	1 µL	0.05 Vol
3M NH <sub>4</sub> Ac	30 µL	0.1 Vol
100% ethanol	660 µL	2.2 Vol

- Incubate on ice for 15 min.
- Centrifuge at 4 °C, 15000xg for 15 min.
- Carefully discard the supernatant.
- Add 200 µL 70 % ethanol (stored at -20 °C). Centrifuge at 4 °C, 15000xg for 15 min.
- Repeat the above 2 steps.
- Remove the ethanol by heating at 50 °C. Dissolve the DNA with 100 µL water.

#### 1.7.5.3. Remove the non-dephosphorylated plasmid

##### (1) Ligation.

100  $\mu\text{L}$  of DNA obtained above was ligated with itself at 16  $^{\circ}\text{C}$  overnight. The reaction system is shown in the table below. In this way, non-dephosphorylated plasmid can form a cycle.

Sample	100 $\mu\text{L}$
Ligase	5 $\mu\text{L}$
Buffer	11.5 $\mu\text{L}$
	2 tubes

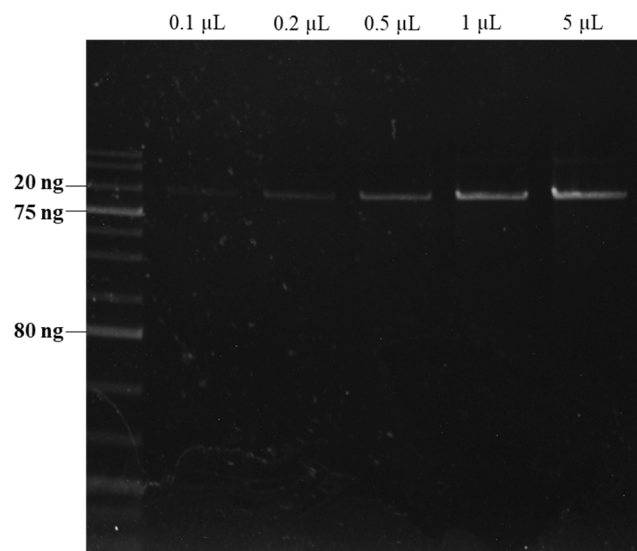
(2) Electrophoresis.

Ligation was loaded in 1% agarose gel and electrophoresed at 100V for 30 min, to separate the linear and cycle plasmid.

(3) Purify DNA from the gel by kit.

(4) Estimate the quality of DNA by comparing it with the known quality of 1kb plus DNA ladder.

A range of volume of the purified DNA and 1  $\mu\text{L}$  of 1kb plus DNA ladder were loaded in 1% agarose gel. The example is shown below. Given the known quality of 1kb plus DNA ladder, the concentration of purified DNA is roughly 90  $\text{ng}/\mu\text{L}$ .



(5) Electroporation.

To determine whether there is a cycle plasmid in the final product, 1  $\mu\text{L}$  of the purified DNA was electroporated into *E. coli*. If there is no transformants on LB plate with AMP, it indicates that the product of linear plasmid is good.

### 1.7.6. Optimization of ligation between pGAD and the partial digestion

We show here the pGAD-N1 and MseI partial digestion as an example.

(1) Ligation.



100 ng of linear and dephosphorylated pGAD-N1 was used to ligate with different amount of MseI partial digestion. The entire experiment setup is shown in the table below. All the samples were incubated at 18 °C over night.

	Not cut	Ratios						
		1:0	1:0.1	1:0.25	1:0.5	1:1	1:2.5	1:5
<b>pGAD-N1</b>	1.07	1.11	1.11	1.11	1.11	1.11	1.11	1.11
<b>Partial digestion</b>	0	1*	2.5*	5*	1**	2.5**	5**	
<b>10X Buffer</b>	2	2	2	2	2	2	2	
<b>Ligase (NEB)</b>	1	1	1	1	1	1	1	
<b>H2O</b>	18.93	15.89	14.89	13.39	10.89	14.89	13.39	10.89

Concentration of MseI digestion: 2 µg/20µL

\*100-fold dilution of MseI partial digestion

\*\*10-fold dilution of MseI partial digestion

## (2) Transformation into *E. coli*.

First, in order to remove the salt in the ligation product, 5 µL of ligation was put on the top of a 0.05 µM filter membrane that was put on the surface of distilled water. Then, 2 µL of ligation was taken from the membrane, electroporated with 50 µL of competent *E. coli*. Finally, the culture was plated on LB medium with ampicillin and incubated at 37 °C over night. The number of transformants is shown below.

LB plate	Dilution	Non cut	Cut/ Cip/Lig pGAD-N1						
			1:0	1:0.1	1:0.25	1:0.5	1:1	1:2.5	1:5
With	N.D.		27	>1000	>1000	>1000	>1000	>1000	>1000
AMP	10 <sup>-1</sup>	>1000	3	261	627	648	529	151	105

### 1.7.7. Construction of libraries in *E. coli*

#### 1.7.7.1. Clone of MseI partial digestion into the linear and dephosphorylated pGAD-N1

Two MseI partial digestions, 15 U or 20 U (both in 200 µL), were chosen. 1500 ng of pGAD-N1 was ligated with the chosen partial digestion by the optimal ratio (1:5) obtained above. The experimental setup is shown below. Reaction was performed at 16 °C over night and stopped at 65 °C for 10 min.

pGAD-N1cut/cip/lig(1500 ng )	16.65
Mse I digestion (15U/20U)	7.5
10 x Buffer	30

Ligase	15
H <sub>2</sub> O	Up to 300
Total volume	300 $\mu$ L

#### 1.7.7.2. DNA extraction from ligation by ethanol precipitation

The procedures is the same as in 5.2. Finally, DNA was dissolved in 30  $\mu$ L H<sub>2</sub>O.

Sample	300 $\mu$ L	1 Vol
Glycogen	15 $\mu$ L	0.05 Vol
3M NH <sub>4</sub> Ac	30 $\mu$ L	0.1 Vol
100 % ethanol	759 $\mu$ L	2.2 Vol

#### 1.7.7.3. Transformation into *E. coli*

In one electroporation setup, 4  $\mu$ L of DNA was transformed with 100  $\mu$ L of competent DH10 $\beta$ . According to this setup, 30  $\mu$ L of DNA obtained above was electroporated into *E. coli*. The electroporation was performed as follows:

- Prepare SOC medium.
- Chill the tubes and 2 mm cuvette on ice.
- Thaw competent cells on ice.
- Mix 100  $\mu$ L of cells and 4  $\mu$ L of DNA gently by pipetting.
- Pulse once at 2.5 kv, immediately add 1 mL of SOC and mix by inverting.
- Repeat above two steps until 30  $\mu$ L of DNA was done.
- Collect all the cell suspension in 50 mL tube and incubate at 37 °C for 1 hr.
- Take 100  $\mu$ L of culture, dilute it by 100-fold, spread 100  $\mu$ L of dilution on LB plate with AMP and incubate at 37 °C over night.
- Add the remaining culture into 1 L of LB medium with AMP. Incubate at 30 °C ,200 rpm over night.

#### 1.7.7.4. Harvest of the cells

After over night incubation, the cells were harvested and stored at -80 °C, until all partial digestion was cloned in pGAD-N1, pGAD-N2 or pGAD-N3. The procedures are shown as follows:

- Chill the culture on ice for 30 min.
- Measure OD (OD=2.7).
- Centrifuge at 6000xg for 15 min.
- Resuspend pellets in 150 mL of buffer (50 mM Tris, 5mM EDTA and 100 mM NaCl. pH=8.0) for each bottle (totally 6 bottles).

- e. Centrifuge at 6000xg for 15 min
- f. Collect all the pellets in one bottle, resuspend pellets in 50 mL of buffer and transfer the cell suspension in 50 mL tube.
- g. Centrifuge at 6000xg for 15 min
- h. Discard supernatant and stored at -80 °C.

#### 1.7.7.5. Library extraction from *E. coli*

We show here the procedure used for N1-library as an example.

(1) Prepare the fresh solutions:

	Reagent	Final concentration	Powder
<b>Solution I</b> (400 mL, pH 8.0)	Glucose	50 mM	3.6 g
	Tris	25 mM	1.2114 g
	EDTA	10 mM	1.48896 g
	Lysosyme	5 mg/mL	2 g
<b>Solution II</b> (400 mL, store at RT)	NaOH	0.2 M	3.2 g
	SDS	1 % (w/v)	4 g
<b>NaAc solution</b> (400mL, pH 4.8)	NaAc	3 M	98.44 g
<b>Lice solution</b> (60 mL)	Lice	9 M	22.89 g

(2) Resuspend three types of cells in Solution I to ensure the OD of each is same.

	OD	Resuspended in	Sol I	Mix	Proportion
N1::Mse I	2.9	To adjust the same	100 mL	100 mL	70%
N1::Csp6 I	2.6	OD	90 mL	20 mL	15%
N1::Bfa I	2.6		90 mL	12 mL	9%

- (3) Mix the three types of cells by the ratio shown in the table above and adjust the volume up to 200 mL with Solution I.
- (4) Incubate at room temperature for 10 min.
- (5) Add 400 mL of Solution II (cold and fresh), mix thoroughly and incubate in ice for 5 min.
- (6) Add 400 mL of 3M NaAc (cold and fresh), mix thoroughly and incubate in ice for 5 min.
- (7) Centrifuge at 9000 rpm for 30 min.
- (8) Collect supernatant in a new flask. If the supernatant has precipitation, centrifuge one more time.  
Add 600 mL of isopropanol into supernatant, mix thoroughly and incubate in ice for seconds.
- (9) Centrifuge at 9000 rpm for 45 min.

- (10) Discard supernatant. Dissolve pellets in each bottle with 8 mL of sterile water (Totally 6 bottles).
- (11) Collect all the suspension in one bottle, add 48 mL of Lice (cold), mix thoroughly and incubate at - 20 °C over night.
- (12) Centrifuge at 9000 rpm for 30 min
- (13) Take supernatant, mix it with 72 mL of isopropanol and distribute the mixture in four 50mL tubes.
- (14) Centrifuge at 8500 rpm for 45 min.
- (15) Wash the pellet with 100 mL of 70% ethanol (cold and fresh) by centrifuge at 8500 rpm for 10 min.
- (16) Dry the pellet in hood and dissolve the pellet in each tube with 1 mL of 10 mM tris.
- (17) Estimate the concentration of library according to the method in 5.3.

### 1.7.8. Optimization of the yeast transformation conditions

#### 1.7.8.1. Determination the optimal ratio between library and yeast

- (1) Prepare medium and solution:

Medium	Component	Quantity	Sterilisation
<b>YEPD</b>	Yeast Extract	10 g	Autoclaved at 121 °C
	Bacto Peptone	20 g	
	Glucose	20 g	
		Q.s. 1 L	
<b>-L medium</b>	Agar	20 g	Autoclaved at 121 °C
	H <sub>2</sub> O	800 mL	
	10X SD	100 mL	
	10X DO	100 mL	
<b>10X SD</b>	Yeast Nitrogen base without amino acids	67 g	Membrane filtered (0.22 µm)
	D(+)-Glucose	200 g	
	Distilled water	Q.s. 1 L	
<b>10X DO</b>	Powder exp. -L (Table S2) )	6.2 g	Membrane filtered (0.22 µm)
	Distilled water	Q.s. 1 L	

- (2) Prepare competent yeast cells
- Streak the yeast strain  $\alpha$  on YEPD plate and incubate at 28 °C for 2-3 days.
  - Inoculate 100 mL of YEPD with 5 colonies in 1 L flask, add 800 µL of tetracycline (12.5 mg/mL stock solution) into the culture, grow at 30 °C shaking at 200 rpm over night to be sure that OD of overnight culture is 1-1.2.

Chapter 1

- c. Inoculate 300 mL of YEPD with a certain amount of overnight culture in 1 L flask to be sure that the initiation OD is 0,15, add 2.4 mL of tetracycline, grow at 30 °C shaking at 200 rpm until OD approximately reached 0.69. The growth roughly took 3 hr 20 min.
- d. Centrifuge at 3500 rpm for 10 min, discard the supernatant and gently resuspend the pellet in 6 mL of sterile water.
- e. Centrifuge at 8000 rpm for 1 min, discard the supernatant and gently resuspend the pellet in 6 mL of sterile water.
- f. Centrifuge at 8000 rpm for 1 min, discard the supernatant and gently resuspend the pellet in 6 mL of LiAc/TE.
- g. Repeat the above step once.
- h. Centrifuge at 8000 rpm for 1 min, discard the supernatant, and gently resuspend the pellet in 600 µL of LiAc/TE.

(3) Transformation

- a. Following the order shown in the table below, gently mix the substances by inverting.

Library-N1	1 µL/2µL/3uL/4µL/5µL/6uL
DNA carrier	5 µL
Yeast	50 µL
50 % PEG	350 µL

- b. Incubate at 30 °C for 30 min and then at 42 °C for 20 min.
- c. Add 700 µL of water and mix by inverting.
- d. Centrifuge at 3000 rpm for 1min.
- e. Resuspend the pellets in 100 µL of water.
- f. Make a 1000-fold dilution and plate 100 µL of duliton on -L plate.
- g. The number of transformants grown on -L plate is shown in the table below.

	1 uL	2 uL	3 uL	4 uL	5 uL	6 uL	50 uL
10 <sup>-3</sup>	40	55	122	212	306/251	110	301/135

1.7.8.2. Test water used to prepare the medium.

- (1) Prepare the mediums and solutions using the water shown in the table below.
- (2) Prepare the cells as the protocol described above.
- (3) Performed transformation using the optimal amount of library obtained above.
- (4) The results are shown in the table below.

	Volvic	Gibco	Disstilled
10 <sup>-3</sup>	239	99	193
	223	94	191
	222	24	121

### 1.7.9. Construct library in Yeast

#### 1.7.9.1. Preparation of competent yeast cells.

The protocol is similar as that in 1.6.8.1, but we extended the volume of culture to 1 L and finally resuspended the cells in 8 mL of LiAc/TE.

#### 1.7.9.2. Transformation

- a. In one tube, mix 500  $\mu$ L of yeast cells and 50  $\mu$ L of library-N1, as shown in the table below.

Totally, there were 15 tubes in one experiment.

Library-N1	50 $\mu$ L
DNA carrier	50 $\mu$ L
Yeast	500 $\mu$ L
40 % PEG	3500 $\mu$ L

- b. Perform the transformation shown above. Finally, resuspend the cells in 25 mL of water.
- c. To estimate the number of transformants, take 100  $\mu$ L of the above cell suspension, dilute it by 100 times and 1000 times and plate 100  $\mu$ L of each dilution on -L plate. The results are shown below.

Dilution	10 <sup>-2</sup>	10 <sup>-3</sup>
Transformants	806	55
	982	73
Average	894	
Estimation	894*10 <sup>2</sup> *10*25=22.35 million	

- d. The remaining cell suspension was distributed on 200 of -L plates and incubated at 28 °C for 5-6 days.

#### 1.7.9.3. Harvest the cell

Contamination were checked after 3 and 5 days of incubation. The sixth day, cells were harvested as described below.

- a. Divide all the plates into 6 batches and finish two batches in one day.

- b. For one plate, use 6 mL of sterile water to scrape off the cells from plate, and use another 6 mL of sterile water to rinse the plate once.
- c. After finishing one batch of plates, collect the cells were by centrifugation at 5000 rpm for 5 min.
- d. Resuspend the pellet in 50 mL of 20% Glycerol, take 2 mL from 50 mL of cell suspension in two independent tubes, and store at -80 °C.

**1.7.10. Evaluate the library in yeast (also the protocol of screening)**

(1) Prepare the media and plates.

The needed media are YEPD, -L, -U liquid medium and agar plates, -LU, -LUH and -LUA agar plates.

(2) Prepare the bait cells

- a. Streak strain 121, which is yeast A containing empty pGBDU, on -U plate. Incubate at 28 °C for 2-3days. (If screening, streak yeast A containing pGBDU-bait)
- b. Inoculate 50 mL of -U medium with 5 colonies of strain 121, add 400 µL of tetracycline (12.5 mg/mL) in medium and grow at 30 °C shaking at 200 rpm over night.
- c. Inoculate 150 mL of -U medium with a certain amount of the over-night culture in 1 L FLASK to be sure that the initiation OD is OD=0.008. Incubate at 30 °C shaking at 200 rpm over night.
- d. Measure OD to ensure that OD is roughly 0.67 (not more than1.19). At this moment, the concentration of cells in the culture is  $1.24 \times 10^7$  cells per mL (calculated by  $2 \cdot X \cdot 10^7$  CFU/mL/OD). Therefore, 16 mL of culture has  $20 \times 10^7$  CFU that are needed for mating.

(3) Prepare the library

- a. Thaw the separated 1 mL of each batch that belongs to the same library by 37 °C water bath. (If screening, directly thaw one aliquot)
- b. Add 1 mL of each batch into 20 mL YEPD. Incubate at 30 °C shaking at 200 rpm for 30 min. (If screening, add an aliquot of library into 20 mL YEPD)
- c. Measure OD of each batch. For example, each batch of library -N1.

Batch	#1	#2	#3	#4	#5	#6
OD	1.55	1.42	1.27	1.34	1.40	1.29

- d. According to the average OD (that is 1.38), roughly 7.2 mL of each batch has  $20 \times 10^7$  cells for mating.

- e. In parallel, take 100  $\mu$ L from each batch, dilute it by 1000 and 10000 times and plate 100  $\mu$ L of each dilution on YEDP plate and -L plate.

#### (4) Mating

- a. According to the calutron above, mix 7.2 mL of each batch with 16 mL of 121, respectively. (If screening, mix 20 mL of library with the yeast A culture that has the same number of cell as library)
- b. Centrifuge at 3500 rpm for 10 min.
- c. Resuspend each batch of pellets in 300  $\mu$ L of water and spread each cell suspension on YEPD plate with tetracycline. (If screening, resuspend the pellet in 2 mL of water, distribute on 5 YEPD plates.)
- d. Leave the beads in the plates, keep the plates facing up and incubate at 30 °C for 5 hr.
- e. Recover cells from each plate by using 10 mL of sterile water to rinse each plate twice.
- f. Collect the cell by centrifugation at 3500 rpm for 10 min.
- g. Resuspend each batch of cells in 1 mL of sterile water ( If screening, resuspend all the cells in 6 mL of sterile water.)
- h. Make 1000-fold dilution and plate 100  $\mu$ L of each dilution on -L, -LU, -LUH or -LUA plate. (If screening, directly distribute all the cells on 40 -LUH plates and 40 -LUA plates)

#### 1.7.11. The “Rescue” protocol

First, considering that the fragments were inserted in pGAD, we had to design primers containing homology with pGAD. Then, we performed hot start PCR on the colonies obtained by screening our libraries. After sequencing the PCR products, we knew whether the fragments were cloned in pGAD in the right phase. If it was the case, we transformed the mixture of 2  $\mu$ L of PCR and 100 ng of open pGAD into yeast strain  $\alpha$  (see 1.6.8.1 (3)). Finally, the transformants were selected by -L medium and verified by PCR.

Once we obtained the “ $\alpha$ ” yeast strain containing the pGAD in which the fragments were cloned, we performed the mating between the above yeast strain and the “a” yeast strain containing the bait used in the corresponding screening (see step 4 of 10). If the interaction is also detected this way, it will be considered as real in yeast.

#### 1.7.12. Construction of bait and prey vectors for Y2H



## Chapter 1

The open reading frames or the sequence encoding protein domains were PCR-amplified from genomic DNA of the wild-type *S. aureus* N315 strain, using the primers shown in Table S3. The various PCR products were cloned into the bait vector pGBDU-C1 and the prey vector pGAD-N1 by digestion of vectors and PCR products by restriction enzymes and transferred in *E. coli* by electroporation. Bait and prey constructs obtained in *E. coli* were purified and introduced by electroporation in the PJ69-4a or PJ69-4 $\alpha$  yeast strains, respectively. The DNA sequences of all cloned fragments were verified using the primers P52-54 (Table S2).

## Supplementary Data

Table S1. Strains

Strains	Genotype
PJ69-4a	<i>MAT<math>\alpha</math>, trp1-901 leu2-3,112 ura3-52 his 3-200 gal4<math>\Delta</math>gal80<math>\Delta</math> LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ</i>
PJ69-4 $\alpha$	<i>MAT<math>\alpha</math>, trp1-901 leu2-3,112 ura3-52 his 3-200 gal4<math>\Delta</math>gal80<math>\Delta</math> LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ</i>

Table S2. Composition of DO(10X)

DO	-L	-U	-LU	-LUH	-LUA	
Uracil	2 g			2 g		
Adenine sulfate	2 g	2 g	2 g	2 g		
L-Histidine HCl	2 g	2 g	2 g	2 g	2 g	
L-Tryptophan	2 g	2 g	2 g	2 g	2 g	
L-Leucine		10 g				
L-Isoleucine	3 g	3 g	3 g	3 g	3 g	
L-Valine	15 g	15 g	15 g	15 g	15 g	
L-Arginine HCl	2 g	2 g	2 g	2 g	2 g	
L-Lysine HCl	3 g	3 g	3 g	3 g	3 g	
L-Methionine	2 g	2 g	2 g	2 g	2 g	
L-Phenylalanine	5 g	5 g	5 g	5 g	5 g	
L-Threonine	20 g	20 g	20 g	20 g	20 g	
L-Tyrosine	3 g	3 g	3 g	3 g	3 g	
Atriazole	-	-	-	8.4 g	-	
	g/L	0.61	0.69	0.59	0.612	0.47

Table S3. Primers

Primers	Name	Nucleotide sequence
P01	Sbprobe-1	TGG TCT GCG ATG GGG CAT CAT ATG
P02	Sbprobe-2	GCA TAT ACT GTG GAT GCT GTC GCA ACA

## Chapter 1

P52	mAD1ext	AAC GGT CCG AAC CTC ATA AC
P53	mBD2ext	AGC TTC TGA ATA AGC CCT CG
P54	mAD1	CGC GTT TGG AAT CAC TAC AGG
P55	mBD1ext	GTC TCC GCT GAC TAG GGC AC
P60	Sal I-sigH-F	GCA GCA GTC GAC TTG AAA TAT GAT TTG ACA ACT C
P61	Bgl II-sigH-R	GCA GCA AGA TCT TCA AAT CAT TTT GAA ATA ACG TTT TAT
P62	Sal I-comK1-F	GCA GCA GTC GAC ATG TAT TCT CAA AAT ATT TAT GTG ATA C
P63	Bgl II-comK1-R	GCA GCA AGA TCT TTA TTT ATC TTC AAA TAA AGA ATA GCG TGT C
P64	Sal I-comK2-F	GCA GCA GTC GAC TTG CAA GAC AAT TCT ACT AAA TAT CTA C
P65	Bgl II-comK2-R	GCA GCA AGA TCT TTA ATT CGA AAT AGT ATT TTC AAT GAA GTG C
P66	Sal I-dprA-F	GCA GCA GTC GAC GTG AAT AGG GAA GAA AAT GT
P67	Bgl II-dprA-R	GCA GCA AGA TCT TTA AAT ATA GTA GTC TTC AAA TAT ATC ATT AGC GTT
P68	Sal I-recA-F	GCA GCA GTC GAC TTG GAT AAC GAT CGT CAA AA
P69	Bgl II-recA-R	GCA GCA AGA TCT CTA TTC TTC GTC AAA TAA TGA CTT TGG T
P70	Sal I-comGA-F	GCA GCA GTC GAC TTG AAG ATT CTA TTT CAA GAA ATA ATT AAT
P71	Bgl II-comGA-R	GCA GCA AGA TCT TTA AAT GTA TTT ATC CAT TGT AGT TTC ACA AAT G
P92	Bgl II-sigH(N)-R	GCA GCA AGA TCT TTA CAA CTC TTT AAT TGC TAA TTT ATT CAA
P93	Sal I-sigH(C)-F	GCA GCA GTC GAC GAA ATA CTT ATA CGG ATG TA
P94	Bgl II-dprA(N)-R	GCA GCA AGA TCT TTA TGT CAT AGG ATT AAA CAT AGA TCC AGG TA
P95	Sal I-dprA(C)-F	GCA GCA GTC GAC GTG CTA TTA AAA GAA ATA TAT CAA
P96	Bgl II-recA(N)-R	GCA GCA AGA TCT TTA CTC TTT AGA AAT ACC TTG TCC ATA
P97	Sal I-recA(C)-F	GCA GCA GTC GAC GCT GAA GTA CAA AGT AAT GG
P98	Sal I-comGA(C)-F	GCA GCA GTC GAC GTT ATC AGA ATT GTA CCT CAA T
P99	Sal I-comEA-F	GCA GCA GTC GAC GTG GTT TTA TTG TAT CAA TTT TTA TTA C
P100	Bgl II-comEA-R	GCA GCA AGA TCT TTA TAT CGT GAA ATA AGA TTT CAG TTT ATC
P101	Sal I-comEA(C)-F	GCA GCA GTC GAC ATG GTA TTA ATT GGT TTT ATA TTC
P102	Sal I-comEB-F	GCA GCA GTC GAC TTG GAA AGA ATC AAA TGG GAA GAA TAT TTT ATG
P103	Bgl II-comEB-R	GCA GCA AGA TCT TTA ACC TTT AGT CAG ATA TTT AGC AAC ATA TTC TG
P104	Sal I-comEC-F	GCA GCA GTC GAC GTG CTC TCT ACA TTT CTT TTC ATT TTA C
P105	Bgl II-comEC-R	GCA GCA AGA TCT CTA TAA ACC ACT TGC ATT TCC ATA AGA G
P106	Bgl II-comEC(N)-R	GCA GCA AGA TCT TTA AAT TGC AGC TAT ATG CGA GCC -3
P107	Sal I-comEC(C)-F	GCA GCA GTC GAC TAT TAC ATA TTT TGG TTA TTG GCT
P108	Sal I-comFA-F	GCA GCA GTC GAC ATG GAT AAT GTA ACA AGA TAT AAA ATA ACA GAG
P109	Bgl II-comFA-R	GCA GCA AGA TCT TTA TTC ATC AAT CCA ACC TCT TTT TAA TGC

P110	Bgl II-comFA(N)-R	GCA GCA AGA TCT TTA AAT CTC TGA CAG AAG TTG TT
P111	Sal I-comFA(C)-F	GCA GCA GTC GAC CAT TTT GTT GTA TGC ACA GTA
P112	Sal I-comFC-F	GCA GCA GTC GAC ATG AAT AAT TGT TTG AGT TGT GGT GC
P113	Bgl II-comFC-R	GCA GCA AGA TCT CTA TCG TGC AAA CGC AAA CAC TT
P114	Sal I-comFC(C)-F	GCA GCA GTC GAC CAC TTA AAT CAA GAT GAA GCG
P115	Sal I-comGC-F	GCA GCA GTC GAC ATG TTT AAA TTT CTT AAG AAA ACT CAA GCG
P116	Bgl II-comGC-R	GCA GCA AGA TCT CTA ATT TGC AAC TGC TTC TCC ATT AC
P117	Sal I-comGB-F	GCA GCA GTC GAC GTG AAA CTA CAA TGG ATA AAT ACA TTT AAA C
P118	Bgl II-comGB-R	GCA GCA AGA TCT TTA TTT TAT ACT TTG CAT CAT TTG AAA CAT TGG
P119	Sal I-comGD-F	GCA GCA GTC GAC ATG GAG AAG CAG TTG CAA ATT AGA AAG
P120	Bgl II-comGD-R	GCA GCA AGA TCT TTA TAG CTT TTC ATA ACG AAT TCT TCC TTT TTC AAT ATG G
P121	Sal I-comGE-F	GCA GCA GTC GAC TTG AAA AAG GAA GAA TTC GTT ATG AAA AGC TAT AAG
P122	Bgl II-comGE-R	GCA GCA AGA TCT TTA ATA CTG TAT ACA TGT CTT TTG ATA AGA AGT GGT
P123	Sal I-comGF-F	GCA GCA GTC GAC ATG ATA TTA AGC AAA GTG ACC AAC
P124	Bgl II-comGF-R	GCA GCA AGA TCT TTA TAC ATA AAT AGT TTT AGT CTG CAG ATT TGT ACC GAC

---



## 1.8. Reference

1. Sany Fds and Ok-kyu Son, *A novel genetic system to detect protein–protein interactions*. Nature, 1989.
2. Mortier-Barrière, I., et al., *A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA*. Cell, 2007. **130**(5): p. 824-836.
3. Annette Fagerlund, P.E.G.a.L.S.H., *Staphylococcus aureus competence genes: mapping of the SigH, ComK1 and ComK2 regulons by transcriptome sequencing*. Molecular Microbiology, 2014. **94**(3): p. 557–579.
4. Kidane, D. and P.L. Graumann, *Intracellular Protein and DNA Dynamics in Competent Bacillus subtilis Cells*. Cell, 2005. **122**(1): p. 73-84.
5. Kramer, N., J. Hahn, and D. Dubnau, *Multiple interactions among the competence proteins of Bacillus subtilis*. Molecular Microbiology, 2007. **65**(2): p. 454-464.
6. Burghard-Schrod, M., S. Altenburger, and P.L. Graumann, *The Bacillus subtilis dCMP deaminase ComEB acts as a dynamic polar localization factor for ComGA within the competence machinery*. Molecular Microbiology, 2020. **113**(5): p. 906-922.
7. De Santis, M., J. Hahn, and D. Dubnau, *ComEB protein is dispensable for the transformation but must be translated for the optimal synthesis of comEC*. Molecular Microbiology, 2021. **116**(1): p. 71-79.
8. Diallo, A., et al., *Bacterial transformation: ComFA is a DNA-dependent ATPase that forms complexes with ComFC and DprA*. Molecular Microbiology, 2017. **105**(5): p. 741-754.
9. Chung, Y. and D. Dubnau, *ComC is required for the processing and translocation of ComGC, a pilin-like competence protein of Bacillus subtilis*. Molecular microbiology, 1995. **15**(3): p. 543-551.
10. Dubrac, S., et al., *A matter of life and death: cell wall homeostasis and the WalkR (YycGF) essential signal transduction pathway*. Molecular microbiology, 2008. **70**(6): p. 1307-1322.
11. Szurmant, H., et al., *YycH and YycI interact to regulate the essential YycFG two-component system in Bacillus subtilis*. Journal of bacteriology, 2007. **189**(8): p. 3280-3289.
12. Briley Jr, K., et al., *Retracted: Maf acts downstream of ComGA to arrest cell division in competent cells of B. subtilis*. Molecular Microbiology, 2011. **81**(1): p. 23-39.
13. Niemann, V., et al., *The NreA protein functions as a nitrate receptor in the staphylococcal nitrate regulation system*. Journal of molecular biology, 2014. **426**(7): p. 1539-1553.
14. Wigley, D.B., *Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB*. Nature Reviews Microbiology, 2013. **11**(1): p. 9-13.
15. Dillingham, M.S. and S.C. Kowalczykowski, *RecBCD enzyme and the repair of double-stranded DNA breaks*. Microbiology and Molecular Biology Reviews, 2008. **72**(4): p. 642-671.
16. Mirouze, N., Bergé, M.A., Soulet, A.L., Mortier-Barrière, I., Quentin, Y., Fichant, G., Granadel, C., Noirot-Gros, M.F., Noirot, P., Polard, P., Martin, B., Claverys, J.P., *Direct involvement of DprA, the transformation-dedicated RecA loader, in the shut-off of pneumococcal competence*. Proc. Natl. Acad. Sci. U.S.A. , 2013. **110**: p. E1035–E1044.
17. Marsin, S., et al., *Genetic dissection of Helicobacter pylori AddAB role in homologous recombination*. FEMS microbiology letters, 2010. **311**: p. 44-50.

18. Kho, K. and T.C. Meredith, *Salt-induced stress stimulates a lipoteichoic acid-specific three-component glycosylation system in Staphylococcus aureus*. Journal of bacteriology, 2018. **200**(12): p. e00017-18.
19. Eugster, M.R., et al., *Bacteriophage predation promotes serovar diversification in Listeria monocytogenes*. Molecular microbiology, 2015. **97**(1): p. 33-46.
20. Mirouze, N., et al., *Antibiotic sensitivity reveals that wall teichoic acids mediate DNA binding during competence in Bacillus subtilis*. Nature communications, 2018. **9**(1): p. 1-11.

## **Chapter 2: Regulation of natural competence-mediated horizontal gene transfer in *S. aureus***

Shi Yuan Feng <sup>¥,1</sup>, Yolande Hauck <sup>¥,1</sup>, Fedy Morgene <sup>1</sup>, Roza Mohammadi <sup>1</sup>, Paul Budisavlevic <sup>1</sup> and Nicolas Mirouze <sup>‡,1</sup>

<sup>¥</sup> Co-first authors

<sup>‡</sup> Corresponding author

<sup>1</sup> Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France





## 2.1. Summary

To perform genetic transformation, one of the three main horizontal gene transfer mechanisms leading to virulence and antibiotic resistance genes acquisition, bacteria need to enter a physiological differentiated state called natural competence. Diverse environmental and cellular cues have been shown to trigger complex regulatory pathways ultimately activating central competence regulators controlling the expression of the genetic transformation genes. Interestingly, new bacteria displaying such aptitude are often discovered, and one of the latest is the human pathogen *Staphylococcus aureus*. However, a clear understanding of the environmental signals, regulatory pathways and central regulators involved in the development of competence in *S. aureus* is not yet available.

Here, using an optimized protocol to naturally induce competence in planktonic cells, leading to improved transformation efficiencies (up to  $1,5 \times 10^{-7}$ ), we showed that the three putative central competence regulators identified in *S. aureus* are all essential for a complete development of competence. Indeed, we found that genes involved in natural transformation can be divided in several classes depending on the regulators controlling their expression. In addition, we showed that oxygen availability is an important signal leading to competence development through the induction of one of the three central regulators. Our results demonstrate the complexity of competence development in *S. aureus*, in comparison to other historical model organisms. We anticipate our results to be a starting point for proving the importance of natural competence and genetic transformation for *S. aureus* genomic plasticity. Furthermore, we also believe that our data will allow us to elucidate the environmental conditions leading to virulence and antibiotic resistance acquisitions *in vivo*, in this important human pathogen.



## 2.2. Introduction

A highly adaptive commensal organism such as *Staphylococcus aureus*, possesses an array of genes that allow the bacterium to survive, grow and infect in a wide variety of ecological niches. This important variety of genes and its close association and evolution with humans have enabled *S. aureus* to become a leading cause of nosocomial and community-acquired infections<sup>1</sup>. *S. aureus* is estimated to colonize up to 30% of the healthy human adult population with approximately 20% being persistently colonized<sup>2</sup>. The anterior nares are generally considered as the native ecological niche of *S. aureus*, though the bacterium can be isolated from other areas of the human body including the skin, the axillae, groin and gastrointestinal tract<sup>3</sup>. While colonization is typically not harmful, *S. aureus* may breach innate host defenses and gain access to deeper tissues, causing a variety of superficial and invasive infections<sup>1</sup>. Furthermore, as *S. aureus* is a facultative anaerobe, the cells have the ability to grow, divide and thwart the host immune system in the presence or absence of oxygen (O<sub>2</sub>). This capacity is particularly important for *S. aureus* as its environment is known to be or become anaerobic during the course of an infection or biofilm formation<sup>4,5</sup>.

In addition to these remarkable adaptive powers, *S. aureus* also became one of the most feared pathogens in the hospital, because of the widespread emergence of antibiotic multi-resistant strains. Horizontal gene transfer (HGT) of antibiotic resistance genes, from other *S. aureus* strains or even from other genera, was though for years to be exclusively mediated through conjugation and transduction. However, few years ago, it was shown that *S. aureus* is capable to become naturally competent for genetic transformation, the third main mechanism leading to HGT in bacteria<sup>6</sup>. These results changed our way to perceive HGT in *S. aureus* and prompted us to characterize these mechanisms and their importance for the evolution of this important human pathogen.

Natural competence is a physiological adaptation that some bacterial species develop in response to various environmental stresses<sup>7</sup>. In response to these stimuli, bacterial cells trigger signal transduction pathways ultimately activating central competence regulators. All these steps are controlled by the so-called early competence genes. Interestingly, central regulators have been identified in several model organisms as transcriptional activators, like ComK in *Bacillus subtilis*<sup>8</sup> or TfoX and QstR in *V. cholerae*<sup>9</sup>, or alternative sigma factors, like ComX in *Streptococcus pneumoniae*<sup>10</sup>. Once activated, central competence regulators initiate the expression of the late competence genes among which are found all the genes essential for genetic transformation. HGT through genetic transformation occurs through three main steps: (i) the binding of exogenous DNA at the competent cell surface<sup>11,12</sup>, (ii) the transport of DNA across the cell wall and the plasmic membrane (ensured by the *comG*, *comF* and *comE* operons and the *comC* gene<sup>13,14</sup>) and (iii) the processing of this



transforming DNA leading to its homologous recombination within the host cell chromosome (allowed by the *dprA*, *recA* and *ssb* genes <sup>15</sup>).

In 2003, *sigH* has been identified in *S. aureus* as an alternative sigma factor, found in the same genetic environment as *sigH* in *B. subtilis*, and capable of inducing the expression of essential genes involved in genetic transformation <sup>16</sup>. Few years later, Morikawa and his colleagues confirmed that genetic competence in *S. aureus* can be naturally induced in different chemically defined media under specific growth conditions <sup>6</sup>.

For example, the authors detected a maximum of 1,6% of competence-inducing cells after 8 hours of aerobic growth in CS2 medium <sup>6</sup>. In addition, the authors proposed two mechanisms leading to SigH activation which in return could control the development of competence. Importantly, they finally showed that *sigH* overexpression leads to *S. aureus* natural transformation, with an efficiency close to  $5 \cdot 10^{-8}$  while in a wild type background the results barely reached the detection limit and were calculated around  $10^{-10}$ .

Interestingly, Morikawa also demonstrated the ability for *S. aureus* to induce natural competence under anaerobic growth conditions in a synthetic medium called GS <sup>6</sup>. It is definitely exciting to consider that *S. aureus* could use natural competence for genetic transformation under different oxygen availability and clearly show the importance to investigate the potential links and regulations at play.

However, even though these cornerstone publications clearly established the importance of SigH as a competence regulator, the natural *sigH* expression was never shown and the direct regulation of the late competence genes by SigH was not verified in all conditions.

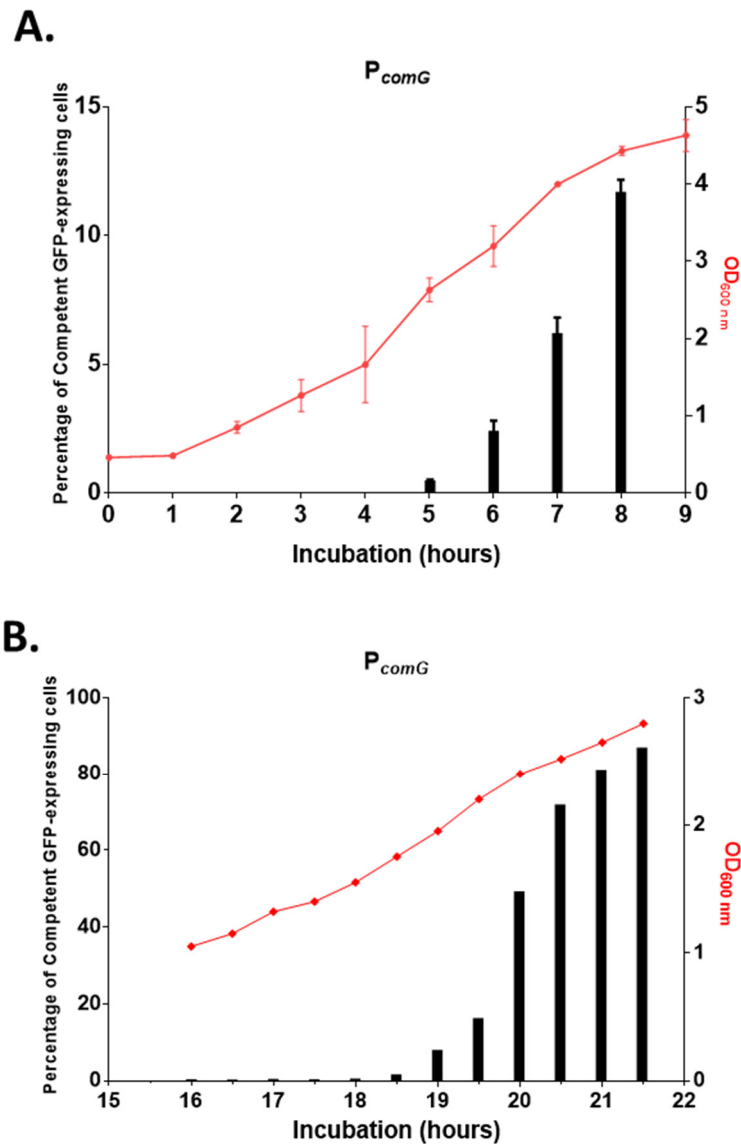
In 2014, two additional putative competence regulators, ComK1 and ComK2, homologous to ComK in *B. subtilis*, have been identified in *S. aureus* <sup>17</sup>. In this study, the authors analyzed by transcriptome-sequencing the impact of *sigH*, *comK1* and *comK2* over-expression. The results showed that *sigH* overexpression leads to the induction of most of the natural transformation genes, with some exceptions. ComK1 can induce the expression of some natural transformation genes, even though less than SigH, and is the only regulator that can control *ssb* expression. No effect has been associated to *comK2* overexpression.

It is important to mention here that this study did not use the genetic background nor the synthetic media published by Morikawa. As a consequence, the authors could not detect any genetic transformation events, even upon *sigH* overexpression, probably because essential genetic transformation genes (*comF* operon and *comC*) were not found induced <sup>17</sup>.



In this study, our main objective was to characterize the development of competence in *S. aureus*, from the environmental stimuli sensed by the cells, the signal transduction pathways transmitting the information to the central competence regulators and their target late competence genes. To reach such goals, we first developed new protocols to optimize the development of competence and genetic transformation in planktonic cultures. Then, using various reporter strains and global transcriptomic analysis, we analyzed the role of the three central regulators during the development of competence in *S. aureus*. We particularly showed that while SigH and ComK1 are both essential for the expression of the genes involved in genetic transformation, the three regulators (SigH, ComK1 and ComK2) are all required for the full development of the competence transcriptional program. Finally, we propose that oxygen availability, sensed by the SrrAB two-component system, activating SigH, partially controls the development of competence in *S. aureus*.





**Figure 1. Protocols to naturally induce competence in CS2.**

Percentage of the cell expressing GFP over growth phases, using the “day” protocol (A\*) or “night” protocol (B). The reporter strain (*P<sub>comG</sub>-gfp*, St29) was grown in CS2 medium. The culture growth (OD<sub>600</sub>) and the percentage GFP-expressing competent cells (determined by flow cytometry) were measured very hour.

\*Measurements were stoppted at 8 hours. I will do the experiment again to provide data all along the growth curve.

## 2.3. Results

### 2.3.1. New protocols to naturally induce competence with high efficiency in CS2

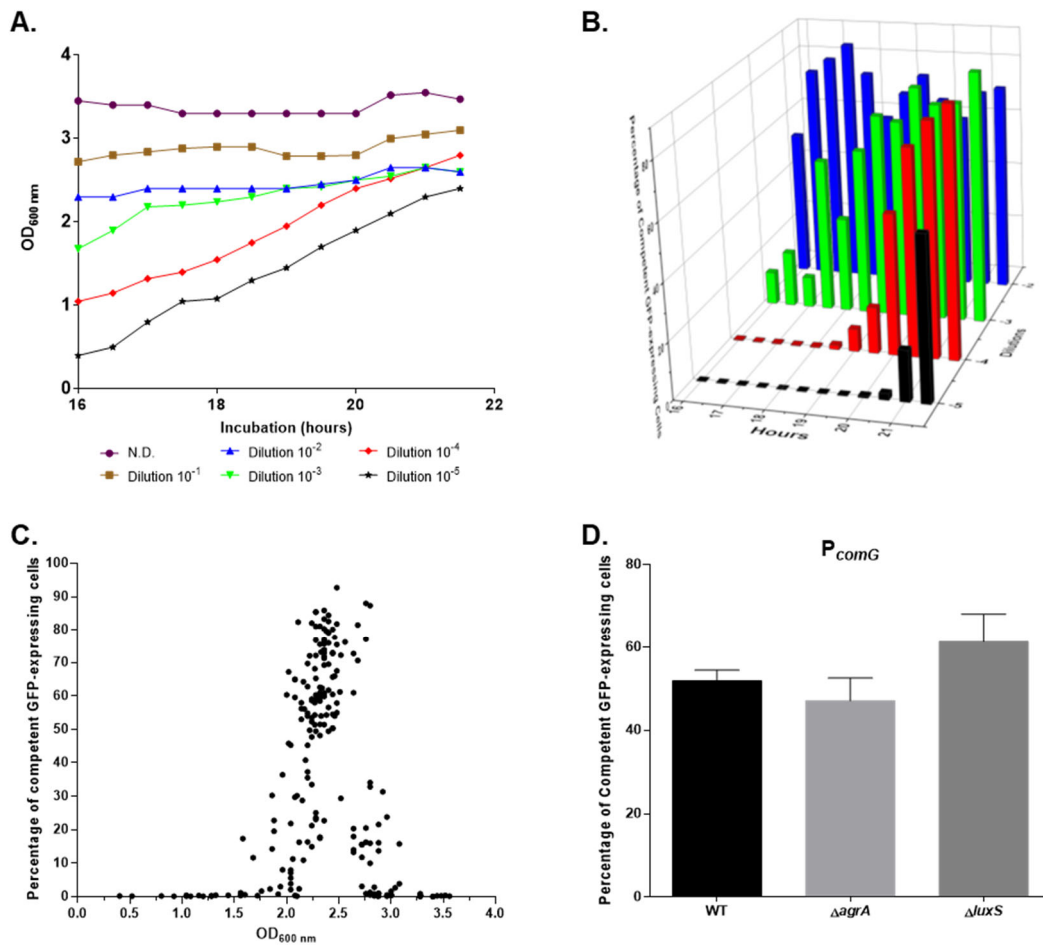
We first decided to optimize the protocol published by Morikawa and his colleagues back in 2012 (see material and methods for details). Using a reporter strain, expressing the *gfp* gene under the control of the *comG* late competence gene' promoter (*PcomG*), we then decided to monitor the development of competence throughout growth in CS2 using our optimized protocol (**Fig. 1A**). During this experiment, we measured every hour, the culture growth (OD<sub>600nm</sub>) and the percentage of competent cells (i.e. the percentage of GFP-expressing cells) by flow cytometry. This percentage was previously calculated by Morikawa through the detection of GFP-expressing cells by microscopy<sup>6</sup>. Therefore, we first verified that measurements by flow cytometry would not introduce any bias and provide similar results than microscopy (**Supplementary Fig. S1**). Interestingly, competence appeared to slowly develop during exponential growth and reach a maximum during the entry in stationary phase (**Fig. 1A**). In our laboratory conditions, a wild-type strain (apart from the *gfp* reporter construct) could naturally induce competence in  $15 \pm 5\%$  of the population. This represents a 10-fold increase in comparison to previous reports (i.e. 1,6%,<sup>6</sup>).

We then wondered if this protocol could be further optimized. This is why we developed a new approach, based on diluted cultures in CS2. Briefly, the strain of interest is first streaked on a BHI agarose plate. Three isolated colonies are then used to inoculate a pre-culture in CS2. This pre-culture is then quickly stopped in exponential growth, centrifuged, washed and used to inoculate a fresh CS2 culture. From this initial culture in fresh CS2, 10-fold serial dilutions were made. **Fig 1B** shows both the growth and the percentage of competent cells in the  $10^{-3}$  dilution culture as an example. Again, competence seems to develop upon entry in stationary phase. Using this new optimized protocol, the percentage of competent cells in the culture reached  $51 \pm 20\%$ .

It is important to mention that GFP is a very stable protein<sup>18</sup>. Therefore, once the maximum percentage of competent cells was reached, this number stayed constant for hours. This feature does not mean that competence stays 'open' for hours but rather that once the maximum is reached, no new competent cells appear. As a consequence, GFP as a reporter protein allows us to describe the induction of competence but not its shutoff.

### 2.3.2. Competence naturally develops at a specific cell density and growth phase in CS2

**Fig. 2A** clearly shows how competence was induced in each diluted cultures using our new optimized protocol (**Fig 1B**). Each additional dilution was characterized by a two-hour delay in growth. It is



**Figure 2. Competence naturally develops at a specific cell density and growth phase in CS2.**

Strain St29 (*P<sub>comG</sub>-gfp*) was grown in CS2 medium using the “night” protocol. A. Cell growth (OD<sub>600nm</sub>) is shown between 18 and 22 hours. Non-diluted (in purple) and 10<sup>-1</sup> (in brown), 10<sup>-2</sup> (in blue), 10<sup>-3</sup> (in green), 10<sup>-4</sup> (in red), 10<sup>-5</sup> (in black) dilutions are represented. A 2 to 3 hours delay in growth can be observed between two consecutive dilutions. B. Percentage of GFP-expressing competent cells (measured by flow cytometry every hour) is shown for the 10<sup>-2</sup> (in blue), 10<sup>-3</sup> (in green), 10<sup>-4</sup> (in red) and 10<sup>-5</sup> (in black) dilutions. No GFP could be detected for the non-diluted and 10<sup>-1</sup> cultures. C. Correlation between competence development (percentage of GFP-expressing cells) and cell density (OD<sub>600nm</sub>). Competence development was maximum at ODs between 2.2 and 2.6 (Data collected from 3 independent experiments, n= 219). D. Comparison of the percentage of cells expressing GFP under control of *comG* promoter in a wild type strain (St29) and in the absence of *agrA* (St107) or *luxS* (St123) (results were collected in the 10<sup>-3</sup> dilution after 22 hours of growth).

important to notice that the first culture (not diluted) and the first diluted culture ( $10^{-1}$ ) reached a higher OD<sub>600nm</sub> value (above 2,5) than the other dilutions. In addition, no competence could be detected in these two first cultures. Based on the 2-hour delay between each consecutive culture and the GFP stability, we concluded that competence was never induced in these two cultures. The four last diluted cultures (i.e.  $10^{-2}$ ,  $10^{-3}$ ,  $10^{-4}$  and  $10^{-5}$ ) reached the same final OD (between 2,3 and 2,4) upon entry in stationary phase. Competence development could be detected in these four cultures. Following the delay in growth, the development of competence was also delayed and reached a maximum once each culture entered stationary phase.

To further verify the existence of a correlation between competence development and cell density, we repeated this experiment three times (**Fig 2B and 2C**). This correlation clearly showed that competence development in *S. aureus* grown in CS2 using our optimized protocol was optimum as the cultures entered stationary phase with an OD<sub>600nm</sub> around 2,4 (in fact between 2,2 and 2,6) (**Fig 2B and 2C**). High cell density, sensed by bacterial cells through quorum sensing (QS), has been proposed and verified in several model organisms as an important stress inducing competence<sup>19–21</sup>. In addition, the two first cultures in each experiment (ND, and  $10^{-1}$ ) reached, and exceeded, this optimal OD value without inducing this physiological adaptation. Therefore, the optimal development of competence could require that the culture reach the right cell density during the right growth phase (in this particular case, the entry in stationary phase).

### 2.3.3. Genetic transformation is improved

We then verified if our optimized protocols, associated to improved development of competence, also led to higher transformation efficiencies. Once maximum competence was reached using our “day” or “dilution” protocols, plasmid DNA was added and incubated for 2h. After being mixed with BHI-agar supplemented with the adequate antibiotic, or not, the culture was finally poured into petri dishes. After one to two days of incubation, the ratio between the number of colonies able to grow in the presence of antibiotic and the number of colonies found on plates in the absence of antibiotic gives access to the strain genetic transformation efficiency.

Interestingly, a wild-type strain reached a transformation efficiency around  $0,95 \cdot 10^{-7} \pm 0,6 \cdot 10^{-8}$  using the “day” protocol (**Table 1**). Such result is 100 times higher than what has been previously published<sup>6</sup> and in some experiments this transformation efficiency even reached  $1,5 \cdot 10^{-7}$ . Importantly, the increase in the percentage of cells inducing competence seemed directly correlated to the increase in transformation efficiency. We also tested the transformation efficiency of a wild type strain grown using the “dilution” protocol (**Work in progress, Table 1**)

Table 1. Optimized natural genetic transformation efficiencies.

Protocols	DNA	N315	$\Delta comGA$	$\Delta sigH$	$\Delta comK1$	$\Delta comK2$
"day"	pCN34 (Kan)	$0,91 \cdot 10^{-7}$ $\pm 6 \cdot 10^{-8}$	$<10^{-10}$	$<10^{-10}$	$<10^{-10}$	$0,93 \cdot 10^{-7}$ $\pm 3 \cdot 10^{-8}$
	pCN38 (Cm)	$0,98 \cdot 10^{-7}$ $\pm 6 \cdot 10^{-8}$	$<10^{-10}$	$<10^{-10}$	$<10^{-10}$	-
"dilution"	pCN34 (Kan)	-	-	-	-	-
	pCN38 (Cm)	-	-	-	-	-

This result was obtained using two different plasmids as exogenous DNA, harboring two different antibiotic resistance genes (i.e. Kan and Cm). To confirm that the colonies able to grow in the presence of antibiotic were real transformants, we also performed this experiment with a strain in which *comGA*, an essential gene for genetic transformation<sup>22</sup>, was deleted. Importantly, in the absence of *comGA*, no colonies could be detected in the presence of antibiotic, confirming that that we really were observing genetic transformation events (Table 1).

#### 2.3.4. Is Quorum Sensing involved in competence development in *S. aureus* ?

Based on the results from Figure 2A, B and C, we first hypothesized that a quorum sensing (QS) system could be involved in the development of competence in *S. aureus*. Indeed, QS is seen as a communication system that enables bacteria to restrict the expression of specific genes to high cells densities thanks to the export of an autoinducer<sup>23</sup>. In addition, QS was often found in the early regulation pathways leading to the development of competence in several model organisms<sup>24–26</sup>. Two QS systems have been identified in *S. aureus* (*Agr* and *Lux*<sup>27</sup>).

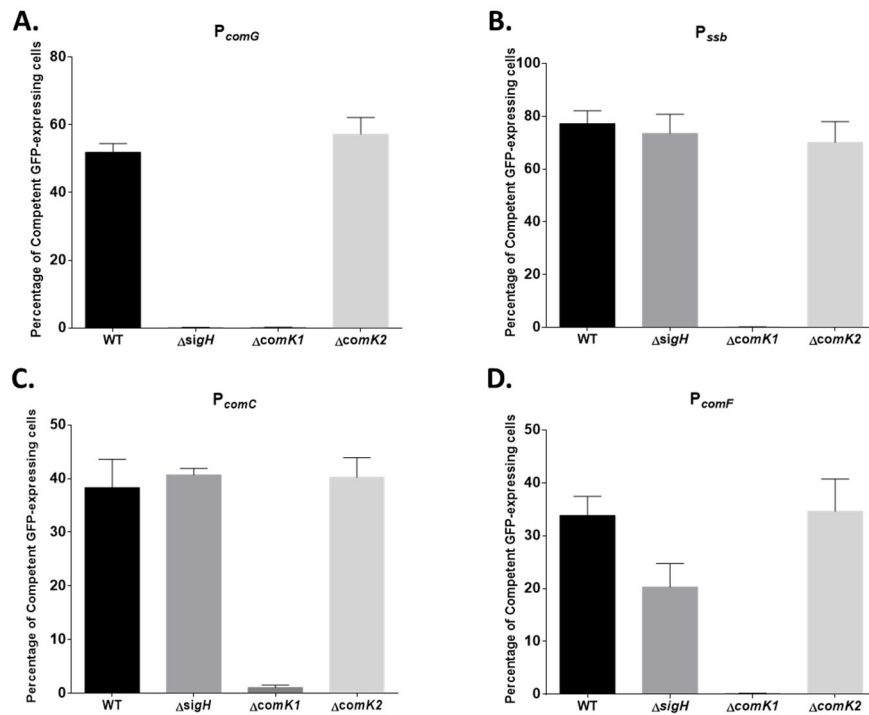
Therefore, we investigated the impact of *agrA* (encoding the *Agr* system transcriptional regulator,<sup>27</sup>) and *luxS* (encoding the *Lux* system regulator,<sup>27</sup>) deletion on the development of competence. Surprisingly, none of these genes were found involved in the induction of the *comG* operon (Fig 2D). This result seems surprising, especially when compared to recent data where the deletion of *agrAC* would decrease competence by a 3-fold factor<sup>28</sup>. Even though the protocols used in this study are different from ours, further investigations will be required to understand why QS does not seem involved in all conditions.

#### 2.3.5. *PcomG*, *Pssb*, *PcomC* and *PcomF* are not controlled by the same regulators

As mentioned in the introduction, three potential competence regulators have been identified in the literature. Even though SigH has been shown to be important to activate the expression from *PcomG*<sup>6</sup>, no role has been clearly assigned to ComK1 and ComK2.

Here, using the dilution protocol (Fig. 1B), we decided to test the effect of *sigH*, *comK1* and *comK2* deletion on the expression from four promoters: *PcomG* (a promoter only activated by SigH over-expression,<sup>6</sup>), *PssB* (a promoter that could only be activated by ComK1 over-expression,<sup>17</sup>), *PcomC* and *PcomF* (two promoters that could not be activated by any regulator,<sup>17</sup>).

As expected, in a wild-type background, the four promoters were found activated with an average percentage of GFP-expressing cells of 51,87 % for *PcomG*, 77,26 % for *PssB*, 38,32 % for *PcomC* and



**Figure 3. SigH and ComK1 regulate the natural genetic transformation genes expression.**

Comparison of the percentage of competent GFP-expressing cells from the *comG* (A), *ssb* (B), *comC* (C) and *comF* (D) promoters in wild type (St29, St50, St48 and St233), *sigH* (St51, St61, St60, St235), *comK1* (St40, St64, St63, St234) and *comK2* (St41, St67, St66, St23) mutant strains, in the  $10^{-3}$  dilution after 22 hours of growth in CS2 medium. All the experiments were repeated at least three times.

Locus tag	Gene name	Gene function	$\Delta sigH$	$\Delta comK1$	$\Delta comK2$
SA1899	<i>ssb</i>	single-strand DNA-binding protein	0.506	0.000	0.494
SA1374	<i>comGA</i>	traffic ATPase	0.059	0.076	0.426
SA1373	<i>comGB</i>	polytopic membrane protein	0.043	0.057	0.435
SA1372	<i>comGC</i>	DNA transport machinery, major pilin	0.051	0.051	0.520
SA1371	<i>comGD</i>	DNA transport machinery, minor pilin	0.031	0.051	0.407
SA1370	<i>comGE</i>	DNA transport machinery, minor pilin	0.057	0.068	0.495
SA1369	<i>comGF</i>	DNA transport machinery, minor pilin	0.030	0.059	0.404
SA0705	<i>comFA</i>	ATP tansloase	0.137	0.056	0.553
SA0706	<i>comFC</i>	amidophosphoribosyltransferases	0.101	0.043	0.536
SA1418	<i>comEA</i>	competence protein, membrane DNA receptor	0.185	0.253	0.424
SA1416	<i>comEC</i>	channel protein for DNA binding and uptake	0.530	0.660	0.795
SA1486	<i>comC</i>	leader peptidase (prepilin peptidase)	0.428	0.305	0.723
SA1128	<i>recA</i>	recombination protein	0.486	0.304	0.764
SA1092	<i>dprA</i>	dprA	0.306	0.426	0.454
SA1485	<i>radC</i>	DNA repair protein RadC	0.212	0.226	0.869
SA0858	<i>coiA</i>	competence protein CoiA	0.176	0.331	0.490

**Table 2. Genetic transformation genes expression is induced during natural competence and controlled by the competence central regulators.**

33,85 % for *PcomF* (Fig 3A-D and Supplementary Fig. S2). When *sigH* was deleted, only the expression from *PcomG* was lost (Fig 3A). This result confirms that SigH does not control the expression of all the genetic transformation genes and that at least one other regulator must be involved. Interestingly, when *comK1* was inactivated, the expression from all the promoters was abolished (Fig 3A-D). Therefore, ComK1 is essential, alongside SigH, for the *comG* operon expression but is required, alone, for *ssb* and *comC* expression. Interestingly, *comF* regulation seems intermediate, as deletion of *comK1* would abolish *comF* expression while the absence of *sigH* would affect it by almost a 2-fold factor (Fig 3D). Finally, in a *comk2* mutant, no effect could be detected.

Importantly, the individual regulatory roles of SigH, ComK1 and ComK2 were confirmed through complementation experiments (work in progress, Supplementary Fig. S3).

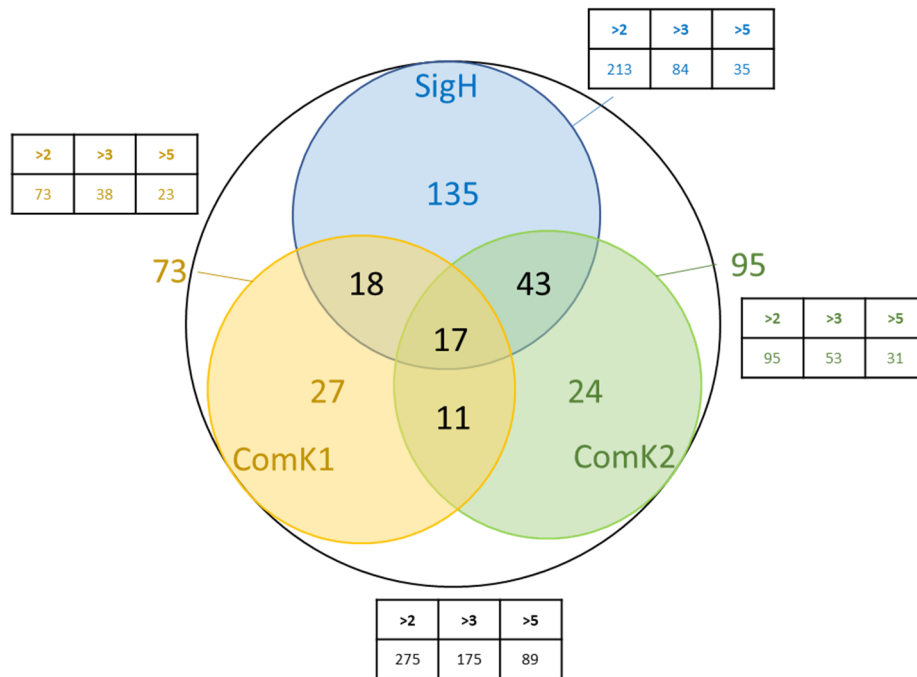
### 2.3.6. SigH and ComK1 control the expression of the natural transformation genes, not ComK2

To complete our results (Fig 3) and exhaustively characterize the SigH, ComK1 and ComK2 regulons, we then performed a global transcriptional analysis by RNA-sequencing. Indeed, we analyzed the impact of the deletion of the genes encoding these individual regulators on the transcription of the late competence genes. First, we focused on the genes involved in natural genetic transformation (Table 2).

Importantly, compared to the GFP reporter strains (Fig 3A-D), the impact of the regulators deletion does not always provide the same result by RNA-sequencing. Indeed, the results are confirmed for the *comG* operon and *ssb* or *comF* expression, controlled either by SigH and ComK1 or only by ComK1. However, if we consider the *comC* expression, ComK1 and SigH seem to have a small impact through RNA-sequencing (Table 2), while only ComK1 appeared essential using GFP-reporter strains (Fig 3C). Importantly, the differences observed between GFP and RNA-sequencing experiments could be explained by the fact that GFP accumulates during time (because of its stability) while mRNA contents evolve rapidly (because of a high turnover<sup>29</sup>). Therefore, samples of GFP-reporter strains could be taken later into stationary phase, while we extracted the mRNA content at a specific time during the transition to stationary phase. Interestingly, these results suggest that all the genetic transformation genes are not expressed exactly at the same time and that specific regulations (involving one or several regulators) are at play.

### 2.3.7. SigH and ComK1 are essential for natural transformation, not ComK2





**Figure 4. Global transcriptomic analysis of competence development in *S. aureus*.**

Each circle represent the genes for which the expression is increased at least two-fold by SigH (in blue), ComK1 (in yellow), by ComK2 (in green) or in total (outside back circle). The number of genes in each category is shown inside the circle or at the intersection between circles, when the same gene is controlled by more than one regulator. The tables outside from the circles show the number of genes which expression is increased by a two-fold, three-fold or five-fold factor by each central regulator and in total.

function	SigH	ComK1	ComK2
Natural transformation	+++	+++	
Stress response	+	+	+++
Toxin / Antitoxin systems	+++	+++	+++
Amino acid and nucleic acids metabolism	+++		+++
Iron transport	+++	+	+

**Table 3. Main function of genes induced during competence and implication of central competence regulators.**

To confirm the results obtained above (Fig 3), we verified the transformation efficiency in strains where the genes encoding the central competence regulators were individually deleted (Table 1). Expectedly, only the absence of *sigH* or *comK1* had an impact on the transformation efficiency. Indeed no transformants could be detected in the absence of *sigH* or *comK1* while a *comK2* mutant displayed a transformation efficiency comparable to that of wild-type strain (Table 1).

### 2.3.8. SigH, ComK1 and ComK2 are all required for a full development of competence

Importantly, the competence transcriptional program always encompasses more genes than just the genes involved in genetic transformation<sup>30-33</sup>. It is clearly also the case in *S. aureus*. Indeed, in comparison to the *sigH*, *comK1* and *comK2* mutant strains, 213/73/95 genes were respectively found overexpressed by at least a two-fold factor in the wild type strain (Fig 4). Furthermore, if we considered higher over-expression cutoffs, the number of genes found overexpressed were still important, with 84/38/53 genes overexpressed by a three-fold factor and even 35/23/31 genes overexpressed by a five-fold factor (Fig 4). The numbers found here, especially for the three-fold expression factor, are quite similar to what has been described in other model organisms, revealing a true core set of competence genes in *S. aureus*.

We then analyzed the functions associated to the genes induced during *S. aureus* competence transcriptional program (Table 3, Supplementary Fig. S4, S5 and S6). In addition to the natural genetic transformation genes, controlled by SigH and ComK1, we found many genes involved in (i) the regulation of the stress response (mainly but not exclusively controlled by ComK2), (ii) genes encoding multiple putative Toxin/Antitoxin systems (controlled by the three regulators), (iii) genes involved in the amino and nucleic acids metabolism (controlled by SigH and ComK2) and (iv) genes involved in iron transport (mainly but not exclusively controlled by SigH).

Interestingly, our results clearly establish that even though only SigH and ComK1 are essential for the expression of the genetic transformation genes, all the central competence regulators (i.e. SigH, ComK1 and ComK2) are absolutely required for the full development of the competence transcriptional program in *S. aureus*.

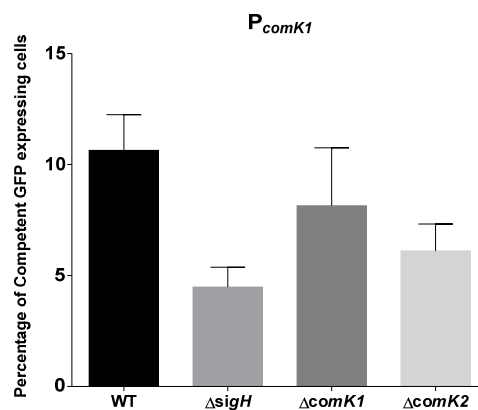
### 2.3.9. Competence development is associated with virulence inhibition

Interestingly, our global transcriptional analysis also revealed that the expression of numerous genes was inhibited during the natural development of competence in *S. aureus*. Indeed, we found that a total of 399 genes were inhibited in the wild type strain compared to the *sigH*, *comK1* or *comK2* indiv

Gene ID	Function	$\Delta sigH$	$\Delta comK1$	$\Delta comK2$
<i>capA</i>	Capsular biosynthesis operon	<b>4,46</b>	2,14	<b>4,11</b>
<i>capB</i>	Capsular biosynthesis operon	<b>4,52</b>	-	2,71
<i>capC</i>	Capsular biosynthesis operon	<b>5,58</b>	-	<b>3,90</b>
<i>capD</i>	Capsular biosynthesis operon	<b>4,22</b>	-	<b>3,26</b>
<i>capE</i>	Capsular biosynthesis operon	<b>3,77</b>	-	<b>3,45</b>
<i>capF</i>	Capsular biosynthesis operon	<b>3,04</b>	-	2,58
<i>capG</i>	Capsular biosynthesis operon	2,84	-	2,52
<i>capH</i>	Capsular biosynthesis operon	2,41	-	-
<i>capI</i>	Capsular biosynthesis operon	2,32	-	2,09
<i>capJ</i>	Capsular biosynthesis operon	2,41	-	2,37
<i>capK</i>	Capsular biosynthesis operon	2,25	-	-
<i>capL</i>	Capsular biosynthesis operon	<b>3,04</b>	-	2,31
<i>capN</i>	Capsular biosynthesis operon	2,34	-	-
<i>capO</i>	Capsular biosynthesis operon	2,24	-	-
<i>clfA</i>	Clumping factor A	<b>6,39</b>	-	<b>3,31</b>
<i>clfB</i>	Clumping factor B	<b>4,21</b>	-	<b>3,46</b>
<i>coa</i>	Staphylocoagulase	-	-	2,49
<i>fnb</i>	Fibronectin-binding protein	2,70	<b>3,85</b>	-
<i>fnbB</i>	Fibronectin-binding protein	<b>5,64</b>	<b>8,02</b>	-
<i>geh</i>	Triacylglycerol lipase	<b>3,53</b>	-	-
<i>hlgB</i>	Toxin (Gamma-hemolysin component B)	-	2,12	-
<i>hlgC</i>	Toxin (Gamma-hemolysin component C)	2,59	2,24	<b>3,01</b>
<i>icaA</i>	Glycosyltransferase	<b>3,23</b>	<b>3,09</b>	-
<i>icaB</i>	Polysaccharide intercellular adhesin deacetylase	2,14	3,15	-
<i>saeR</i>	Transcriptional regulator, involved in the regulation of virulence factors	-	-	2,90
<i>saeS</i>	Histidine kinase, involved in the regulation of virulence factors	2,31	-	2,73
<i>sdrC</i>	Surface serine-aspartate repeat-containing protein C	2,24	-	-
<i>sdrD</i>	Surface serine-aspartate repeat-containing protein D	<b>4,15</b>	-	<b>3,58</b>
<i>spa</i>	Staphylococcal protein A (virulence factor)	<b>10,28</b>	2,11	<b>5,38</b>
<i>splA</i>	Serine protease	-	-	2,52
<i>splB</i>	Serine protease	-	-	2,48
<i>splC</i>	Serine protease	-	-	2,27
<i>splD</i>	Serine protease	-	-	2,39
<i>splF</i>	Serine protease	-	-	2,36
<i>sspA</i>	Serine protease	2,39	-	-
<i>sspB</i>	Serine protease	2,03	-	-
<i>sspC</i>	Serine protease	2,32	-	-

**Table 4. Expression of virulence-related genes is inhibited during competence.**

For each gene, the ratio between the expression in the wild type and mutant strains is presented in the last three columns. Ratios higher than 3 are in bold.



**Figure 5. *comK1* expression is partially controlled by SigH.**

Comparison of the percentage of competent GFP-expressing cells from the *comK1* promoter in wild type (St94), *sigH* (St110), *comK1* (St97) and *comK2* (St100) mutant strains. Samples were taken in the  $10^{-3}$  dilution after 22 hours of growth in CS2 medium. All the experiments were repeated at least three times.

-idual mutant strains (**Supplementary Fig. S7**). This represents more than 15% of all the genes present on the N315 *S. aureus*' chromosome.

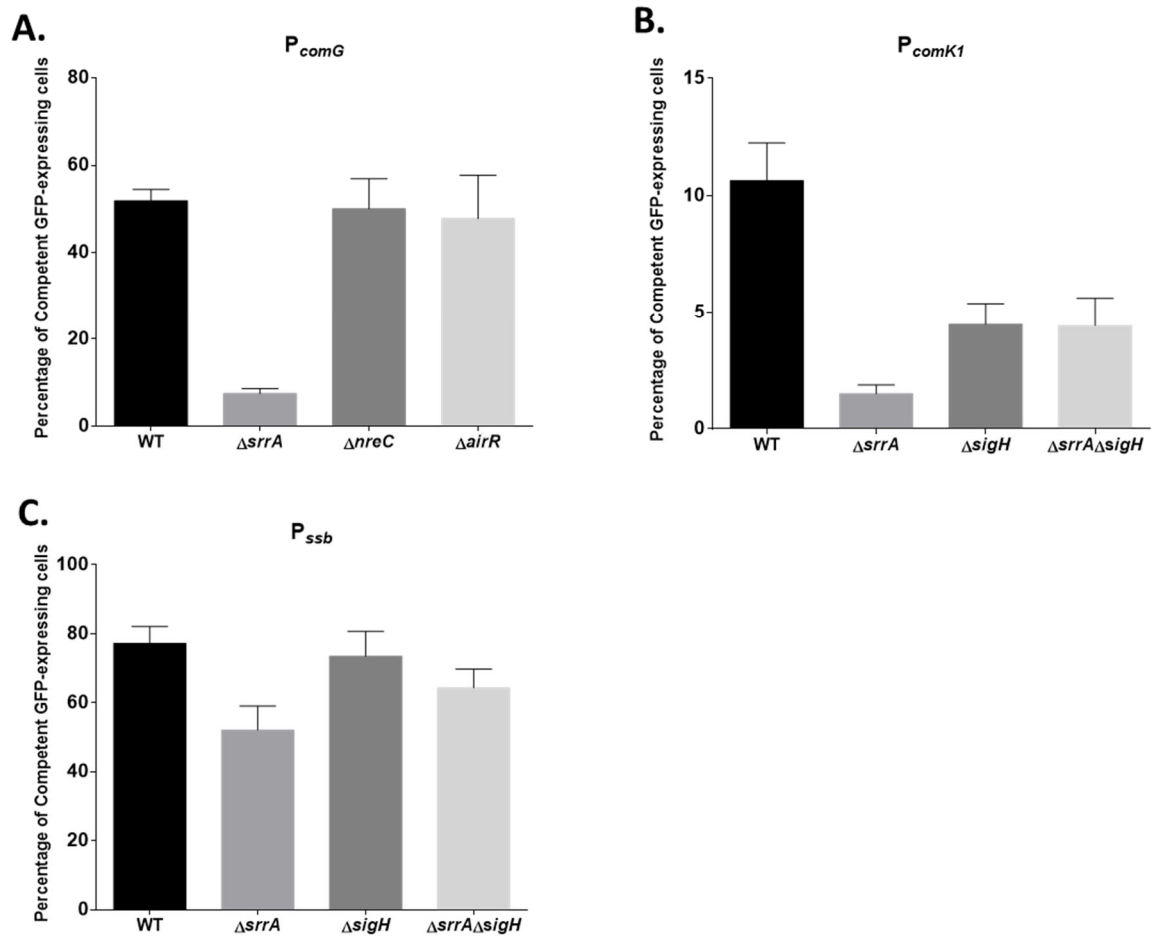
Even more interestingly, more than 10% of the repressed genes (37 to be exact), are directly involved in virulence or virulence regulation (**Table 4**). Among these genes were found many well characterized virulence factors, including genes encoding for the capsule (CapA-O), serine proteases (SplA-F, SspA-C), intercellular adhesins (IcaAB), exoproteins (Hlg, Coa), surface proteins (ClfAB, fnb, FnbB, geh, sdrCD, Spa) and a lipase (Geh). In addition, the expression of an important two-component system, *saeRS*, was found repressed during competence. SaeRS is known to drive the expression of over 20 virulence factors, *in vivo*, during infection<sup>34</sup>. Even though all the central competence regulators were found involved in virulence repression, SigH and ComK2 played a major role with an important overlap between the genes they each repress (**Table 4**).

### 2.3.10. In CS2, PcomK1 is regulated by SigH

While the two historical model organisms, *B. subtilis* and *S. pneumoniae* only use a single central competence regulator, *V. cholerae* controls the development of competence using two, namely TfoX and QstR. Importantly, it was shown that TfoX and QstR each control a different set of late competence genes<sup>35</sup>, a specificity reminiscent to what we just described in *S. aureus*. Furthermore, it appeared that TfoX could control the expression of QstR and complexify even more the regulation of competence in *V. cholerae*<sup>26</sup>.

Because of the multiple similarities that exist between the regulation of competence in *S. aureus* and *V. cholerae*, we decided to verify if one, or several, competence regulator(s) could control the expression of another regulator. Therefore, we constructed transcriptional fusions between the promoter of *sigH*, *comK1* and *comK2* and the gene encoding the *gfp*. In a wild type strain grown in CS2 using the optimized protocol, we could only detect the *comK1* expression (**Fig. 5**). In the case of *sigH*, we know that in addition of its transcriptional regulation, the SigH production is also regulated at the translational level. Therefore, we don't really know the amount of protein produced and based on our result, we expect it to be low. In the case of *comK2*, we know from our global transcriptomic study that the transcription of this gene is quite low which could explain why we could not detect it.

We then tested if the deletion of the genes encoding the three competence regulators could affect the expression of *comK1*. First, we can notice that the absence of *comK2* is characterized by a small, but reproducible inhibitory effect on *comK1* expression (**Fig. 5**). Interestingly, when *sigH* was absent, the percentage of cells expressing *comK1* decreased by more than a 3-fold factor (**Fig. 5**). As this percentage did not go to zero, we can conclude that *comK1* expression is partially controlled by SigH and by at least another pathway that still needs to be discovered. Indeed, if SigH was essential for



**Figure 6. The SrrAB TCS controls the development of competence in *S. aureus*.**

Percentage of the population expressing GFP under control of the promoter of *comG* (A), *comK1* (B) and *ssb* (C). The expression of *comG* is compared in wild type (St29) and *srrA* (St145), *nreC* (St158) and *airR* (St177) mutant strains (A). The expression of *comK1* is compared in wild type (St94), *srrA* (St247), *sigH* (St110), and *srrA/sigH* (St253) mutant strains (B). The expression of *ssb* is compared in wild type (St50), *srrA* (St147), *sigH* (St61), and *srrA/sigH* (St252) mutant strains (C). All the strains were grown in CS2 for 22 hours and the result from the  $10^{-3}$  dilution is shown. All experiments were repeated at least three times.

*comK1* expression, the two regulators would control the expression of the same set of genes, and we know that this is not the case, at least for *ssb* (Fig 3B).

### 2.3.11. Oxygen concentration is an important factor to induce competence development

Finally, competence development in *S. aureus* has been demonstrated in different conditions, associated to different oxygen availability<sup>6</sup>. Thus, we wondered if oxygen-sensing two-component systems (TCS) present in *S. aureus* (i.e. *SrrAB*<sup>36</sup>, *NreBC*<sup>37</sup> and *AirRS*<sup>38</sup>) were involved during the early competence regulation steps in the course of planktonic growth in CS2 medium. To test such hypothesis, we compared the expression of the *comG*, *ssb*, *comC* and *comK1* genes in wild type and *srrA*, *nreC* and *airR* mutant strains. In this experiment, deletion of *nreC* and *airR* did not affect the expression from the *comG* promoter (Fig 6A). However, when *srrA* was absent, *comG* expression was decreased by a 7-fold factor (Fig 6A).

Since both SigH and ComK1 are involved in the *comG* operon expression (Fig 3A), we then decided to determine which central competence regulator was under the control of the oxygen availability, through *SrrAB*. Interestingly, many competence genes have a specific regulation profile: *comK1* expression is only controlled by SigH while *ssb* expression is only controlled by ComK1. Therefore, we then tested the effect of *srrA* deletion on the expression of *comK1* and *ssb*. Indeed, in the absence of *srrA*, the expression of *comK1* was affected but not that of *ssb* (Fig 6B and 6C), a regulatory profile resembling that of SigH.

Importantly, deletion of both *sigH* and *srrA* had the same impact on the *comK1* and *ssb* expression as the individual *sigH* mutant strain, even though these expressions seemed lower in the single *srrA* mutant (Fig 6B and 6C). This reproducible result confirms that *SrrAB* and SigH might work in the same pathway but also indicate the possible existence of additional regulatory pathway(s) controlling the expression of *comK1*, in response to low oxygen concentrations, alongside SigH.

We can therefore propose that the SigH central competence regulator is, at least partially, controlled by the oxygen-sensing *SrrAB* TCS. Since we cannot yet monitor *sigH* expression, it will be important to test this potential regulation pathway in the future. As explained above, oxygen-sensing TCS allow *S. aureus* to survey and respond to hypoxic conditions. In particular, *SrrAB* have been shown to be a global regulator of virulence factors under low-oxygen conditions<sup>36</sup>. It is therefore tempting to speculate that while *S. aureus* reached stationary phase in CS2 medium, the oxygen concentration dropped in the culture. This hypothesis would be different than what has been previously published where the authors described the culture in CS2 medium as fully anaerobic<sup>6</sup> but consistent with the fact that they performed their cultures in closed falcon tubes<sup>6</sup>.



## 2.4. Discussion

### **Optimized protocols to naturally induce competence and improve genetic transformation in *S. aureus*.**

In this study, we presented new protocols allowing the optimized induction of competence in *S. aureus*. Such protocols were essential to lead the genetic study proposed here. Previously, the work of Fagerlund and her colleagues<sup>17</sup> clearly established that central competence regulators individual or combined over-expression was not enough to induce the full development of competence in *S. aureus*. Indeed, under such conditions, they could not see the induction of essential genes for genetic transformation, such as *comC* or *comF*.

In addition, the environmental conditions described here not only allowed an optimized expression of competence but also improved genetic transformation. Importantly, the transformation efficiencies detected with a wild type strain, naturally inducing competence, clearly demonstrate the potential of this horizontal gene transfer mechanism in *S. aureus*.

### **Environmental conditions and Competence development**

Quorum sensing. We showed here that competence develops during the transition into stationary phase in *S. aureus*. This specific growth stage is also used by *B. subtilis* to induce competence<sup>39</sup> by sensing high cell density through quorum sensing (QS)<sup>40,41</sup>. In fact, this bacterial communication mechanism is highly present in the early pathways of competence development in most model organisms. This is why we hypothesized that the Agr and/or Lux systems could be involved in *S. aureus*. However, no effect associated to the absence of these QS systems could be detected.

Interestingly, the Agr global regulator of the staphylococcal virulon has been shown to partially control the expression of the *comG* operon upon competence induction in *S. aureus*<sup>28</sup>. However, the 3-fold decrease in the percentage of *comGA*-expressing cells associated to the absence of Agr does not fit with the essentiality expected with such important systems. In addition, the Agr mutant strain used in this study was associated to a minor growth defect and we know that the optical density reached in stationary phase is an important parameter to ensure full competence development. However, the idea that QS systems (known or to be discovered) could be involved in the development of competence in *S. aureus* should be pursued and probably tested in other environmental conditions.

Virulence. Interestingly, many virulence genes were found repressed during the development of competence. Virulence regulatory pathways and virulence factors are known as accessory genes, since





they are not essential for growth. All these factors are essential for pathogenicity and contribute to adherence to host cells, evasion of host defenses, nutrient degradation and acquisition. Competence is often considered as a dormancy-like state where the competent cells stop growing, dividing and replicating<sup>42</sup>. Therefore, it is not surprising to find that competent cells would inhibit a process as energy consuming as virulence.

This result does not imply that competence and virulence are antinomic. We could imagine that, similarly to *B. subtilis* which can “choose” between several environmental adaptations, different *S. aureus* cells could enter different pathways to promote either virulence or genetic transformation to adapt to unfavorable conditions.

Toxin/Antitoxin systems. Importantly, our global transcriptomic study revealed the induction of many putative Toxin/Antitoxin (TA) systems upon competence induction. In the light of such result, it would be tempting to propose that *S. aureus* could use such systems to target neighboring bacteria, induce their lysis and recover their DNA through genetic transformation. Comparable mechanisms have already been described, notably in *S. pneumoniae*, and called fratricide<sup>43</sup>. During competence in *S. pneumoniae*, killing/lytic factors (bacteriocins and bacteriolysins) are not involved in interspecies competition, as usually, but rather as an alternative mechanism to ensure the presence of DNA in the environment while the cells are competent<sup>44</sup>. It will be important to further investigate these TA systems and attempt to prove that such genes should not be considered as dispensable during competence, but rather as essential mechanisms to favor genetic plasticity through genetic transformation.

Oxygen availability. Finally, we demonstrated how oxygen availability, sensed by the SrrAB two-component system, controls the development of competence in *S. aureus*. Results from the literature already indicated that *S. aureus* has the ability to induce competence under different oxygen concentrations<sup>6,28</sup>. Here we hypothesized that rather than being a fully aerobic condition, growth in CS2 leads to a limitation in oxygen that is essential for the development of competence. Induction of competence under anaerobic conditions in GS medium<sup>6</sup> or during biofilm formation<sup>28</sup> further reinforce this idea. We also propose that SrrAB would activate SigH (through transcription, traduction or stability) in response to variations in the environmental oxygen concentration. Additional work will be required to fully understand how a potential drop in the oxygen concentration regulates SigH through SrrAB and at which exact oxygen concentration competence is truly induced in the different conditions described above.

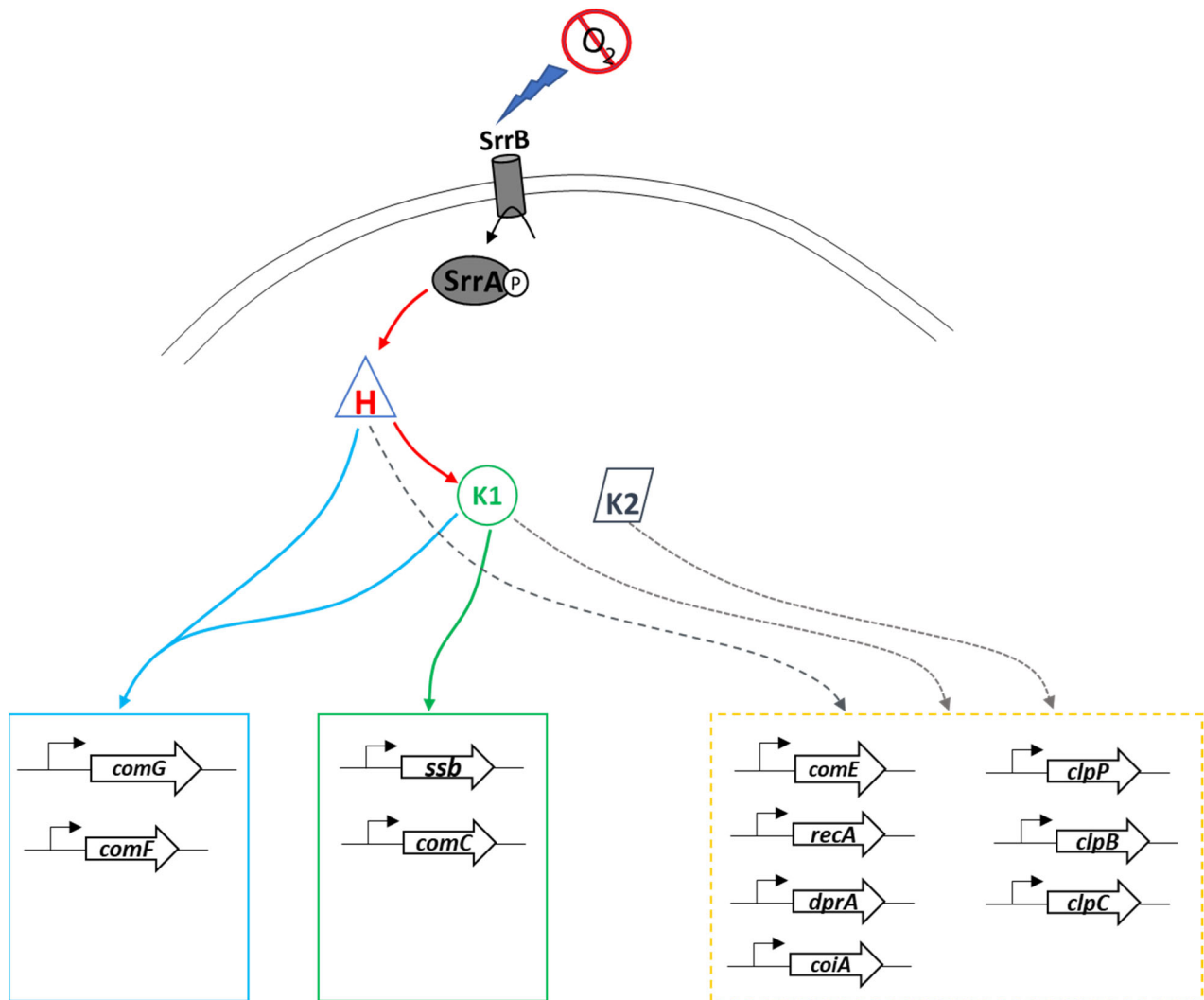


Figure 7. Suggested model of competence development regulation in *S. aureus*.

### **Different stimuli, controlling multiple regulators, lead to competence development in *S. aureus***

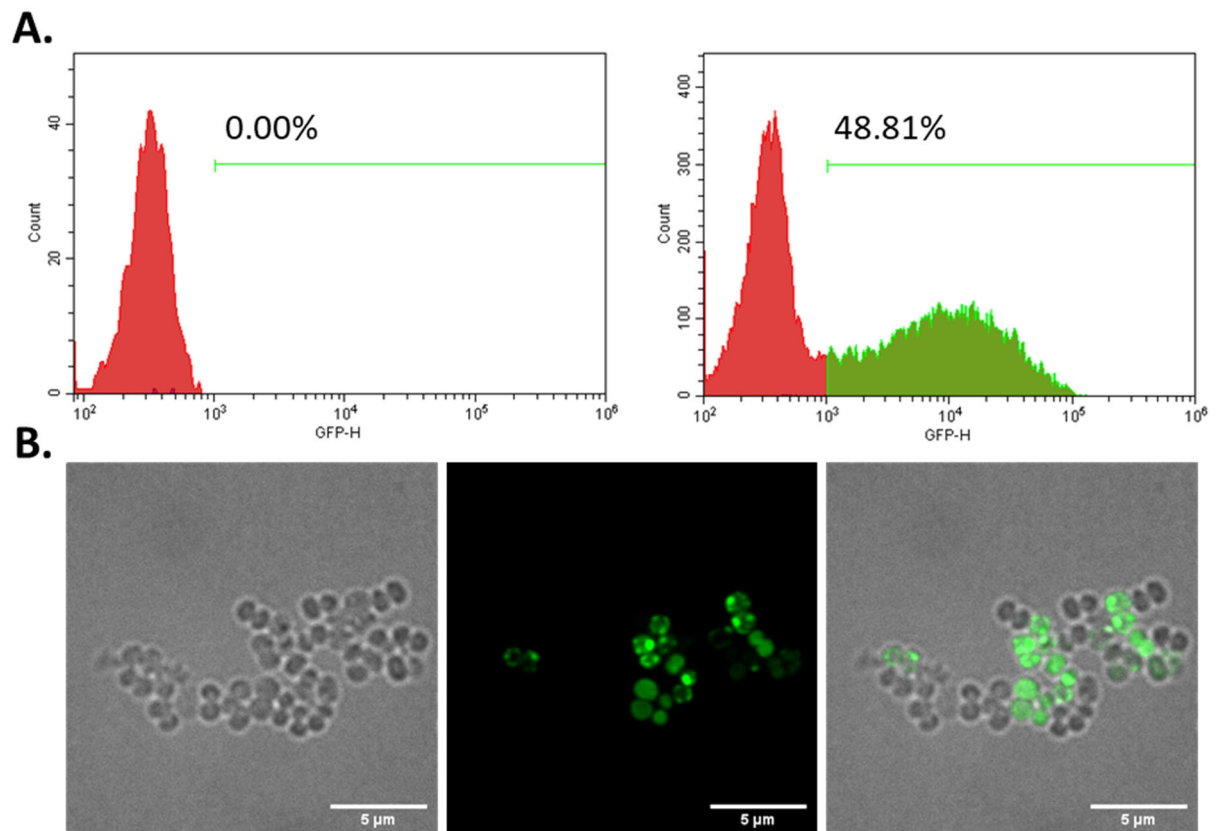
Here, we have established that three central competence regulators are essential for a complete development of the competence transcriptional program in *S. aureus*. In addition, several results from the literature or obtained in this study clearly indicate that each regulator might be controlled by different stimuli and signal transduction pathways (**Fig 7**): (i) SigH production is controlled by a chromosomal gene duplication rearrangement and post-transcriptional regulation through upstream repeated and inverted sequences <sup>6</sup>, (ii) *sigH* expression is probably controlled by oxygen availability through the SrrAB TCS, (iii) *comK1* expression is partially controlled by SigH but also by an unknown second pathway and (iv) the potential involvement of QS <sup>28</sup>. Finally, even though we could not detect and study *comK2* expression, we can propose that since it does not control the expression of the genes involved in genetic transformation, *comK2* could be controlled by a third pathway.

Therefore, it is legitimate to speculate that *S. aureus* only develops natural competence for genetic transformation under specific and stressful environmental conditions, that need to be simultaneously present. The multiplication of signaling pathways, already described in *B. subtilis* (high cell density, nutrient exhaustion), consequently limit the development of competence. The choice to induce competence is often associated with important “drawbacks”. Indeed, cell division, elongation and replication are usually stopped during competence <sup>42</sup>, to ensure the success of exogenous DNA import and recombination during transformation. The presence of multiple signaling pathways and central competence regulators could therefore guarantee that acquisition of new DNA sequences is restricted to the adaptation to stressful environmental conditions threatening *S. aureus* survival.

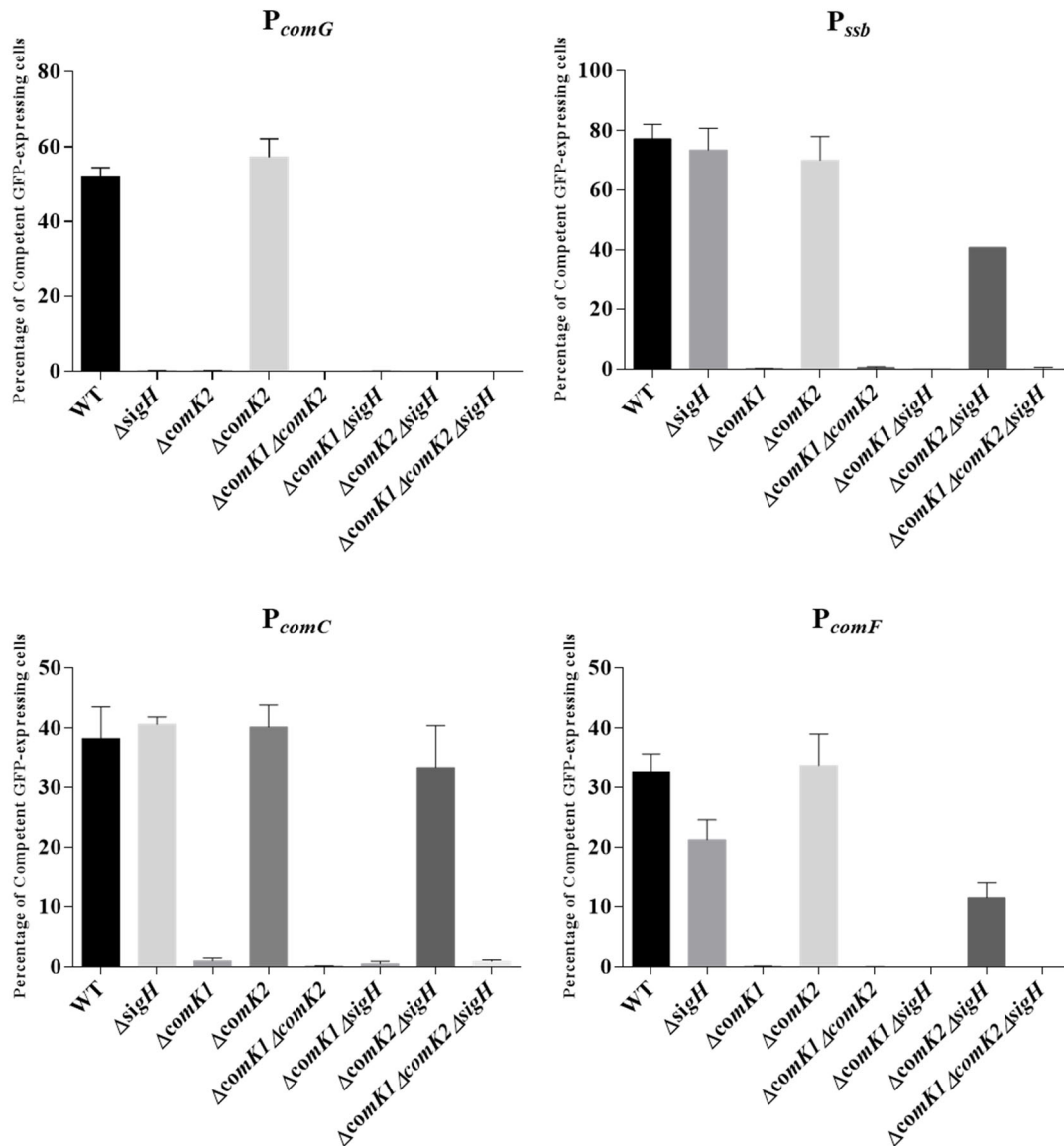
Finally, such organization of competence regulation resemble what happens in *V. cholerae* in multiple ways: (i) several stimuli are integrated to perform genetic transformation, (ii) several regulators (TfoX and QstR in *V. cholerae*) are present and each control a different set of late competence genes and (iii) one regulator (TfoX) partially controls the expression of the other (QstR) <sup>21</sup>. *S. aureus* would therefore share traits with three important model organisms: *V. cholerae* through the characteristics described above, *S. pneumoniae* and its alternative sigma factor (ComX), phylogenetically close from SigH <sup>16</sup> and *B. subtilis* and its central competence regulator, ComK, homologous to ComK1 and ComK2 <sup>17</sup>. Thus, it would be interesting and probably important to test in the future if other regulation features, present in these three model organisms, could be also used by *S. aureus* to control the development of competence.



## Supplementary Figures



Supplementary Figure. 1. Flow cytometry (A) and microscopy (B) provide identical percentage of GFP-expressing competent cells within the same culture with respectively 48.81% and 44.7%.



**Supplementary Figure. 2. P<sub>comG</sub>, P<sub>ssb</sub>, P<sub>comC</sub> and P<sub>comF</sub> expression in all mutant strains**  
 Comparison of the percentage of competent GFP-expressing cells from the *comG* (A), *ssb* (B), *comC* (C) and *comF* (D) promoters in wild type (St29, St50, St48 and St233), *sigH* (St51, St61, St60, St235), *comK1* (St40, St64, St63, St234), *comK2* (St41, St67, St66, St23) *comK1/comK2* (St47, St74, St75, St268), *comK1/sigH* (St81, St82, St83, St269), *comK2/sigH* (St77, St78, St79, St270) and *comK1/comK2/sigH* (St69, St70, St71, St271) mutant strains, in the 10<sup>-3</sup> dilution after 22 hours of growth in CS2 medium. All the experiments were repeated at least three times.

**Supplementary Figure. 3. PcomG expression in *comK1*, *sigH* and *comK2* complementation strains**

**(work in progress)**



## Chapter 2

Gene ID	Function	Ratio	Gene ID	Function	Ratio
SA2091	SarY, transcriptional activator	3,01	SA1152	hypothetical protein, hypothetical endotoxin? (see SA1151, 1153)	4,50
purD	Purine metabolism	3,01	truncated-SA	Lytic regulatory protein truncated with Tn554	4,52
SA0415	conserved hypothetical protein (yjbB)	3,02	SA1418	comEA, DNA binding and uptake protein	4,53
feoB	major iron permease	3,02	SA1249	hypothetical protein, resistance to vancomycin?	4,87
SA1824	arsC Arsenate reductase (resistance to arsenic)	3,06	purL	Purine metabolism	4,83
lytN	cell wall hydrolase	3,07	SA0858	Empbp, secretory extracellular matrix and plasma binding protein, fibrinogen-binding protein (see fnbB)	4,88
SA0632	Putative lipoprotein	3,12	SA2338	ferrous iron transport protein B homolog (see SA1978)	5,51
SA2150	Putative hemin transport system permease protein HrtB	3,14		Toxin/antitoxin system? (see Mat&Met of <a href="https://europepmc.org/article/pmc/pmc6805077#free-full-text">https://europepmc.org/article/pmc/pmc6805077#free-full-text</a> )	5,53
SA0136	similar to phosphonates transport permease	3,14	SA1821	hypothetical protein	5,74
SA0408	hypothetical protein	3,17	SA1151	hypothetical protein	5,74
clpB	Proteolysis in bacteria, stress	3,20	lpl4	lipoprotein	5,74
SA0704	hypothetical protein, DegV domain containing prot (operon with comF)	3,22	SA0705	comFA; competence protein ComFA	5,77
SA0530	alkaline phosphatase, aa (histidine) metabolism	3,24	uhpT	Sugar phosphate antiporter	5,95
SA2090	fmbB Fibronectin-binding protein, expression affected by ClpP deletion	3,24	purQ	Purine metabolism	6,04
SA1878	epiA, lantibiotic epidermin precursor	3,24	SA2012	hypothetical protein	6,38
purK	purine biosynthesis pathway	3,26	SA1153	DNA-dependent DNA polymerase family X	6,62
SA0333	Conserved hypothetical protein	3,30	valS	valyl-tRNA synthetase, Protein Metabolism biosynthesis	6,65
SA1378	Conserved hypothetical protein	3,34	SA0919	purS, phosphoribosylformylglycinamide synthase, PurS protein	6,65
purF	purine biosynthesis pathway	3,37	SAS040	Putative uncharacterized protein	6,96
SA1822	Toxin/antitoxin system?	3,37	folC	putative folylpolyglutamate synthase/dihydrofolate synthase	6,92
SA0838	glyceraldehyde-3-phosphate dehydrogenase	3,41	purC	Purine metabolism	7,08
SA0379	probable transposase	3,38	SA0282	conserved hypothetical protein, Protein secretion system?	7,36
modA	Molybdate ABC transporter substrate-binding protein	3,46	SA0706	comFC; competence protein ComFC	7,38
purM	Purine metabolism	3,47	SA2197	putative DsbA homologue, protein disulfide bond formation	7,60
SA1978	similar to ferrichrome ABC transporter (permease)	3,49	lpl7	lipoprotein	8,24
SAS028	sRNA (Bouloc)	3,51	SA0576	toxin/antitoxin system?	11,69
SA1435	similar to acetyl-CoA carboxylase accB, fatty acid (lipid) metabolism	3,55	SA2198	putative DsbA homologue, protein disulfide bond formation	11,80
SA0283	serine protease?	3,59	trpB	tryptophan synthase subunit beta	12,44
SA0407	CHP (dipeptidyl aminopeptidases/acylaminoacyl-peptidases)	3,70	SA1199	hypothetical protein, similar to anthranilate synthase component I	12,90
fmhC	secreted bacteriocin	3,72	SA1370	ComGE, Late competence protein in ComGE	13,49
SA1636	pathogenicity island saPI3	3,71	SA1374	ComGA; competence protein ComGA	13,53
butA	Acetoin Reductase	3,77	SA1372	comGC; competence protein ComGC	14,69
radC	DNA repair protein	3,83	SA1373	comGB; competence protein ComGB	17,92
pyrP	uracil permease	3,90	SA0575	toxin/antitoxin system? (see SA0576)	20,20
SA0331	Efem/EfeO family lipoprotein	3,96	trpD	Anthranilate phosphoribosyltransferase	21,41
SA0837	Amino acid metabolism/transport	3,99	trpG	anthranilate synthase component II	22,18
ndhF	NADH dehydrogenase subunit 5	4,14	trpC	Indole-3-glycerol phosphate synthase	22,52
purN	Purine metabolism	4,22	SA1371	comGD; competence protein ComGD	23,53
SA0412	hypothetical protein	4,28	trpF	N-(5'-phosphoribosyl)anthranilate isomerase	23,80
SA0138	similar to alkylphosphonate ABC transporter	4,30	SA1369	ComGF, Late competence protein in ComGF	24,56
SAS056	sRNA	4,34		conserved hypothetical protein, contain Thioredox_DsbH and acetyl-transferase domains	25,23
SA0191	HlyD family secretion protein	4,42	SA2355		25,23
SA0413	hypothetical protein (see SA0412, 0408, 0415)	4,48	SA1979	similar to ferrichrome ABC transporter	34,69

### Supplementary Figure. 4. SigH regulon

Are listed all the genes for which the expression is activated at least 3 times by SigH. The real factor is mentioned in the last column.

Gene ID	Function	Ratio
SA1833	annotated as a transcriptional activator (Str), Toxin/antitoxin system?	3,24
SA1486	ComC, Late competence protein ComC	3,24
groES	heat shock protein	3,27
grpE	response to hyperosmotic and heat shock by preventing the aggregation of stress-denatured proteins	3,28
SA1418	comEA, DNA binding and uptake protein	3,30
recA	Recombinase, catalyze the hydrolysis of ATP in the presence of single-stranded DNA	3,35
xprT	Xanthine phosphoribosyltransferase	3,48
truncated( radC)-2	DNA repair protein	3,59
SA1822	Toxin/antitoxin system?	3,78
SA1831	Toxin/antitoxin system?	3,82
SA2353	autolysins	4,16
SA1978	similar to ferrichrome ABC transporter	4,20
lacR	Repressor of the lactose catabolism operon	4,42
SA1709	ftnA, Ferritin-like protein	4,61
SA1820	operon with SA1821-1822, toxin/antitoxin?	4,80
hutH	Histidine ammonia-lyase	5,26
SA1821	Toxin/antitoxin system?	5,29
clpB	Proteolysis in bacteria, stress	6,64
folC	putative folypolyglutamate synthase/dihydrofolate synthase	7,23
valS	valyl-tRNA synthetase	7,53
SA0704	conserved hypothetical protein, DegV domain containing prot	8,86
SA2198	Toxin/antitoxin system?	9,65
SA1635	hypothetical protein, Pathogenicity island SaPln3	10,28
SA2197	Toxin/antitoxin system?	10,16
SA1374	ComGA; competence protein ComGA	10,55
SA1370	ComGE, Late competence protein ComGE	11,22
SA1369	ComGF, Late competence protein ComGF	12,51
SA1373	comGB; competence protein ComGB	13,60
SA0705	comFA; competence protein ComFA	14,10
SA1372	comGC; competence protein ComGC	14,60
SA1371	comGD; competence protein ComGD	14,57
SA2355	conserved hypothetical protein, contain Thioredox_DsbH and acetyl-transferase domains	15,32
SA0706	comFC; competence protein ComFC	17,47
SA1979	similar to ferrichrome ABC transporter	20,20
SA0575	Toxin/antitoxin system?	31,02
SA0576	Toxin/antitoxin system?	36,03
SA1898	hypothetical protein, similar to SceD precursor, able to cleave peptidoglycan and affects clumping and separation of bacterial cells	721,03
SA1899	ssb, Single-strand DNA-binding protein	2533,34

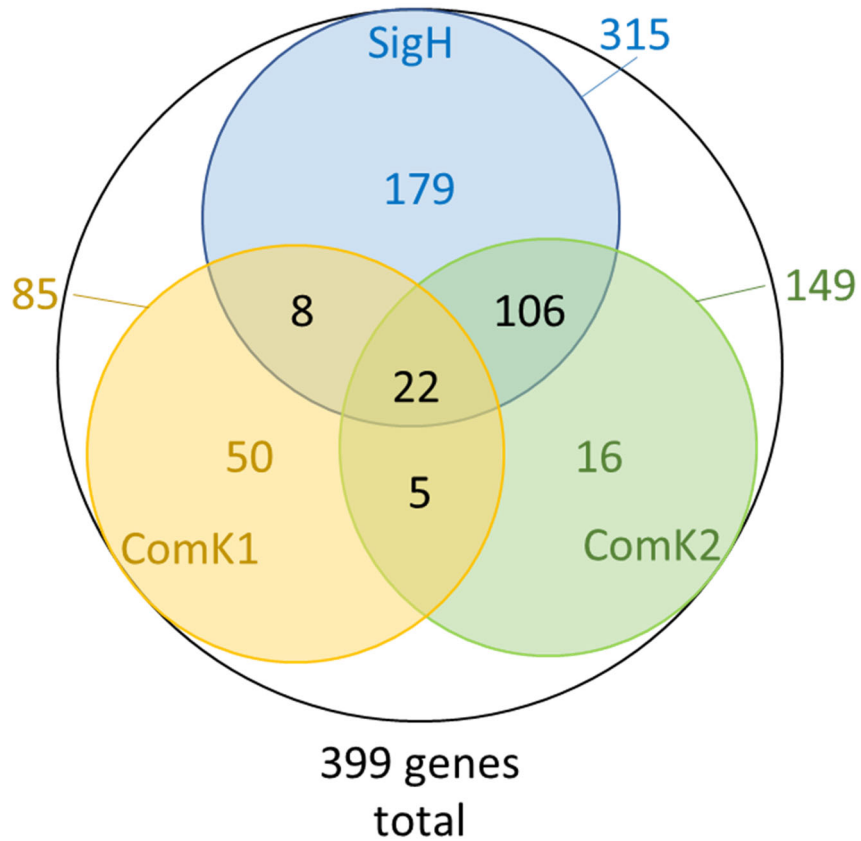
### Supplementary Figure. 5. ComK1 regulon

Are listed all the genes for which the expression is activated at least 3 times by ComK1. The real factor is mentioned in the last column.

Gene ID	Function	Ratio
SA1825	Toxin/antitoxin system?	3,17
SA0412	hypothetical protein	3,30
butA	Acetoin Reductase	3,42
SA2479	Polyisoprenoid-binding protein	3,47
SA1407	Conserved hypothetical protein, expression affected by ClpP deletion	3,51
SAS041	Hypothetical 62 aa protein	3,53
groES	heat shock protein, 10 kDa chaperonin, CtsR regulon (heat shock)	3,59
hutH	Histidine ammonia-lyase	3,63
groEL	heat shock protein, 60 kDa chaperonin, CtsR regulon (heat shock)	3,67
SA1824	Toxin/antitoxin system?	3,78
SA1978	similar to ferrichrome ABC transporter	3,88
SA1820	ribD, Riboflavin biosynthesis protein	3,89
dnaJ	Chaperone protein DnaJ	3,90
cysM	cysteine biosynthetic process from serine	3,91
pyrR	Regulates transcriptional attenuation of the pyrimidine nucleotide	4,00
SA1823	Arsenical pump membrane protein	4,18
SA0916	purE, phosphoribosylaminoimidazole carboxylase catalytic subunit	4,22
SA1821	Toxin/antitoxin system?	4,27
uhpT	Hexose phosphate transport protein	4,33
SA1822	Toxin/antitoxin system?	4,41
pyrC	Dihydroorotase	4,84
SA0481	hypothetical protein, CtsR regulon (heat shock)	4,91
purK	purine biosynthesis protein	5,09
purD	purine biosynthesis protein	5,10
purN	purine biosynthesis protein	5,29
clpC	CtsR regulon (heat shock)	5,30
SA0482	CtsR regulon (heat shock)	5,33
purH	purine biosynthesis protein	5,36
dnaK	Chaperone protein DnaK, CtsR regulon (heat shock)	5,41
purM	purine biosynthesis protein	5,60
ctsR	regulator of stress response, CtsR regulon (heat shock)	5,83
purF	purine biosynthesis protein response to hyperosmotic and heat shock by preventing the aggregation of stress-denatured proteins, CtsR regulon (heat shock)	5,85
grpE	response to hyperosmotic and heat shock by preventing the aggregation of stress-denatured proteins, CtsR regulon (heat shock)	5,96
purC	purine biosynthesis protein	6,00
purL	purine biosynthesis protein	6,01
SA0919	Purine metabolism	6,11
purQ	purine biosynthesis protein	6,33
SA1596	aroK, Amino acid biosynthesis	6,50
hrcA	Heat-inducible transcriptional repressor, CtsR regulon (heat shock)	6,83
valS	valyl-tRNA synthetase, Protein biosynthesis	7,54
folC	putative folylpolyglutamate synthase/dihydrofolate synthase	8,07
trpA	tryptophan biosynthesis	8,45
SA1199	hypothetical protein	8,51
trpB	tryptophan biosynthesis	10,37
trpC	tryptophan biosynthesis	12,91
trpF	tryptophan biosynthesis	13,70
trpG	tryptophan biosynthesis	14,94
trpD	tryptophan biosynthesis	15,68
SA2355	conserved hypothetical protein, contain Thioredox_DsbH and acetyl-transferase domains	16,04
clpB	Chaperone protein, CtsR regulon (heat shock)	19,99
SA1979	similar to ferrichrome ABC transporter	26,78
SA0576	Toxin/antitoxin system?	307,37
SA0575	Toxin/antitoxin system?	356,23

### Supplementary Figure. 6. ComK2 regulon

Are listed all the genes for which the expression is activated at least 3 times by ComK2. The real factor is mentioned in the last column.



**Supplementary Figure. 7. Inhibited genes during competence development in *S. aureus***

Each circle represent the genes for which the expression is inhibited by at least a two-fold factor by SigH (in blue), ComK1 (in yellow), ComK2 (in green) or in total (outside back circle). The number of genes in each category is shown inside the circle or at the intersection between circles, when the same gene is controlled by more than one regulator.



## 2.5. Materials and Methods

### Bacterial strains and culture conditions

*S. aureus* strains used in this project are listed in Table S2. Staphylococcal strains were grown in BHI medium or a complete synthetic medium, called CS2 <sup>58</sup> (Table S3) depending on the experiment.

#### “day” protocol

Cells were isolated from -80 °C stock on BHI plate. Four clones were inoculated in 10 mL of BHI and incubated at 37 °C with shaking at 180 rpm overnight (Preculture was prepared from -80 °C stock by overnight growth in TSB medium in Morikawa *et. al.* 2012 <sup>58</sup>). After calculating the required volume of overnight culture to have an OD of 0.4 in 10 mL of CS2 medium (Cells were directly harvested from five hundred µL of overnight culture in Morikawa *et. al.* 2012 <sup>58</sup>), cells were collected by brief centrifugation, washed with CS2 medium, resuspended in 10 mL of CS2 medium, and incubated at 37 °C with shaking at 180 rpm for 8 hours.

#### “night” or “dilution” protocol

Cells were isolated from -80 °C stock on BHI plate. Four clones were inoculated in 10 mL of BHI and incubated at 37 °C with shaking at 180 rpm roughly for 3 hr until OD reached 2.4 to 2.6. Cells were collected from 1 mL of pre-culture by brief centrifugation and resuspended in 4 mL of CS2 medium. After adjusting OD of initial suspension to 0.5, the cell suspension was diluted to  $10^{-1}$ ,  $10^{-2}$ ,  $10^{-3}$ ,  $10^{-4}$  and more if needed. Dilutions were incubated overnight at 37 °C with shaking at 120 rpm. Cells were collected when the OD was between 2 and 2.2.

### Construction of GFP reporter plasmids and strains

Oligonucleotides used in this study are listed Table S1. The promoters and pRIT plasmids were amplified with the dedicated primers. Thanks to these primers, there is a 25-50 bp homologous overlap regions between obtained fragments of promoters and linear plasmid. Subsequent to PCR with Phusion High-Fidelity DNA Polymerase (purchased from Thermo Scientific), each of the PCR fragments were purified using standard commercial silicon column (PCR cleanup kit). Each of the

fragments to be joined were then electrophoresed to verify the absence of non-specific or minor PCR fragments.

The promoters were inserted and circularized in pRIT plasmid using the Gibson assembly method. Gibson assembly master mix was prepared by adding 320  $\mu\text{L}$  of 5x ISO buffer, 0.64  $\mu\text{L}$  10U/ $\mu\text{L}$  T5 exonuclease, 20  $\mu\text{L}$  of 2U/ $\mu\text{L}$  Phusion polymerase, 160  $\mu\text{L}$  of 40U/ $\mu\text{L}$  Taq DNA ligase and 699.36  $\mu\text{L}$  of water. All reagents were purchased from New England Biolabs. 5  $\mu\text{L}$  of mixture of the two DNA fragments containing 100 ng of linear pRIT plasmid and 3-fold of inserts were added to 15  $\mu\text{L}$  of Gibson assembly master mix. The tubes were therefore incubated at 50 °C for 1 hr. Finally, 1  $\mu\text{L}$  of the assembly reaction was transformed into IM08B competent cells. Transformed cells were incubated at 37 °C on LB agar with 100  $\mu\text{g}/\text{mL}$  ampicillin.

In order to select the transformants containing the expected plasmid, colony PCRs were performed on each transformant. The resulting plasmid was extracted from overnight culture of the positive transformant and purified by commercial kit (NucleoSpin Plasmid extraction kit). All the plasmids were verified by sequencing (GATC company). Finally, in order to obtain the reporter stains, *S. aureus* electrocompetent cells were transformed with the constructed plasmids.

### **Construction of *S. aureus* deletion mutants**

In order to investigate the role of the main genes predicted to be involved in the regulation of competence development in *S. aureus*, listed in Table S1, allelic replacement constructs were cloned into pIMAY, which is a plasmid vector for allelic exchange with a temperature sensitive replicon. All the primers used for genetic manipulation in the current study are listed in Table S1. Fragments corresponding to 1 kbp flanking regions of the genes to be deleted were amplified from *S. aureus* N315 genomic DNA using the primers flanked by restriction enzyme sites chosen from multiple cloning site (MCS) of pIMAY, thus generating the upstream and downstream fragments of the genes. The upstream or downstream regions were digested by the two chosen restriction enzymes and ligated into pIMAY cut with the same enzymes. The resulting plasmids were transformed into *E. coli* IM08B. Colony PCR was used to verify the transformants. Then, the plasmids were purified from the positive transformants. After verifying the plasmids sequence, *S. aureus* N315ex *wo*p strain was transformed by electroporation and then plated on BHI agar with 10 mg/mL chloramphenicol and incubated at 28 °C.

In order to allow pIMAY integration into the chromosome, a single colony from the transformation plate was resuspended in 200  $\mu\text{L}$  of BHI. The suspension was 10-fold diluted down to  $10^{-3}$  and 100  $\mu\text{L}$  of each dilution was spread on BHI with 10 mg/mL chloramphenicol and incubated at 37 °C overnight.

Large colonies were subsequently streaked in the same conditions. Meanwhile colony PCR analysis was performed to check the absence of extrachromosomal pIMAY and whether plasmid integration had occurred upstream or downstream.

Based on the colony PCR results, an overnight culture in BHI at 28 °C without chloramphenicol was made from the positive clones. Then, the overnight culture was 10-fold diluted down to  $10^{-7}$ . 100  $\mu$ L of  $10^{-4}$  to  $10^{-7}$  dilutions were plated onto BHI containing 1  $\mu$ g/mL anhydrotetracycline (aTc). The plates were therefore incubated at 28 °C for 2-3 days. The induction of *secY* antisense RNA produced from pIMAY would prevent growth of cells that retain the integrated plasmid and select for cells that have lost the plasmid [203](#). Large colonies were patched on BHI without antibiotic and BHI with 10 mg/mL chloramphenicol and grown at 37 °C overnight. Chloramphenicol-sensitive colonies were screened by colony PCR with oligonucleotides to identify clones containing the desired mutation. Putative mutants were verified by PCR and DNA sequencing.

### **Electroporation in *E. coli***

*E. coli* IM08B was streaked on LB plate and incubated at 37 °C overnight. Three colonies were resuspended in 20 mL of SLB and the cell suspension was diluted 10-fold and 100-fold then grown in 20 mL of SLB overnight at 37 °C with shaking at 180 rpm. The next day, 10 mL of the overnight culture with cell density between 4 and 6 were added into 1 L of SLB and incubated at 37 °C, 180 rpm until OD reached 0.8-0.9. The culture was immediately chilled in ice water for 30 min. All subsequent steps were performed at 4 °C. Cells were harvested by centrifugation at 4500 g for 15 min. Pellets were resuspended in an equal volume of sterile cold water and centrifuged at 4500 g for 15 min, then resuspended in half the volume of sterile cold water and centrifuged a second time at 4500 g for 15 min. Cells were resuspended again in 20 mL of sterile cold 10% glycerol (w/v) and centrifuged 4500xg for 15 min and finally resuspended in 3 mL of cold 10% glycerol. Aliquots of 100  $\mu$ L were frozen and stocked at -80 °C. For electroporation, electrocompetent cells were thawed on ice for 3 min, then 50  $\mu$ L of cells was mixed with 1  $\mu$ L of ligation product (or Gibson assembly product), transferred into 2-mm electroporation cuvette (pre-cold on ice) and, immediately, pulsed at 2.5 kV for 5 ms. The cells were immediately recovered in 1 mL of SOC and incubated at 37 °C for 1 hr. Finally, 100  $\mu$ L of transformation were plated on LB with the appropriate antibiotic depending on plasmids.

### **Electroporation in *S. aureus***



*S. aureus* strains were isolated on BHI plate at 37 °C. An overnight culture with 3 colonies was prepared in 10 mL of BHI and incubated at 37 °C with shaking at 180 rpm, and then diluted 200-fold in 200 mL of BHI. The culture was incubated at 37 °C with shaking at 180 rpm until the OD was between 0.6 and 0.8. Then, the culture was chilled in ice for 30 min. Cells were harvested by centrifugation at 4500 g for 15 min and resuspended in an equal volume of sterile cold 0.5 M sucrose. After successive centrifugations at 4500 g for 15 min, pellets were resuspended in half the volume of sterile 0.5 M sucrose, then in 20 mL of sterile cold 0.5 M sucrose with 10% glycerol, and finally in 1 mL of 0.5 M sucrose and 10% glycerol. Aliquots of 100 µL each were frozen at -80 °C. For electroporation, electrocompetent cells were thawed at room temperature for 10 min, heated at 56 °C for 2 min, and then immediately resuspended in 500 µL of 0.5 M sucrose with 10% glycerol. Cell suspension was centrifuged at 8000 rpm for 1 min and resuspended in 50 µL of 0.5 M sucrose and 10% glycerol. 5 µL of plasmid (1-2 µg) was mixed with the cells, transferred into 2 mm electroporation cuvette at room temperature and pulsed at 3 kV, 1 ms. The cells were immediately resuspended in 1 mL of BHI supplemented with 0.5% glucose and incubated at 37 °C for 1 hr or 28 °C for 3 hr.

### **Flow cytometry to determine the percentage of the population expressing GFP**

Subsequent to growth in CS, 500 µL of the cells was harvested by centrifugation at 11000 g for 1 min. Pellets were resuspended in 500 µL of cold 70% ethanol and incubated on ice for 20 min. Then, the cells were resuspended in 500 µL of PBS after centrifugation at 11000 g for 1 min. Finally, the population with GFP signals was detected by Flow cytometry.

### **Microscopy to determine the percentage of the population expressing GFP**

In order to observe competent cells expressing GFP microscopically, 500 µL were collected from the overnight CS2 cultures then centrifuged at 11000 g for 30 seconds. Cell pellets were rinsed with sterile PBS, fixed with 1.6% formaldehyde (Invitrogen™) for 20 minutes at room temperature then rinsed twice in equal volume of sterile PBS and finally re-suspended in 100 µL 45% glycerol/PBS and stored at -20°C. 10 µL of fixed cells were mounted on 1% agarose coated slide with a drop of mounting medium (Fluoromount-G®). Sildes were therefore observed under confocal microscope Nikon Eclipse Ti-E equipped with a spinning disk module (Yokogawa CSU-X1-A1, ImagerieGif platform) and a super resolution system (Live SR 3D, Gataka systems, ImagerieGif platform). Images were acquired with x100 oil objective (Nikon; NA = 1.49) and analyzed with Fiji software (ImageJ 1.53c).

### Natural transformation of *S. aureus* cells

Wild type strain (St12), *comK1* mutant (St27), *comK2* mutant (St38) and *sigH* mutant (St45) were isolated on BHI plate. Four colonies of each strain were inoculated in 10 mL of BHI medium and incubated at 37 °C with shaking at 180 rpm roughly for 3 hr until OD reached 2.4 to 2.6. After calculating the volume of preculture to have 0,4 of OD600 in 10 mL of CS2, cells were harvested by centrifugation at 10000xg for 1min, washed with CS2, resuspend in 10 mL of CS2, and incubated at 37 °C for 8 hours with shaking at 180 rpm. The OD600 of culture was monitored every hour from 5 hours to 8 hours. 100 µL of the suspension was diluted 10<sup>-6</sup>. 100 µL of the dilution was plated on BHI plate without antibiotic to estimate the total number of cells. Cells were harvested by centrifugation at 5000xg for 5 min at 4 °C, resuspend in 10 mL of fresh CS2. Ten µg of DNA was added to the suspension and incubated at 37 °C for 2.5 hours with agitation at 180 rpm. 5 mL of the culture was mixed into 125 mL of melted BHI agar pre-cooled to 55 °C together with antibiotic, and the mixture was aliquoted in 5 plates. After solidification, the plates were incubated at 37 °C for 48 hours. Colonies were counted to calculate genetic transformation efficiency.

### RNA-sequencing

Sampling and isolation. Cultures of *S. aureus* were grown in CS2 medium, at 37°C and 200 rpm until the OD600 reached. To quench cellular metabolism and transcription and to stabilize RNAs, cells were harvested by incubating cultures in an equal volume of icecold methanol for 3 min before centrifugation at 4000 g for 10 min at 4°C. Pellets were frozen in liquid nitrogen and stored at -80°C. Three independent biological replicates were collected for each of the four strains (wild type, St29;  $\Delta comK1$ , St40;  $\Delta comK2$ , St41;  $\Delta sigH$ , St61). For extraction of RNA, cells were lysed using Lysing Matrix B and a FastPrep instrument (both MP Biomedicals), and RNA was isolated using the RNeasy Mini Kit (Qiagen). RNA was treated with TURBO DNase (Ambion) as described, purified using the RNA Cleanup protocol from the RNeasy Mini Kit (Qiagen), and stored at -80°C. The integrity of the RNA was analysed using an Agilent Bioanalyzer (Agilent Technologies).

rRNA depletion, library construction and sequencing. Removal of 23S, 16S and 5S rRNA using the RiboZero rRNA Removal Kit (Epicentre) (two times), strand-specific library construction yielding fragments of size range 100–500 bp, pooling of 12 indexed libraries, sequencing in one flow-cell lane on a Illumina HiSeq2000 instrument with a 75 nt paired end protocol, and demultiplexing of the 12 samples of indexed reads was performed by the “Next Generation Sequencing (NGS) Core Facility” from the Institute for the Integrative Biology of the Cell (I2BC, Gif sur Yvette, France). The complete dataset from this study will be available.

Read mapping and analysis of differential expression. Differential expression of all annotated features was assessed using the R statistical programming environment. Differential expression was determined between the wild type strain samples (St29, n = 3) and each of the 9 samples (each n = 3) in which *sigH*, *comK1*, or *comK2* were absent. The output from the differential expression analysis is presented in Tables S4, S5 and S6. Genes with False Discovery Rate (FDR)-corrected P-values < 0.01 and a ratio of differential expression superior to two were considered significantly differentially expressed.

## Supplementary data

Table S1. Primers

Name	Nucleotide sequence	Descriptions
IM151	TAC ATG TCA AGA ATA AAC TGC CAA AGC	Used for verifying and sequencing the allelic replacement pIMAY constructs
IM152	AAT ACC TGT GAC GGA AGA TCA CTT CG	
KpnI-sigH KO-F-F	GCA GGT ACC GAC CCG CAT AAC TTG GGA TCA ATT TTA AGA	To amplify upstream of <i>sigH</i> used for the deletion of <i>sigH</i>
Sall-sigH KO-F-R	GCA GTC GAC CCC CTT CTA TCT AAA ATT TAA GGT TAG TTT AAT ATT GTT ACA TTC	
EcoRI-sigH KO-R-F	GCA GAA TTC AGC GCC TTA GGA CGT GAA TTG AAT TAT AAC GTG	To amplify downstream of <i>sigH</i> used for the deletion of <i>sigH</i>
XmaI-sigH KO-R-R	GCA CCC GGG CTG GTG TTT CTC GGC CAA ACA TAT CTA CTA ATA C	
sigH-OUT-F	TGC AGC GCT TAT TGC ACC ATA TGA ATA TGC	To verify the <i>sigH</i> deletion mutant
sigH-OUT-R	TTG CCC TCC CAC TCT TAA CAT TTC GTG A	
Kpn I-SA2017-up-F	GCA GGT ACC CAT ATG ACA CTC CCA ATG C	To amplify upstream of <i>sa2107</i> used for the deletion of <i>sa2017</i>
Sal I-SA2017-up-R	GCA GTC GAC GTA GAA GTA CCT CCA AAA ATC AAT	
EcoRI-SA2107-down-F	GCA GAA TTC GCC CAT TAA CCT ATT TTT CAT A	To amplify downstream of <i>sa2107</i> used for the deletion of <i>sa2017</i>
Sac I-SA2107-down-R	GCA GAG CTC ACA ATA TTT GAT GCC TGT GCT A	
check-mutant-SA2107-F	TGA AAG TCA GTC GTA CTC GAC AT	To verify the <i>sa2107</i> deletion mutant
check-mutant-SA2107-R	GTT GCT CCC ATA TGC ATC TCA	
Kpn I-K1-up-F	GGT ACC GAA TGT TTA GTC ATG GTA CGT TGA TG	To amplify upstream of <i>comK1</i> used for the deletion of <i>comK1</i>
Sal I-K1-up-R	GCA GTC GAC AAG CAA AAC CTC GCT TTA TTA AGT TT	
Sma I-K1-down-F	GCA CCC GGG TAT GCG GAA TGT TTT ATA	To amplify downstream of <i>comK1</i> used for the deletion of <i>comK1</i>
Sac I-K1-down-R	GCA GAG CTC AAG GTT TGA TGT TAT CAG TGA ATG	
Kpn I-K2-up-F	GCA GGT ACC ATC TAT GCA TTG TTA CAT ATT CAT	To amplify upstream of <i>comK2</i> used for the deletion of <i>comK2</i>
Sal I-RK2-up-	GCA GTC GAC ATA TAG TAC TCC TCG TAT AAT AAG	
EcoR I-K2-down-F	GCA GAA TTC GAA GCA CTT CAT TGA AAA TAC	To amplify downstream of <i>comK2</i> used for the deletion of <i>comK2</i>
Sac I-K2-down-R	GCA GAG CTC AAG TGT CAG ATT CGT GTA AT	
check-mutant-comK1-F	TCT CAG GGA ATG CTA TGG TTA AAG T	To verify the <i>comK1</i> deletion mutant
check-mutant-comK1-R	CAG TGG CTT ATT GGT CAC AGG	
check-mutant-comK2-F	GAG ACA ACA GCA AAT ATG ACA ACA AG	To verify the <i>comK2</i> deletion mutant
check-mutant-comK2-R	GTG CAT GAA TAT TAC CAC TTC TAA TCG	
Kpn I-up- <i>agrA</i> -F	GCA GGTACC TTA GTG ACC ATG ATC ATA ATG TAT TTG AGT	To amplify upstream of <i>agrA</i> used for the deletion of <i>agrA</i>
Sal I-up- <i>agrA</i> -R	GCA GTCGAC TTC ACA TCC TTA TGG CTA GTT GTT AA	
EcoR I-down- <i>agrA</i> -F	GCA GAATTC AAT AAG ATA ATA AAG TCA GTT AAC GGC GTA	To amplify downstream of <i>agrA</i> used for the deletion of <i>agrA</i>
Sac I-down- <i>agrA</i> -R	GCA GAGCTC TCT GCT GAT ATG TTA TTT GAA CCA AG	
Check-mutant- <i>agrA</i> -F	GTA TTA CTT CTA TGG AAG TAG AGC CGT ATT	To verify the <i>agrA</i> deletion mutant
Check-mutant- <i>agrA</i> -R	GTG TAT CGC ACG AAT GAA GCA	
Kpn I-up <i>srrA</i> -F	GCA GGT ACC GAC ATT GAT TGC CAA AGG ACT AGT TGA G	To amplify upstream of <i>srrA</i> used for the deletion of <i>srrA</i>
EcoR I-up <i>srrA</i> -R	GCA GAA TTC TAC CTC CCA CAC ATG CTT TTC TTT ACA	
sma I-down <i>srrA</i> -F	GCA CCC GGG GGG CGT TGG GTA TAA ATT TGA GGT TAA ATC TA	To amplify downstream of <i>srrA</i> used for the deletion of <i>srrA</i>
Sac I-down <i>srrA</i> -R	GCA GAG CTC GCT CAT TAG TCA TAT CAC GAA CTG TCA C	
Check-mutant- <i>srrA</i> -F	GAA ATT ATC ACA AGC AGC AAT GGA AGT ACT	To verify the <i>srrA</i> deletion mutant
Check-mutant- <i>srrA</i> -R	TGG TAT TGT TAC AGA ACC GGA TGA AAT AAA AG	
Sal I-up <i>nreC</i> -F	GCA GTC GAC GAG CAT TCA ATC GAA AAG ATA GTT TTT GCT GAT G	To amplify upstream of <i>nreC</i> used for the deletion of <i>nreC</i>
Sma I-up <i>nreC</i> -R	GCA CCC GGG TCA AAT TGG AAT GTT CAA TGT AAC ATT GGT AC	
Not I-down <i>nreC</i> -F	GCA GCG GCC GC CGA GAG GAA ATG ATT ATC TTC TGG C	To amplify downstream of <i>nreC</i> used for the deletion of <i>nreC</i>
Sac I-down <i>nreC</i> -R	GCA GAG CTC GGTGACTGAATTTTGGCATAAC	

## Chapter 2

Check-mutant-nreC-F	GAT GAA TTA GGG TGT AAG TCA TGA TTA ATG AGG AC	To verify the <i>nreC</i> deletion mutant
Check-mutant-nreC-R	CTATTGAAGTTGCTACAACACTTCCAGCAC	
Sal I-up airR-F	GCA GTC GAC CAA GGT GCG TTA AAA TAT TTA ATT GAG GGC	To amplify upstream of <i>airR</i> used for the deletion of <i>airR</i>
EcoR I-up airR-R	GCA GAA TTC CAT GGG TTA TCT CCT TAA ATC AAG CTA TT	
sma I-down airR-F	GCA CCC GGG GTG TAA GTT TTA TCA AAT CTA GTG ATT GTG	To amplify downstream of <i>airR</i> used for the deletion of <i>airR</i>
Sac I-down airR-R	GCA GAG CTC GGT GAA ACA GCA TTC TAA TTG CAA TTA	
Check-mutant-airR-F	GGA GCG ATT TGA TGG AAC AAA GG	To verify the <i>airR</i> deletion mutant
Check-mutant-airR-R	GGC CAA TTT TAG TTG CAA TTT CTT CGC	
KpnI-up-luxS-F	GCA GGTACC AAT ATG TTG TCG TAG TTG TGC TAC	To amplify upstream of <i>luxS</i> used for the deletion of <i>luxS</i>
Sal I-up-luxS-R	GCA GTCGAC TTT GAA TTT CCT CCT ATT TAC TAC TCA A	
Sma I-down-luxS-F	GCA CCCGGG ATC TTA GTC AAT CAA GTT AAT CAG AAA AGC	To amplify downstream of <i>luxS</i> used for the deletion of <i>luxS</i>
Sac I-down-luxS-R	GCA GAGCTC AAA GCT GAA ACG CTT GAA GAA GC	
Check mutant-luxS-F	GCC ATC TAA AAT CAC ACC ATG CAC	To verify the <i>luxS</i> deletion mutant
Check mutant-luxS-R	CAT TAG GTT CTC AAA TGG TTG TAC TTG	
KpnI-up-comGA-F	GCA GGTACC CGG AAG GTC CGA GTG TTA GTA AAT	To amplify upstream of <i>comGA</i> used for the deletion of <i>comGA</i>
Sal I-up-comGA-R	GCA GTCGAC CAC CTC CTA CAT ATA ATC ACG TAG GAG	
Sma I-down-comGA-F	GCA CCCGGG ACT ACA TTC TAA GAA GCG ACA ATT AAG TAA	To amplify downstream of <i>comGA</i> used for the deletion of <i>comGA</i>
Sac I-down-comGA-R	GCA GAGCTC CAT CTC TAT CAA TGT AAA CGC TTG AG	
Check mutant-comGA-F	CTA TTG TTG ATG TGG TTG TTA TTC CAG	To verify the <i>comGA</i> deletion mutant
Check mutant-comGA-R	CAA CCT GTT GAT TGT ATG TGA GCA G	
KpnI-up-comC-F	GCA GGTACC GGT ACT TAT CTA GAT ATT GCT AGG GAT	To amplify upstream of <i>comC</i> used for the deletion of <i>comC</i>
Sal I-up-comC-R	GCA GTCGAC ATG ACA ACC TCC TTA TGT AAA TTA TAG T	
Sma I-down-comC-F	GCA CCCGGG GTC ATT GCT GGT TTA GTT GCT TTA A	To amplify downstream of <i>comC</i> used for the deletion of <i>comC</i>
Sac I-down-comC-R	GCA GAGCTC TAC GAA GTC TGC TAA CAA CTC AA	
Check mutant-comC-F	CTA GTA TCG GTC TAG ACC ATA CAG ATA	To verify the <i>comC</i> deletion mutant
Check mutant-comC-R	TTT GAT GCT GAA GGA TAT CTC AAC C	
KpnI-up-ssb-F	GCA GGTACC TAA TGC CAG GCG TAT TAA TTA CTG	To amplify upstream of <i>ssb</i> used for the deletion of <i>ssb</i>
Sal I-up-ssb-R	GCA GTCGAC GTA GTG TGT GAT TCA CCT CCT ATG	
Sma I-down-ssb-F	GCA CCCGGG ATC CAA TTA TCC TAA ACA TCC TTA ATA TAC	To amplify downstream of <i>ssb</i> used for the deletion of <i>ssb</i>
Sac I-down-ssb-R	GCA GAGCTC ACT ATT GAT TAG CTA TGC ATA AAT GGC	
Check mutant-ssb-F	AGT ATC AGG AAA CGA ACC ATT CTT TCA A	To verify the <i>ssb</i> deletion mutant
Check mutant-ssb-R	GGA ACA TAT CGA GAA TTC CCC GAT ATA T	
Gibson-pRIT-gfp-F	GGA TCC GGG AGG CCG TTT	To amplify the linear pRIT used for the reporter plasmid construction by Gibson
Gibson-pRIT-gfp-R	CTC TCA GTA CAA TCT GCT CTG ATG CCG	
pRIT-PcomF-F	ATA TGT GTT CGT GAA TTC CAG TTT GAT CAT AAT TCA GTG TTA CTA TAC ATG GTA CTG	To amplify the promoter of <i>comF</i> used for the reporter plasmid constructoin by Gibson
gfp-PcomF-R	AAA CGG CCT CCC GGA TCC TGT TAT TTT ATA TCT TGT TAC ATT ATC CAT TCG ACC CAG TGA TAT	
pRIT-Pssb-F	CGG CAT CAG AGC AGA TTG TAC TGA GAG TGC GGT AAT GCC AGG CGT ATT AAT TAC TG	To amplify the promoter of <i>ssb</i> used for the reporter plasmid construction by Gibson
gfp-Pssb-R	AAA CGG CCT CCC GGA TCC TCT CCC GAC AAT TAC GAT TTT ATT TAG CAT GTA GTG TG	
pRIT-PcomC-F	CGG CAT CAG AGC AGA TTG TAC TGA GAG AAG CCT AAC GTT CCA GTG ATA TAT GCT GTT AAA AAT G	To amplify the promoter of <i>comC</i> used for the reporter plasmid construction by Gibson
gfp-PcomC-R	AAA CGG CCT CCC GGA TCC TGC AGC TAT AAG ATA ACA ATA CTA CCA AAT GAC AAC CT	
pRIT-PcomK1-F	CGG CAT CAG AGC AGA TTG TAC TGA GAG GTA ACG ACA AAT ATA GAA CTT GGG CAT GGA	

gfp-PcomK1-R	AAA CGG CCT CCC GGA TCC AAG CAA AAC CTC GCT TTA TTA AGT TTT AAA CC	To amplify the promoter of <i>comK1</i> used for the reporter plasmid construction by Gibson
pRIT-PcomK2-F	CGG CAT CAG AGC AGA TTG TAC TGA GAG GAC CAT CAC GAT ACA TCA TTT CAT TAG TGA	To amplify the promoter of <i>comK2</i> used for the reporter plasmid construction by Gibson
gfp-PcomK2-R	AAA CGG CCT CCC GGA TCC ATA TAG TAC TCC TCG TAT AAT AAG TTG TTA ATT TAA TGG	
pRIT-PsigH-F	CGG CAT CAG AGC AGA TTG TAC TGA GAG AAC AGG TGC AAT TGA ACA TGT ACC AGT TA	To amplify the promoter of <i>sigH</i> used for the reporter plasmid construction by Gibson
gfp-PsigH-R	AAA CGG CCT CCC GGA TCC GTA TTA AAC TAA CCC CTT CTA TCT AAA ATT TAA GGT TAG	
pRIT-X-F	GGC TTA ACT ATG CGG CAT CAG AGC	Used for verifying and sequencing the pRIT constructs
pRIT-X-R	GAA TTG TGA GCG GAT AAC AAT TTC ACA CAG G	

Table S2 Strains

Strains	Genotype/Construction	Source
St012	N315ex woφ	Morikawa et. al, 2012
St029	N315ex woφ / pRIT-PcomG-gfp	Morikawa et. al, 2012
St048	N315ex woφ / pRIT-PcomC-gfp	This study
St050	N315ex woφ / pRIT-Pssb-gfp	This study
St094	N315ex woφ / pRIT-PcomK1-GFP	This study
St095	N315ex woφ / pRIT-PcomK2-GFP	This study
St096	N315ex woφ / pRIT-PsigH-GFP	This study
St129	N315ex woφ / pRIT-PcomE-GFP	This study
St233	N315ex woφ / pRIT-PcomF-gfp	This study
St037	N315ex woφ ΔcomK1	This study
St040	ΔcomK1 / pRIT-PcomG-gfp	This study
St063	ΔcomK1 / pRIT-PcomC-gfp	This study
St064	ΔcomK1 / pRIT-Pssb-gfp	This study
St234	ΔcomK1 / pRIT-PcomF-gfp	This study
St097	ΔcomK1 / pRIT-PcomK1-GFP	This study
St038	N315ex woφ ΔcomK2	This study
St041	ΔcomK2 / pRIT-PcomG-gfp	This study
St066	ΔcomK2 / pRIT-PcomC-gfp	This study
St067	ΔcomK2 / pRIT-Pssb-gfp	This study
St236	ΔcomK2 / pRIT-PcomF-gfp	This study
St100	ΔcomK2 / pRIT-PcomK1-GFP	This study
St045	N315ex woφ ΔsigH	This study
St051	ΔsigH / pRIT-PcomG-gfp	This study
St060	ΔsigH / pRIT-PcomC-gfp	This study
St061	ΔsigH / pRIT-Pssb-gfp	This study
St235	ΔsigH / pRIT-PcomF-gfp	This study
St110	ΔsigH / pRIT-PcomK1-GFP	This study
St044	N315ex woφ ΔcomK1+ΔcomK2	This study
St047	ΔcomK1+ΔcomK2 / pRIT-PcomG-gfp	This study
St074	ΔcomK1+ΔcomK2 / pRIT-PcomC-gfp	This study

## Chapter 2

St075	<i>ΔcomK1+ΔcomK2 / pRIT-Pssb-gfp</i>	This study
St268	<i>ΔcomK1+ΔcomK2 / pRIT-PcomF-GFP</i>	This study
St055	<i>N315ex woφ ΔcomK1+ΔsigH</i>	This study
St081	<i>ΔcomK1+ΔsigH / pRIT-PcomG-gfp</i>	This study
St082	<i>ΔcomK1+ΔsigH / pRIT-PcomC-gfp</i>	This study
St083	<i>ΔcomK1+ΔsigH / pRIT-Pssb-gfp</i>	This study
St269	<i>ΔcomK1+ΔsigH / pRIT-PcomF-GFP</i>	This study
St053	<i>N315ex woφ ΔcomK2+ΔsigH</i>	This study
St077	<i>ΔcomK2+ΔsigH / pRIT-PcomG-gfp</i>	This study
St078	<i>ΔcomK2+ΔsigH / pRIT-PcomC-gfp</i>	This study
St079	<i>ΔcomK2+ΔsigH / pRIT-Pssb-gfp</i>	This study
St270	<i>ΔcomK2+ΔsigH / pRIT-PcomF-GFP</i>	This study
St054	<i>N315ex woφ ΔcomK1+ΔcomK2+ΔsigH</i>	This study
St069	<i>ΔcomK1+ΔcomK2+ΔsigH / pRIT-PcomG-gfp</i>	This study
St070	<i>ΔcomK1+ΔcomK2+ΔsigH / pRIT-PcomC-gfp</i>	This study
St071	<i>ΔcomK1+ΔcomK2+ΔsigH / pRIT-Pssb-gfp</i>	This study
St271	<i>ΔcomK1+ΔcomK2+ΔsigH / pRIT-PcomF-GFP</i>	This study
St103	<i>N315ex woφ ΔagrA</i>	This study
St107	<i>ΔagrA / pRIT-PcomG-GFP</i>	This study
St122	<i>N315ex woφ ΔluxS</i>	This study
St123	<i>ΔluxS / pRIT-PcomG-GFP</i>	This study
St117	<i>N315ex woφ ΔsrrA</i>	This study
St145	<i>ΔsrrA / pRIT-PcomG-gfp</i>	This study
St147	<i>ΔsrrA / pRIT-Pssb-gfp</i>	This study
St247	<i>ΔsrrA / pRIT-PcomK1-gfp</i>	This study
St118	<i>N315ex woφ ΔnreC</i>	This study
St158	<i>ΔnreC / pRIT-PcomG-gfp</i>	This study
St142	<i>N315ex woφ ΔairR</i>	This study
St177	<i>ΔairR / pRIT-PcomG-gfp</i>	This study
St250	<i>N315ex woφ ΔsrrA+ΔsigH</i>	This study
St251	<i>ΔsrrA+ΔsigH / pRIT-PcomG-gfp</i>	This study
St252	<i>ΔsrrA+ΔsigH / pRIT-Pssb-gfp</i>	This study
St253	<i>ΔsrrA+ΔsigH / pRIT-PcomK1-gfp</i>	This study

Table S3. The components of synthetic CS2 medium

	Ingredient	stock conc.	1000 mL
H2O			480 mL
Solution A (10x)	Na <sub>2</sub> HPO <sub>4</sub> ·2H <sub>2</sub> O	89.52 g/L	100 mL
	KH <sub>2</sub> PO <sub>4</sub>	30 g/L	
	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	20 g/L	
Solution B1 (100x)	MgSO <sub>4</sub> ·7H <sub>2</sub> O	5 g/L	10 mL
	MnSO <sub>4</sub> ·H <sub>2</sub> O	2 mM	
Solution B2 (100x)	FeSO <sub>4</sub> ·7H <sub>2</sub> O	0.28 g/L	10 mL
Solution C (50x)	Glucose	50% (w/v)	20 mL
Solution D (10x)	D-Biotine	1 mg/L	100 mL
	Acide Nicotinique	20 mg/L	
	Acide D Panthothenique	20 mg/L	
	Riboflavine	20 mg/L	
	Thiamine hydrochloride	20 mg/L	
	Pyridoxine hydrochloride	40 mg/L	
Solution Adenine (20x)	Adénine 1/2 H <sub>2</sub> SO <sub>4</sub>	300 mg/L	50 mL
Solution Guanine (20x)	NaOH	0.05 M	50 mL
	L Guanine	592 mg/L	
Solution Glutamine (20x)	L Glutamine	5.84 g/L	50 mL
Solution AA (50x)	RPMI 1640		20 mL
Solution Traces (100x)	ZnSO <sub>4</sub>	1 mM	10 mL
	CuSO <sub>4</sub>	0.05 mM	
	CoCl <sub>2</sub> ·6H <sub>2</sub> O	12 mg/L	
	NaMoO <sub>4</sub> ·2H <sub>2</sub> O	12 mg/L	
Solutin Ca	CaCl <sub>2</sub>	100 mM	0.7 mL





## 2.6. References

1. Klevens, R. M. Invasive Methicillin-Resistant <EMPH TYPE="ITAL">Staphylococcus aureus</EMPH> Infections in the United States. *JAMA* **298**, 1763 (2007).
2. Lowy, F. D. Staphylococcus aureus Infections. *N. Engl. J. Med.* **339**, 520–532 (1998).
3. Gordon, R. J. & Lowy, F. D. Pathogenesis of Methicillin-Resistant Staphylococcus aureus Infection. *Clin. Infect. Dis.* **46**, S350–S359 (2008).
4. Dietz, I., Jerchel, S., Szaszák, M., Shima, K. & Rupp, J. When oxygen runs short: the microenvironment drives host–pathogen interactions. *Microbes Infect.* **14**, 311–316 (2012).
5. Beenken, K. E. *et al.* Global Gene Expression in Staphylococcus aureus Biofilms. *J. Bacteriol.* **186**, 4665–4684 (2004).
6. Morikawa, K. *et al.* Expression of a cryptic secondary sigma factor gene unveils natural competence for DNA transformation in Staphylococcus aureus. *PLoS Pathog* **8**, e1003003 (2012).
7. Solomon, J. M. & Grossman, A. D. Who’s competent and when: regulation of natural genetic competence in bacteria. *Trends Genet.* **12**, 150–155 (1996).
8. van Sinderen, D. *et al.* comK encodes the competence transcription factor, the key regulatory protein for competence development in Bacillus subtilis. *Mol Microbiol* **15**, 455–462 (1995).
9. Meibom, K. L., Blokesch, M., Dolganov, N. A., Wu, C. Y. & Schoolnik, G. K. Chitin Induces Natural Competence in Vibrio cholerae. *Science (80-. )*. **310**, 1824–1827 (2005).
10. Lee, M. S. & Morrison, D. A. Identification of a new regulator in Streptococcus pneumoniae linking quorum sensing to competence for genetic transformation. *J Bacteriol* **181**, 5004–5016 (1999).
11. Laurenceau, R. *et al.* A type IV pilus mediates DNA binding during natural transformation in Streptococcus pneumoniae. *PLoS Pathog* **9**, e1003473 (2013).
12. Mirouze, N. . F. C. . C.-L. R. Antibiotic sensitivity reveals that wall teichoic acids mediate DNA binding during competence in Bacillus subtilis. *Nat. Commun.* **Under revi**, (2018).
13. Briley Jr., K. *et al.* The secretion ATPase ComGA is required for the binding and transport of transforming DNA. *Mol Microbiol* **81**, 818–830 (2011).
14. Meima, R. *et al.* The bdbDC operon of Bacillus subtilis encodes thiol-disulfide oxidoreductases required for competence development. *J Biol Chem* **277**, 6994–7001 (2002).
15. Mortier-Barriere, I. *et al.* A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. *Cell* **130**, 824–836 (2007).
16. Morikawa, K. *et al.* A new staphylococcal sigma factor in the conserved gene cassette: functional significance and implication for the evolutionary processes. *Genes to Cells* **8**, 699–712 (2003).
17. Fagerlund, A., Granum, P. E. & Havarstein, L. S. Staphylococcus aureus competence genes: mapping of the SigH, ComK1 and ComK2 regulons by transcriptome sequencing. *Mol Microbiol* **94**, 557–579 (2014).
18. Andersen, J. B. *et al.* New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl. Environ. Microbiol.* **64**, 2240–6 (1998).
19. Solomon, J. M., Magnuson, R., Srivastava, A. & Grossman, A. D. Convergent sensing pathways mediate response to two extracellular competence factors in Bacillus subtilis. *Genes Dev.* **9**, 547–558 (1995).
20. Moreno-Gómez, S. *et al.* Quorum sensing integrates environmental cues, cell density and cell history to control bacterial competence. *Nat. Commun.* **8**, 854 (2017).
21. Lo Scudato, M. & Blokesch, M. The Regulatory Network of Natural Competence and Transformation of Vibrio cholerae. *PLoS Genet.* **8**, e1002778 (2012).
22. Chung, Y. S., Breidt, F. & Dubnau, D. Cell surface localization and processing of the ComG proteins, required for DNA binding during transformation of Bacillus subtilis. *Mol Microbiol* **29**, 905–913 (1998).

23. Miller, M. B. & Bassler, B. L. Quorum Sensing in Bacteria. *Annu. Rev. Microbiol.* **55**, 165–199 (2001).
24. Alloing, G., Martin, B., Granadel, C. & Claverys, J. P. Development of competence in *Streptococcus pneumoniae*: pheromone autoinduction and control of quorum sensing by the oligopeptide permease. *Mol Microbiol* **29**, 75–83 (1998).
25. Turgay, K., Hahn, J., Burghoorn, J. & Dubnau, D. Competence in *Bacillus subtilis* is controlled by regulated proteolysis of a transcription factor. *EMBO J* **17**, 6730–6738 (1998).
26. Lo Scudato, M. & Blokesch, M. A transcriptional regulator linking quorum sensing and chitin induction to render *Vibrio cholerae* naturally transformable. *Nucleic Acids Res.* **41**, 3644–3658 (2013).
27. Le, K. Y. & Otto, M. Quorum-sensing regulation in staphylococci—an overview. *Frontiers in Microbiology* **6**, (2015).
28. Maree, M. *et al.* MRSA emerges through natural transformation. *bioRxiv* 2020.12.08.415893 (2020). doi:10.1101/2020.12.08.415893
29. Rauhut, R. & Klug, G. mRNA degradation in bacteria. *FEMS Microbiol. Rev.* **23**, 353–370 (1999).
30. Berka, R. M. *et al.* Microarray analysis of the *Bacillus subtilis* K-state: genome-wide expression changes dependent on ComK. *Mol Microbiol* **43**, 1331–1345 (2002).
31. Ogura, M. *et al.* Whole-genome analysis of genes regulated by the *Bacillus subtilis* competence transcription factor ComK. *J Bacteriol* **184**, 2344–2351 (2002).
32. Dagkessamanskaia, A. *et al.* Interconnection of competence, stress and CiaR regulons in *Streptococcus pneumoniae*: competence triggers stationary phase autolysis of *ciaR* mutant cells. *Mol Microbiol* **51**, 1071–1086 (2004).
33. Peterson, S. N. *et al.* Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. *Mol. Microbiol.* **51**, 1051–1070 (2004).
34. Liu, Q., Yeo, W. & Bae, T. The SaeRS Two-Component System of *Staphylococcus aureus*. *Genes (Basel)*. **7**, 81 (2016).
35. Jaskólska, M., Stutzmann, S., Stoudmann, C. & Blokesch, M. QstR-dependent regulation of natural competence and type VI secretion in *Vibrio cholerae*. *Nucleic Acids Res.* (2018). doi:10.1093/nar/gky717
36. Yarwood, J. M., McCormick, J. K. & Schlievert, P. M. Identification of a Novel Two-Component Regulatory System That Acts in Global Regulation of Virulence Factors of *Staphylococcus aureus*. *J. Bacteriol.* **183**, 1113–1123 (2001).
37. Schlag, S. *et al.* Characterization of the Oxygen-Responsive NreABC Regulon of *Staphylococcus aureus*. *J. Bacteriol.* **190**, 7847–7858 (2008).
38. Sun, F. *et al.* AirSR, a [2Fe-2S] Cluster-Containing Two-Component System, Mediates Global Oxygen Sensing and Redox Signaling in *Staphylococcus aureus*. *J. Am. Chem. Soc.* **134**, 305–314 (2012).
39. Albano, M., Hahn, J. & Dubnau, D. Expression of competence genes in *Bacillus subtilis*. *J Bacteriol* **169**, 3110–3117 (1987).
40. Magnuson, R., Solomon, J. & Grossman, A. D. Biochemical and genetic characterization of a competence pheromone from *B. subtilis*. *Cell* **77**, 207–216 (1994).
41. Lazazzera, B. A., Solomon, J. M. & Grossman, A. D. An Exported Peptide Functions Intracellularly to Contribute to Cell Density Signaling in *B. subtilis*. *Cell* **89**, 917–925 (1997).
42. Briley Jr., K., Prepiak, P., Dias, M. J., Hahn, J. & Dubnau, D. Maf acts downstream of ComGA to arrest cell division in competent cells of *B. subtilis*. *Mol Microbiol* **81**, 23–39 (2011).
43. Claverys, J. P. & Havarstein, L. S. Cannibalism and fratricide: mechanisms and reasons d’être. *Nat Rev Microbiol* **5**, 219–229 (2007).
44. Claverys, J. P., Martin, B. & Havarstein, L. S. Competence-induced fratricide in streptococci. *Mol Microbiol* **64**, 1423–1433 (2007).

**Chapter 3: Natural development of competence in *S. aureus*  
under anaerobe conditions**



### 3.1. Introduction

Molecular oxygen ( $O_2$ ) is critical for *S. aureus* growth both *in vitro* and in the host tissues <sup>1</sup>. *In vivo*,  $O_2$  varies by tissue sites but also during the course of infection in a given site <sup>2</sup>. For example, the arterial blood  $O_2$  content is 68-95 mmHg, while venous blood contains 40 mmHg  $O_2$ . The skin possesses a wide range of  $O_2$  concentrations depending on the depth from the surface (8-35 mm Hg). The intestinal lumen is completely anaerobic and contains <2 mmHg, whereas critical organs such as the kidneys and liver contain relatively high levels of  $O_2$  (50-72 mmHg and 30-40 mmHg, respectively). In addition, during infection, rapid recruitment of energy-consuming immune cells and induced inflammation can locally reduce  $O_2$  levels dramatically. In the case of osteomyelitis (infection of the bones), infection is characterized by a low oxygen biofilm formation <sup>1</sup>. Moreover, *S. aureus* can induce hypoxia even in tissues that have relatively high levels of  $O_2$ , like the kidneys, leading to the formation of  $O_2$ -restricted environments, such as abscesses, from which *S. aureus* can disseminate and seed a variety of vital organs <sup>3</sup>. Taken together, these observations suggest that *S. aureus* promotes hypoxia in tissues, which is a key signal for biofilm formation and enhanced virulence.

In order to grow by oxidative phosphorylation, dissimilatory nitrate reduction or mixed-acid fermentation when  $O_2$  is limited or absent, *S. aureus* employs a multitude of regulatory  $O_2$ -sensing systems <sup>4</sup>. Among these are found three two-component transduction systems (TCSs): NreBC, SrrAB and AirSR. A brief overview of the three TCSs has been summarized below.

NreBC is a TCS involved in the direct sensing of  $O_2$ . The sensor histidine kinase NreB is a cytoplasmic protein containing four conserved cysteine residues that together comprise a Fe-S cluster <sup>5</sup>. The presence of  $O_2$  renders the NreBC TCS inactive due to oxidation of the Fe-S cluster, while the absence of  $O_2$  leads to the reduction of the Fe-S cluster, causing NreB dimerization and phosphorylation of the NreC response regulator. Once activated, NreC activates the nitrite and nitrate reductase genes, forcing the bacterium to up-regulate the fermentation pathways in order to survive.

Then, the staphylococcal respiratory response AB (*srrAB*) TCS is critical for anaerobic growth of *S. aureus in vitro* <sup>6</sup>. In the case of SrrAB,  $O_2$  is indirectly sensed. Although the ligand responsible for its activation is currently unknown, menaquinone has been proposed as the most likely candidate. *In vitro* studies have shown that under anaerobic conditions, SrrAB controls the up-regulation of fermentation enzymes, TCA cycle enzymes and the repression of the *agr* (accessory gene regulator) system, a prototype of quorum-sensing regulator. As a result, inhibition of RNAIII (the *agr* effector) is associated to the repression of many virulence genes, under anaerobic conditions.

Finally, the AirSR (anaerobic iron-sulphur cluster-containing redox sensor regulator, also known as YhcSR) TCS is involved in the regulation of nitrate respiration under anaerobic conditions <sup>7</sup>. Interestingly, AirSR regulates the expression of NreBC. Indeed, the response regulator AirR is found to



directly bind to the promoter regions of *narG* and *nreCBA*, thereby also altering *nreCBA* regulon <sup>8</sup>. AirS is a sensor histidine kinase consisting of an Fe-S cluster coordinated by four cystine residues. It responds to oxidation signals such as O<sub>2</sub> itself, reactive oxygen species (ROS) and reactive nitrogen species (RNS). Oxidized AirS shows maximal enzymatic kinase activity and subsequently phosphorylates AirR, while both reduced or overoxidized AirS loses this activity <sup>9</sup>. As shown by microarray analysis, AirSR considerably alters *S. aureus* gene regulation under anaerobic growth by affecting the expression and regulation of more than 350 genes including that of the *agr* system, the *saeRS* TCS and stress-associated factors (*rsbU* and *rsbW*) <sup>10</sup>.

Importantly, the original publication demonstrating the ability of *S. aureus* to be naturally transformable already showed that natural competence could be induced under different O<sub>2</sub> availability <sup>11</sup>. Indeed, the authors developed two protocols, and two media, named CS2 and GS, to induce competence under aerobic and anaerobic conditions respectively. In addition, we previously showed that the so-called aerobic condition, when *S. aureus* is grown in CS2 medium, (Results: Chapter 2), might not be fully aerobic and competence is potentially partially controlled by the decrease of O<sub>2</sub> during the transition into stationary phase, through the SrrAB TCS. Finally, more recently, Morikawa and his colleagues have shown that competence for natural transformation could also be induced when *S. aureus* forms biofilms, an environmental adaptation associated to low oxygen <sup>12</sup>. All these results clearly indicate that O<sub>2</sub> availability regulates the development of competence in *S. aureus* but considering all the different conditions described above these regulations must be complex and integrated at several levels.

Furthermore, three potential central competence regulators: SigH, ComK1 and ComK2 have been identified in *S. aureus*. SigH, an alternative sigma factor, is homologous to ComX in *S. pneumoniae*; ComK1 and ComK2, two transcriptional activators, are homologous to ComK in *B. subtilis*. It is therefore legitimate to question why several competence regulators have been selected in *S. aureus*? One interesting hypothesis could be that several regulators are needed to answer and integrate a large set of different environmental stimuli and that oxygen availability could be integrated at different levels to control the development of competence in *S. aureus*.

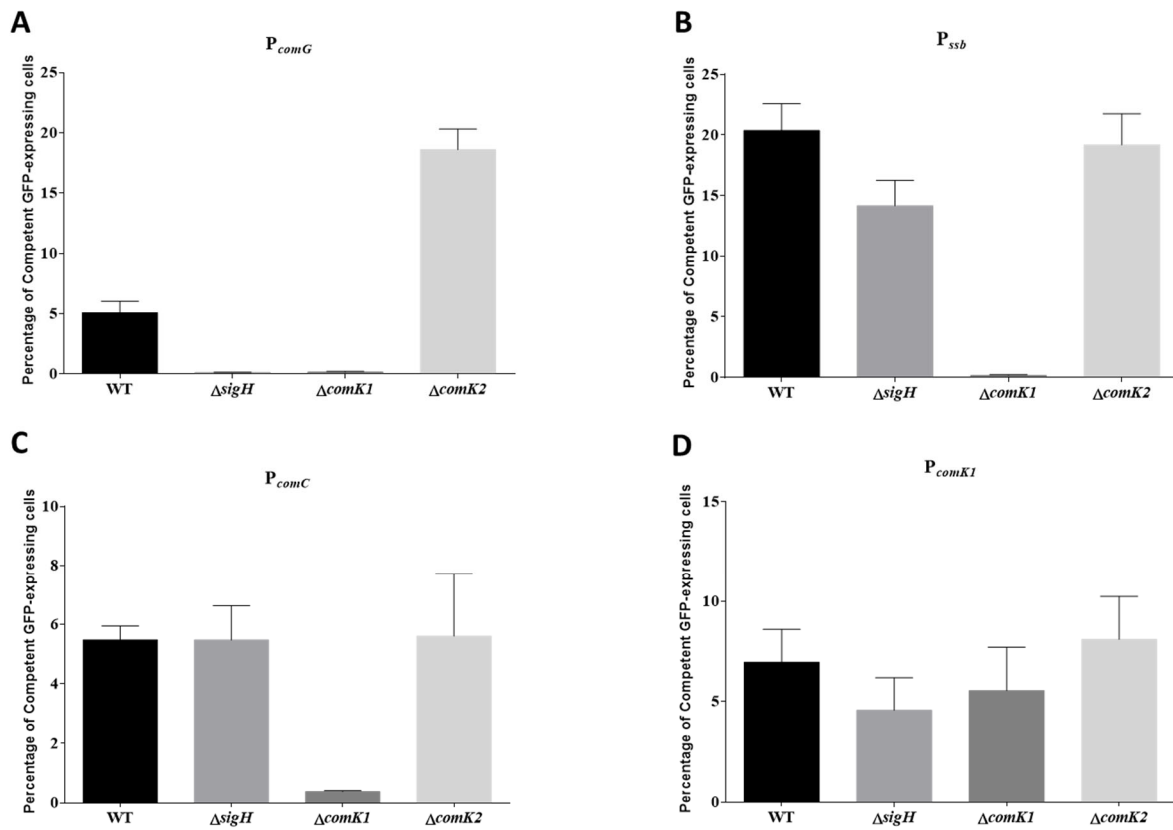
This is why, in the present study, we decided to investigate how the absence of O<sub>2</sub> could affect the development of competence when *S. aureus* is grown under anaerobic conditions in GS medium <sup>11</sup>. To monitor the development of competence, we constructed transcriptional fusions where the gene encoding the GFP is fused to the promoter of late competence genes involved in genetic transformation. The percentage of GFP-expressing competent cells was then compared in wild type and mutant strains in which the gene encoding important central competence regulators (SigH, ComK1





and ComK2), oxygen sensing systems (NreC, SrrA and AirR) or proteins of interest (SA2107) were deleted.

Interestingly, we showed that, as in CS2 medium, SigH and ComK1 activate the expression of the genetic transformation genes during competence development in GS medium. Furthermore, we show that, under anaerobic conditions, ComK2 is also involved in the development of competence, but as an inhibitor of some genetic transformation genes expression. We also proposed that this inhibition activity could be linked to an interaction between ComK2 and a protein with no-predicted function, SA2107. Then, we demonstrated how the lack of O<sub>2</sub> in GS medium inhibits the development of competence through the NreBC TCS. Finally, we gathered evidence suggesting that the NreBC oxygen-sensing TCS and the ComK2 central competence regulator could operate in the same pathway to inhibit competence development in response to low O<sub>2</sub> concentration.



**Figure 1. SigH, ComK1 and ComK2 regulate the natural genetic transformation genes expression.**

Percentage of the population expressing GFP under control of the promoters of *comG* (A), *comC* (B), *ssb* (C) and *comK1* (D) was compared in a wild type (St29, St50, St48 and St94) and *sigH* (St51, St61, St60, St110), *comK1* (St40, St64, St63, St97), or *comK2* (St41, St67, St66, St100) mutant strains. The measurements were realized after 48 hours of growth in GS medium without shaking under anaerobic condition. All the experiments were repeated at least three times.

## 3.2. Results

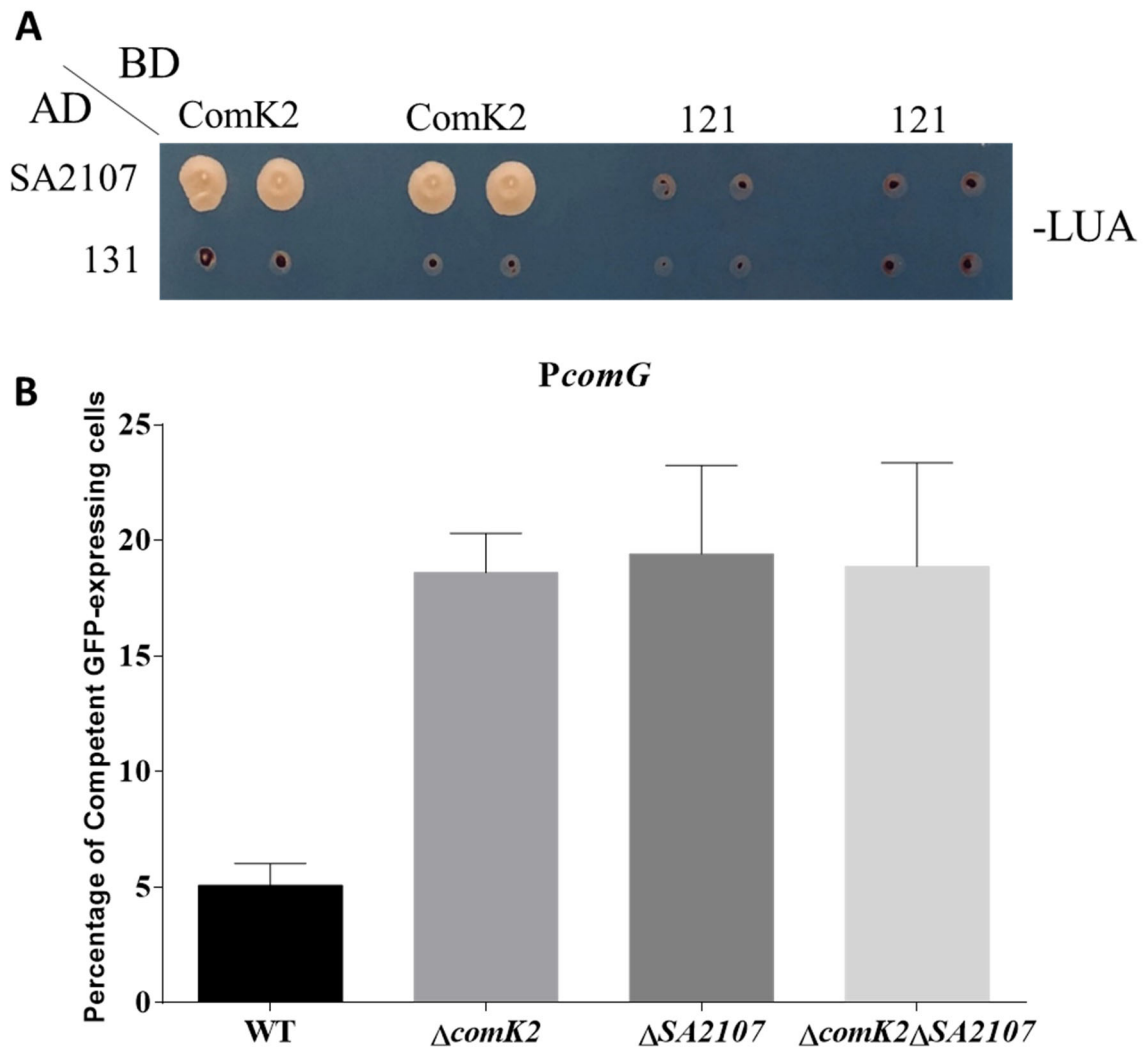
### 3.2.1. Anaerobic conditions can induce the expression of *comK1*, *comGA*, *comC* and *ssb*

First, in order to verify the validity of the anaerobic conditions published by Morikawa *et. al.* [11](#), we attempted to reproduce their protocol by measuring the expression of GFP under the control of the *comGA* promoter throughout a culture of a wild type strain (the same strain as published in [11](#)) in GS for 24 hours. Interestingly, 10 % of the cells were found to express GFP under the control of the *comGA* promoter (**Fig. 1A**), confirming that *S. aureus* is able to induce competence under anaerobic conditions. Furthermore, given the fact that, during competence, all the late competence genes are supposed to be induced, we analyzed other transcriptional fusions. We chose the *ssb* gene to represent the genes only controlled by ComK1, and *comC* to represent the genes that could not be previously detected (see [13](#)). Under these same conditions, roughly 20 % of the cells could express GFP from the *ssb* promoter (**Fig. 1C**), and 6 % of cells from *comC* promoter (**Fig. 1B**). Finally, we found that roughly 6 % of the cells could express GFP from the *comK1* promoter (**Fig. 1D**). Taken together, these results allowed us to confirm that anaerobic conditions are favorable for *S. aureus* to induce competence through the expression of at least one master competence regulator and some late competence genes involved in genetic transformation.

### 3.2.2. ComK2 inhibits the development of competence under anaerobic conditions

As mentioned above, three potential central competence regulators have been identified in *S. aureus*. We then aimed to investigate the role of each regulator under anaerobic conditions, by studying the natural development of competence in single central competence regulator knock-outs.

We first focused on *comGA* expression. *comGA* encodes a traffic ATPase found in association with the transformosome complex that allows the binding and transport of exogenous DNA across the cell wall. Orthologues can be found in all the transformable species and *comGA* is considered as a key late competence gene involved in genetic transformation. Interestingly, deletion of *sigH* or *comK1* completely abolished the expression of GFP under the control of the *comGA* promoter under anaerobic conditions (**Fig. 1A**). This result indicates that, as in CS2 medium (see Results, Chapter 2), both SigH and ComK1 are essential for *comGA* expression in GS medium. However, in the absence of *comK2*, the expression of GFP from the *comGA* promoter was increased by almost a 4-fold factor (**Fig. 1A**). Under anaerobic conditions, ComK2 is therefore considered as an inhibitor of the expression of *comGA*, a new role that was not identified in CS2 medium.



**Figure 2. SA2107, a protein of unknown function interacting with ComK2, is involved in the regulation of competence development.**

A. Y2H experiment showing a physical interaction between SA2107 and ComK2. Binary interaction was revealed by growth of diploid cells on selective medium lacking adenine (-LUA). 121 : yeast strain containing empty pGBDU. 131 : yeast strain containing empty pGAD.

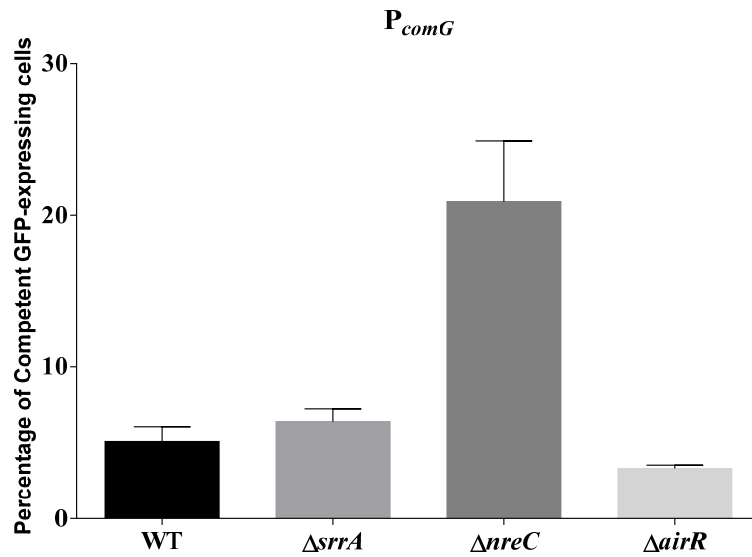
B. Comparison of the percentage of *S. aureus* cells expressing GFP under the control of *comG* promoter in a wild type (St29), *comK2* (St41), *sa2107* (St52) and *comK2/sa2107* (St161) mutant strains. The measurements were realized after 48 hours of growth in GS medium without shaking under anaerobic condition. All the experiments were repeated at least three times.

We then investigated the role of the central competence regulators on the expression of two other late competence genes, *ssb* and *comC*. Interestingly, only the absence of *comK1* drastically affected the expression from the *ssb* and *comC* promoters (Fig. 1B&C), even though the absence of *sigH* had a very small effect on *ssb* expression (Fig. 1B), confirming again the data obtained in CS2 medium (see Results, Chapter 2). Finally, no central competence regulator was found essential for the control of the *comK1* expression under anaerobic growth, even though again the absence of *sigH* displayed a small effect (Figure 1.D).

### 3.2.3. SA2107 interacts with ComK2 and inhibits the development of competence under anaerobic conditions

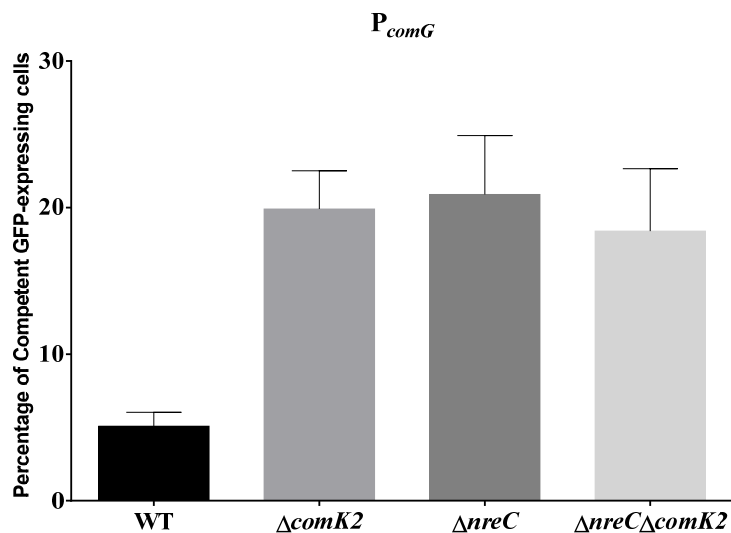
As we just identified a new role for ComK2, which is able to inhibit the expression of the *comG* operon promoter, we wondered if this new function would require a regulatory partner. To determine if we could identify new competence regulators, interacting with ComK2, we screened a *S. aureus* genomic library using the yeast two-hybrid (Y2H) method. Interestingly, we found that a protein, with unknown function, encoded by the gene *sa2107*, could interact with ComK2 in yeast (Fig. 2A).

Considering that a transcriptional regulator interacting partner could modulate its activity, we then wondered if SA2107 could affect the expression of some late competence genes in GS medium. Given the fact that ComK2 can inhibit the expression of *comGA*, we then tested the impact of the deletion of *sa2107*, alone or in combination with the deletion of *comK2*. Interestingly, deletion of *sa2107* also increased the percentage of cells expressing the GFP under the control of the *comGA* promoter (Fig. 2B). To test if the interaction between ComK2 and SA2107 was important for these phenotypes, we finally investigated the effect of the simultaneous absence of *comK2* and *sa2107*. Indeed, if the two proteins work in separate pathways to inhibit the expression of *comGA*, the phenotypes associated to each protein should be cumulative. At the opposite, if the two proteins work together, probably through a direct protein-protein interaction, the phenotype of the double mutant strain should be identical to single mutants. As expected for interacting partners, the double mutant strain *comK2/sa2107*, displayed the same increase in the percentage of cells expressing GFP under the control of the *comGA* promoter as the single mutant strains (Fig. 2B). Therefore, we can hypothesize that ComK2, in interaction with the SA2107 protein, inhibits the expression of *comGA* under anaerobic conditions.



**Figure 3. Different O<sub>2</sub>-detection TCSs impacts on competence development under various O<sub>2</sub> concentrations.**

Comparison of the percentage of *S. aureus* cells expressing GFP under the control of *comG* promoter in a wild type (St29), *srrA* (St145), *nreC* (St58) and *air* (St177) mutant strains. The measurements were realized after 48 hours of growth in GS medium without shaking under anaerobic condition. All the experiments were repeated at least three times.



**Figure 4. ComK2 and NreC function in the same pathway to inhibit *comG* expression.**

Comparison of the percentage of *S. aureus* cells expressing GFP under the control of *comG* promoter in a wild type (St29), *comK2* (St41), *nreC* (St158) and *comK2/nreC* (St273) mutant strains. The measurements were realized after 48 hours of growth in GS medium without shaking under anaerobic condition. All the experiments were repeated at least three times.

### 3.2.4. NreC inhibits *comGA* expression under anaerobic conditions

As explained above, in the absence of O<sub>2</sub> *S. aureus* can alter its central metabolism by switching to fermentation or by using alternative terminal electron acceptor, such as nitrate <sup>10</sup>. The ability of *S. aureus* to adapt to changes in oxygen availability and to regulate the expression of genes involved in this transition from aerobic to anaerobic growth are controlled by three TCSs, named SrrAB, NreCBA and AirRS (Figure 4). To determine if these three systems are involved in the development of competence under anaerobic conditions, we analyzed the expression of *comGA* in the absence of *srrA*, *nreC* or *airR*, respectively encoding for the response regulator of these three TCSs.

First, deletion of *srrA* or *airR* had no significant impact on the *comGA* expression (Fig. 3). This result is interesting if compared to the data obtained in CS2 (see Results: Chapter 2). Indeed, while deletion of *srrA* inhibited the expression from the *comGA* promoter in CS2 medium, it had no impact in GS. This result suggest that even though we hypothesized that the oxygen concentration would drop while reaching stationary phase in CS2 medium, we do not reach the anaerobic environment encountered in GS.

On the opposite, deletion of *nreC* increased the percentage of cells expressing the GFP under anaerobic conditions by a 4-fold factor (Fig. 3). This result indicates that NreC normally inhibits the expression of *comGA* in the absence of O<sub>2</sub>.

### 3.2.5. NreC inhibits *comGA* expression through ComK2 under anaerobic conditions

As the phenotypes associated to the individual deletion of *comK2* and *nreC* seem identical, we hypothesized that these regulators could operate in the same pathway. Similarly to what we have done previously, we combined in the same strain the deletion of these two genes and compared the expression of *comGA* in this double-mutant strain to that in the single mutant strains. If the double mutant strain displays the same phenotype as the single mutant strains, it strongly suggests that the regulators work in the same pathways while if the phenotypes are cumulative, it suggests that the regulators function independently.

Interestingly, the double *comK2/nreC* double mutant strain induced the *comG* expression in the same percentage of cells as the single *comK2* and *nreC* mutants (Fig. 4). Therefore, we can postulate that under anaerobic growth in GS medium *S. aureus* cells sense low O<sub>2</sub> concentration through the NreBC TCS, which in turn activate the ComK2 central competence regulator that is able to inhibit the expression from the *comG* promoter. To further complete this model, we are currently trying to evaluate the development of competence in *sa2107/nreC* and *comK2/2107/nreC* mutant strains (work in progress, Fig. 4)



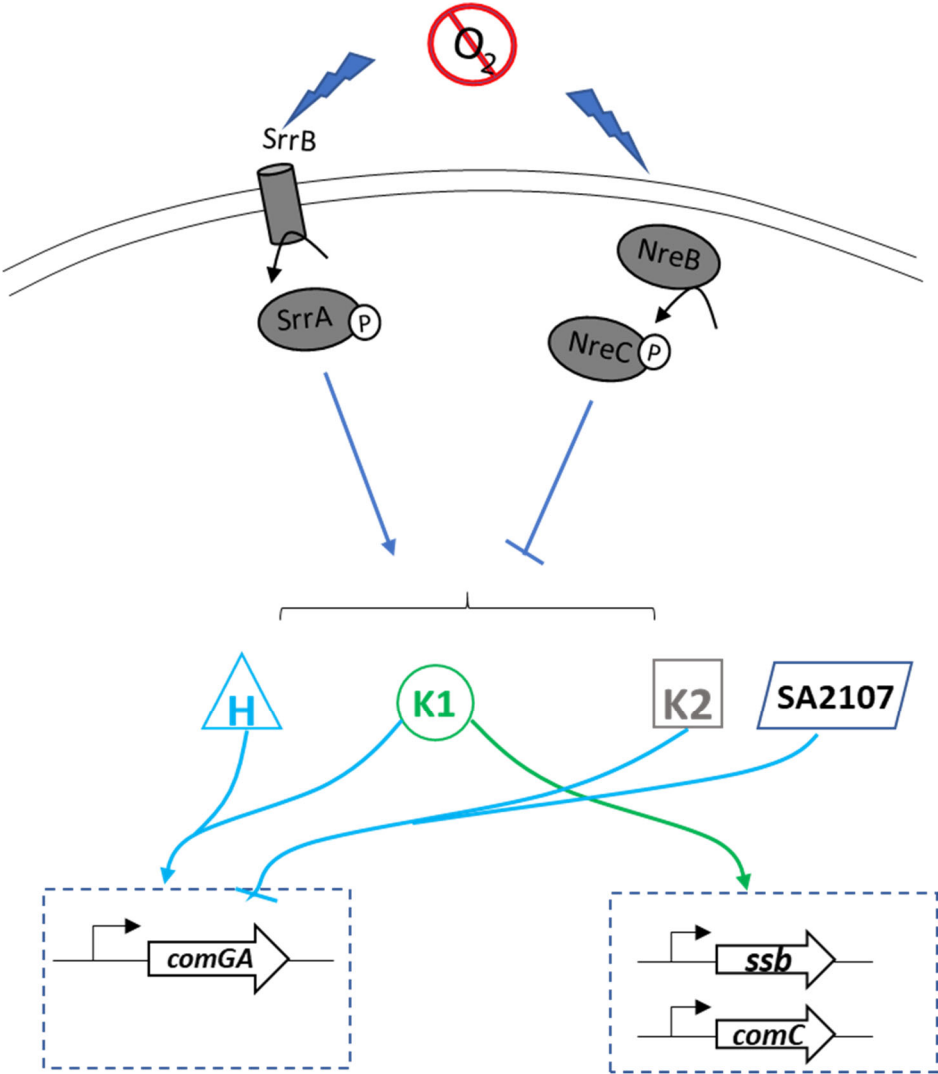


Figure 5. Suggested model of competence development regulation in *S. aureus*.

### 3.3. Discussion

#### **The three central competence regulators control the expression of the genetic transformation genes.**

In Chapter 2, we concluded that only SigH and ComK1 were essential for the expression of the genetic transformation genes in CS2 medium. In parallel, we also showed that ComK2 was also important during competence by controlling the expression of other late competence genes, with complementary functions.

However, in this third Chapter, we showed that while SigH and ComK1 are still essential, ComK2 also controls the expression of some genetic transformation genes, in GS medium under anaerobic conditions. Importantly, while SigH and ComK1 are essential transcriptional activators, in both CS2 and GS medium, ComK2 can be considered as a transcriptional inhibitor, in GS medium. It is important to mention that we still do not know if this ComK2 effect on the *comG* promoter' transcription is direct or indirect, through intermediate regulatory steps.

Inhibition of competence has already been reported in the literature. Usually, such inhibitory mechanisms are concentrated on the expression or stability of central competence regulators. This might be also the case for ComK2, which could indirectly inhibit the expression of the *comG* operon through the modulation of another central competence regulator. As we have shown that deletion of *comK2* does not inhibit *comK1* and *ssb* expression, this hypothetical pathway might rather implicate SigH. Alternatively, ComK2 might be a direct transcriptional inhibitor of the *comG* operon expression. This second hypothesis is very seducing as ComK2 would be, to our knowledge, the first of its kind in *S. aureus*. The simple demonstration of ComK2 ability to directly bind to the *comG* promoter would clearly add some interest to this last hypothesis.

#### **ComK2/SA2107 interaction might be important for the late competence genes inhibition.**

Interestingly, we have also showed that the ComK2 transcriptional inhibitory activity is closely linked to a protein of unknown function, SA2107, which directly interacts with ComK2 in Y2H experiments. Unfortunately, we could not identify any conserved domain or structural motif enabling us to understand how SA2107 might work. Deeper characterization of this protein will be important in the future to be able to decipher this aspect of competence regulation under anaerobic conditions.

Additional regulatory pathways, modulating the activity of central competence regulators have already been shown in other model organisms. We can cite for example:

- The ComK transcriptional regulator stability, in *B. subtilis*, is modulated through multiple



protein-protein interactions [14](#).

- The stability and activity of the ComX alternative sigma factor in *S. pneumoniae* is ensured through protein-protein interaction with the ComW early competence protein [15](#).
- The QstR central competence regulator activity in *V. cholerae* requires a competence-induced cofactor (not identified yet) [16](#).

In some cases, such regulations, through protein-protein interactions, have been considered to be essential for the central competence regulator activity. However, we can also hypothesize that these additional regulation pathways can be used to modulate the activity of these regulators, only under specific conditions and in response to alternative stimuli. Therefore, further investigations will be required to understand how the ComK2/SA2107 interaction is important to inhibit the late competence genes, and how their expression and/or activity is regulated by anaerobic conditions.

### **Oxygen is an essential, but complex, signal for the development of competence in *S.***

#### ***aureus***

I have clearly established, throughout my different PhD projects, that oxygen plays a key role in the general regulation of competence in *S. aureus*. First, I have shown that an oxygen-associated signal can be integrated through multiple pathways, each transmitting the signal to a different central competence regulator. Indeed, a decrease in the oxygen concentration upon entry in stationary phase in CS2 medium (micro-aerobic environment?) might activate the development of competence through the SrrAB TCS which in turn would promote the expression of SigH/ComK1 (Chapter 2). In parallel, under anaerobic conditions, the lack of oxygen is probably sensed by the NreBC TCS which inhibits the development of competence through ComK2 (**Fig 5**).

Identifying the oxygen concentration as an important signal controlling the development of competence is not surprising for an organism with such adaptative powers. *S. aureus* ability to grow under a wide range of oxygen and the presence of multiple regulatory pathways sensing these conditions provide the ideal setup to control an environmental adaptation as important as natural competence for genetic transformation.

Importantly, oxygen has already been proposed as a signal controlling the development of competence in *S. pneumoniae* [17](#). Indeed, different studies have shown how aerobic conditions promote the development of competence while micro-aerobic environments inhibit it. These different oxygen availabilities are all integrated through the Nox NADH reductase.

It appears that a similar general control of competence by the oxygen availability is employed by *S. aureus*. However, all my results show that these regulations are multiple, complex and integrated by



several pathways and actors. Additional experiments will be required to correlate growth under different protocols, competence induction through the transcriptional regulation of early and late competence genes and oxygen availability thanks to precise measurements of oxygen concentration throughout growth.



## 3.4. Materials and Methods

### 1. Bacterial strains and culture conditions

*S. aureus* strains used in this project are listed in Table S2. Staphylococcal strains were grown in BHI medium or a complete synthetic medium (called GS) depending on the experiment. The components of GS medium are shown in Table S3.

For static anaerobic growth in GS medium, after the preculture was ready, cells were collected from a certain amount of preculture and resuspended in 1 mL of GS medium to ensure that OD was 0.5. Then the cell suspension was serially diluted in GS medium to  $10^{-1}$ ,  $10^{-2}$  and  $10^{-3}$ , fully filled in 2-mL Eppendorf tube and incubated at 37 °C without shaking.

### 2. Construction of GFP reporter plasmids and strains

Oligonucleotides used in this study are listed Table S1. The promoters and pRIT plasmids were amplified with the dedicated primers. Thanks to these primers, there is a 25-50 bp homologous overlap regions between obtained fragments of promoters and linear plasmid. Subsequent to PCR with Phusion High-Fidelity DNA Polymerase (purchased from Thermo Scientific), each of the PCR fragments were purified using standard commercial silicon column (PCR cleanup kit). Each of the fragments to be joined were then electrophoresed to verify the absence of non-specific or minor PCR fragments.

The promoters were inserted and circularized in pRIT plasmid using the Gibson assembly method. Gibson assembly master mix was prepared by adding 320  $\mu$ L of 5x ISO buffer, 0.64  $\mu$ L 10U/ $\mu$ L T5 exonuclease, 20  $\mu$ L of 2U/ $\mu$ L Phusion polymerase, 160  $\mu$ L of 40U/ $\mu$ L Taq DNA ligase and 699.36  $\mu$ L of water. All reagents were purchased from New England Biolabs. 5  $\mu$ L of mixture of the two DNA fragments containing 100 ng of linear pRIT plasmid and 3-fold of inserts were added to 15  $\mu$ L of Gibson assembly master mix. The tubes were therefore incubated at 50 °C for 1 hr. Finally, 1  $\mu$ L of the assembly reaction was transformed into IM08B competent cells. Transformed cells were incubated at 37 °C on LB agar with 100  $\mu$ g/mL ampicillin.

In order to select the transformants containing the expected plasmid, colony PCRs were performed on each transformant. The resulting plasmid was extracted from overnight culture of the positive transformant and purified by commercial kit (NucleoSpin Plasmid extraction kit). All the plasmids were verified by sequencing (GATC company). Finally, in order to obtain the reporter stains, *S. aureus* electrocompetent cells were transformed with the constructed plasmids.



### 3. Construction of *S. aureus* deletion mutants

In order to investigate the role of the main genes predicted to be involved in the regulation of competence development in *S. aureus*, listed in Table 4, allelic replacement constructs were cloned into pIMAY, which is a plasmid vector for allelic exchange with a temperature sensitive replicon. All the primers used for genetic manipulation in the current study are listed in Table 3. Fragments corresponding to 1 kbp flanking regions of the genes to be deleted were amplified from *S. aureus* N315 genomic DNA using the primers flanked by restriction enzyme sites chosen from multiple cloning site (MCS) of pIMAY, thus generating the upstream and downstream of the genes. The upstream or downstream regions were digested by two chosen restriction enzymes and ligated into pIMAY cut with the same enzymes. The resulting plasmids were transformed into *E. coli* IM08B. Colony PCR was used to verify the transformants. Then, plasmids were purified from the positive transformants. After verifying the plasmids sequences, *S. aureus* N315ex  $\omega\phi$  strain was transformed by electroporation then plated on BHI agar with 10 mg/mL chloramphenicol and incubated at 28 °C.

In order to allow pIMAY integration into the chromosome, a single colony from the transformation plate was resuspended in 200  $\mu$ L of BHI. The suspension was 10-fold diluted down to  $10^{-3}$  and 100  $\mu$ L of each dilution was spread on BHI with 10 mg/mL chloramphenicol and incubated at 37 °C overnight. Large colonies were subsequently streaked in the same conditions. Meanwhile colony PCR analysis was performed to check the absence of extrachromosomal pIMAY and whether plasmid integration had occurred upstream or downstream.

Based on the colony PCR results, an overnight culture in BHI at 28 °C without chloramphenicol was made from the positive clones. Then, the overnight culture was 10-fold diluted down to  $10^{-7}$ . 100  $\mu$ L of  $10^{-4}$  to  $10^{-7}$  dilutions were plated onto BHI containing 1  $\mu$ g/mL anhydrotetracycline (aTc). The plates were therefore incubated at 28 °C for 2-3 days. The induction of *secY* antisense RNA produced from pIMAY would prevent growth of cells that retain the integrated plasmid and select for cells that have lost the plasmid [18](#). Large colonies were patched on BHI without antibiotic and BHI with 10 mg/mL chloramphenicol and grown at 37 °C overnight. Chloramphenicol-sensitive colonies were screened by colony PCR with oligonucleotides to identify clones containing the desired mutation. Putative mutants were verified by PCR and DNA sequencing.

### 4. Electroporation in *E. coli*

*E. coli* IM08B was streaked on LB plate and incubated at 37 °C overnight. Three colonies were resuspended in 20 mL of SLB and the cell suspension was diluted 10-fold and 100-fold then grown in 20 mL of SLB overnight at 37 °C with shaking at 180 rpm. The next day, 10 mL of the overnight culture with cell density between 4 and 6 were added into 1 L of SLB and incubated at 37 °C, 180 rpm until OD reached 0.8-0.9. The culture was immediately chilled in ice water for 30 min. All subsequent steps were performed at 4 °C. Cells were harvested by centrifugation at 4500 g for 15 min. Pellets were resuspended in an equal volume of sterile cold water and centrifuged at 4500 g for 15 min, then resuspended in half the volume of sterile cold water and centrifuged a second time at 4500 g for 15 min. Cells were resuspended again in 20 mL of sterile cold 10% glycerol (w/v) and centrifuged 4500xg for 15 min and finally resuspended in 3 mL of cold 10% glycerol. Aliquots of 100 µL were frozen and stocked at -80 °C. For electroporation, electrocompetent cells were thawed on ice for 3 min, then 50 µL of cells was mixed with 1 µL of ligation product (or Gibson assembly product), transferred into 2-mm electroporation cuvette (pre-cold on ice) and, immediately, pulsed at 2.5 kV for 5 ms. The cells were immediately recovered in 1 mL of SOC and incubated at 37 °C for 1 hr. Finally, 100 µL of transformation were plated on LB with the appropriate antibiotic depending on plasmids.

### **5. Electroporation in *S. aureus***

*S. aureus* strains were isolated on BHI plate at 37 °C. An overnight culture with 3 colonies was prepared in 10 mL of BHI and incubated at 37 °C with shaking at 180 rpm, and then diluted 200-fold in 200 mL of BHI. The culture was incubated at 37 °C with shaking at 180 rpm until the OD was between 0.6 and 0.8. Then, the culture was chilled in ice for 30 min. Cells were harvested by centrifugation at 4500 g for 15 min and resuspended in an equal volume of sterile cold 0.5 M sucrose. After successive centrifugations at 4500 g for 15 min, pellets were resuspended in half the volume of sterile 0.5 M sucrose, then in 20 mL of sterile cold 0.5 M sucrose with 10% glycerol, and finally in 1 mL of 0.5 M sucrose and 10% glycerol. Aliquots of 100 µL each were frozen at -80 °C. For electroporation, electrocompetent cells were thawed at room temperature for 10 min, heated at 56 °C for 2 min, and then immediately resuspended in 500 µL of 0.5 M sucrose with 10% glycerol. Cell suspension was centrifuged at 8000 rpm for 1 min and resuspended in 50 µL of 0.5 M sucrose and 10% glycerol. 5 µL of plasmid (1-2 µg) was mixed with the cells, transferred into 2 mm electroporation cuvette at room temperature and pulsed at 3 kV, 1 ms. The cells were immediately resuspended in 1 mL of BHI supplemented with 0.5% glucose and incubated at 37 °C for 1 hr or 28 °C for 3 hr.

### **Flow cytometry to determine the percentage of the population expressing GFP**

### Chapter 3

Subsequent to growth in CS<sub>2</sub>, 500  $\mu$ L of the cells was harvested by centrifugation at 11000 g for 1 min. Pellets were resuspended in 500  $\mu$ L of cold 70% ethanol and incubated on ice for 20 min. Then, the cells were resuspended in 500  $\mu$ L of PBS after centrifugation at 11000 g for 1 min. Finally, the population with GFP signals was detected by Flow cytometry.

## Supplementary data

Table S1. Primers

Name	Nucleotide sequence	Descriptions
IM151	TAC ATG TCA AGA ATA AAC TGC CAA AGC	Used for verifying and sequencing the allelic replacement pIMAY constructs
IM152	AAT ACC TGT GAC GGA AGA TCA CTT CG	
KpnI-sigH KO-F-F	GCA GGT ACC GAC CCG CAT AAC TTG GGA TCA ATT TTA AGA	To amplify upstream of <i>sigH</i> used for the deletion of <i>sigH</i>
Sall-sigH KO-F-R	GCA GTC GAC CCC CTT CTA TCT AAA ATT TAA GGT TAG TTT AAT ATT GTT ACA TTC	
EcoRI-sigH KO-R-F	GCA GAA TTC AGC GCC TTA GGA CGT GAA TTG AAT TAT AAC GTG	To amplify downstream of <i>sigH</i> used for the deletion of <i>sigH</i>
XmaI-sigH KO-R-R	GCA CCC GGG CTG GTG TTT CTC GGC CAA ACA TAT CTA CTA ATA C	
sigH-OUT-F	TGC AGC GCT TAT TGC ACC ATA TGA ATA TGC	To verify the <i>sigH</i> deletion mutant
sigH-OUT-R	TTG CCC TCC CAC TCT TAA CAT TTC GTG A	
Kpn I-SA2017-up-F	GCA GGT ACC CAT ATG ACA CTC CCA ATG C	To amplify upstream of <i>sa2107</i> used for the deletion of <i>sa2107</i>
Sal I-SA2017-up-R	GCA GTC GAC GTA GAA GTA CCT CCA AAA ATC AAT	
EcoRI-SA2107-down-F	GCA GAA TTC GCC CAT TAA CCT ATT TTT CAT A	To amplify downstream of <i>sa2107</i> used for the deletion of <i>sa2107</i>
Sac I-SA2107-down-R	GCA GAG CTC ACA ATA TTT GAT GCC TGT GCT A	
check-mutant-SA2107-F	TGA AAG TCA GTC GTA CTC GAC AT	To verify the <i>sa2107</i> deletion mutant
check-mutant-SA2107-R	GTT GCT CCC ATA TGC ATC TCA	
Kpn I-K1-up-F	GGT ACC GAA TGT TTA GTC ATG GTA CGT TGA TG	To amplify upstream of <i>comK1</i> used for the deletion of <i>comK1</i>
Sal I-K1-up-R	GCA GTC GAC AAG CAA AAC CTC GCT TTA TTA AGT TT	
Sma I-K1-down-F	GCA CCC GGG TAT GCG GAA TGT TTT ATA	To amplify downstream of <i>comK1</i> used for the deletion of <i>comK1</i>
Sac I-K1-down-R	GCA GAG CTC AAG GTT TGA TGT TAT CAG TGA ATG	
Kpn I-K2-up-F	GCA GGT ACC ATC TAT GCA TTG TTA CAT ATT CAT	To amplify upstream of <i>comK2</i> used for the deletion of <i>comK2</i>
Sal I-RK2-up-	GCA GTC GAC ATA TAG TAC TCC TCG TAT AAT AAG	
EcoR I-K2-down-F	GCA GAA TTC GAA GCA CTT CAT TGA AAA TAC	To amplify downstream of <i>comK2</i> used for the deletion of <i>comK2</i>
Sac I-K2-down-R	GCA GAG CTC AAG TGT CAG ATT CGT GTA AT	
check-mutant-comK1-F	TCT CAG GGA ATG CTA TGG TTA AAG T	To verify the <i>comK1</i> deletion mutant
check-mutant-comK1-R	CAG TGG CTT ATT GGT CAC AGG	
check-mutant-comK2-F	GAG ACA ACA GCA AAT ATG ACA ACA AG	To verify the <i>comK2</i> deletion mutant
check-mutant-comK2-R	GTG CAT GAA TAT TAC CAC TTC TAA TCG	
Kpn I-up- <i>agrA</i> -F	GCA GGTACC TTA GTG ACC ATG ATC ATA ATG TAT TTG AGT	To amplify upstream of <i>agrA</i> used for the deletion of <i>agrA</i>
Sal I-up- <i>agrA</i> -R	GCA GTCGAC TTC ACA TCC TTA TGG CTA GTT GTT AA	
EcoR I-down- <i>agrA</i> -F	GCA GAATTC AAT AAG ATA ATA AAG TCA GTT AAC GGC GTA	To amplify downstream of <i>agrA</i> used for the deletion of <i>agrA</i>
Sac I-down- <i>agrA</i> -R	GCA GAGCTC TCT GCT GAT ATG TTA TTT GAA CCA AG	
Check-mutant- <i>agrA</i> -F	GTA TTA CTT CTA TGG AAG TAG AGC CGT ATT	To verify the <i>agrA</i> deletion mutant
Check-mutant- <i>agrA</i> -R	GTG TAT CGC ACG AAT GAA GCA	
Kpn I-up <i>srrA</i> -F	GCA GGT ACC GAC ATT GAT TGC CAA AGG ACT AGT TGA G	To amplify upstream of <i>srrA</i> used for the deletion of <i>srrA</i>
EcoR I-up <i>srrA</i> -R	GCA GAA TTC TAC CTC CCA CAC ATG CTT TTC TTT ACA	
<i>sma</i> I-down <i>srrA</i> -F	GCA CCC GGG GGG CGT TGG GTA TAA ATT TGA GGT TAA ATC TA	To amplify downstream of <i>srrA</i> used for the deletion of <i>srrA</i>
Sac I-down <i>srrA</i> -R	GCA GAG CTC GCT CAT TAG TCA TAT CAC GAA CTG TCA C	
Check-mutant- <i>srrA</i> -F	GAA ATT ATC ACA AGC AGC AAT GGA AGT ACT	To verify the <i>srrA</i> deletion mutant
Check-mutant- <i>srrA</i> -R	TGG TAT TGT TAC AGA ACC GGA TGA AAT AAA AG	
Sal I-up <i>nreC</i> -F	GCA GTC GAC GAG CAT TCA ATC GAA AAG ATA GTT TTT GCT GAT G	To amplify upstream of <i>nreC</i> used for the deletion of <i>nreC</i>
<i>Sma</i> I-up <i>nreC</i> -R	GCA CCC GGG TCA AAT TGG AAT GTT CAA TGT AAC ATT GGT AC	
Not I-down <i>nreC</i> -F	GCA GCG GCC GC CGA GAG GAA ATG ATT ATC TTC TGG C	

## Chapter 3

Sac I-down <i>nreC</i> -R	GCA GAG CTC GGTGACTGAATTTGGCATAAC	To amplify downstream of <i>nreC</i> used for the deletion of <i>nreC</i>
Check-mutant- <i>nreC</i> -F	GAT GAA TTA GGG TGT AAG TCA TGA TTA ATG AGG AC	To verify the <i>nreC</i> deletion mutant
Check-mutant- <i>nreC</i> -R	CTATTGAAGTTGCTACAACACTTCCAGCAC	
Sal I-up <i>airR</i> -F	GCA GTC GAC CAA GGT GCG TTA AAA TAT TTA ATT GAG GGC	To amplify upstream of <i>airR</i> used for the deletion of <i>airR</i>
EcoR I-up <i>airR</i> -R	GCA GAA TTC CAT GGG TTA TCT CCT TAA ATC AAG CTA TT	
<i>sma</i> I-down <i>airR</i> -F	GCA CCC GGG GTG TAA GTT TTA TCA AAT CTA GTG ATT GTG	To amplify downstream of <i>airR</i> used for the deletion of <i>airR</i>
Sac I-down <i>airR</i> -R	GCA GAG CTC GGT GAA ACA GCA TTC TAA TTG CAA TTA	
Check-mutant- <i>airR</i> -F	GGA GCG ATT TGA TGG AAC AAA GG	To verify the <i>airR</i> deletion mutant
Check-mutant- <i>airR</i> -R	GGC CAA TTT TAG TTG CAA TTT CTT CGC	
Gibson-pRIT- <i>gfp</i> -F	GGA TCC GGG AGG CCG TTT	To amplify the linear pRIT used for the reporter plasmid construction by Gibson
Gibson-pRIT- <i>gfp</i> -R	CTC TCA GTA CAA TCT GCT CTG ATG CCG	
pRIT-P <i>comF</i> -F	ATA TGT GTT CGT GAA TTC CAG TTT GAT CAT AAT TCA GTG TTA CTA TAC ATG GTA CTG	To amplify the promoter of <i>comF</i> used for the reporter plasmid constructoin by Gibson
<i>gfp</i> -P <i>comF</i> -R	AAA CGG CCT CCC GGA TCC TGT TAT TTT ATA TCT TGT TAC ATT ATC CAT TCG ACC CAG TGA TAT	
pRIT-P <i>ssb</i> -F	CGG CAT CAG AGC AGA TTG TAC TGA GAG TGC GGT AAT GCC AGG CGT ATT AAT TAC TG	To amplify the promoter of <i>ssb</i> used for the reporter plasmid construction by Gibson
<i>gfp</i> -P <i>ssb</i> -R	AAA CGG CCT CCC GGA TCC TCT CCC GAC AAT TAC GAT TTT ATT TAG CAT GTA GTG TG	
pRIT-P <i>comC</i> -F	CGG CAT CAG AGC AGA TTG TAC TGA GAG AAG CCT AAC GTT CCA GTG ATA TAT GCT GTT AAA AAT G	To amplify the promoter of <i>comC</i> used for the reporter plasmid construction by Gibson
<i>gfp</i> -P <i>comC</i> -R	AAA CGG CCT CCC GGA TCC TGC AGC TAT AAG ATA ACA ATA CTA CCA AAT GAC AAC CT	
pRIT-P <i>comK1</i> -F	CGG CAT CAG AGC AGA TTG TAC TGA GAG GTA ACG ACA AAT ATA GAA CTT GGG CAT GGA	To amplify the promoter of <i>comK1</i> used for the reporter plasmid constructoin by Gibson
<i>gfp</i> -P <i>comK1</i> -R	AAA CGG CCT CCC GGA TCC AAG CAA AAC CTC GCT TTA TTA AGT TTT AAA CC	
pRIT-P <i>comK2</i> -F	CGG CAT CAG AGC AGA TTG TAC TGA GAG GAC CAT CAC GAT ACA TCA TTT CAT TAG TGA	To amplify the promoter of <i>comK2</i> used for the reporter plasmid construction by Gibson
<i>gfp</i> -P <i>comK2</i> -R	AAA CGG CCT CCC GGA TCC ATA TAG TAC TCC TCG TAT AAT AAG TTG TTA ATT TAA TGG	
pRIT-P <i>sigH</i> -F	CGG CAT CAG AGC AGA TTG TAC TGA GAG AAC AGG TGC AAT TGA ACA TGT ACC AGT TA	To amplify the promoter of <i>sigH</i> used for the reporter plasmid construction by Gibson
<i>gfp</i> -P <i>sigH</i> -R	AAA CGG CCT CCC GGA TCC GTA TTA AAC TAA CCC CTT CTA TCT AAA ATT TAA GGT TAG	
pRIT-X-F	GGC TTA ACT ATG CGG CAT CAG AGC	Used for verifying and sequencing the pRIT constructs
pRIT-X-R	GAA TTG TGA GCG GAT AAC AAT TTC ACA CAG G	

Table S2 Strains

Strains	Genotype/Construction	Source
St012	N315ex woφ	Morikawa et. al, 2012
St029	N315ex woφ / pRIT-PcomG-gfp	Morikawa et. al, 2012
St048	N315ex woφ / pRIT-PcomC-gfp	This study
St050	<i>N315ex woφ / pRIT-Pssb-gfp</i>	This study
St094	<i>N315ex woφ / pRIT-PcomK1-GFP</i>	This study
St095	<i>N315ex woφ / pRIT-PcomK2-GFP</i>	This study
St096	<i>N315ex woφ / pRIT-PsigH-GFP</i>	This study
St129	N315ex woφ / pRIT-PcomE-GFP	This study
St233	N315ex woφ / pRIT-PcomF-gfp	This study
St037	<i>N315ex woφ ΔcomK1</i>	This study
St040	<i>ΔcomK1 / pRIT-PcomG-gfp</i>	This study
St063	<i>ΔcomK1 / pRIT-PcomC-gfp</i>	This study
St064	<i>ΔcomK1 / pRIT-Pssb-gfp</i>	This study
St097	<i>ΔcomK1 / pRIT-PcomK1-GFP</i>	This study
St038	<i>N315ex woφ ΔcomK2</i>	This study
St041	<i>ΔcomK2 / pRIT-PcomG-gfp</i>	This study
St066	<i>ΔcomK2 / pRIT-PcomC-gfp</i>	This study
St067	<i>ΔcomK2 / pRIT-Pssb-gfp</i>	This study
St100	<i>ΔcomK2 / pRIT-PcomK1-GFP</i>	This study
St045	<i>N315ex woφ ΔsigH</i>	This study
St051	<i>ΔsigH / pRIT-PcomG-gfp</i>	This study
St060	<i>ΔsigH / pRIT-PcomC-gfp</i>	This study
St061	<i>ΔsigH / pRIT-Pssb-gfp</i>	This study
St110	<i>ΔsigH / pRIT-PcomK1-GFP</i>	This study
St044	<i>N315ex woφ ΔcomK1+ΔcomK2</i>	This study
St047	<i>ΔcomK1+ΔcomK2 / pRIT-PcomG-gfp</i>	This study
St074	<i>ΔcomK1+ΔcomK2 / pRIT-PcomC-gfp</i>	This study
St075	<i>ΔcomK1+ΔcomK2 / pRIT-Pssb-gfp</i>	This study
St055	<i>N315ex woφ ΔcomK1+ΔsigH</i>	This study
St081	<i>ΔcomK1+ΔsigH / pRIT-PcomG-gfp</i>	This study
St082	<i>ΔcomK1+ΔsigH / pRIT-PcomC-gfp</i>	This study
St083	<i>ΔcomK1+ΔsigH / pRIT-Pssb-gfp</i>	This study
St053	<i>N315ex woφ ΔcomK2+ΔsigH</i>	This study
St077	<i>ΔcomK2+ΔsigH / pRIT-PcomG-gfp</i>	This study
St078	<i>ΔcomK2+ΔsigH / pRIT-PcomC-gfp</i>	This study
St079	<i>ΔcomK2+ΔsigH / pRIT-Pssb-gfp</i>	This study
St054	<i>N315ex woφ ΔcomK1+ΔcomK2+ΔsigH</i>	This study
St069	<i>ΔcomK1+ΔcomK2+ΔsigH / pRIT-PcomG-gfp</i>	This study
St070	<i>ΔcomK1+ΔcomK2+ΔsigH / pRIT-PcomC-gfp</i>	This study
St071	<i>ΔcomK1+ΔcomK2+ΔsigH / pRIT-Pssb-gfp</i>	This study
St039	N315ex woφ ΔSA2107	This study
St042	ΔSA2107/ pRIT-PcomG-gfp	This study
St046	N315ex woφ ΔcomK2+SA2017	This study

### Chapter 3

St161	$\Delta$ comK2+SA2017/pRIT-PcomG-gfp	This study
St117	N315ex wo $\phi$ $\Delta$ srrA	This study
St145	$\Delta$ srrA / pRIT-PcomG-gfp	This study
St118	N315ex wo $\phi$ $\Delta$ nreC	This study
St158	$\Delta$ nreC / pRIT-PcomG-gfp	This study
St142	N315ex wo $\phi$ $\Delta$ airR	This study
St177	$\Delta$ airR / pRIT-PcomG-gfp	This study
St272	N315 ex.wo $\phi$ $\Delta$ nreC+ $\Delta$ comK2	This study
St273	$\Delta$ nreC+ $\Delta$ comK2/pRIT-PcomG-GFP	This study

Table S3. Composition of GS medium

Ingredient	Amount	Ingredient	Amount
Glucose	5 g	Amino acids	
Salt		L-glutamic acid	100 mg
K <sub>2</sub> HPO <sub>4</sub>	7 g	L-methionine	3 mg
KH <sub>2</sub> PO <sub>4</sub>	2 g	L-tyrosine	50 mg
Na <sub>3</sub> citrate·2H <sub>2</sub> O	0.4 g	L-alanine	60 mg
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	1 g	L-lysine	50 mg
MgSO <sub>4</sub> ·7H <sub>2</sub> O	0.05 g	L-threonine	30 mg
Vitamins		L-phenylalanine	40 mg
Thiamine	1 mg	L-histidine	20 mg
Niacin	1.2 mg	L-tryptophan	10 mg
Biotin	0.005 mg	L-isoleucine	30 mg
D Pantothenate	0.25 mg	L-valine	80 mg
Purines and pyrimidines		L-leucine	90 mg
Adenine	5 mg	L-aspartic acid	90 mg
Guanine	5 mg	L-arginine	50 mg
Cytosine	5 mg	L-proline	80 mg
Uracil	5 mg	L-cystine	20 mg
Thymine	20 mg		
		H <sub>2</sub> O	1000 mL

### 3.5. References

1. Balasubramanian, D., et al., *Staphylococcus aureus* pathogenesis in diverse host environments. *Pathogens and Disease*, 2017. **75**(1).
2. Carreau, A., et al., *Why is the partial oxygen pressure of human tissues a crucial parameter? Small molecules and hypoxia*. *Journal of cellular and molecular medicine*, 2011. **15**(6): p. 1239-1253.
3. Rubinstein, E., *Staphylococcus aureus* bacteraemia with known sources. *International journal of antimicrobial agents*, 2008. **32**: p. S18-S20.
4. Haag, A. and F. Bagnoli, *The Role of Two-Component Signal Transduction Systems in Staphylococcus aureus Virulence Regulation*. *Current topics in microbiology and immunology*, 2016. **409**.
5. Müllner, M., et al., *A PAS Domain with an Oxygen Labile [4Fe-4S]<sub>2</sub><sup>+</sup> Cluster in the Oxygen Sensor Kinase NreB of Staphylococcus carnosus*. *Biochemistry*, 2008. **47**(52): p. 13921-13932.
6. Ulrich, M., et al., *The staphylococcal respiratory response regulator SrrAB induces ica gene transcription and polysaccharide intercellular adhesin expression, protecting Staphylococcus aureus from neutrophil killing under anaerobic growth conditions*. *Molecular Microbiology*, 2007. **65**(5): p. 1276-1287.
7. Yan, M., et al., *The Essential Two-Component System YhcSR Is Involved in Regulation of the Nitrate Respiratory Pathway of Staphylococcus aureus*. *Journal of Bacteriology*, 2011. **193**(8): p. 1799-1805.
8. Yan, M., et al., *The essential two-component system YhcSR is involved in regulation of the nitrate respiratory pathway of Staphylococcus aureus*. *J Bacteriol*, 2011. **193**(8): p. 1799-805.
9. Sun, F., J. Chen, and M.-Q. Liu, *Connections between uniformity and aberration in general multi-level factorials*. *Metrika*, 2011. **73**(3): p. 305-315.
10. Haag, A.F. and F. Bagnoli, *The role of two-component signal transduction systems in Staphylococcus aureus virulence regulation*. *Staphylococcus aureus*, 2015: p. 145-198.
11. Kazuya Morikawa, A.J.T., Yumiko Inose, Melody Tsai, Le Thuy Nguyen Thi, Toshiko Ohta, Tarek Msadek, *Expression of a Cryptic Secondary Sigma Factor Gene Unveils Natural Competence for DNA Transformation in Staphylococcus aureus*. *PLoS Pathog*, 2012. **8**(11): p. e1003003.
12. Maree, M., et al., *MRSA emerges through natural transformation*. *bioRxiv*, 2020: p. 2020.12.08.415893.
13. Annette Fagerlund, P.E.G.a.L.S.H., *Staphylococcus aureus* competence genes: mapping of the SigH, ComK1 and ComK2 regulons by transcriptome sequencing. *Molecular Microbiology*, 2014. **94**(3): p. 557–579.
14. Turgay, K., et al., *Competence in Bacillus subtilis is controlled by regulated proteolysis of a transcription factor*. *The EMBO journal*, 1998. **17**(22): p. 6730-6738.
15. Piotrowski, A., P. Luo, and D.A. Morrison, *Competence for genetic transformation in Streptococcus pneumoniae: termination of activity of the alternative sigma factor ComX is independent of proteolysis of ComX and ComW*. *Journal of bacteriology*, 2009. **191**(10): p. 3359-3366.
16. Lo Scudato, M. and M. Blokesch, *A transcriptional regulator linking quorum sensing and chitin induction to render Vibrio cholerae naturally transformable*. *Nucleic acids research*, 2013. **41**(6): p. 3644-3658.



17. Auzat, I., et al., *The NADH oxidase of Streptococcus pneumoniae: its involvement in competence and virulence*. *Molecular microbiology*, 1999. **34**(5): p. 1018-1028.
18. Monk, I.R., et al., *Transforming the untransformable: application of direct transformation to manipulate genetically Staphylococcus aureus and Staphylococcus epidermidis*. *mBio*, 2012. **3**(2).

# Conclusions and Perspectives



## 1. The protein-to-protein interaction network centered on natural competence and genetic transformation

During the first year of my PhD, I generated a genomic library of *S. aureus* in yeast. During the next year, I screened this library, using the yeast two-hybrid technique, with different baits involved in the development of competence and natural transformation in *S. aureus*. In parallel, I also tested, in a “matrix” experiment, if the proteins chosen as baits could directly interact with each other. Ultimately, the interactions detected during the screenings and in the “matrix” experiments are all summarized in our protein-protein interaction network. Importantly, in the different experiments we conducted, some very interesting interactions were detected that could be classified as either:

- interactions already validated in other organisms (such as the ComGA-RecA interaction),
- interactions never seen before but making sense because they involve proteins working in the same pathway (such as the ComEA-ComEB interaction)
- completely new interactions with partners that were not expected to interact with each other before this study (such as the ComK2-SA2107 interaction)

Even though we detected some very interesting interactions, we feel a little bit disappointed with our “library screening” results. Indeed, the total number of potential interactions seems low in comparison to previous studies. Two different strategies could be conducted by the host laboratory in the future in order to improve our results:

1/ In comparison to previous libraries constructed (*B. subtilis* or *S. pneumoniae*, INRA, Jouy en Josas), the most abundant size of genomic fragments present in our libraries seem a little high (200 to 400bp in comparison to 100-200pb in previous libraries). Therefore, one solution could be to construct again our libraries by increasing the amount of restriction enzymes used in the partial digestions in order to reduce the average size of the fragments generated. This approach could produce fragments encoding smaller peptides that could be better to detect an interaction. As this size “issue” only is for now a hypothesis, this would have to be demonstrated.

2/ At this point, it is also important to mention that we have been collaborating with another team from the institute, directed by F.X. Barre, with which we have been working on a new protocol to generate genomic libraries. Prey banks are constructed based on molecular biology protocols associated with deep sequencing. In brief, genomic DNA is fragmented by sonication. Fragment ends are repaired by the same techniques used for Marker Frequency Analysis (MFA) or Chromosome Conformation Capture (3C). The genomic fragments are then cloned in our three vectors so that each



fragment is expressed in the right phase. The advantages are multiple:

- (i) we can generate prey banks with homogeneous fragment sizes thanks to the sonication parameters;
- (ii) genomic fragments generated are not limited by the presence of a restriction site and every base on the genome can be a starting or ending extremity of a genomic fragment multiplying the potential complexity of our libraries;
- (iii) the quality of the prey banks can be assessed at different steps during their creation using deep sequencing to analyze the representativity of the repaired fragments, of the genomic library resulting from their cloning in *E. coli* and after the transfer of the library in yeast.

Finally, many expected interactions, because already described in the literature, were not confirmed in our study, neither in the screening nor in the “matrix” experiments [188](#). Several explanations could be proposed to elucidate this:

- (i) The fragments generated by partial digestion did not encode for peptides that could reveal an important interaction
- (ii) Homologous proteins from different species, fused to the GAL4 activating domain, do not adopt the exact same tridimensional conformation, revealing or not an interaction
- (iii) The expected interaction does not exist *in vivo* in the new model organism considered

To verify which case scenario explains the non-detection of an expected interaction, we will have to test again the interaction in new libraries in yeast (see above) or directly *in vivo*, in *S. aureus*, through alternative techniques such as microscopy.

## **2. Regulation of natural competence-mediated horizontal gene transfer in *S. aureus***

In Chapter 2 and 3 of this thesis manuscript, I investigated the regulation of the natural development of competence in the human pathogen *S. aureus*.

First, we developed in the team different protocols to optimize the induction of natural competence and genetic transformation in *S. aureus*. These protocols were key in my projects as they provided an important induction of competence to perform genetic studies: indeed, the percentage of cells naturally inducing competence in a wild type strain was now sufficient to be compared with the percentage in mutant strains (as a reminder the higher percentage of competent cells ever published barely reached 2%, [58](#)).

Then, I particularly showed that in CS2 medium the three potential central competence regulators identified in *S. aureus* (i.e. SigH, ComK1 and ComK2) are all important for a full development of comp-



-ence. Indeed, I found that SigH and ComK1 are both essential for the expression of the genetic transformation late competence genes. It appears that they control different set of genes, with some overlap. Meanwhile, ComK2 is also required for the expression of additional late competence genes involved in stress response, toxin/antitoxin systems and amino acid and nucleic acid metabolism. I also showed how SigH partially controls the expression of *comK1*, which must also be controlled by at least one additional pathway.

Finally, I revealed how different environmental conditions (i.e. CS2 or GS media), associated to different oxygen concentrations, can activate or inhibit the development of competence through different central competence regulators (this last part will be further discussed in the next paragraph).

## 2.1 *In vivo* conditions and the complex competence regulatory network

All the results compiled in this manuscript clearly show how complex the development of natural competence is in *S. aureus*. This complexity is often found in important model organisms such as *B. subtilis* or *V. cholerae* [216, 217](#). Multiple pathways, integrating multiple environmental signals, regulate a single or several central competence regulators. As already discussed in Chapter 2, the multiplicity of signals/pathways can be used to restrict the development of competence to specific stressful environmental conditions. Interestingly, all the conditions and regulations characterized *in vitro*, could be transposed *in vivo* while *S. aureus* interacts with its human host:

- After epithelial breach, *S. aureus* invades organs and tissues where it forms encapsulated abscesses. Once an abscess is established, bacterial proliferation occurs [218](#). Importantly, high bacterial density means that nutrients become scarce, a situation reminiscent of what happens during the entry in stationary phase *in vitro*. As we showed that cells inducing competence also repress the expression of many virulence genes (among which certain are necessary for the abscess formation and maintenance), we could imagine that some cells inside the abscess could differentiate, by repressing virulence and inducing competence.
- The ability for *S. aureus* to grow in the presence or absence of oxygen becomes particularly important at the onset of inflammation and recruitment of phagocytic immune cells which rapidly produce a hypoxic environment [219](#). Therefore, *S. aureus* could encounter situations, *in vivo*, during which the oxygen concentration drops, that would be suitable for the development of competence. Once again, some cells would have to differentiate from virulence to competence, once specific stimuli are sensed.
- Interestingly, the conditions listed above (high cell density, low oxygen concentration) are all found during biofilm formation. In addition, biofilms can form on tissue surfaces such as heart





valves in the case of endocarditis [220](#). Therefore it would be important to verify if, as it does *in vitro* [184](#), *S. aureus* can induce natural competence for genetic transformation *in vivo*.

To conclude, the conditions described *in vitro* could perfectly fit with the ecological niches and the physiology of *S. aureus* interacting with its host. Therefore, we can assume that there is a high probability that, *in vivo*, *S. aureus* induces competence to perform genetic transformation, promote genomic plasticity and acquire antibiotic resistance genes.

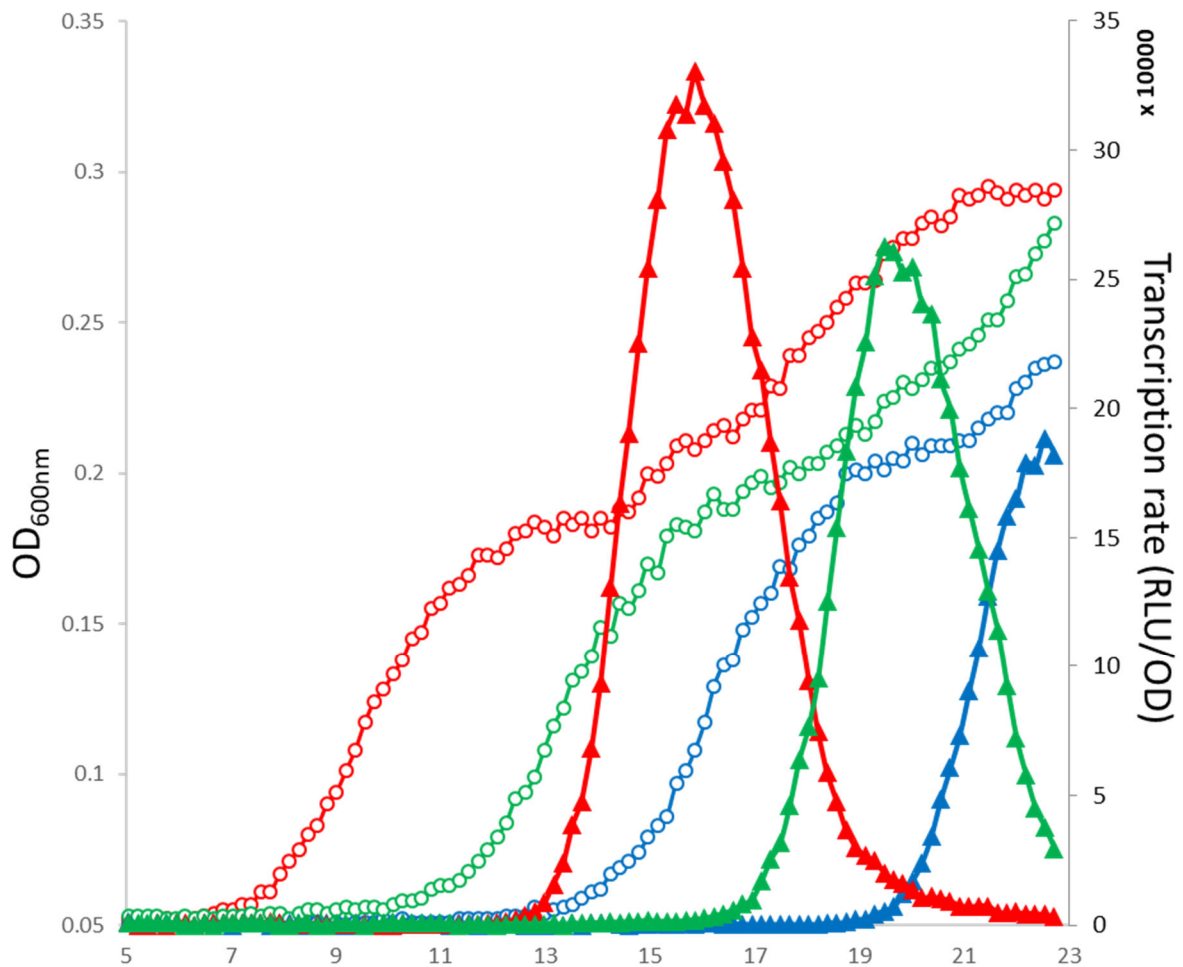
## 2.2 The luciferase as a new tool

An important part of the results generated during my PhD was obtained using transcriptional fusions between promoters of interest and the gene encoding the *gfp*. GFP as a transcriptional reporter has many advantages: (i) it can be easily detected by microscopy or flow cytometry, (ii) it provides a percentage of cells inducing competence, (iii) it is stable so that a measurement during an experiment gives access to the total number of cells that expressed a given gene at any moment, even if this expression is finished at the measurement time.

Unfortunately, this last point can also be considered as a drawback. Indeed, GFP stability makes it difficult to perform dynamic studies and identify when a gene is expressed and when it is not. The other method used to in our studies to analyze the competence transcriptional program, RNA-sequencing, could complete the GFP results as it gives access to the total cellular mRNA content at a given time. However, to investigate any transcriptional dynamic, RNA-sequencing would require the analysis of multiple samples, taken throughout the experiment, which would be a very expensive and time-consuming approach.

To address this problem, I initiated the use of a new reporter gene, the luciferase from *Photinus pyralis*, to study the dynamic development of natural competence. It is important to mention that this strategy was already used by [182](#), but as they overexpressed regulators to detect late competence genes expression, they did not have access to the competence natural dynamics. In comparison to the GFP, the luciferase combines several key characteristics: (i) a low stability (half-life around 6 min in *B. subtilis*, [221](#)) providing a measurement close from a rate of transcription, (ii) a relatively high signal/background ratio, allowing the detection of expressions that was impossible with the GFP, (iii) a method already used and validated in multiple model organisms for the characterization of the development of competence [221, 222](#).

Through the management of Master 1 and 2 students, we therefore constructed transcriptional fusions between the same promoters of interest as for the GFP, and the gene encoding the luciferase cloned



**Fig. 1. Growth and transcription from the *comG* operon promoter fused to luciferase in CS2 medium.**

The St249 *S. aureus* strain was grown in our plate reader, in a 96-well plate, in CS2 medium using the « dilution protocol ». Three dilutions are presented, -3 in red, -4 in green and -5 in blue. The growth curves for each dilution is represented by open circles while the plain triangles show the relative luminescence readings corrected for OD for the *comG* operon promoter. Measurements were realized (OD and RLU) every 10 minutes.

on a multi-copy plasmid (we tried to clone these fusions as a single copy on the chromosome but we could not detect any expression). The strains containing our different luciferase reporters were then cultured in 96-well plates, incubated in a plate reader able to control the temperature and the agitation.

Here, we show the example of the dynamic expression detected from the *comG* operon promoter, fused to the luciferase in a strain grown using the “dilution protocol” in our plate reader (Fig.1). The *comG* expression started to increase when the growth paused in the middle of the curve. The expression reached a maximum, 3 hours later. Strikingly, after this maximum was reached, the *comG* expression then clearly decreased to go back to a very low basal level 4 hours later, at the transition to stationary phase. Importantly, this dynamic expression defines a temporal window, that lasts around 7 hours, during which the competence transcriptional program is allowed. This transient induction of competence has been described in numerous model organisms [221](#), [222](#) and is a defining characteristic of this physiological adaptation.

Importantly, the luciferase should allow us to repeat some experiments already performed with the GFP, continue the investigation of competence development in *S. aureus* but also study the transient characteristic of the specific competence transcriptional program. For example, it will be important to investigate the competence shutoff in *S. aureus*, a highly regulated phase in other model organisms such as *S. pneumoniae* [146](#).

In parallel to the luciferase that should become the standard in the team, we will continue to use the GFP to get access to the percentage of cells inducing competence and the RNA-sequencing method to be able to measure the expression of all *S. aureus* genes in a single experiment.

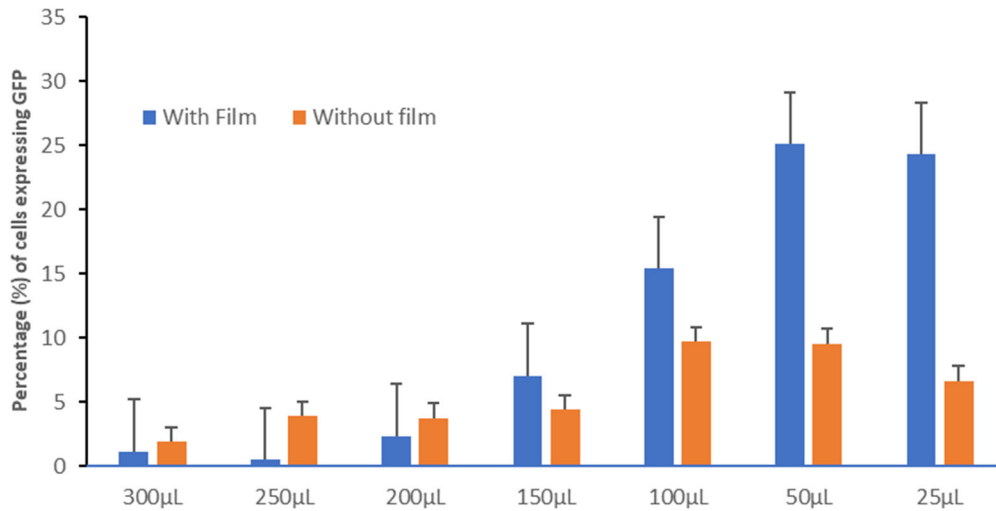
### 3. Oxygen availability and natural competence development in *S. aureus*

During my PhD, I showed in different projects how the development of competence in *S. aureus* can be influenced (activated or inhibited) by different growth conditions associated to different oxygen availability.

First, I proposed that during growth in CS2 (using the night protocol), the concentration in oxygen progressively decreases. This change in oxygen availability is sensed by the SrrAB TCS, which promotes (potentially through SigH) the development of competence.

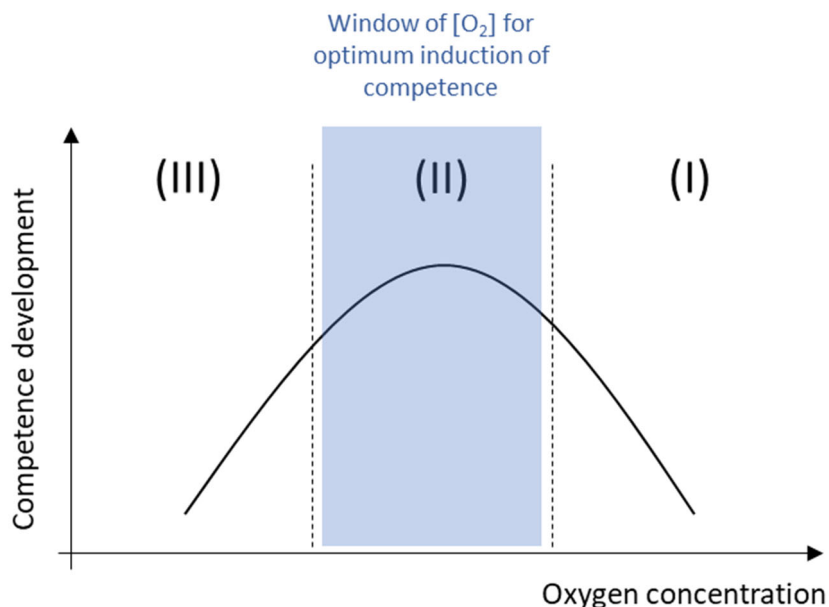
In parallel, I showed that under clear anaerobic conditions (i.e. static growth in GS medium), the lack of oxygen is sensed by a second TCS, NreBC, which inhibits competence, potentially through the third central competence regulator, ComK2.

Finally, colleagues in the team have performed experiments using *S. aureus* strains expressing the GFP



**Fig.2. Oxygen availability and development of competence.**

A strain harboring a PcomG-gfp transcriptional fusion (St29) was grown in our plate reader, in CS2 medium, using the day “protocol”. Different volumes of the culture were poured in the wells that were covered (in blue), or not (in orange) with a film preventing gas exchange. After 8h of growth, the cultures were stopped and the percentage of cells expressing GFP was measured by flow cytometry.



**Fig. 3. Model: how does oxygen influence competence development?**

as a transcriptional reporter and grown in 96-well plates inside our plate reader. This experimental setup allowed us to modulate the oxygen availability by (i) applying, or not, a film at the top of the wells preventing any gas exchange and (ii) by changing the volume of the culture in the well, which affects the volume of air (and oxygen) available, at least in the presence of the film (Fig. 2).

Interestingly, when the film was absent, the development of competence was low and almost not affected by the volume of the culture (a 3-fold ratio was observed between the lowest value: 3% of competent cells for 300 $\mu$ l of culture, and the highest value around 10% for 100 $\mu$ l of culture). Importantly, when the film was present, the development of competence was greatly influenced by the volume of the culture: wells with an important volume of culture (250-300  $\mu$ l) would poorly induce competence (1% of competent cells), while wells with a low volume (25 to 50 $\mu$ l) would fully induce competence (up to 25% of competent cells). These results clearly indicated that *S. aureus* probably consume a lot of oxygen during growth and that oxygen availability can modulate the development of competence, even for a wild-type strain.

We can therefore imagine a model (Fig 3) in which optimal environmental conditions for the development of competence require a specific oxygen concentration:

- (I) At high oxygen availability (well without film), the development of competence is inhibited by an unknown mechanism
- (II) When the culture in CS2 reaches stationary phase (closed falcon tube or well with film), the oxygen concentration starts to decrease and competence is induced
- (III) Anaerobic conditions (static culture in GS medium) inhibit the development of competence, potentially through the NreBC TCS and the ComK2 central regulator

As presented before (see introduction, paragraph 2.6.1.7), Morikawa's team has recently described a new protocol to efficiently induce competence during biofilm formation. Interestingly, during growth in biofilm structures, physiological heterogeneity happens due to the occurrence of oxygen and other nutrient gradients. These gradients are created because cells that are close to the surface of the biofilm consume obtainable nutrient sources and oxygen before they disperse into the depth of the biofilm [223](#). According to our model, the competent cells detected by Morikawa would be part of specific layers inside the biofilm where the oxygen concentration decreases enough but not too much in order to be in stage II (Fig. 3).

To test such general hypothesis, precise measurements of the oxygen concentration correlated with the development of the competence transcriptional program throughout growth in various environmental conditions will be essential. To do so, we propose to use a new system developed by



the Presens® company: indeed, they developed non-invasive optical oxygen sensors able to measure the partial pressure of both dissolved and gaseous oxygen. These sensors are fixed on the inner surface of the transparent glass or plastic material. The oxygen concentration can then be measured in a contact-less and non-destructive manner from outside, through the wall of the vessel.

Such equipment would be essential for our future experiments as it would allow us to measure the oxygen concentration throughout growth in CS<sub>2</sub> in closed falcon tubes, in GS in closed and static Eppendorf tubes and in both CS<sub>2</sub> and GS in filmed 96-well-plates. These oxygen measurements will also be compatible with GFP and Luciferase experiments.





# References

References

1. G., L., *Etymologia: Staphylococcus*. Emerging Infectious Diseases 2013. **19** (9): p. 1553.
2. C. Ryan, K.R., *Sherrie's medical microbiology*. McGraw Hill, 2004.
3. Lowy, F.D., *Medical progress - Staphylococcus aureus infections*. . New England Journal of Medicine 1998. **339**: p. 520–532
4. Cheung, G.Y.C., J.S. Bae, and M. Otto, *Pathogenicity and virulence of Staphylococcus aureus*. Virulence, 2021. **12**(1): p. 547-569.
5. Wertheim, H.F.L., et al., *Nose Picking and Nasal Carriage of Staphylococcus aureus*. Infection Control & Hospital Epidemiology, 2016. **27**(8): p. 863-867.
6. Sherertz, R.J., S. Bassetti, and B. Bassetti-Wyss, "*Cloud*" health-care workers. Emerging infectious diseases, 2001. **7**(2): p. 241-244.
7. Argudín, M.Á., M.C. Mendoza, and M.R. Rodicio, *Food Poisoning and Staphylococcus aureus Enterotoxins*. Toxins, 2010. **2**(7): p. 1751-1773.
8. Metzger, D.W. and K. Sun, *Immune dysfunction and bacterial coinfections following influenza*. Journal of immunology (Baltimore, Md. : 1950), 2013. **191**(5): p. 2047-2052.
9. Del Giudice, P., *Skin Infections Caused by Staphylococcus aureus*. Acta dermato-venereologica, 2020. **100**.
10. Balasubramanian, D., et al., *Staphylococcus aureus pathogenesis in diverse host environments*. Pathogens and Disease, 2017. **75**(1).
11. Archer, N.K., et al., *Staphylococcus aureus biofilms: properties, regulation, and roles in human disease*. Virulence, 2011. **2**(5): p. 445-459.
12. Cramton, S.E., et al., *The intercellular adhesion (ica) locus is present in Staphylococcus aureus and is required for biofilm formation*. Infect Immun, 1999. **67**(10): p. 5427-33.
13. Ulrich, M., et al., *The staphylococcal respiratory response regulator SrrAB induces ica gene transcription and polysaccharide intercellular adhesin expression, protecting Staphylococcus aureus from neutrophil killing under anaerobic growth conditions*. Mol Microbiol, 2007. **65**(5): p. 1276-87.
14. Otto, M., *Staphylococcal Infections: Mechanisms of Biofilm Maturation and Detachment as Critical Determinants of Pathogenicity*. Annual Review of Medicine, 2013. **64**(1): p. 175-188.
15. Archer, G.L., *Staphylococcus aureus: a well-armed pathogen*. Clin. Infect. Dis, 1998. **26**: p. 1179-1181.
16. Skinner D, K.C., *Significance of bacteremia caused by Staphylococcus aureus*. Arch Intern Med 1941. **68**: p. 851-75.
17. Jennifer M Junecko, A.K.Z., Lara N Mrak, Douglas C Ryan, Justin W Graham, Mark S Smeltzer, Chia Y Lee, *Transcribing virulence in Staphylococcus aureus*. World J Clin Infect Dis, 2012. **2**(4): p. 63-76.
18. Uchechukwu U. Nwodo, E.G., and Anthony I. Okoh, *Bacterial Exopolysaccharides: Functionality and Prospects*. Int J Mol Sci, 2012. **13**(11): p. 14002–14015.
19. Naglaa Mohamed, Y.T., Dorota Jamrozy, Eduardo Rojas, Li Hao, Natalie C. Silmon de Monerri, Julio Hawkins, Guy Singh, Bing Cai, Paul Liberator, Shite Sebastian, Robert G. K. Donald, Ingrid L. Scull, Annaliesa S. Anderson *Molecular epidemiology and expression of capsular polysaccharides in Staphylococcus aureus clinical isolates in the United States*. PLoS One, 2019. **14**(1): p. e0208356.
20. Gross, M., et al., *Key role of teichoic acid net charge in Staphylococcus aureus colonization of artificial surfaces*. Infection and immunity, 2001. **69**(5): p. 3423-3426.

## References

21. Sulica, A., et al., *Effect of protein A of Staphylococcus aureus on the binding of monomeric and polymeric IgG to Fc receptor-bearing cells*. Immunology, 1979. **38**(1): p. 173.
22. Claro, T.n., et al., *Staphylococcus aureus protein A binding to osteoblast tumour necrosis factor receptor 1 results in activation of nuclear factor kappa B and release of interleukin-6 in bone infection*. Microbiology, 2013. **159**(Pt\_1): p. 147-154.
23. Viela, F., et al., *Binding of Staphylococcus aureus protein A to von Willebrand factor is regulated by mechanical force*. MBio, 2019. **10**(2): p. e00555-19.
24. Speziale P, P.G., Rindi S, Provenzano M, Provenza G, Di Poto A, *Structural and functional role of Staphylococcus aureus surface components recognizing adhesive matrix molecules of the host*. Future Microbiol, 2009. **4**(10): p. 1337-1352.
25. Chavakis T, W.K., Preissner KT, Herrmann M, *Staphylococcus aureus interactions with the endothelium: the role of bacterial "secretable expanded repertoire adhesive molecules" (SERAM) in disturbing host defense systems*. Thrombosis and Haemostasis, 2005. **94**(2): p. 278–285.
26. Hoffmann C, O.K., Hauck CR, *integrin-mediated uptake of fibronectinbinding bacteria*. Eur J Cell Biol 2011. **90**: p. 891–896.
27. Herman-Bausier, P., et al., *Mechanical strength and inhibition of the Staphylococcus aureus collagen-binding protein Cna*. MBio, 2016. **7**(5): p. e01529-16.
28. Ganesh, V.K., et al., *A structural model of the Staphylococcus aureus ClfA–fibrinogen interaction opens new avenues for the design of anti-staphylococcal therapeutics*. PLoS pathogens, 2008. **4**(11): p. e1000226.
29. Otto, M., *Staphylococcus aureus toxins*. Curr Opin Microbiol, 2014. **0**: p. 32–37.
30. Grumann, D.N., U.; Bröker, B.M., *Staphylococcus aureus toxins—Their functions and genetics*. Infect. Genet. Evol. , 2014. **21**: p. 583–592.
31. Katayama, Y.B., T.; Sekine, M.; Fukuda, M.; Hiramatsu, K., *Beta-hemolysin promotes skin colonization by Staphylococcus aureus*. J. Bacteriol., 2013. **195**: p. 1194–1203.
32. Huseby, M.S., K.; Brown, C.K.; Digre, J.; Mengistu, F.; Seo, K.S.; Bohach, G.A.; Schlievert, P.M.; Ohlendorf, D.H.; Earhart, C.A., *Structure and biological activities of beta toxin from Staphylococcus aureus*. J. Bacteriol., 2007. **189**: p. 8719–8726.
33. Otto, M., *Phenol-soluble modulins*. International journal of medical microbiology : IJMM, 2014. **304**(2): p. 164-169.
34. Aman, M.J.K., H.; Bowden, M.G.; Nguyen, T.L., *Structural model of the pre-pore ring-like structure of Panton-Valentine leukocidin: Providing dimensionality to biophysical and mutational data*. J. Biomol. Struct. Dyn., 2010. **28**: p. 1–12.
35. Bera A, H.S., Jakob A, Vollmer W, Gotz F, *Why are pathogenic staphylococci so lysozyme resistant? The peptidoglycan O-acetyltransferase OatA is the major determinant for lysozyme resistance of Staphylococcus aureus*. Mol Microbiol 2005. **55**: p. 778–87.
36. Bokarewa MI, J.T., Tarkowski A., *Staphylococcus aureus: staphylokinase*. Int J Biochem Cell Biol, 2006. **38**(504–9).
37. Li M, L.Y., Villaruz AE, Cha DJ, Sturdevant DE, Otto M. , *Gram-positive threecomponent antimicrobial peptide-sensing system*. Proc Natl Acad Sci USA, 2007. **104**: p. 9469–74.
38. Villanueva, M., et al., *Sensory deprivation in Staphylococcus aureus*. Nature Communications, 2018. **9**(1): p. 523.
39. Jenul, C. and A.R. Horswill, *Regulation of Staphylococcus aureus Virulence*. Microbiology spectrum, 2019. **7**(2): p. 10.1128/microbiolspec.GPP3-0031-2018.

40. Haag, A.F. and F. Bagnoli, *The role of two-component signal transduction systems in Staphylococcus aureus virulence regulation*. *Staphylococcus aureus*, 2015: p. 145-198.
41. Mark Roberts, G.R., KormanecMaria, Elisabeth and Johanna Zalm, *The Role of Alternative Sigma Factors in Pathogen Virulence*. *Foodborne Pathogens* pp 2017: p. 229-303.
42. Recsei P, K.B., Oreilly M, Schlievert P, Gruss A, Novick RP, *Regulation of exoprotein gene expression in Staphylococcus aureus by agr*. *Mol Gen Genet* 1986. **202**(1): p. 58–61.
43. Burgess, R.R., *Encyclopedia of Genetics*, 2001.
44. Mark J. Kazmierczak, M.W., and Kathryn J. Boor, *Alternative Sigma Factors and Their Roles in Bacterial Virulence*. *Microbiol Mol Biol Rev*, 2005. **69**(4): p. 527-543.
45. Gross, T.M.G.a.C.A., *Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space*. *Annual Review of Microbiology*, 2003. **57**: p. 441-466
46. Misra, R.D.a.T.K., *Characterization of the Primary  $\sigma$  Factor of Staphylococcus aureus*. *Journal of Biological Chemistry*, 1996. **271**: p. 21828-21834.
47. Rajkrishna Mondal, T.G., Palas K Chanda, Amitava Bandhu, Jana Biswanath, Keya Sau, Chia Y Lee, and Subrata Sau, *Stabilization of the primary sigma factor of Staphylococcus aureus by core RNA polymerase*. *BMB Rep.*, 2010. **43**(3): p. 176-181.
48. Marco Palma, A.L.C.,  *$\zeta$ B Activity in Staphylococcus aureus Is Controlled by RsbU and an Additional Factor(s) during Bacterial Growth*. *J Bacteriol*, 2001. **69**(12): p. 7858-7865.
49. Willem van Schaik, M.H.T., Marcel H. Zwietering, Willem M. de Vos, and Tjakko Abee, *Analysis of the Role of RsbV, RsbW, and RsbY in Regulating  $\sigma$ B Activity in Bacillus cereus*. *JOURNAL OF BACTERIOLOGY*, 2005. **187**(16): p. 5846-5851.
50. Lauderdale KJ, B.B., Cheung AL, Horswill AR, *Interconnections between Sigma B, agr, and proteolytic activity in Staphylococcus aureus biofilm maturation*. *Infect Immun*, 2009. **77**(4): p. 1623-35.
51. Horsburgh MJ, A.J., White IJ, Shaw L, Lithgow JK, Foster SJ, *sigmaB modulates virulence determinant expression and stress resistance: characterization of a functional rsbU strain derived from Staphylococcus aureus 8325-4*. *J Bacteriol*, 2002. **184**(19): p. 5457-67.
52. M. Bischoff, J.M.E., P. Giachino, *Influence of a Functional sigB Operon on the Global Regulators sar and agr in Staphylococcus aureus*. *J Bacteriol*, 2001. **183**(17): p. 5171-5179.
53. Markus Bischoff, P.D., Jan Kormanec, Daphne Macapagal, Ellen Murphy, William Mounts, Brigitte Berger-Bächi, Steven Projan, *Microarray-Based Analysis of the Staphylococcus aureus  $\sigma$ B Regulon*. *J Bacteriol*, 2004. **186**(13): p. 4085-4099.
54. Novick, R.P., & Geisinger, E. , *Quorum sensing in staphylococci*. *Annual Review of Genetics*, 2008. **42**(1): p. 541-564.
55. Lindsey N. Shaw , C.L., Tomasz K. Prajsnar, Halie K. Miller, Melanie C. Brown, Ewa Golonka, George C. Stewart, Andrej Tarkowski, and Jan Potempa, *Identification and Characterization of  $\sigma$ S, a Novel Component of the Staphylococcus aureus Stress and Virulence Responses*. *PLoS One*, 2008. **3**(12): p. e3844.
56. Miller HK, C.R., Burda WN, Krute CN, Davenport JE, Shaw LN, *The extracytoplasmic function sigma factor  $\sigma$ S protects against both intracellular and extracytoplasmic stresses in Staphylococcus aureus*. *J Bacteriol*, 2012. **194**(16): p. 4342-54.
57. Burda WN, M.H., Krute CN, Leighton SL, Carroll RK, Shaw LN, *Investigating the genetic regulation of the ECF sigma factor  $\sigma$ S in Staphylococcus aureus*. *BMC Microbiol*, 2014. **30**(14): p. 280.

## References

58. Kazuya Morikawa, A.J.T., Yumiko Inose, Melody Tsai, Le Thuy Nguyen Thi, Toshiko Ohta, Tarek Msadek, *Expression of a Cryptic Secondary Sigma Factor Gene Unveils Natural Competence for DNA Transformation in Staphylococcus aureus*. PLoS Pathog, 2012. **8**(11): p. e1003003.
59. Weir, J., et al., *Regulation of spoOH, a gene coding for the Bacillus subtilis sigma H factor*. Journal of Bacteriology, 1991. **173**(2): p. 521-529.
60. Tao L, W.X., Sun B, *Alternative sigma factor  $\sigma$ H modulates prophage integration and excision in staphylococcus aureus*. PLoS Pathogens, 2010. **6**(5): p. 1-11.
61. YVONNE BENITO, F.A.K., PASCALE ROMBY , GERARD LINA *Probing the structure of RNAIII, the Staphylococcus aureus agr regulatory RNA, and identification of the RNA domain involved in repression of protein A expression*. EMBO J, 2000. **6**(5): p. 668-679.
62. Pichon C, F.B., *Small RNA genes expressed from Staphylococcus aureus genomic and pathogenicity islands with specific expression among pathogenic strains*. Proc Natl Acad Sci U S A, 2005. **102**(40): p. 14249-54.
63. Waters LS, S.G., *Regulatory RNAs in bacteria*. Cell, 2009 **136**(4): p. 615-28.
64. Liu JM, C.A., *A broadening world of bacterial small RNAs*. Curr Opin Microbiol, 2010. **13**(1): p. 18-23.
65. Felden B, V.F., Bouloc P, Romby P. , *The Staphylococcus aureus RNome and its commitment to virulence*. PLoS Pathog, 2011. **7**: p. e1002006.
66. Boisset S, G.T., Huntzinger E, Fechter P, Bendridi N, Possedko M, Chevalier C, Helfer AC, Benito Y, Jacquier A, Gaspin C, Vandenesch F, Romby P, *Staphylococcus aureus RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism*. Genes Dev, 2007. **21**(11): p. 1353-66.
67. Nina Sesto, O.W., Cristel Archambaud, Rotem Sorek and Pascale Cossa, *The excludon: a new concept in bacterial antisense RNA-mediated gene regulation*. Nature Reviews Microbiology, 2013. **11**(2): p. 75-82.
68. Novick RP, R.H., Projan SJ, Kornblum J, Kreiswirth B, Moghazeh S., *Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule*. EMBO J, 1993. **12**(10): p. 3967-75.
69. Lasa, I.e.a., *Genome-wide antisense transcription drives mRNA processing in bacteria*. Proc. Natl Acad. Sci. USA, 2011. **108**: p. 20172-20177
70. A., F., *On the antibacterial action of cultures of a Penicillium, with special reference to their use in the isolation of B. influenza*. British Journal of Experimental Pathology 1929. **10**(3): p. 226-236.
71. Chain E., e.a., *Penicillin as chemotherapeutic agent*. Lancet, 1940. **1**: p. 1172.
72. Abraham EP., e.a., *An enzyme from bacteria able to destroy penicillin*. Nature, 1940. **146**: p. 387.
73. M., J., *Celbenin-resistant staphylococci*. British Medical Journal 1961. **1**(5219 ): p. 124-12.
74. Casewell MW., e.a., *Epidemiology and control of the 'modern' methicillin-resistant Staphylococcus aureus*. Journal of Hospital Infection 1986. **7** p. 1-11.
75. Hiramatsu K., e.a., *Methicillin resistant Staphylococcus aureus clinical strain with reduced vancomycin susceptibility*. Journal of Antimicrobial Chemotherapy 1997. **40** p. 135-136.
76. Prevention., C.f.D.C.a., *Staphylococcus aureus resistant to vancomycin – United States*. Morbidity and Mortality Weekly Report MMWR 2002. **51** p. 565-567.

77. Khan MF., e.a., *Emergence of linezolid resistant Staphylococcus aureus in Bastar tribal region, India*. Journal of Microbiology and Infectious Diseases 2012. **2**.(3): p. 127-128.
78. Frost LS, L.R., Summers AO, Toussaint A *Mobile genetic elements: the agents of open source evolution*. Nat Rev Microbiol 2005. **3**: p. 722–732.
79. Katayama, Y., Ito, T. & Hiramatsu, K. , *A new class of genetic element, staphylococcus cassette chromosome mec, encodes methicillin resistance in Staphylococcus aureus*. Antimicrob. Agents Chemother, 2000. **44**(1549-1555 ).
80. Liu, J.e.a., *Staphylococcal chromosomal cassettes mec (SCCmec): A mobile genetic element in methicillin-resistant Staphylococcus aureus*. Microbial Pathogenesis 2016. **101**(56-67).
81. Joshua P. Ramsay, S.M.K., Riley J. T. Murphy, Karina Yui Eto, Karina J. Price, Quang T. Nguyen, Frances G. O'Brien, Warren B. Grubb, Geoffrey W. Coombs and Neville Firth, *An updated view of plasmid conjugation and mobilization in Staphylococcus*. MOBILE GENETIC ELEMENTS, 2016. **6**(4): p. e1208317.
82. Thomas, W.D., Jr. and G.L. Archer, *Identification and cloning of the conjugative transfer region of Staphylococcus aureus plasmid pGO1*. J Bacteriol, 1989. **171**(2): p. 684-91.
83. Ramsay, J.P., S.M. Kwong, Murphy, R.J.T., Eto, K.Y., Price, K.J., Nguyen, Q.T., O'Brien, F.G., Grubb, W.B., Coombs, G.W., *Neville Firth An updated view of plasmid conjugation and mobilization in Staphylococcus*. Mob Genet Elements, 2016. **6**(4): p. 1-11.
84. Bender J, S.B., Steglich M, Zimmermann O, Fenner I, Lensing C, Dagwadordsch U, Kekulé AS, Werner G, Layer F., *Linezolid resistance in clinical isolates of Staphylococcus epidermidis from German hospitals and characterization of two cfr-carrying plasmids*. J Antimicrob Chemother 2015. **70**: p. 1630-8.
85. Weigel LM, C.D., Gill SR, Clark NC, McDougal LK, Flannagan SE, Kolonay JF, Shetty J, Killgore GE, Tenover FC, *Genetic analysis of a high-level vancomycin-resistant isolate of Staphylococcus aureus*. Science 2003. **302**: p. 1569-71.
86. Udo E, T.D., Grubb WB. , *A conjugative staphylococcal plasmid with no resistance phenotype*. FEMS Microbiol Lett 1987. **40**: p. 279-83.
87. Townsend DE, A.N., Annear DI, Grubb WB., *A conjugative plasmid encoding production of a diffusible pigment and resistance to aminoglycosides and macrolides in Staphylococcus aureus*. Aust J Exp Biol Med Sci, 1985. **63**: p. 573-86.
88. Udo EE, W.M., Grubb WB., *Conjugative trimethoprim resistance in Staphylococcus aureus*. FEMS Microbiol Lett 1992. **76**: p. 243-8.
89. DeLeo, N.M.a.F.R., *Mobile genetic elements of Staphylococcus aureus*. Cellular and Molecular Life Sciences volume, 2010. **67**: p. 3057–3071.
90. Phillips S, N.R., *Tn554-a site-specific repressorcontrolled transposon in Staphylococcus aureus*. Nature, 1979. **278**: p. 476–478.
91. Ito T, O.K., Ma XX, Yuzawa H, Hiramatsu K, *Insights on antibiotic resistance of Staphylococcus aureus from its whole genome: Genomic island SCC*. Drug Resist Update, 2003. **6**: p. 41–52.
92. Soge OO, B.N., White TM, No DB, Roberts MC, *A novel transposon, Tn6009, composed of a Tn916 element linked with a Staphylococcus aureus mer operon*. J Antimicrob Chemother, 2008. **62**: p. 674–680.
93. Byrne ME, R.D., Skurray RA, *Nucleotide sequence analysis of IS256 from the Staphylococcus aureus gentamicin-tobramycin-kanamycin-resistance transposon Tn4001*. Gene, 1989. **81**: p. 361–367.



## References

94. Jaurin B, N.S., *Insertion of IS2 creates a novel ampC promoter in Escherichia coli*. Cell, 1983. **32**: p. 809–816.
95. Maryury Brown-Jaque, W.C.-C., Maite Muniesa, *Transfer of antibiotic-resistance genes via phage-related mobile elements*. Plasmid 2015. **79** p. 1–7.
96. American Academy of Microbiology, Washington, D., *Antibiotic Resistance: An Ecological Perspective on an Old Problem*. 2009.
97. Billard-Pomares, T., Fouteau, S., Jacquet, M.E., Roche, D., Barbe, V., Castellanos, M., et al., *Characterization of a P1-like bacteriophage carrying an SHV-2 extended-spectrum  $\beta$ -lactamase from an Escherichia coli strain*. Antimicrob. Agents Chemother, 2014. **58**: p. 6550–6557.
98. Schmieger, H., Schicklmaier, P., *Transduction of multiple drug resistance of Salmonella enterica serovar typhimurium DT104*. FEMS Microbiol, 1999. **170**: p. 251–256.
99. Hyder, S.L., Streitfeld, M.M., *Transfer of erythromycin resistance from clinically isolated lysogenic strains of Streptococcus pyogenes via their endogenous phage*. J. Infect. Dis., 1978. **138**: p. 281–286.
100. Ubukata, K., Konno, M., Fujii, R., *Transduction of drug resistance to tetracycline, chloramphenicol, macrolides, lincomycin and clindamycin with phages induced from Streptococcus pyogenes*. J. Antibiot. (Tokyo), 1975. **28**: p. 681–688.
101. Schuch, R., Fischetti, V.A., *Detailed genomic analysis of the Wbeta and gamma phages infecting Bacillus anthracis: implications for evolution of environmental fitness and antibiotic resistance*. J. Bacteriol, 2006. **188**: p. 3037–3051.
102. Utter, B., Deutsch, D.R., Schuch, R., Winer, B.Y., Verratti, K., Bishop-Lilly, K., *Beyond the chromosome: the prevalence of unique extra-chromosomal bacteriophages with integrated virulence genes in pathogenic Staphylococcus aureus*. PLoS ONE, 2014. **9**: p. e100502.
103. Roy Curtiss, I., *BACTERIAL CONJUGATION*. Annual Review of Microbiology, 1969. **23**(1): p. 69-136.
104. Anthony, K.G., et al., *The role of the pilus in recipient cell recognition during bacterial conjugation mediated by F-like plasmids*. Molecular Microbiology, 1994. **13**(6): p. 939-953.
105. Novick, R.P., G.E. Christie, and J.R. Penades, *The phage-related chromosomal islands of Gram-positive bacteria*. Nat Rev Microbiol, 2010. **8**(8): p. 541-51.
106. Bushman, F., *Lateral DNA Transfer: Mechanisms and Consequences*. Cold Spring Harbor Laboratory Press, New York, 2002.
107. Calum Johnston, B.M., Gwennaele Fichant<sup>1</sup>, Patrice Polard and J.-P. Claverys, *Bacterial transformation distribution shared mechanisms and divergent*. NATURE REVIEWS, 2014.
108. Dubnau, D. and R. Provvedi, *Internalizing DNA*. Research in microbiology, 2000. **151**(6): p. 475-480.
109. Meima, R., et al., *The bdbDC operon of Bacillus subtilis encodes thiol-disulfide oxidoreductases required for competence development*. Journal of Biological Chemistry, 2002. **277**(9): p. 6994-7001.
110. Laurenceau, R., et al., *A type IV pilus mediates DNA binding during natural transformation in Streptococcus pneumoniae*. PLoS pathogens, 2013. **9**(6): p. e1003473.

111. Mirouze, N., et al., *Antibiotic sensitivity reveals that wall teichoic acids mediate DNA binding during competence in Bacillus subtilis*. Nature communications, 2018. **9**(1): p. 1-11.
112. Chen, I., R. Provvedi, and D. Dubnau, *A macromolecular complex formed by a pilin-like protein in competent Bacillus subtilis*. Journal of Biological Chemistry, 2006. **281**(31): p. 21720-21727.
113. Puyet, A., B. Greenberg, and S.A. Lacks, *Genetic and structural characterization of endA: A membrane-bound nuclease required for transformation of Streptococcus pneumoniae*. Journal of molecular biology, 1990. **213**(4): p. 727-738.
114. Chen, I. and D. Dubnau, *DNA uptake during bacterial transformation*. Nature Reviews Microbiology, 2004. **2**(3): p. 241-249.
115. Dubnau, D., *DNA uptake in bacteria*. Annual Reviews in Microbiology, 1999. **53**(1): p. 217-244.
116. Attaiech L, e.a., *Role of the single-stranded DNA-binding protein SsbB in pneumococcal transformation: Maintenance of a reservoir for genetic plasticity*. PLoS Genet 2011. **7**: p. e1002156.
117. Dubnau, H.M.a.D., *Bistability in the Bacillus subtilis K-state (competence) system requires a positive feedback loop*. Molecular Microbiology 2005. **56**(3): p. 615–624.
118. Lee, M.S.M., D. A., *Identification of a new regulator in Streptococcus pneumoniae linking quorum sensing to competence for genetic transformation*. J. Bacteriol, 1999. **181**: p. 5004-5016
119. A., W., J. N., Paton, J. C., and Andrew, P. W., *The role of Streptococcus pneumoniae virulence factors in host respiratory colonization and disease*. Nat. Rev. Microbiol. , 2008. **6**: p. 288-301.
120. Raphaël Laurenceau, P.V.K., Amy Diallo, Sahra Ouarti, Magalie Duchateau, Christian Malosse<sup>2</sup>, Julia Chamot-Rooke, Rémi Fronzes, *Conserved Streptococcus pneumoniae Spirosomes Suggest a Single Type of Transformation Pilus in Competence*. PLOS PATHOGENS, 2015.
121. Straume, D., Stamsås, G. A., and Håvarstein, L. S. , *Natural transformation and genome evolution in Streptococcus pneumoniae*. Infect. Genet. Evol. , 2015. **33**: p. 371–380.
122. Richards, V.P., Palmer, S.R., Pavinski Bitar, P.D., Qin, X., Weinstock, G.M., Highlander, S.K, *Phylogenomics and the dynamic genome evolution of the genus Streptococcus*. Genome Biology and Evolution, 2014. **6**: p. 741–753.
123. Salvadori, G., et al., *Competence in Streptococcus pneumoniae and Close Commensal Relatives: Mechanisms and Implications*. Frontiers in cellular and infection microbiology, 2019. **9**: p. 94-94.
124. Griffith, F., *The significance of pneumococcal types*. J. Hyg, 1928. **27**: p. 108–159.
125. Johnston C, M.B., Fichant G, Polard P, Claverys JP., *Bacterial transformation: distribution, shared mechanisms and divergent control*. Nat Rev Microbiol. , 2014 **12**(3): p. 181-96.
126. Lee, M.S. and D.A. Morrison, *Identification of a new regulator in Streptococcus pneumoniae linking quorum sensing to competence for genetic transformation*. Journal of Bacteriology, 1999. **181**(16): p. 5004-5016.
127. Prudhomme, M., Attaiech, L., Sanchez, G., Martin, B., Claverys, J.P., *Antibiotic stress induces genetic transformability in the human pathogen Streptococcus pneumoniae*. Science, 2006. **313**: p. 89-92.

## References

128. Stevens, K.E., Chang, D., Zwack, E.E., Sebert, M.E., , *Competence in Streptococcus pneumoniae is regulated by the rate of ribosomal decoding errors*. mBio, 2011. **2**: p. e00071–11.
129. Cassone, M., Gagne, A.L., Spruce, L.A., Seeholzer, S.H., Sebert, M.E., *The HtrA protease from Streptococcus pneumoniae digests both denatured proteins and the competence-stimulating peptide*. J. Biol. Chem. , 2012. **287**: p. 38449–38459.
130. Gagne, A.L., Stevens, K.E., Cassone, M., Pujari, A., Abiola, O.E., Chang, D.J., Sebert, M.E.,, *Competence in Streptococcus pneumoniae is a response to an increasing mutational burden*. PLoS One, 2013. **8**: p. e72613.
131. Slager, J., Kjos, M., Attaiech, L., Veening, J.W., *Antibiotic-induced replication stress triggers bacterial competence by increasing gene dosage near the origin*. Cell, 2014. **157**(395–406).
132. Federle, E.S.a.M.J., *Quorum Sensing Regulation of Competence and Bacteriocins in Streptococcus pneumoniae and mutants*. genes, 2017.
133. Havarstein, L.S.D., D.B.; Nes, I.F. , *A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export*. Mol. Microbiol. , 1995. **16**: p. 229–240.
134. Håvarstein, L.S., Gaustad, P., Nes, I.F., Morrison, D.A., *Identification of the streptococcal competence-pheromone receptor*. Mol. Microbiol. , 1996. **21**: p. 863–869.
135. Ween, O.G., P.; Havarstein, L.S. , *Identification of DNA binding sites for ComE, a key regulator of natural competence in Streptococcus pneumoniae*. . Mol. Microbiol. , 1999. **33**: p. 817–827.
136. Martin, B.S., A.L.; Mirouze, N.; Prudhomme, M.; Mortier-Barriere, I.; Granadel, C.; Noiro-Gros, M.F.; Noiro, P.; Polard, P.; Claverys, J.P. , *ComE/ComE~P interplay dictates activation or extinction status of pneumococcal X-state (competence)*. Mol. Microbiol, 2013. **87**: p. 394–411.
137. Havarstein, J.-P.C.a.L.S., *EXTRACELLULAR-PEPTIDE CONTROL OF COMPETENCE FOR GENETIC TRANSFORMATION IN STREPTOCOCCUS PNEUMONIAE*. Frontiers in Bioscience, 2002. **7**: p. 1798-1814.
138. Peterson SN, S.C., Cline R, Desai BV, Snesrud EC, Luo P, Walling J, Li H, Mintz M, Tsegaye G, Burr PC, Do Y, Ahn S, Gilbert J, Fleischmann RD, Morrison DA., *Identification of competence pheromone responsive genes in Streptococcus pneumoniae by use of DNA microarrays*. Mol Microbiol., 2004. **4**: p. 1051-70.
139. Nicole L. Inniss, G.P., and Donald A. Morrison, *The pneumococcal  $\sigma$ X activator, ComW, is a DNA-binding protein critical for natural transformation*. J. Biol. Chem., 2019.
140. Håvarstein, L.S., Martin, B., Johnsborg, O., Granadel, C., Claverys, J.P., *New insights into the pneumococcal fratricide: relationship to clumping and identification of a novel immunity factor*. Mol. Microbiol. , 2006. **59**,: p. 1297–1307.
141. Wholey, W.Y.K., T.J.; Storck, D.N.; Dawid, S. , *Coordinated Bacteriocin Expression and Competence in Streptococcus pneumoniae Contributes to Genetic Adaptation through Neighbor Predation*. PLoS Pathog, 2016. **12**: p. e1005413.
142. Kjos M, M.E., Slager J, Lake FB, Gericke O, Roberts IS, Rozen DE, Veening JW., *Expression of Streptococcus pneumoniae Bacteriocins Is Induced by Antibiotics via Regulatory Interplay with the Competence System*. PLoS Pathog, 2019. **2**: p. e1005422.
143. de Saizieu, A.G., C.; Flint, N.; Wagner, C.; Kamber, M.; Mitchell, T.J.; Keck, W.; Amrein, K.E.; Lange, R., *Microarray-based identification of a novel Streptococcus pneumoniae regulon controlled by an autoinduced peptide*. J. Bacteriol. , 2000. **182**: p. 4696–4703.

144. Pinchas MD, L.N., Dawid S, *An electrostatic interaction between BlpC and BlpH dictates pheromone specificity in the control of bacteriocin production and immunity in Streptococcus pneumoniae*. J Bacteriol, 2015. **197**(7): p. 1236-48.
145. Lee, M.S.M., D. A., *Identification of a new regulator in Streptococcus pneumoniae linking quorum sensing to competence for genetic transformation*, in J. Bacteriol. 1999. p. 5004–5016
146. Mirouze, N., Bergé, M.A., Soulet, A.L., Mortier-Barrière, I., Quentin, Y., Fichant, G., Granadel, C., Noirot-Gros, M.F., Noirot, P., Polard, P., Martin, B., Claverys, J.P., *Direct involvement of DprA, the transformation-dedicated RecA loader, in the shut-off of pneumococcal competence*. Proc. Natl. Acad. Sci. U.S.A. , 2013. **110**: p. E1035–E1044.
147. Rahmer, R., K. Morabbi Heravi, and J. Altenbuchner, *Construction of a Super-Competent Bacillus subtilis 168 Using the PmtIA-comKS Inducible Cassette*. Frontiers in Microbiology, 2015. **6**(1431).
148. Spizizen, J., *Transformation of biochemically deficient strains of Bacillus subtilis by deoxyribonucleate*. Proc. Natl. Acad. Sci. U.S.A., 1958. **44**: p. 1072–1078.
149. Haijema, B.J., Hahn, J., Haynes, J., and Dubnau, D., *A ComGA-dependent checkpoint limits growth during the escape from competence*. Mol Microbiol 2001. **40**: p. 52–64.
150. Berka, R.M., et al., *Microarray analysis of the Bacillus subtilis K-state: genome-wide expression changes dependent on ComK*. Mol Microbiol, 2002. **43**(5): p. 1331-45.
151. Dubnau, D., *The regulation of genetic competence in Bacillus subtilis*. Molecular Microbiology, 1991. **5**(1): p. 11-18.
152. Ansaldi, M., Marolt, D., Stebe, T., Mandic-Mulec, I. & Dubnau, D., *Specific activation of the Bacillus quorum-sensing systems by isoprenylated pheromone variants*. . Mol Microbiol 2002. **44**: p. 1561–1573.
153. Solomon, J.M., et al., *Convergent sensing pathways mediate response to two extracellular competence factors in Bacillus subtilis*. Genes & Development, 1995. **9**(5): p. 547-558.
154. Ratnayake-Lecamwasam, M., Serror, P., Wong, K. W. & Sonenshein, A. L., *Bacillus subtilis CodY represses early stationary-phase genes by sensing GTP levels*. Genes Dev, 2001. **15**: p. 1093–1103.
155. Hoa, T.T., Tortosa, P., Albano, M. & Dubnau, D., *Rok (YkuW) regulates genetic competence in Bacillus subtilis by directly repressing comK*. Mol Microbiol 2002. **43**: p. 15–26.
156. Turgay, K., Hamoen, L. W., Venema, G. & Dubnau, D. , *Biochemical characterization of a molecular switch involving the heat shock protein ClpC, which controls the activity of ComK, the competence transcription factor of Bacillus subtilis*. Genes Dev, 1997. **11**: p. 119–128.
157. Turgay, K., Hahn, J., Burghoorn, J. & Dubnau, D. , *Competence in Bacillus subtilis is controlled by regulated proteolysis of a transcription factor*. . EMBO J 1998. **17**: p. 6730–6738.
158. Nicolas Mirouze, Y.D., Arjun Raj, David Dubnau, *Spo0A-P Imposes a Temporal Gate for the Bimodal Expression of Competence in Bacillus subtilis*. PLoS Genet, 2012. **8**(3): p. e1002586.
159. Hamoen, L.W., Van Werkhoven, A.F., Venema, G., and Dubnau, D. , *The pleiotropic response regulator DegU functions as a priming protein in competence development in bacillus subtilis*. . Proc Natl Acad Sci USA, 2000. **97**: p. 9246–9251.

## References

160. Solomon, J., Magnuson, R., Srivastava, A., and Grossman, A.D., *Convergent sensing pathways mediate response to two extracellular competence factors in Bacillus subtilis*. Genes Dev 1995. **9**: p. 547–558.
161. Claverys, J.P., Prudhomme, M. & Martin, B., *Induction of competence regulons as general stress responses in Gram positive bacteria*. Annu. Rev. Microbiol. , 2006. **60**: p. 451–475
162. Roggiani, M., Dubnau, D. , *ComA, a phosphorylated response regulator protein of Bacillus subtilis, binds to the promoter region of srfA*. . J Bacteriol 1993. **175**: p. 3182–3187.
163. Stephenson, S., et al., *Molecular Analysis of Phr Peptide Processing in Bacillus subtilis*. Journal of Bacteriology, 2003. **185**(16): p. 4861-4871.
164. Lanigan-Gerdes S, D.A., Faull KF, Lazazzera BA., *Identification of subtilisin, Epr and Vpr as enzymes that produce CSF, an extracellular signalling peptide of Bacillus subtilis*. Mol Microbiol, 2007. **65**(5): p. 1321-33.
165. Lazazzera, B.A., Solomon, J. M. & Grossman, A. D. , *An exported peptide functions intracellularly to contribute to cell density signaling in B. subtilis*. Cell 1997. **89**: p. 917–925.
166. Core L, P.M., *TPR-mediated interaction of RapC with ComA inhibits response regulator-DNA binding for competence development in Bacillus subtilis*. Mol Microbiol, 2003.
167. Briley Jr, K., et al., *Retracted: Maf acts downstream of ComGA to arrest cell division in competent cells of B. subtilis*. Molecular Microbiology, 2011. **81**(1): p. 23-39.
168. Lipp EK, H.A., and Colwell RR, *Effects of global climate on infectious disease: the cholera model*. Clin Microbiol Rev, 2002. **15**: p. 757-770.
169. Blokesch M, S.G., *Serogroup conversion of Vibrio cholerae in aquatic reservoirs*. . PLoS Pathog 2007. **3**: p. e81.
170. Meibom KL, B.M., Dolganov NA, Wu C-Y, Schoolnik GK, *Chitin induces natural competence in Vibrio cholerae*. Science, 2005. **310**: p. 1824-1827.
171. Metzger, L.C. and M. Blokesch, *Regulation of competence-mediated horizontal gene transfer in the natural habitat of Vibrio cholerae*. Curr Opin Microbiol, 2016. **30**: p. 1-7.
172. Dalia, A.B., D.W. Lazinski, and A. Camilli, *Identification of a membrane-bound transcriptional regulator that links chitin and natural competence in Vibrio cholerae*. mBio, 2014. **5**(1): p. e01028.
173. Lo Scudato M, B.M., *A transcriptional regulator linking quorum sensing and chitin induction to render Vibrio cholerae naturally transformable*. Nucleic Acids Res, 2013. **41**: p. 3644-3658.
174. Higgins, D.A., *The major Vibrio cholerae autoinducer and its role in virulence factor production*. Nature 2007. **450**: p. 883–886.
175. Chen, X., *Structural identification of a bacterial quorum-sensing signal containing boron*. Nature 2002. **415**: p. 545–549.
176. Henke, J.M.a.B., B. L., *Three parallel quorumsensing systems regulate gene expression in Vibrio harveyi*. J. Bacteriol. , 2004. **186**: p. 6902–6914.
177. Neiditch, M.B., *Regulation of LuxPQ receptor activity by the quorum-sensing signal autoinducer-2*. Mol. Cell, 2005. **18**: p. 507–518.
178. BL, L.B.a.B., *Regulation of quorum sensing in Vibrio harveyi by LuxO and sigma-54*. Mol Microbiol, 2000. **36**: p. 940–954.
179. Rutherford ST, B.B., *Bacterial quorum sensing: its role in virulence and possibilities for its control*. Cold Spring Harb Perspect Med., 2012. **2**(11): p. a012427.

180. Blokesch M, S.G., *The extracellular nuclease Dns and its role in natural transformation of Vibrio cholerae*. J Bacteriol, 2008. **190**: p. 7232-7240.
181. Morikawa K, I.Y., Okamura H, Maruyama A, Hayashi H, Takeyasu K, Ohta T, *A new staphylococcal sigma factor in the conserved gene cassette: functional significance and implication for the evolutionary processes*. Genes Cells, 2003 **8**(8): p. 699-712.
182. Annette Fagerlund, P.E.G.a.L.S.H., *Staphylococcus aureus competence genes: mapping of the SigH, ComK1 and ComK2 regulons by transcriptome sequencing*. Molecular Microbiology, 2014. **94**(3): p. 557–579.
183. Kuroda, M., et al., *Whole genome sequencing of meticillin-resistant Staphylococcus aureus*. The Lancet, 2001. **357**(9264): p. 1225-1240.
184. Maree, M., et al., *MRSA emerges through natural transformation*. bioRxiv, 2020: p. 2020.12.08.415893.
185. Kuroda, M., K. Kuwahara-Arai, and K. Hiramatsu, *Identification of the Up- and Down-Regulated Genes in Vancomycin-Resistant Staphylococcus aureus Strains Mu3 and Mu50 by cDNA Differential Hybridization Method*. Biochemical and Biophysical Research Communications, 2000. **269**(2): p. 485-490.
186. Aurélie Hiron, M.F., Jaione Valle, Michel Débarbouillé, Tarek Msadek, *Bacitracin and nisin resistance in Staphylococcus aureus: a novel pathway involving the BraS/BraR two-component system (SA2417/SA2418) and both the BraD/BraE and VraD/VraE ABC transporters*. Mol Microbiol, 2011. **81**(3): p. 602–622.
187. Sany Fds and Ok-kyu Son, *A novel genetic system to detect protein–protein interactions*. Nature, 1989.
188. Mortier-Barrière, I., et al., *A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA*. Cell, 2007. **130**(5): p. 824-836.
189. Kidane, D. and P.L. Graumann, *Intracellular Protein and DNA Dynamics in Competent Bacillus subtilis Cells*. Cell, 2005. **122**(1): p. 73-84.
190. Kramer, N., J. Hahn, and D. Dubnau, *Multiple interactions among the competence proteins of Bacillus subtilis*. Molecular Microbiology, 2007. **65**(2): p. 454-464.
191. Burghard-Schrod, M., S. Altenburger, and P.L. Graumann, *The Bacillus subtilis dCMP deaminase ComEB acts as a dynamic polar localization factor for ComGA within the competence machinery*. Molecular Microbiology, 2020. **113**(5): p. 906-922.
192. De Santis, M., J. Hahn, and D. Dubnau, *ComEB protein is dispensable for the transformation but must be translated for the optimal synthesis of comEC*. Molecular Microbiology, 2021. **116**(1): p. 71-79.
193. Diallo, A., et al., *Bacterial transformation: ComFA is a DNA-dependent ATPase that forms complexes with ComFC and DprA*. Molecular Microbiology, 2017. **105**(5): p. 741-754.
194. Chung, Y. and D. Dubnau, *ComC is required for the processing and translocation of ComGC, a pilin-like competence protein of Bacillus subtilis*. Molecular microbiology, 1995. **15**(3): p. 543-551.
195. Dubrac, S., et al., *A matter of life and death: cell wall homeostasis and the WalkR (YycGF) essential signal transduction pathway*. Molecular microbiology, 2008. **70**(6): p. 1307-1322.
196. Szurmant, H., et al., *YycH and YycI interact to regulate the essential YycFG two-component system in Bacillus subtilis*. Journal of bacteriology, 2007. **189**(8): p. 3280-3289.

## References

197. Niemann, V., et al., *The NreA protein functions as a nitrate receptor in the staphylococcal nitrate regulation system*. Journal of molecular biology, 2014. **426**(7): p. 1539-1553.
198. Wigley, D.B., *Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB*. Nature Reviews Microbiology, 2013. **11**(1): p. 9-13.
199. Dillingham, M.S. and S.C. Kowalczykowski, *RecBCD enzyme and the repair of double-stranded DNA breaks*. Microbiology and Molecular Biology Reviews, 2008. **72**(4): p. 642-671.
200. Marsin, S., et al., *Genetic dissection of Helicobacter pylori AddAB role in homologous recombination*. FEMS microbiology letters, 2010. **311**: p. 44-50.
201. Kho, K. and T.C. Meredith, *Salt-induced stress stimulates a lipoteichoic acid-specific three-component glycosylation system in Staphylococcus aureus*. Journal of bacteriology, 2018. **200**(12): p. e00017-18.
202. Eugster, M.R., et al., *Bacteriophage predation promotes serovar diversification in Listeria monocytogenes*. Molecular microbiology, 2015. **97**(1): p. 33-46.
203. Monk, I.R., et al., *Transforming the untransformable: application of direct transformation to manipulate genetically Staphylococcus aureus and Staphylococcus epidermidis*. mBio, 2012. **3**(2).
204. Carreau, A., et al., *Why is the partial oxygen pressure of human tissues a crucial parameter? Small molecules and hypoxia*. Journal of cellular and molecular medicine, 2011. **15**(6): p. 1239-1253.
205. Rubinstein, E., *Staphylococcus aureus bacteraemia with known sources*. International journal of antimicrobial agents, 2008. **32**: p. S18-S20.
206. Haag, A. and F. Bagnoli, *The Role of Two-Component Signal Transduction Systems in Staphylococcus aureus Virulence Regulation*. Current topics in microbiology and immunology, 2016. **409**.
207. Müllner, M., et al., *A PAS Domain with an Oxygen Labile [4Fe-4S]<sub>2</sub><sup>+</sup> Cluster in the Oxygen Sensor Kinase NreB of Staphylococcus carnosus*. Biochemistry, 2008. **47**(52): p. 13921-13932.
208. Ulrich, M., et al., *The staphylococcal respiratory response regulator SrrAB induces ica gene transcription and polysaccharide intercellular adhesin expression, protecting Staphylococcus aureus from neutrophil killing under anaerobic growth conditions*. Molecular Microbiology, 2007. **65**(5): p. 1276-1287.
209. Yan, M., et al., *The Essential Two-Component System YhcSR Is Involved in Regulation of the Nitrate Respiratory Pathway of Staphylococcus aureus*. Journal of Bacteriology, 2011. **193**(8): p. 1799-1805.
210. Yan, M., et al., *The essential two-component system YhcSR is involved in regulation of the nitrate respiratory pathway of Staphylococcus aureus*. J Bacteriol, 2011. **193**(8): p. 1799-805.
211. Sun, F., J. Chen, and M.-Q. Liu, *Connections between uniformity and aberration in general multi-level factorials*. Metrika, 2011. **73**(3): p. 305-315.
212. Turgay, K., et al., *Competence in Bacillus subtilis is controlled by regulated proteolysis of a transcription factor*. The EMBO journal, 1998. **17**(22): p. 6730-6738.
213. Piotrowski, A., P. Luo, and D.A. Morrison, *Competence for genetic transformation in Streptococcus pneumoniae: termination of activity of the alternative sigma factor ComX is independent of proteolysis of ComX and ComW*. Journal of bacteriology, 2009. **191**(10): p. 3359-3366.

214. Lo Scudato, M. and M. Blokesch, *A transcriptional regulator linking quorum sensing and chitin induction to render Vibrio cholerae naturally transformable*. Nucleic acids research, 2013. **41**(6): p. 3644-3658.
215. Auzat, I., et al., *The NADH oxidase of Streptococcus pneumoniae: its involvement in competence and virulence*. Molecular microbiology, 1999. **34**(5): p. 1018-1028.
216. Antonova, E.S. and B.K. Hammer, *Genetics of natural competence in Vibrio cholerae and other vibrios*. Microbiology spectrum, 2015. **3**(3): p. 3.3. 20.
217. Rahmer, R., K. Morabbi Heravi, and J. Altenbuchner, *Construction of a super-competent Bacillus subtilis 168 using the PmtIA-comKS inducible cassette*. Frontiers in microbiology, 2015. **6**: p. 1431.
218. Kobayashi, S.D., N. Malachowa, and F.R. DeLeo, *Pathogenesis of Staphylococcus aureus abscesses*. The American journal of pathology, 2015. **185**(6): p. 1518-1527.
219. Hajdamowicz, N.H., et al., *The Impact of Hypoxia on the Host-Pathogen Interaction between Neutrophils and Staphylococcus aureus*. International journal of molecular sciences, 2019. **20**(22): p. 5561.
220. Bosio, S., et al., *Mycobacterium fortuitum prosthetic valve endocarditis: a case for the pathogenetic role of biofilms*. Cardiovascular Pathology, 2012. **21**(4): p. 361-364.
221. Mirouze, N., P. Prepiak, and D. Dubnau, *Fluctuations in spo0A transcription control rare developmental transitions in Bacillus subtilis*. PLoS genetics, 2011. **7**(4): p. e1002048.
222. Martin, B., et al., *ComE/ComE ~ P interplay dictates activation or extinction status of pneumococcal X-state (competence)*. Molecular microbiology, 2013. **87**(2): p. 394-411.
223. Stewart, P.S. and M.J. Franklin, *Physiological heterogeneity in biofilms*. Nature Reviews Microbiology, 2008. **6**(3): p. 199-210.



