VNIVERSITAT DE VALÈNCIA

Doctoral Programme in Biomedicine and Biotechnology

Staphylococcus aureus Pathogenicity islands mobilisation and their role in chromosomally-encoded virulence gene expression

Doctoral Thesis

Author

Mercedes Cervera Alamar

Supervisor

Mª Ángeles Tormo Mas

April, 2019

Dña. María Ángeles Tormo Más, investigadora en el grupo de Infección Grave del Instituto de Investigación Sanitaria La Fe de Valencia

INFORMA QUE:

Doña Mercedes Cervera Alamar, licenciada en Biotecnología por la Universidad Politécnica de Valencia, ha realizado en el grupo de Infección Grave del Instituto de Investigación Sanitaria La Fe, bajo su dirección, el trabajo titulado "*Staphylococcus aureus* Pathogenicity Islands mobilisation and their role in chromosomally-encoded virulence gene expression" y que hallándose concluida, reúne todos los requisitos para su juicio y calificación, por tanto, autoriza su presentación, a fin de que pueda ser juzgado por el tribunal correspondiente para la obtención del grado de Doctor por la Universidad de Valencia.

Y para que conste, en cumplimiento con la legislación, firma el presente informe en Valencia, a 5 de abril de 2019.

Dra. María Ángeles Tormo Más

Para la realización de la presente Tesis Doctoral, la autora ha sido beneficiaria de una beca predoctoral del Programa de Formación de Personal Investigador (FPI BES-2015-073651) adscrita al proyecto "Transferencia horizontal de factores de virulencia contenidos en elementos genéticos móviles de *Staphylococcus aureus* en pacientes con Fibrosis Quística" Plan Nacional de I+D+I MINECO, SAF2014-56986-R.

Esta tesis también ha sido financiada con los proyectos:

Programa I+D Ramón y Cajal RYC-2012-12246. Ministerio de Economía y Competitividad (IP: María Ángeles Tormo Mas).

"Nuevas estrategias de prevención y diagnóstico de las infecciones relacionadas con dispositivos biomédicos causados por *Staphylococcus* spp." Plan Nacional I+D+I MINECO, SAF2017-82251-R (IP: Maria Ángeles Tormo Mas).

A Gonzalo

A mis padres

A Irene y Jaime

Agradecimientos

Este es el final de una importante etapa de mi vida y no quiero cerrarla sin agradecer a todas las personas que han hecho posible que esta etapa sea más sencilla y feliz.

En primer lugar, a mi directora de tesis, Geles, ella me ha dado la oportunidad de empezar en el mundo de la investigación. He aprendido mucho con ella y siempre ha confiado en mí.

Gracias a aquellas personas que han pasado por el laboratorio y me han dejado huella. Al principio las Marias, Belén y Miguel que son los que me enseñaron a manejarme en el laboratorio. Posteriormente Marcelo, mi consejero personal, siempre dispuesto a ayudarme. Por último, Elías, la alegría y el optimismo que transmite me ha dado energía para afrontar esta última etapa.

Gracias a Ana Pilar, nos conocimos al principio de la tesis y siempre hemos estado apoyándonos, el buen rollo que transmite es contagioso. Gracias también a Paula que, aunque ajena a este mundo, también me ha apoyado y ha confiado en mí. Las visitas de las dos al despachito me alegraban el día.

Gracias a Vince Fischetti y Edmondo por acogerme en su laboratorio, la estancia en Nueva York ha sido una experiencia muy bonita en mi vida. Aprendí mucho de ellos y me hicieron sentir como en casa.

A mis vecinos favoritos, los descubridores de fármacos, estos años han sido mucho más amenos gracias a ellos, compartiendo comidas y almuerzos. Gracias a Arturo, Marta y María. Gracias también a Antonio que se implicó y colaboró para que el artículo saliera adelante. Especialmente quiero dar las gracias a Leti, que desde que nos conocimos conectamos muy bien, me ha enseñado muchísimo del mundo de las proteínas y siempre ha estado dispuesta a escucharme.

Muchas gracias a mis dos grandes descubrimientos, Ayelen y Nuria, convivir en el despachito de 2m^2 estos últimos meses ha sido genial. ¡Hemos compartido muchas crisis, mucho chocolate y muchas confesiones y espero seguir compartiendo todo eso y mucho más!

Infinitas gracias a Patri, este último año me ha ayudado muchísimo. Una de las personas más trabajadoras que conozco. Siempre con una sonrisa, dispuesta a ayudarme con los experimentos para que me pudiera sentar a escribir.

Katy y Migle, las dos personas más importantes que me llevo de esta etapa. Hemos compartido el día a día en el laboratorio y nos hemos convertido en muy buenas amigas. Gracias por haberme regalado tantos momentos de risas, hemos estado muy unidas y trabajar con vosotras ha sido un placer. Además, vuestro apoyo tanto cuando estabais como cuando os fuisteis ha sido vital para mí, ¡Gracias!

Quiero dar las gracias a mi familia; a mis padres y a mi hermana por su apoyo incondicional y su cariño diario. Todos estos años habéis compartido mis alegrías y angustias y esto no hubiera sido posible sin vosotros.

A Gonzalo, ¡¡¡GRACIAS!!!, por confiar en mí, por tu generosidad, paciencia y ayuda. Por las palabras de apoyo y ánimo constantes, por la entrega absoluta, por la positividad que me transmites. Gracias por estar aquí, por haberme encontrado y por formar parte de mi vida.

¡¡GRACIAS!!

Table of content

Abbreviations

Φ	Phage
Δ	Deletion
Ø	Empty vector
λ	Lambda
σ^B	Sigma B
att	Attachment site
blaZ	β-lactamase gene
bp	Base pair(s)
°C	Degrees Celsius
CC	Clonal Complex
$CaCl_2$	Calcium chloride
CDSs	Coding sequences
cDNA	complementary deoxyribonucleic acid
CFU	Colony Forming Unit
DEGs	Differentially Expressed Genes
DIG	Digoxyribonucleic acid
DNA	Deoxyribonucleic acid
ERP	Excision-Replication-Packaging cycle
g	Gram(s)
h	Hour(s)
HPLC	High performance liquid chromatography
HGT	Horizontal Gene Transfer
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kb	Kilo base pair(s)
LB	Luria-bertani broth
Mb	Mega base pair(s)
MC	Mitomycin C
MGE	Mobile Genetic Element
MLST	Multilocus Sequence Typing
MRSA	Methicillin-resistant *S. aureus*
µL	Microlitre(s)

mL	Millilitre(s)
mM	Millimolar
min	Minute (s)
MW	Molecular weight
ng	Nanogram (s)
nm	nanometer
nt	Nucleotide
OD	Optical density at a certain wavelength
ORF	Open reading frame
P	Promoter
PA	Processive antitermination
P_{cad}	Cadmium-inducible promoter
PCR	Polymerase Chain Reaction
PICIs	Phage Inducible Chromosomal Islands
ROS	Reactive Oxygen Species
SDS	Sodium dodecyl sulfate
SaPIs	*Staphylococcus aureus* Pathogenicity Islands
SePIs	*Staphylococcus epidermidis* Pathogenicity islands
SOS	Stress bacterial response
SP	Scaffolding protein
SSAP	Single Strand Annealing Proteins
STX	Staphyloxanthin
TSB	Tryptic soy broth
TSA	Tryptic soy Agar
VISA	Vancomycin-resistant *S. aureus*
WT	Wild Type

List of Figures

Introduction

Results

Discussion

List of Tables

Summary

Introducción

Staphylococcus aureus

Staphylococcus aureus es una bacteria Gram positiva, comensal de la mucosa nasal y de la piel de los seres humanos. Aproximadamente el 30% de la población está colonizada con *S. aureus*, sin embargo, es también un patógeno oportunista, causante de infecciones principalmente hospitalarias, pero también comunitarias.

La enorme versatilidad de *S. aureus* como bacteria patógena se debe a su habilidad para persistir y multiplicarse en diferentes ambientes junto con su capacidad para producir una gran variedad de factores de virulencia. Muchos de estos factores de virulencia como superantígenos, toxinas, proteínas implicadas en la formación de biofilm o en la adaptación del hospedador descritas hasta la fecha se encuentran en Elementos Genéticos Móviles (EGMs).

Virulencia y regulación

Durante una infección, *S. aureus* debe expresar correctamente los genes de virulencia para poder sobrevivir a las condiciones impuestas por el hospedador. Dado el número tan elevado de genes de virulencia que posee este microorganismo, la expresión debe ser finamente regulada para asegurar una eficiente coordinación. Se han identificado multitud de ARNs y proteínas con función reguladora de estos procesos, pero vamos a centrarnos en la regulación por antiterminación ya que es la que se desarrolla posteriormente en la tesis. La antiterminación suprime la acción de los terminadores y permite la expresión de los genes que se encuentran aguas abajo. Los antiterminadores pueden manipular la RNA polimerasa para que evite las señales de terminación y transcribe un conjunto de genes en un solo transcrito. Se han descrito ya multitud de proteínas antiterminadoras, por ejemplo, las proteínas N y Q del bacteriófago lambda (λ) o la proteína NusG y sus parálogos.

Elementos Genéticos Móviles

Los EGMs son regiones del cromosoma que pueden moverse dentro del genoma y pueden transferirse de manera horizontal a otras bacterias. En el caso de *Staphylococcus* spp, los EGMs son entre el 15-20% del genoma y suelen contener diferente porcentaje de G+C en comparación con el que se observa en el resto del genoma. Se movilizan de manera horizontal siendo el principal paso evolutivo que acelera las variaciones fenotípicas y genotípicas en las poblaciones de *Staphylococcus* spp. permitiendo la adaptación rápida a distintos ambientes y a diferentes hospedadores.

Entre los EGMs destacan los plásmidos, transposones, casetes cromosómicos y en los que nos vamos a centrar en este estudio que son los bacteriófagos y las islas de patogenicidad de *S. aureus* (SaPIs).

Bacteriófagos

Los EGMs más ampliamente estudiados son los bacteriófagos. También se les conoce como fagos y son agentes infecciosos que se replican como agentes intracelulares obligados dentro de las bacterias y que para hacerlo utilizan la maquinaria biosintética de su hospedador. En *Staphylococcus* spp, los bacteriófagos son mayoritariamente lisogénicos y pertenecen al orden *Caudovirales*, principalmente a la familia *Syphoviridae.*

Los bacteriófagos pueden clasificarse dependiendo de su ciclo de vida en:

Fagos líticos virulentos que tras su multiplicación provocan la lisis de la célula hospedadora, liberando una nueva progenie fágica en un proceso conocido como ciclo lítico.

Fagos atemperados en los que el DNA viral se integra en el genoma de la bacteria infectada y se replica junto a éste, transmitiéndose de una generación a otra (transferencia vertical) sin que se produzca la lisis celular. Este fenómeno se conoce como lisogenia, a la cepa así infectada como lisógeno y al DNA viral integrado como profago.

Los bacteriófagos de *S. aureus* tienen una estructura conservada en módulos. Todos ellos tienen dos genes encargados de controlar el ciclo lítico y el lisogénico. El gen *cI* es requerido para el establecimiento de la lisogenia y el gen *cro* es necesario para promover el desarrollo lítico de los fagos mediante la inhibición de la expresión de CI. Bajo condiciones de daño celular, se desencadena la respuesta SOS como mecanismo de supervivencia por parte de la bacteria, en presencia de $RecA^{*}$ se promueve la autoescisión del represor CI seguida de una degradación proteolítica, lo cual libera la represión de *cro* y permite la expresión de los genes implicados en el ciclo lítico que conducirá a la lisis celular.

Para llevar a cabo la lisogenización de una bacteria, el bacteriófago se integra de manera específica en el cromosoma bacteriano, en el sitio de integración denominado *att*B. La recombinación se realiza mediante una integrasa (*int*). Adyacente a *int* se halla *xis*, que codifica para una escisionasa, que junto a la integrasa es la que cataliza la escisión del profago. Una vez se ha escindido del cromosoma se lleva a cabo la replicación.

En paralelo, una vez el fago empieza a replicarse, se empiezan a sintetizar las proteínas estructurales que forman la cabeza y el tallo fágico, así como las proteínas encargadas del ensamblaje de las partículas víricas y las proteínas encargadas de la lisis celular.

Una vez se ha formado la cápside, la proteína fágica terminasa forma un complejo entre sus dos subunidades (subunidad grande y subunidad pequeña) que interacciona con el DNA concatemérico formado durante la replicación del bacteriófago.

Existen dos modelos de empaquetamiento del DNA fágico. En ambos, la subunidad pequeña de la terminasa reconoce una secuencia nucleotídica específica donde se une, la subunidad grande realiza el primer corte y ambas interactúan con la proteína portal para introducir el DNA dentro de la procápside. En el caso de fagos *cos*, el segundo corte que realiza la subunidad grande es específico de secuencia (secuencia *cos*) y se empaqueta una única copia del genoma del fago comprendida

entre los dos sitios *cos*. Por el contrario, en los fagos *pac*, el segundo corte es inespecífico de secuencia, empaquetándose DNA hasta que la cápside esté llena.

La gran mayoría de proteínas implicadas en virulencia se codifican en la región más alejada de la integrasa, después del módulo de lisis. Las proteínas más frecuentes en esta región son enterotoxinas estafilocócicas, como SEA o SEP, que presentan actividad emética y son los agentes causantes de casos de intoxicación alimentaria en humanos, estas toxinas son también superantígenos, que tienen la capacidad de estimular una gran población de células T. Además, encontramos proteínas inhibidoras del sistema inmune del hospedador, como pueden ser Sak, una estafiloquinasa moduladora responsable de la destrucción tisular del hospedador; CHIPS, una proteína inhibitoria de la quimiotaxis o SCIN que inhibe el sistema del complemento. Todos estos genes pertenecen al cluster de evasión del sistema inmune (IEC), localizado principalmente en los bacteriófagos que se integran en la β-hemolisina.

Cuando un bacteriófago se atempera en una cepa, lisogenizándola, le proporciona un arsenal de genes de virulencia que no disponía. Por tanto, los bacteriófagos juegan un papel importante en la transmisión de factores de virulencia, bien porque están codificados en su genoma o también de forma indirecta ya que son vehículos de información genética no codificada por el fago. En concreto son los únicos responsables de la movilización de islas de patogenicidad.

La transducción es la transferencia horizontal de genes mediada por bacteriófagos. La transducción impulsa la evolución de los procariotas y puede darse principalmente por tres mecanismos sustancialmente diferentes.

Transducción generalizada: es el resultado de “errores” continuos en el empaquetamiento del ADN por parte del fago, pudiendo empaquetar aleatoriamente tanto ADN cromosómico como su propio ADN en partículas fágicas. Estas partículas pueden infectar a una nueva bacteria y el ADN podrá incorporarse o no por recombinación, pero no podrá volver a entrar en ciclo lítico ya que no posee los genes necesarios para hacerlo.

Transducción especializada: resulta de un proceso de escisión inexacta del profago insertado en el cromosoma, que conduce al empaquetamiento del genoma del profago (total o parcialmente) y una pequeña parte de ADN cromosómico contiguo al profago.

Transducción lateral: es un mecanismo recientemente descubierto en bacteriófagos de *S. aureus*. Los profagos retardan la escisión e inician la replicación del DNA estando todavía integrados en el cromosoma. El empaquetamiento se inicia con los fagos todavía integrados y se empaquetan cientos de pares de bases del genoma de *S. aureus* con alta frecuencia.

Islas de Patogenicidad de *S. aureus* (SaPIs)

Las islas de patogenicidad de *S. aureus* (SaPIs) son segmentos cromosomales discretos, de un tamaño variable de entre 15-20 Kb que generalmente codifican para toxinas, como la del síndrome del shock tóxico (TSST-1), superantígenos, o para otros factores de virulencia. Las SaPIs fueron las primeras islas de patogenicidad descritas en Gram positivos.

Las SaPIs presentan una relación muy estrecha con los bacteriófagos de *S. aureus*. Así pues, para la movilización de las SaPIs es necesaria la infección de la cepa portadora de la isla por parte de un fago o la inducción de un profago residente.

Las SaPIs presentan una organización genética muy conservada que recuerda a la organización modular de los bacteriófagos, de manera que genes con funciones relacionadas pertenecen al mismo módulo funcional. El primer gen de la isla codifica para una integrasa (*int*) que es necesaria para la integración y escisión de la isla de sitios específicos del cromosoma. A continuación, se encuentra el módulo de regulación que está constituido por dos genes: *stl* y *str,* codificados en dirección divergente. Las SaPIs se encuentran de manera estable integradas en el cromosoma debido a la acción del represor Stl, el cual, al unirse a la región intergénica de los genes *stl* y *str* bloquea la transcripción de los genes de la SaPI. La desrepresión de la isla es llevada a cabo por proteínas del fago que actúan como inductores y mediante una interacción proteína-proteína, se unen a Stl impidiendo su unión a la región intergénica y activando así la transcripción de los genes de la isla. Hay una gran

diversidad de secuencia entre los represores Stl codificados por las distintas islas y dado que la interacción es específica, cada represor interacciona con distintas proteínas fágicas. Los bacteriófagos de *S. aureus* difieren en su capacidad para inducir diferentes SaPIs, dependiendo de si codifican para las proteínas fágicas inductoras y de la afinidad que presentan dichas proteínas por su unión al Stl.

Una vez se ha eliminado la represión, comienza la expresión de la integrasa (*int*) y la escisionasa (*xis*) para la correcta escisión de la isla. A continuación, se inicia la replicación en forma de círculo rodante, formando un concatémero de gran tamaño que posteriormente es procesado durante la encapsidación de la isla.

Las SaPIs carecen de genes que codifiquen para las proteínas estructurales que forman las cápsides y utilizan las que forman los fagos para su propia encapsidación y así poder movilizarse. No obstante, la mayoría de SaPIs codifican algunas proteínas que interfieren con los fagos y forman el llamado operón I. Este operón, controlado por LexA, contiene una subunidad pequeña de terminasa (TerS) que interacciona con la subunidad grande del fago y favorece la encapsidación del genoma de la SaPI en detrimento de la encapsidación de genoma fágico. Además, en este operón, también se encuentran las proteínas CpmA y CpmB que redirigen la producción de cápsides disminuyendo su tamaño, donde únicamente el genoma de la isla puede empaquetarse. PtiA y PtiM también se localizan en el Operón I y dificultan la transcripción de los genes fágicos relacionados con morfogénesis y lisis.

Recientemente se ha identificado una nueva familia de SaPIs que no contienen el clásico operón I. Éste, ha sido sustituido por un segmento con un sitio *cos* y una proteína homóloga a una proteína de cápside de fago (Ccm). Estas SaPIs, además, no contienen una subunidad pequeña de la terminasa (TerS). Pueden ser encapsidadas por fagos *pac* y *cos* utilizando únicamente la terminasa pequeña del fago. Además, la proteína Ccm causa interferencia con los fagos y parece necesaria para la formación de cápsides pequeñas.

Además de la amplia distribución de las islas dentro de los estafilococos, se ha observado la transferencia de estos elementos entre distintos géneros y especies

bacterianas. En concreto se demostró la transferencia de SaPIs a *Staphylococcus epidermidis*, *Staphylococcus xylosus* y *Listeria monocytogenes.*

Recientemente, se han identificado elementos similares a las SaPIs en otras especies de *Staphylococcus* spp y en otros géneros bacterianos como *Streptococcus* spp*, Enterococcus* spp y *Lactococcus* spp. Incluso, se han identificado en bacterias Gram negativas como *Escherichia coli* o *Pasteurella multocida*. Esto permitió la creación de una nueva familia de EGMs denominada PICIs (phage-inducible chromosomal islands).

Objetivos

Basado en lo que se ha expuesto en la introducción, es evidente que los bacteriófagos y las SaPIs son dos piezas fundamentales en la transferencia horizontal de genes y en la diseminación de factores de virulencia entre poblaciones bacterianas. Por tanto, el presente trabajo amplia el conocimiento sobre estos EGMs lo que permitirá abordar las infecciones bacterianas desde otra perspectiva. En base a dichos antecedentes se propusieron tres objetivos específicos de la tesis doctoral.

1º Caracterización de SaPIs en cepas clínicas de *S. aureus*

Las SaPIs están implicadas en diferentes procesos relacionados con virulencia como la producción de toxinas, la formación de biofilm o la adaptación al hospedador. Sin embargo, la relevancia de estos EGMs en cepas clínicas no ha sido estudiada en profundidad. El primer objetivo de la tesis es determinar la presencia de SaPIs en cepas clínicas. Además, debido a que las SaPIs tienen una estructura conservada organizada en módulos funcionales, se determinó la presencia de estos módulos dando especial relevancia al módulo de integración y regulación, así como al contenido de genes de virulencia.

2º Estudio de la movilización de SaPIs por fagos endógenos en cepas clínicas de *S. aureus.*

Hasta la fecha, el estudio de la movilización de SaPIs siempre se había realizado en cepas de laboratorio. En cambio, recientes evidencias han demostrado que

antibióticos de uso clínico (ciprofloxacina o trimetoprima) pueden provocar la inducción de profagos. La activación de los fagos puede producir la inducción de SaPIs y la consecuente diseminación de los factores de virulencia que contienen. Como resultado de este proceso, antibióticos que se usan para tratar infecciones pueden agravar los riesgos asociados con la patogenicidad de la bacteria. Por tanto, el segundo objetivo de la tesis es estudiar la movilización de SaPIs en el ambiente clínico. Este estudio es fundamental para entender la patogénesis de este microorganismo y considerar el efecto de los antibióticos en la propagación de factores de virulencia.

3º Estudiar el papel de las SaPIs como reguladoras de genes de virulencia cromosómicos.

Diversos estudios han reflejado que las SaPIs no son solo portadoras de genes de virulencia. Además, son capaces de mediar la transferencia de otras zonas cromosómicas, interferir con la reproducción de los bacteriófagos, incluso pueden determinar la especificidad del huésped. El hecho de que las SaPIs estén implicadas en otros procesos además de los descritos hasta ahora las hace todavía más relevantes en la evolución y adaptación del microorganismo. El tercer objetivo de la tesis es determinar si las SaPIs están implicadas en la regulación de genes cromosómicos.

Metodología y Resultados

Caracterización de SaPIs en cepas clínicas

Para el estudio de caracterización de los EGMs en cepas clínicas, se utilizaron cepas de *S. aureus* aisladas de pacientes con fibrosis quística (FQ) y cepas de *S. epidermidis* procedentes de infecciones protésicas y de catéteres. Éstas fueron cedidas por el servicio de microbiología del Hospital Universitario y Politécnico la Fe de Valencia.

Se extrajo el ADN de las 200 cepas de *S. aureus* procedentes de pacientes con FQ y se determinó la presencia de SaPIs y bacteriófagos mediante PCR y Southern blot. Las SaPIs se pueden clasificar en función de su integrasa en cinco familias

principalmente, las cuales están altamente conservadas y permiten que estos EGMs se integren en sitios específicos. Al igual que las SaPIs, los bacteriófagos también se clasifican en función de sus integrasas en 7 tipos.

Los resultados obtenidos mostraron que el 100% de las cepas poseían fagos en el genoma y el 58,3% de las cepas contenían más de uno tipo de bacteriófago, siendo más frecuentes los fagos con integrasa tipo 3, 2 y 5 (77,5%, 37% y 28%). En el caso de las SaPIs, el 50% de las cepas contenían al menos una SaPI integrada en el genoma y 23,8% contenían más de un tipo de isla. Las SaPIs más frecuentes fueron las que contenían integrasa tipo I, V y III (42%, 36% y 35%).

Por otro lado, se estudió la presencia de estos EGMs en otras especies, para ello, se utilizaron 48 cepas de *S. epidermidis* aisladas de infecciones relacionadas con dispositivos médicos. Siguiendo la misma metodología, se identificaron las SePIs que contenían en el genoma. El 83,3% de las cepas contenían una SePI integrada en su genoma y el 43,75% contenían más de un tipo de SePI. Las SePIs más frecuentes fueron las que codificaban para integrasas tipo V (85%).

Algunas de las cepas fueron secuenciadas y mediante el uso de herramientas bioinformáticas se confirmó la presencia de SaPIs, SePIs y bacteriófagos. Por otro lado, se caracterizaron los distintos módulos funcionales de las SaPIs y SePIs como son los módulos de integración y regulación y se identificaron los genes relacionados con virulencia.

Movilización de SaPIs por fagos endógenos en cepas clínicas de *S. aureus*

Teniendo en cuenta la clasificación de las SaPIs en el apartado anterior, se determinó la inducción de SaPIs por bacteriófagos endógenos mediante Southern blot con sondas específicas de las distintas integrasas. Las cepas fueron sometidas a tratamiento con mitomicina C (MC) para activar la respuesta SOS, inducir los profagos y estudiar si estos fagos son capaces de inducir a las SaPIs residentes. Para ello, se tomó muestra a tiempo 0 y tras 90 minutos del tratamiento con MC, las muestras fueron procesadas, se corrieron en un gel de agarosa y se transfirieron a una membrana de nylon. Posteriormente, fueron hibridadas con sondas de cada una de las integrasas y se pudo determinar si la SaPI estaba integrada en el cromosoma

(SaPI no inducida) o si por el contrario se había escindido del cromosoma (SaPI inducida). De las 100 cepas de *S. aureus* analizadas, 22 cepas poseían SaPIs inducidas por fagos endógenos.

Con el fin de determinar si los resultados obtenidos con cepas procedentes de nuestro Hospital y de un nicho concreto como eran los aislados procedentes de pacientes con fibrosis quística eran extrapolables a otros nichos y con cepas de distinta procedencia, se analizaron 14 cepas clínicas descritas en la bibliografía. Utilizando la misma metodología se determinó que 7 cepas poseían SaPIs que eran inducidas por fagos endógenos por lo que se confirmó que las SaPIs son inducidas en un alto porcentaje por fagos endógenos en cepas clínicas

Asimismo, la inducción de las islas también fue estudiada en *S. epidermidis*. Se determinó que, de 40 cepas, únicamente una cepa contenía tres SePIs y dos de ellas eran inducidas por fagos endógenos.

La proteína hipotética con dominio de función desconocida DUF3113 del fago Sa2mw es responsable de la inducción de SaPImw2

Una vez se confirmó la movilización de SaPIs en cepas clínicas, se estudió en profundidad el mecanismo de movilización de SaPImw2 contenida en la cepa MW2. La cepa MW2 contiene dos profagos (ΦSa2mw y ΦSa3mw) y una SaPI (SaPImw2) la cual codifica para 3 factores de virulencia.

Para estudiar la movilización de la isla SaPImw2, el primer paso fue introducir un marcador en la isla. Para ello se sustituyó un gen que codifica para una enterotoxina por un casete de resistencia a kanamicina (SaPI*entC::cat*). Se obtuvo un lisógeno del ΦSa2mw procedente de la cepa MW2 en la cepa de laboratorio no lisogénica RN4220 y se transdujo la isla marcada SaPI*entC::cat.* Mediante Southern blot se demostró que el fago ΦSa2mw era el responsable de la inducción de SaPImw2.

Para identificar que proteína del fago era la responsable de interaccionar con $Stl_{SaPImw2}$, se realizó una fusión transcripcional con la región reguladora de la isla. Para ello se utilizó el plásmido pCN41 y se clonaron los genes *stl* y *str* de SaPImw2

fusionados con el gen de la β-lactamasa. Por otro lado, se utilizó el plásmido pCU1 en el que se había introducido previamente un promotor inducible por cadmio y se clonaron las distintas regiones del fago. Ambos plásmidos fueron transformados en la cepa no lisogénica RN4220, se indujo con $CdCl_2$ y se cuantificó la actividad β-lactamasa. El gen MW1424 que codifica para una proteína de función desconocida con dominio DUF3113 parecía ser el responsable de la desrepresión de la isla SaPImw2.

Para confirmar el papel como inductor de SaPImw2, se debía demostrar que la proteína hipotética DUF3113 se unía al represor $Stl_{SaPImw2.}$ Para ello, se clonaron ambos genes en el vector pET28a fusionados con una cola de histidinas. Se llevó a cabo la purificación mediante cromatografía de afinidad utilizando una columna con resina de níquel y se realizaron experimentos de pull-down. Además, las proteínas purificadas se sometieron a cromatografía por exclusión de tamaño. Se introdujo His_6-$Stl_{SaPImw2}$ y His_6-$DUF3113_{ΦSa2mw}$ por separado y mezcladas en la columna y el patrón de picos fue diferente. His_6-$Stl_{SaPImw2}$ eluía en dos picos, 116 y 165 mL y His_6-$DUF3113_{ΦSa2mw}$ eluía en un pico a 187 mL. En cambio, la mezcla His_6-$Stl_{SaPImw2}$ con His_6-$DUF3113_{ΦSa2mw}$ además de los picos enumerados previamente, eluía en un nuevo pico a 135 mL que correspondía a la interacción de las dos proteínas. El contenido proteico de cada uno de los picos eluídos fue comprobado mediante electroforesis en gel de poliacrilamida (SDS-PAGE) y se confirmó la interacción de $Stl_{SaPImw2}$ con $DUF3113_{ΦSa2mw.}$

Finalmente, se estudió la capacidad de otras variantes alélicas del gen DUF3113 presentes en otros fagos para inducir SaPImw2. Primero se comprobó mediante fusión transcripcional la capacidad de los fagos 80α, Φ11 y Φ12 para desreprimir SaPImw2. El plásmido pCN41*stl-str*$_{SaPImw2}$ fue introducido en los diferentes lisógenos (RN4220 80α, RN4220 Φ11, RN4220 Φ12) y se cuantificó la actividad β-lactamasa tras la inducción con MC. Tanto el lisógeno del fago 80α como el del Φ12 mostraron actividad β-lactamasa. Estos resultados se confirmaron por Southern blot.

En paralelo, se clonaron las distintas variantes alélicas del gen MW1424 de los fagos 80α, Φ11, Φ2c y Φ6c marcadas con la etiqueta 3xflag en el plásmido pCN51.

Todas ellas se clonaron con la misma zona de unión al ribosoma con la intención de que la expresión de estos genes fuera la misma en todos los casos. Estos plásmidos se transformaron en la cepa RN4220 SaPImw2*entC::cat* y se indujeron con $CdCl_2$. Se realizó un Western blot para corroborar que la expresión de las distintas proteínas era similar. Mediante Southern blot se mostró la distinta capacidad de las proteínas DUF3113 para inducir SaPImw2. DUF3113 del fago 80α y Φ6c fueron capaces de inducir SaPImw2, en cambio, DUF313 del fago Φ11 y Φ2c no eran capaces de inducirla.

Por último, se realizaron análisis *in silico* para determinar si DUF3113 estaba presente en otras cepas del género *Staphylococcus*.

Parte de estos resultados han sido publicados en Cervera-Alamar *et al.*, 2018.

Las SaPIs controlan la expresión de genes cromosómicos

Con el fin de estudiar la posible implicación de las SaPIs en la expresión de genes cromosómicos, se realizó un "tiling array" de distintas cepas de *S. aureus*. Las cepas escogidas fueron RN450, RN450 + 80α, RN450 + SaPIbov1 (SB1) + SaPI1, RN450 + 80α + SaPIbov1 (SB1) + SaPI1 y RN450 SaPIbov1Δ*stl*. Se estudiaron dos condiciones: sin inducir y tras 60 minutos de inducción con MC. El análisis de los resultados mostró que las SaPIs estaban implicadas en la expresión de varios genes cromosómicos. Pero, en particular, se producía un aumento de la expresión del operón *crtOPQMN* y del gen SAOUHSC_02886 en todas las condiciones en las que la SaPI estaba inducida.

El operón *crtOPQMN* contiene 5 genes que son responsables de la biosíntesis del carotenoide estafiloxantina (STX). STX es el pigmento que le proporciona el color dorado característico a *S. aureus,* pero también es considerado un importante factor de virulencia. Éste actúa como un antioxidante que permite la desintoxicación de las especies reactivas de oxígeno (EROs) generadas por el sistema inmune del hospedador como O_2, H_2O_2 o HOCl. Las bacterias que carecen de los pigmentos carotenoides crecen normalmente, pero son eliminadas rápidamente por las EROs de los neutrófilos del hospedador. Estudiar en profundidad el mecanismo de

biosíntesis de estafiloxantina para poder bloquearlo lo hace una diana terapéutica potencialmente atractiva.

El gen SAOUHSC_02886 codifica para una proteína con función desconocida con un dominio acetiltransferasa (gnat). Miembros de esta familia están relacionados con la resistencia a antibióticos aminoglicósidos como la kanamicina, gentamicina y tobramicina. La sobreexpresión de estos dos loci se validó mediante qRT-PCR.

Las proteinas de la isla SaPIbov1 PtiA y PtiM son responsables de la producción de STX

Previamente se había observado que un mutante en el represor de SaPIbov1 (Δ*stl*) en el que la isla está replicando constantemente tenía una coloración amarilla intensa. Para comprobar que era debido a la expresión de STX se realizó un mutante en el operón *crtOPQMN* y se demostró que la coloración era debida a STX, ya que dicho mutante era blanco a pesar de la sobrexpresión de la isla. Por otro lado, para identificar que genes de SaPIbov1 eran responsables de la sobrexpresión de STX, nos centramos en el Operón I, ya que este se encuentra regulado por LexA y por tanto es sensible a la respuesta SOS. El mutante del operón I de SaPIbov1 mostró coloración blanca tras la inducción con MC, por tanto, era el operón responsable de la inducción de STX.

Con el fin de conocer que ORFs del operón I SaPIbov1 estaban implicadas en la inducción de STX, se obtuvieron mutantes de cada uno de los genes y se observó cómo afectaba a la coloración de los precipitados tras la inducción con MC. Estos resultados mostraron que las ORFs 6 y 7 de SaPIbov1 eran las principales implicadas en el proceso. Para confirmar este resultado, se complementó la cepa RN450 (que no contiene ninguna SaPI) con plásmidos de expresión (pCN51) que contenían las ORF 6, ORF 7 y las ORF 6-7 y se analizó la coloración. Los resultados indicaron que la ORF7 era la responsable de la sobreexpresión de STX y que la ORF6 contribuía en el proceso aumentando dicho efecto. Los resultados de sobreexpresión de STX por parte de las ORF6-7 de SaPIbov1 fueron confirmados mediante qRT-PCR.

En SaPI2 se han descrito homólogos a ORF 6 y 7 de SaPIbov1 y fueron nombrados como *ptiA* y *ptiM* respectivamente. Estos genes están relacionados con mecanismos de interferencia con fagos. PtiA (ORF 6) se une al activador transcripcional fágico LtrC y bloquea la producción de proteínas del virión y de lisis. PtiM (ORF 7) se une directamente a PtiA y modula su actividad. Por tanto, ambas proteínas son multifuncionales. Además de interferir con el ciclo de los bacteriófagos, son reguladoras de genes cromosómicos implicados en virulencia.

Un análisis comparativo *in silico* mostró que había proteínas similares a PtiA y PtiM en otras SaPIs. Se clonaron diferentes parejas de genes *ptiA-M* procedentes de distintas SaPIs (SaPI2, SaPIeq y SaPIpt1028) en el plásmido de expresión pCN51 inducible por $CdCl_2$ y se transformó por electroporación en la cepa RN4220. Tras la inducción con $CdCl_2$ se produjo un aumento en la pigmentación en todas las cepas.

Las proteínas PtiA y PtiM regulan la inducción de STX mediante un mecanismo de antiterminación

Para determinar el mecanismo por el cual las proteínas PtiA y PtiM controlaban la expresión de STX, se estudió la hipótesis de que ambas proteínas actuaban como factores de transcripción y se unían al promotor del operón *crtOPQMN* para iniciar el proceso de transcripción. Para ello se realizó un mutante en el promotor del operón (Δ*PcrtO*) en la cepa RN450 y se complementaron con el plásmido de expresión (pCN51) en el que se clonaron los genes *ptiA-M*. Tras inducir la cepa con $CdCl_2$ se cuantificó la pigmentación. La deleción en el promotor no tuvo ningún efecto negativo en la producción de STX y se descartó la posibilidad de que PtiA-M actuaban como factores de transcripción mediante la unión directa al promotor.

Con el fin de elucidar el mecanismo de regulación, se estudió cuidadosamente el perfil de expresión del operón *crtOPQMN* y las zonas flanqueantes entre la cepa que poseía las islas integradas en el genoma (RN450 +SaPIbov1 + SaPI1) y la cepa en la que la isla estaba continuamente inducida replicándose (RN450 SaPIbov1Δ*stl).* La principal observación fue que la expresión de la cepa RN450 SaPIbov1Δ*stl* empezaba en el gen *isaA* (SAOUHSC_02887) y no decaía hasta el gen *copA*

(SAOUHSC_02874). Parecía que los genes se transcribían conjuntamente en un solo transcrito de alrededor de 13400 pares de bases. En cambio, el perfil de expresión de la cepa RN450 +SaPIbov1 + SaPI1 era el esperado. Los niveles de expresión eran altos en las zonas codificante del gen y bajos en las regiones intergénicas. Los genes parecían expresarse de manera independiente formando distintos transcritos. El comportamiento en relación a la expresión de la cepa RN450 SaPIbov1Δ*stl* sugería que la inducción de la isla provocaba que los genes flanquenates al operón *crtOPQMN* se cotranscribiesen. Un mecanismo posible podía ser la antiterminación procesiva en el que se evita que la RNA polimerasa interactúe con secuencias terminadoras y transcriba varios genes en un solo transcrito.

Para confirmar el mecanismo de antiterminación como consecuencia de la expresión de PtiA-M, se llevaron a cabo varias RT-PCRs. Primero se extrajo el RNA de las cepas SH1000 $\Delta\sigma^B$ pCN51 y SH1000 $\Delta\sigma^B$ pCN51_*ptiA-M* y se retrotranscribió a cDNA. Este cDNA se utilizó como molde para las PCRs. Se diseñaron 5 parejas de cebadores que hibridaban con las regiones intergénicas presentes a lo largo de toda la región estudiada. En el caso de que existiese más de un transcrito no se obtendría producto de PCR con las parejas de cebadores de las regiones intergénicas, mientras que si por el contrario sólo existiese un transcrito (desde SAOUHSC_02887 hasta SAOUHSC_02872) se obtendría producto de PCR en todas las reacciones. Todas las PCRs correspondientes a las regiones intergénicas estudiadas fueron positivas cuando se sobreexpresaba *ptiA-M* siendo negativas en ausencia de dichas proteínas. Estos resultados confirmaron la formación de un solo transcrito cuando PtiA-M estaban presentes y sugerían un posible mecanismo de antiterminación que se iniciaba en el *isaA*.

Para confirmar que la antiterminación se producía en el gen *isaA*, en la cepa RN450 SaPIbov1Δ*stl*, se llevó a cabo la mutación de su promotor, así como de otros promotores que se encontraban en el posible transcrito: *gnat, oatA* y *ssaA.* Las cepas mutantes fueron complementadas con el plásmido pCN51_*ptiA-M* y se calculó la producción de carotenoide tras el tratamiento con $CdCl_2$. Únicamente la deleción del promotor del gen *isaA* causó un descenso en la pigmentación de los precipitados. Por tanto, se determinó que la antiterminación se producía en la región correspondiente a *isaA.*

Para entender el papel del gen *isaA* en la expresión del operón *crtOPQMN* se llevó a cabo la mutación de todo el gen (RN450 Δ*isaA*) y se sobreexpresaron los genes *ptiA-M.* Esta mutación se correlacionó con una ausencia total de pigmentación. Sin embargo, en un mutante puntual en IsaA (L10stop), si se produjo aumento de pigmentación cuando se sobreexpresaron los genes *ptiA-M*, por tanto, la proteína IsaA no estaba implicada en dicho proceso.

Una característica importante de los operones es que todos los genes se transcriben en la misma orientación. En este estudio se revirtió la orientación del gen *oatA* (aguas arriba del operón *crtOPQMN*) en la cepa RN450 y se estudió como afectaba a la producción de STX cuando se sobreexpresaban los genes *ptiA-M*. El cambio de orientación bloqueó el mecanismo de antiterminación por el que los genes se encontraban en un solo transcrito y no hubo pigmentación pese a la sobreexpresión de PtiA-M

Conclusiones

1. Las PICIs están presentes en el 50% de los aislados clínicos de *S. aureus* analizados provenientes de pacientes con fibrosis quística y en el 83.3% de los aislados clínicos de *S. epidermidis* procedentes de infecciones asociadas a catéteres.
2. Las SaPIs pueden ser movilizadas por bacteriófagos endógenos en cepas clínicas. En concreto, 18,75% los aislados procedentes de pacientes con FQ y el 50% de los aislados clínicos de *S. aureus* de orígenes muy diversos que han sido analizados contenían SaPIs inducibles por sus profagos residentes.
3. Dos SePIs de un aislado clínico de *S. epidermidis* (SCN45) procedente de una punta de catéter infectada, son movilizadas por alguno de sus bacteriófagos endógenos.
4. El gen MW1424 del fago ΦSa2mw codifica para una proteína con función desconocida y dominio DUF3113. Esta proteína es la responsable de la inducción de SaPImw2 mediante la interacción directa con el represor Stl.
5. Variantes alélicas del gen MW1424 codifican para proteínas con distinta capacidad para inducir SaPImw2.

6. Las SaPIs parecen ser capaces de secuestrar más de un bacteriófago para poder movilizarse. SaPIc, procedente de la cepa C, es inducida por el bacteriófago Φ6c y encapsidada por el Φ2c. SaPImw2, integrada en la cepa MW2, es movilizada e inducida por el fago ΦSa2mw, pero la presencia de ΦSa3mw incrementa la frecuencia de transferencia.
7. Las SaPIs pueden regular la expresión de genes relacionados con virulencia contenidos en el cromosoma. En concreto, son responsables de la sobreexpresión del operón *crtOPQMN* y el gen SAOUHSC_02886.
8. La proteína PtiM es la principal responsable de la sobreexpresión del operón *crtOPQMN* y, por tanto, de la producción de STX cuando la isla está inducida. La proteína PtiA interacciona con PtiM y contribuye al aumento de producción de STX. Proteínas homólogas a PtiA y PtiM presentes en otras islas también son capaces de regular la expresión del operón *crtOPQMN*.
9. El mecanismo por el que PtiA y PtiM regulan la expresión de STX es independiente del mecanismo de regulación de SigB. También es independiente de la unión al promotor del operón *crtOPQMN* y corresponde con un mecanismo de antiterminación procesiva que se inicia en el gen *isaA* y termina en el gen *copA.*

1. Introduction

1.1 *Staphylococcus aureus*

1.1.1 Description

Staphylococcus aureus is a Gram-positive opportunistic bacterium that colonises the skin and mucosae of human beings and several animal species. The anterior nares are the primary reservoirs of *S. aureus*. About 30% of the human population carry *S. aureus* asymptomatically but it is also able to cause superficial to serious infections of almost all tissues (Peacock, De Silva, & Lowy, 2001) (Wertheim *et al*., 2005).

S. aureus is considered a versatile opportunistic pathogen due to its ability to cause disease in humans, mainly during moments of vulnerability associated with a decrease in immune system competition. The infection usually begins with the previous colonisation of *S. aureus* and when the natural skin barrier is disturbed (e.g., cut, surgical incision, wound or bite), then the bacteria can spread to deeper tissues. *S. aureus* causes a wide range of diseases, from superficial skin infections such as furuncles and atopic dermatitis, to generalised and severe infections such as endocarditis, necrotising pneumonia, osteomyelitis or septicaemia (Tong, Davis, Eichenberger, Holland, & Fowler, 2015). It also causes toxin-mediated diseases such as toxic shock syndrome, scalded skin syndrome or food poisoning (Otto, 2014).

The number of both hospital and community *S. aureus* infections is increasing and this microorganism becomes more virulent and resistant to different antibiotics (Foster, 2017). A global epidemic of methicillin-resistant *S. aureus* (MRSA) has spread over the last four decades and since 1997, there has been an increase in the number of cases with vancomycin-resistant *S. aureus* (VISA) (Hiramatsu *et al*., 1997) (Grundmann, Aires-de-Sousa, Boyce, & Tiemersma, 2006) (Chambers & DeLeo, 2009).

According to an EPINE study in 2017 (Study of Prevalence of Nosocomial Infections in Spain), *S. aureus* is the second microorganism responsible for community infections behind *Escherichia coli* as well as the third in nosocomial infections behind *E. coli* and *Pseudomonas aeruginosa*.

1.1.2 Virulence

A successful S. *aureus* infection relies on the pathogen´s ability to colonise the initial infection site, to adapt to the stresses of the host environment, to resist the host defences and to spread to other sites in the host. To archive the infection, *S. aureus* encodes for a myriad of virulence factors, such as adhesins, immune evasions proteins, stress resistance proteins, degradative enzymes or toxins.

The genome size of *S. aureus* varies between 2.8 and 2.9 Mb and contains around 2500 predicted protein-coding sequences (CDSs). *In silico* analysis suggests that the *S. aureus* genome is comprised of around 75% core genome and around 25% of accessory genome (Lindsay & Holden, 2004). The core genome is conserved and represents the genes present in all strains of a species. It typically includes housekeeping genes for cell envelope and regulatory functions (>97%), as well as genes related to virulence. In contrast, the accessory or dispensable genome is highly variable and leads to defining intraspecies diversity. The accessory genome contains genes with non-essential functions such as genes related to virulence or resistance to metals and drugs (Lindsay & Holden, 2006).

The core genome contains virulence genes that are thoroughly studied, such as the coagulase gene (*coa*) which facilitates the infection via its procoagulant and fibrinogen-binding activity (Moreillon *et al*., 1995), the lipase gene (*lip*) which interferes with the host granulocyte function (Hu, Xiong, Zhang, Rayner, & Chen, 2012) or the α and β haemolysins (*hla* and *hlb*,) responsible for the red blood cell lysis. In addition, new genes related to virulence are emerging, e.g., the N-acetyltransferase gene involved in aminoglycoside antibiotic resistance (Srivastava *et al*., 2014). Another example is the *crtOPQMN* operon that encodes for staphyloxanthin (STX), the yellowish-orange (golden) pigment of *S. aureus*. STX is linked to virulence due to its antioxidant properties. In the absence of STX, the microorganism becomes more susceptible to death by reactive oxygen species (H_2O_2, OH or ClO^-) produced by neutrophils (Clauditz, Resch, Wieland, Peschel, & Götz, 2006).

The accessory genome often includes factors that provide a survival advantage in specific environment and defense mechanisms. This includes cytotoxins such as Panton-Valentine leukocidin (PVL) or immune evasion cluster (IEC) that include molecules such as the staphylococcal complement inhibitor (encoded by *scn*), staphylokinase (Sak) or chemotaxis inhibitory protein (CHIPS). Enterotoxins also play a broader role as virulence factors, which are powerful non-specific T-cell stimulators that cause unregulated activation of the immune response. This family includes SEA, SEB, SEQ or SEL among others.

Besides the secretion of diverse virulence determinants, *S. aureus* forms biofilm. A biofilm can be defined as a community, typified by cells that are attached to a substratum and embedded in a matrix of extracellular polymeric substance. Biofilm formation protects *S. aureus* from opsonophagocytosis and antibiotic treatment, and is therefore considered an important virulence factor. Two major surface components have been implicated: the PIA/PNAG encoded by operon *icaADBC* and the Bap protein (Cramton, Gerke, Schnell, Nichols, & Götz, 1999)(Cucarella *et al*., 2004).

1.1.3 Regulation of virulence genes expression

During infection, pathogenic bacteria must be able to express their virulence genes properly, as well as survive in the environmental conditions imposed by the host (Toledo-Arana, Repoila, & Cossart, 2007). Different systems have been developed to enable the bacteria to sense the environment and subsequently pair a stimulus to a specific response. A plethora of RNAs and proteins with regulatory function have been discovered in bacteria to control these processes (Fechter, Caldelari, Lioliou, & Romby, 2014).

S. aureus rapidly adapts to different niches in diverse hosts. Given the high number of virulence factors present in these bacteria, gene expression must be finely tuned in order to efficiently coordinate their expression (Ibarra, Pérez-Rueda, Carroll, & Shaw, 2013). In this thesis, the regulatory control mechanisms are separated into transcriptional and posttranscriptional regulation.

Transcription needs the binding of RNA polymerase (RNAP) to a promoter DNA to form a complex, the reversible opening of DNA region, the incorporation of nucleoside triphosphates, the progressive elongation complex and the release of the complex at a termination signal. Each step is critical for transcription and constitute possible targets for transcriptional regulation.

A large group of proteins called activators have been shown to activate transcription by facilitating the binding of the RNAP to the promoter. They can interact with the RNAP complex or alter the promoter conformation. Transcription can also become negatively regulated by repressors. They can exert their activity binding to the same site as RNAP and thereby hinder the binding to the promoter. *S. aureus* genome encodes a number of transcriptional factors, including SarS, Rot or SigB among others. SarS is a positive regulator of *spa* transcription and Rot is a negative transcription regulator of α-hemolysin and protease expression (Mcnamara, Milligan-Monroe, Khalili, & Proctor, 2000). *lexA* and *recA* are regulatory genes of SOS response. LexA is a global repressor of SOS genes binding to their operator region but, DNA damage can induce the SOS response and activate the RecA protein. The activated RecA interacts with LexA repressor to facilitate the LexA repressor's self-cleavage from the operator which activates DNA-repair genes expression (Bisognano *et al*., 2004) (Cirz *et al*., 2007).

Gene expression can be controlled at the transcriptional elongation level through premature termination or antitermination mechanisms. The termination signals can be grouped into two classes: intrinsic and factor-dependent (Merino & Yanofsky, 2005). In the former, a specific sequence triggers the dissociation of the RNAP complex, and they do not require any factors but can be enhanced by accessory proteins (Greenblatt, McLimont, & Hanly, 1981) (Mondal, Yakhnin, Sebastian, Albert, & Babitzke, 2016). The termination of factor-dependent signals depends on the action of a regulatory protein such as Rho (Epshtein, Dutta, Wade, & Nudler, 2010). Instead, the antitermination mechanisms suppress the action of terminators in order to increase the expression of downstream genes (Santangelo & Artsimovitch, 2011). Antiterminators can target a terminator region and also manipulate the RNAP complex to avoid termination signals throughout an individual transcript (Weisberg & Gottesman, 1999). This type of antitermination

mechanism is called processive antitermination (PA) and causes the RNAP read-through of multiple consecutive termination sites, even over longer genomic distances.

Phage proteins from *E. coli* λ bacteriophage N and Q were the first reported examples of antiterminators. They control the expression of late phage genes related to morphogenesis and lysis. N requires several host factors, called Nus factors to bind to a specific structure in the mRNA and interact with the RNAP. N suppresses termination at both intrinsic and factor-dependent termination signals. Instead, Q interacts with double-stranded DNA, travels with the RNAP over long distances and modifies the enzyme into a termination-resistant form. Further examples include NusG and NusA, which increase transcription processivity by suppressing the RNAP pausing mechanism (Alexander V. Yakhnin, Murakami, & Babitzke, 2016). Paralogues to NusG, such as RfaH, Spt5 or LoaP interact with the DNA when associated with the RNAP resulting in the suppression of transcription pausing (Sevostyanova, Belogurov, Mooney, Landick, & Artsimovitch, 2011) (Goodson, Klupt, Zhang, Straight, & Winkler, 2017).

The steady-state level of a mRNA is dependent on transcription initiation as well as its stability (Le Scornet & Redder, 2018). Post-transcriptional regulatory factors include proteins or antisense RNAs which can interact with mRNA and promote stability or facilitate degradation. One of the most intriguing post-transcriptional regulation mechanisms is RNAIII from *S. aureus* which interact with a variety of mRNAs involved in virulence (R P Novick *et al*., 1993).

A different mechanism that controls the expression of certain proteins is phase variation which is one of the many strategies employed by pathogenic bacteria to avoid detection by the host immune system. It involves the ability to switch on the expression of proteins when they are needed, and to switch them off when they are likely to trigger immune responses. In *S. aureus*, a phase variation process controls expression of the *ica* and *bap* genes, and this process also modulates the functionality of different staphylococcal adhesins (M. Ángeles Tormo *et al.*, 2007) (Valle, Vergara-Irigaray, Merino, Penadés, & Lasa, 2007).

1.1.4 Mobile Genetic Elements

The accessory genome is primarily composed of mobile genetic elements (MGEs). Horizontal transmission of these MGEs is fundamental to the evolution of *S. aureus* and mainly contributes to its plasticity (Malachowa & Deleo, 2010). The MGEs described to date include plasmids, bacteriophages, pathogenicity and genomic islands, chromosomal cassettes and transposons. They all usually carry clinically relevant genes involved in virulence and antibiotic resistance as well as the enzymes that mediate their own transfer and integration into new DNA (Frost, Leplae, Summers, & Toussaint, 2005). MGEs are transferred horizontally and they are essential to rapid adaptation to different environments and hosts as well as the epidemiological success of the microorganism (Alibayov, Baba-Moussa, Sina, Zdeňková, & Demnerová, 2014). Our research group is focused on the study of *S. aureus* MGEs and their horizontal gene transfer (HGT) due to their fundamental role regarding adaptation and pathogenicity, mainly in bacteriophages and pathogenicity islands (SaPIs).

1.2 Bacteriophages

1.2.1 Overview

Bacteriophages, also known as phages, are viruses that exclusively infect bacteria. There are an estimated $>10^{30}$ phages in the biosphere and because phages typically outnumber bacteria by about 10-fold, they probably represent the most abundant “organisms” on earth in sheer numbers (Brüssow & Hendrix, 2002). Not surprisingly, phages constitute major players shaping bacteria communities in most ecosystems and are associated with virulence and the evolution of several important bacterial pathogens (Canchaya, Fournous, & Brüssow, 2004). Phages are the major vehicles of horizontal gene transfer (HGT) and contribute to virulence by encoding fitness and virulence genes, by the direct transfer of MGEs or by lysogenic conversion. Phages play an important role in regulating microbial balance due to their high level of specificity, long-term survivability and ability to quickly reproduce in a suitable host (Guttman, Raya, & Kutter, 2004).

1.2.2 Bacteriophage life cycle

Like all viruses, phages are obligate intracellular parasites that depend on specific host bacteria for reproduction. Upon infection, phages bind to specific targets present in the bacterial surface and inject their genome into the cytoplasm (Clokie, Millard, Letarov, & Heaphy, 2011). They can then begin two different lifecycles: the lytic cycle (lytic phages) or the lysogenic cycle (lysogenic phages).

The lytic phage replicates inside the cell and utilises the ribosomes of the host to manufacture its proteins. It then, lyses the bacteria and the virions are liberated and ready to infect susceptible cells in the environment. The lysogenic phage (also called prophage) also attaches to a bacterium and introduces the genome into the host cell. However, in this case, the phage genome is integrated into a specific location of the host chromosome. The prophage genome passively replicates along with the host genome and is transmitted to daughter cells at each cell division without killing the bacterium. However, exposure to altered environmental conditions such as antibiotic treatment, oxidative stress or DNA damage can turn the prophage into a lytic phage, allowing it to undergo the lytic cycle and escape from the host (Figure 1.1).

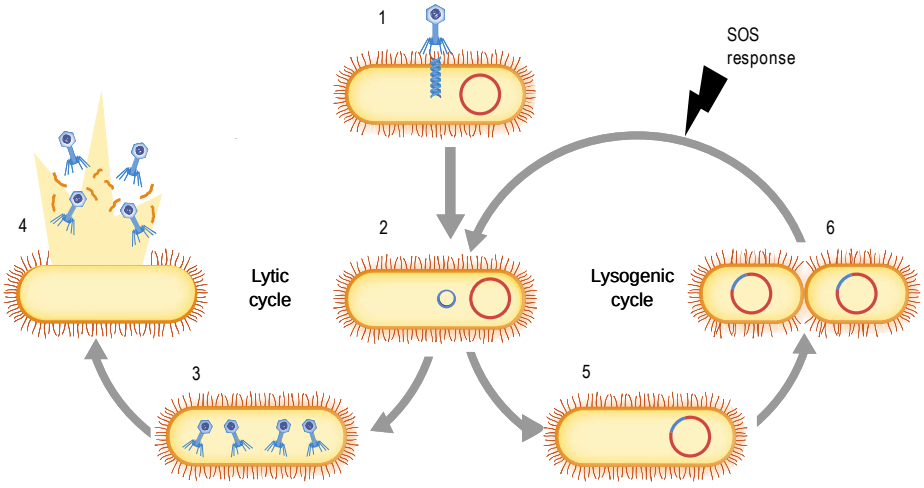

Figure 1.1 Life cycle of bacteriophages. The lytic cycle of a bacteriophage consists of the adsorption to the bacterial cell wall (1), injection of the phage genome into the cytoplasm (2), replication of the genome, formation of phage structural components, assembly of phage parts (3) and release from the host (4). Instead, in the lysogenic cycle, the phage also attaches to the cell surface (1) and injects the DNA into the cell (2) but the phage DNA is integrated into the bacterial chromosome (5) and replicates along with it (6). Adapted from Keen & Dantas, 2018.

1.2.3 Staphylococcal bacteriophages

S. aureus bacteriophages are extremely diverse concerning size, morphology and genomic organisation. However, they all consist of two main protein structures: DNA-filled head-like capsid and a syringe-like tail fibre (Ackermann, 1975). *S. aureus* bacteriophages are usually composed of double-stranded DNA (dsDNA), are located in an icosahedral capsid and belong to the *Caudovirales* order (Maniloff & Ackermann, 1998). According to the tail morphology, they can be classified into three families: *Myoviridae, Siphoviridae* and *Podoviridae*. *Myoviridae* have a long, rigid and contractile tail. *Siphoviridae* are non-contractile tail phages ended by a base-plate structure. *Podoviridae* are composed of a short, non-flexible and non-contractile tail (Figure 1.2). *S. aureus* bacteriophages' genome size ranges from 20 kb to 125 kb and encode from 27 to over 600 genes clustered in functions. Three genome size categories can be established that correlate with morphological classification. The *Myoviridae* family has the largest genomes (>125 kb), the *Podoviridae* family the smallest (<20 kb), while the *Siphoviridae* family regards an intermediate size (around 40 kb) (Kutter & Sulakvelidze, 2005).

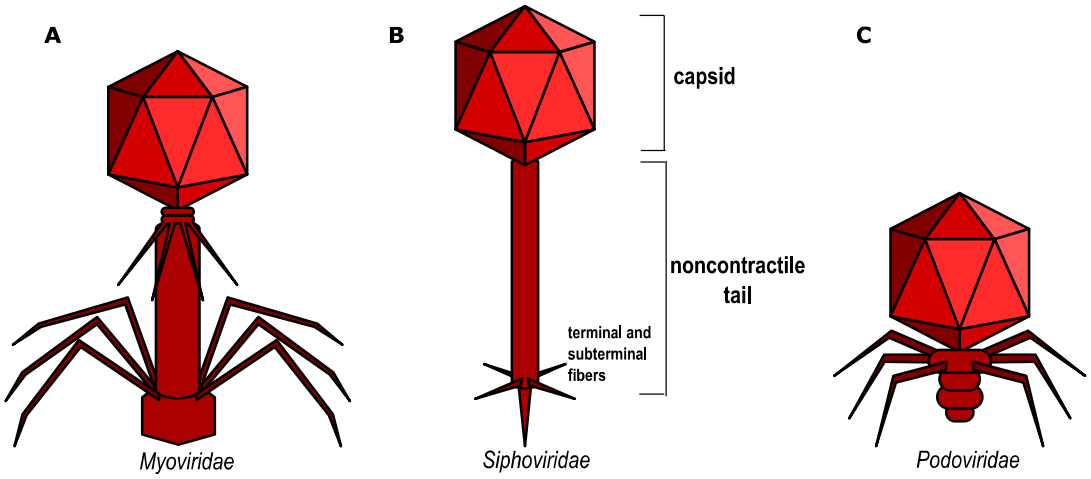

Figure 1.2 Prototypical structure of *Caudovirales* order bacteriophages. *Myoviridae* (*E. coli* T4), *Siphoviridae* (lambda λ) and *Podoviridae* (P22). Modified from Wikipedia.

The vast majority of *S. aureus* phages belongs to the *Siphoviridae* family, and thus this study is exclusively limited to them (Deghorain & Van Melderen, 2012).

Temperate phages can form a stable association with bacterial hosts, through a process known as lysogeny. *Siphoviridae* phages are responsible for different human diseases by encoding important toxins, such as exfoliatin A (scalded skin

syndrome), enterotoxin A (food poisoning) or PVL (necrotising pneumonia). The expression of temperate phages is able to confer novel properties to the bacteria and leads to phenotypic changes in the host. This phenomenon is called lysogenic conversion (Touchon, Moura De Sousa, & PC Rocha, 2017).

1.2.4 Genome organisation of *S. aureus* phages

Phage genomes display a modular organisation with genes grouped according to biological function, and therefore, modular organisation has been proposed as a parameter of phage classification. The lambda (λ) phage infects *E. coli* and represents one of the best studied phages. It is the prototype of lambdoid bacteriophages group in which the majority of *S. aureus* bacteriophages are restricted (Casjens & Hendrix, 2015). *S. aureus* phage genomes are usually organised into interchangeable functional modules and their expression proceeds in a controlled manner (Figure 1.3). Each module controls a distinct functional activity of the phage. The siphovirus genomes encode around 50-80 ORFs and are organised into the functional modules: lysogeny, DNA replication, virion morphogenesis (packaging, head, tail) and lysis (Brussow, Canchaya, & Hardt, 2004) (Xia & Wolz, 2014). The modular organisation facilitates the temporal control of transcription and gene expression in an established pattern in order to mediate phage life cycle. The modules, in turn, can be classified in three stages according to the expression sequence: early (lysogeny module), middle (replication and DNA metabolism module) and late (morphogenesis and lysis modules).

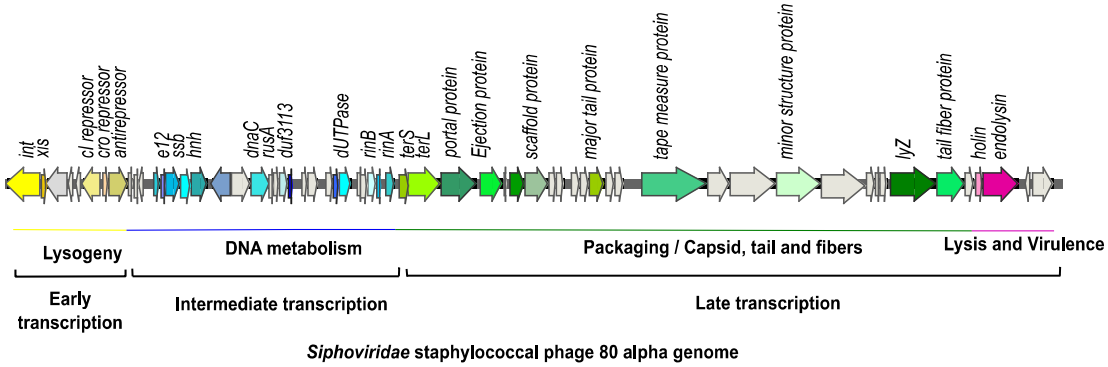

Figure 1.3 Genomic map of *Siphoviridae* staphylococcal 80α bacteriophage. The phage genome can be divided into different modules: lysogeny (yellow), DNA metabolism (blue), packaging (green) and lysis / virulence (pink).

1.2.5 Control of lysogeny

The lysogeny module contains two divergent ORFs that are characteristic of temperate phages and resemble the *E. coli* phage λ: the Ci-like repressor and the Cro-like anti-repressor (Oppenheim, Kobiler, Stavans, Court, & Adhya, 2005). They represent the key players in determining the life cycle of a temperate phage. The *cI*-like gene is responsible for maintaining the phage integrated in the chromosome during the lysogenic cycle, while the *cro*-like gene is a repressor of Ci and leads the phage to undergo the lytic cycle. Three homologous operators, O1, O2 and O3 are located between the two regulator genes.

CI contains a helix-turn-helix motif and a RecA cleavage motif. In the lysogenic phage, the CI repressor forms a dimer and binds to the adjacent operator in order to prevent the expression of the phage genes. The SOS response activates the bacterial protease RecA which produces the CI autolysis. Cleaved CI can no longer dimerise and lose the affinity for DNA binding. Without CI repression, Cro is expressed and inhibits CI expression (Benson & Youderian, 1989).

1.2.6 Integration and excision

Integration of the phage genome into the chromosome is vital to the progression of the lysogenic cycle. The phage genome integrates into the host through site-specific recombination. The integrase gene (*int*), phage attachment site (*attP*) and excisionase gene (*xis*) are involved in phage integration and excision. The integrase dictates the insertion site of the phage into the bacterial chromosome by catalysing the recombination between *attP* and an identical sequence found in the bacterial chromosome (*attB*) (G. E. Christie *et al*., 2010). This mechanism is so specific that a reliable classification for staphylococcal phages was developed based on the integrase sequence which is currently being used to cluster the prophages into seven groups (Cho, Gumport, & Gardner, 2002) (Goerke *et al*., 2009).

Once the phage genome is integrated, hybrid attachment sites border the prophage (*attL* and *attR*) and act as recognition sites for phage genome excision. Following either spontaneous or SOS-mediated prophage induction, Int and Xis act in a coordinated manner to catalyse phage DNA excision (Schindler & Echols, 1981).

1.2.7 DNA replication

Based on the classical λ phage model, phage DNA replication initiates at a single origin (*ori*) and proceeds bidirectionally from the *ori*, generating θ structures (Skalka, Poonian, & Bartl, 1972). Besides *ori*, the replication module also encodes two essential DNA replication genes: o-like and p-like genes. The O protein binds to the *ori* site while the P protein binds to the host replication machinery (DnaB) forming a binary complex that binds to O:*ori* (Zylicz, Liberek, Wawrzynow, & Georgopoulos, 1998). This complex effectively commandeers the host DNA polymerase which initiates the replication process. Finally, the phage DNA switches to rolling circle replication to generate concatemers that serve as substrates to package into phage proheads.

Prophages frequently excise from the chromosome after induction and then replicate to form concatemers. However, it was recently discovered that some *S. aureus* prophages have an atypical program and do not excise until later resulting in DNA replication and packaging from integrated viral genomes (J. Chen *et al.*, 2018).

1.2.8 Virion structure and assembly

Genes associated with the assembly of virions and lysis are clustered together and are expressed at the late stage of phage infection. Regulation of these modules is tightly controlled by phage specific regulatory mechanisms. One of the best studied regulatory system involves the positive regulator Q in phage λ, which is a transcription antitermination protein. Q interacts with a specific DNA site (*qut*), located just upstream of the late promoter, and with the elongation complex. Q travels with RNAP over long distances and modifies the enzyme into a termination-resistant form. This modification cases that the transcription does not cease at termination point and extends to the entire morphogenesis and lysis modules.

For *S. aureus* phages, a wide range of late transcriptional regulators (Ltr) has been described. Ltr activators bind to a promoter region upstream of the *terS* gene and activate the late gene transcription of morphogenesis and lysis gene clusters.

The bacteriophages assemble their capsids and tails with the help of internal scaffolding proteins, and genes for head and tail assembly are arranged together. The head genes typically include one or two terminase subunits, the portal protein, a prohead protease, a scaffold protein and the major capsid subunit. The tail genes include the major tail subunit, tapemeasure protein and minor tail proteins (Hatfull, 2008).

An initial step of the phage packaging process concerns the formation of a procapsid which is composed of multiple copies of the major capsid proteins. Whilst assembly of the procapsid occurs, the phage genome undergoes rolling-circle replication and generates dsDNA concatemers.

Following phage DNA replication, phage packaging begins with the cleavage of the recently formed concatemeric DNA. The genome packaging is initiated by the Terminase protein complex (Ter_{phage}) which comprises two subunits. The small subunit (TerS) recognises the DNA and recruits the large subunit (TerL) to cut the DNA concatemers (Fokine & Rossmann, 2014). The bacteriophages can be classified into two groups regarding packaging mechanism: *pac* phages and *cos* phages.

In the former, TerS binds close to the *pac* site on the concatemer and TerL makes the initiating cut at a nearby sequence. The complex TerS-TerL (Ter) remains bound to the DNA and also captures a virion prohead and translocate the DNA inside. Ter makes a second, nonspecific cut once the head is full (slightly longer than genome unit length). On the other hand, in the alternative mechanism by *cos*-phages, the Ter complex recognizes and cleavages a second *cos* site sequence, therefore packaging precise genome monomers (Figure 1.4). Once the package of phage DNA is completed, the terminase complex dissociates and the tail structure associates to the capsid complete the phage particle.

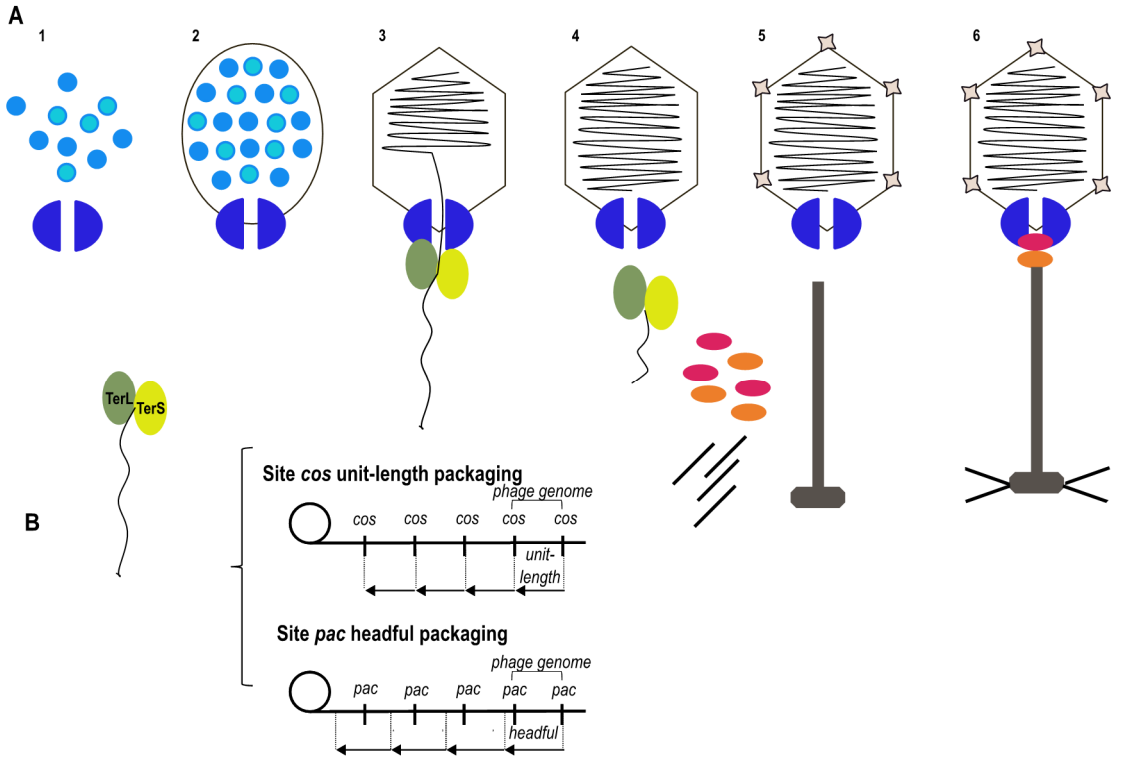

Figure 1.4 Assembly and packaging of virus particles. A (1) Initiation of head assembly by the portal, (2) assembly of major, minor and scaffold capsid proteins into immature prohead, (3) maturation of the prohead, (4) packaging of phage DNA, (5) neck assembly and (6) tail attachment. **B**. DNA packaging. In both models (*pac* and *cos*), TerS recognises and binds to a specific site and then, TerL makes the first cut at a nearby sequence. In *pac* phages, TerL cuts in a nonspecific sequence and cleavages after the capsid is completely filled with DNA, resulting in slightly more than one genome length (headful packaging). In *cos* phages, TerL makes the second cut in other *cos* site sequence and packages one genome length (unit-length packaging) in the phage particle. Adapted from Rao & Feiss, 2015.

1.2.9 Lysis

Once the phage particles are formed, the newly formed phages leave the host to infect new bacterial cells. The lysis module encodes a holin and an endolysin which are responsible for rupturing the host cell and releasing the phage progeny. Endolysins are accumulated in the cytoplasm of the host bacterium and are conformationally ready to degrade the bacterial cell wall. However, holins are produced and accumulated in the membrane until a specific time. Once a critical point is reached, holin creates pores that permeabilise the bacterial membrane and release the endolysins causing the complete bacterium lysis (R Young, 1992) (Ryland Young, 2014).

1.2.10 Bacteriophage transduction

A powerful contributor to bacterial evolution is horizontal gene transfer (HGT) through three mechanisms: transformation, conjugation and transduction. Transduction is mediated by phages and has the greatest impact on HGT. Phages can use different mechanisms to transduce host genome including generalised transduction, specialised transduction and lateral transduction.

Generalised transduction is a rare event in which bacterial DNA is packaged, instead of phage DNA, into phages heads by a headful mechanism. The process is called generalised transduction because the bacterial DNA fragment can be derived from any part of the host genome (Thierauf & Maloy, 2011). This is in contrast to specialised transduction mechanism, in which the prophage imprecisely excises from the host chromosome and the genes adjacent to the prophage are included in the new virus particle (Brammar, 2013).

Lateral transduction is a recently identified mechanism and represents the most powerful means of transduction yet discovered. In this process, the prophages delay excision and initiate DNA replication while integrated into the host genome. DNA packaging also initiates from the integrated prophages and several hundred Kb of the *S. aureus* genome are packaged in phage heads at a high frequency (J. Chen *et al*., 2018). Lateral transduction is unique regarding its efficiency and could explain how bacteria can acquire antibiotic resistance or virulence factors that accelerate the emergence of new pathogenic strains.

1.3 *Staphylococcus aureus* Pathogenicity Islands

1.3.1 Overview

Staphylococcus aureus pathogenicity islands (SaPIs) are a family of highly evolved MGEs that are transiently integrated into the bacterial chromosome. They are a cohesive subfamily within a larger family known as phage-inducible chromosomal islands (PICIs). SaPIs carry virulence genes and can be transferred at a high frequency due to certain bacteriophages whose life cycles parasitise (Richard P. Novick, Christie, & Penadés, 2010) (Martínez-Rubio *et al*., 2017).

SaPIs, the first and most thoroughly studied PICIs, were discovered in 1998 while investigating the etiology of toxic shock syndrome (TSS). This infection is caused by a toxin (TSST-1) initially believed to be part of a transposon, but then it was demonstrated to be mobilised thanks to bacteriophages and was carried to a new MGE family known as SaPIs (Lindsay, Ruzin, Ross, Kurepina, & Novick, 1998).

Almost all SaPIs share some common characteristics: (i) the genome size is around 15 kb, (ii) the G-C content differ from the rest of the chromosome, (iii) they are inserted at specific sites in a single orientation (*attC*) and flanked by direct repeats (*attL* and *attR*), (iv) like phages, SaPIs have a conserved and recognisable structure organised in functional modules.

SaPIs have a high impact on the virulence of *S. aureus* since these elements are involved in its adaptation to new hosts and in the development of new virulent clones. Most SaPIs contain at least one gene related to virulence. For example, SaPI1 encodes three superantigens; toxic shock syndrome toxin (*tst*), enterotoxin Q and K (*seq, sek*) or SaPImw2 contains three enterotoxins (*ear, se2l* and *sec4*) (Richard P. Novick & Ram, 2016). They are highly evolved MGEs descended from ancestral prophages and their special characteristic role is the relation to certain functional prophages, known as helper phages. These helper phages induce the SaPIs and transduce them with extremely high efficiency to other bacteria.

1.3.2 SaPIs genome organisation

SaPIs encode approximately between 18 to 25 ORFs and their genome is organised into five different modules. Each of them provides a different function for the SaPI such as mediation of SaPI regulation, integration and excision, autonomous replication, packaging-interference and virulence (Figure 1.5).

The integration-excision module consists of a site-specific integrase gene (*int*) and an excisionase gene (*xis)* (Ruzin, Lindsay, & Novick, 2001) (Maiques *et al*., 2007). Like prophages, SaPIs reside integrated into the host chromosome until they are induced. Five different SaPI insertion sites have been identified, each characterised by a core *attC* sequence of 15-22 nt. The *attC* are completely different from each other and are present only once in the genome. The integrase catalyses the

integration of the SaPI into the bacterial chromosome. According to the integrase, the SaPI integrates into a specific *attC* site, and to date no more than one SaPI has been detected in each *attC* site. After induction, *xis* gene is necessary for SaPI excision from the chromosome (Mir-Sanchis *et al.*, 2012).

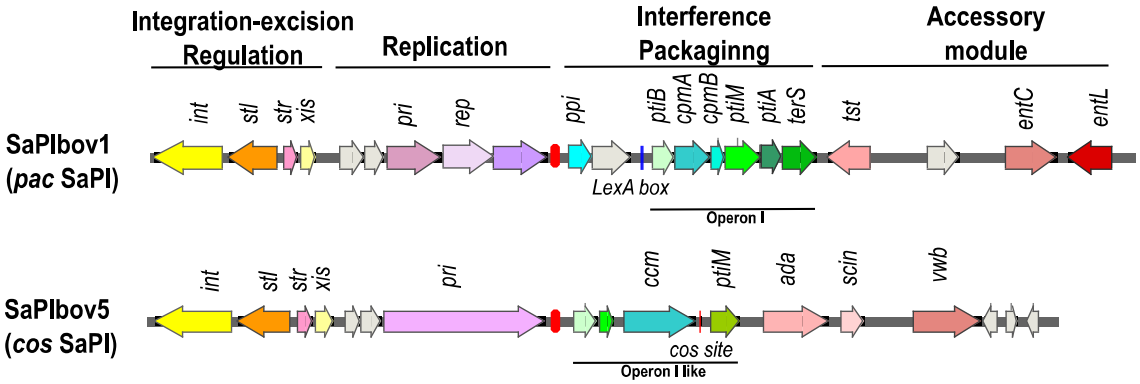

Figure 1.5 Genomic map of prototypical SaPIs. SaPIbov1 as *pac* SaPI and SaPIbov5 as *cos* SaPI. Genes are coloured according to their sequence and function: *int* and *xis* are yellow, repressor gene *stl* is orange and *str* pink, replication genes (including primase (*pri*) or replication initiator gene (*rep*)) are purple, encapsidation genes are green (operon I), accessory genes are red and genes with hypothetical functions are grey.

Another important module is the transcriptional regulation module that is composed of two genes: the global repressor *stl* and the derepressor *str*. Both are codified in divergent orientation, *str* in rightward and *stl* in leftward direction. If there are no stimuli, the SaPI resides quiescently in the bacterial genome due to the repressor Stl (M. Á. Tormo-Más *et al.*, 2010). During this condition, Stl binds to a region between *stl* and *str* which inhibits the expression of the SaPI genome and promotes the maintenance into the bacterial chromosome. (Papp-Kádár, Szabó, Nyíri, & Vertessy, 2016). On the other hand, Str seems to mediate the SaPI gene transcription, but the function has not yet been clearly determined since the *str* deletion had no influence on the SaPI cycle (Úbeda *et al.*, 2008).

The replication module includes a primase (*pri*) and a helicase gene (*rep*) which are sometimes fused and recognise a specific replication origin (*ori*) and bind to it. The binding initiates melting within the origin and Rep begins to unwind. Replication is catalysed by the bacterial DNA polymerase and generates hundreds of copies of the SaPI (Ubeda, Tormo-Más, Penadés, & Novick, 2012).

The interference/packaging module is divided into two clusters separated by an intergenic region that comprises the LexA-regulated promoter and the specific *pac* site (in the case of *pac* SaPIs) recognised by TerS. The first cluster includes the phage packaging interference gene (*ppi*) responsible for the interference with the DNA packaging of the helper bacteriophages. Ppi ensures advantageous packaging of the SaPI instead of the phage. Followed by this module and distal to the LexA promoter, there is an operon known as "Operon I" that contains genes essential for the SaPI packaging. Among them, the CpmA and CpmB direct the phage virion to form small procapsids commensurate with the SaPI genome (Poliakov *et al*., 2008) (Damle *et al*., 2012). Other proteins such as PtiA, PtiB and PtiM interfere with the phage by blocking the late gene expression operon (Geeta Ram, Chen, Ross, & Novick, 2014). This module also includes a *terS* gene which recognises the SaPI *pac* site and facilitates the SaPI encapsidation in the phage virions. The operon I can be expressed independently since it is regulated by the LexA promoter, even though the SaPI is not induced (Úbeda *et al*., 2007). Recently, it has been identified a new family of SaPIs, known as *cos* SaPIs, that lack the classical "Operon I". They have substituted *terS* gene for a 90 bp fragment (*cos* site) and *cpmA-B* for *ccm* gene.

SaPIs are related to pathogenicity because they encode many virulence factors in the accessory module. Although this module is usually located after Operon I but virulence genes can sometimes be located between *int* and *stl*. Expression of these genes seems to occur independently of the SaPI and depends on their own promoters. The accessory genes include superantigens such as enterotoxins B, C, K, L or Q, TSST or the staphylococcal complement inhibitor (SCIN).

1.3.3 SaPIs mobilisation

Similar to prophages, SaPIs remain stably integrated into the chromosome, but with the difference that excision and replication processes are not directly induced by the SOS response. Instead, they become mobilised by specific helper phages resulting in the packaging of the SaPI genome into phage-transducing particles. SaPIs do not codify the morphogenic genes necessary for the capsid formation, but they hijack the phage capsids and pack inside them to be mobilised (Úbeda *et al*.,

2005). The resulting particles can transfer the virulence factors contained in SaPIs to other cells, even across genera, at high frequency.

The SaPI life cycle begins with the infection of a helper phage or by the activation of an endogenous helper prophage through the SOS response. This response can be caused by oxidative stress, exposure to UV light or antibiotic treatment. Once the bacteriophage is active, an interaction between a phage protein and the Stl causes the initiation of the excision-replication-packaging (ERP) cycle of the SaPI (Figure 1.6).

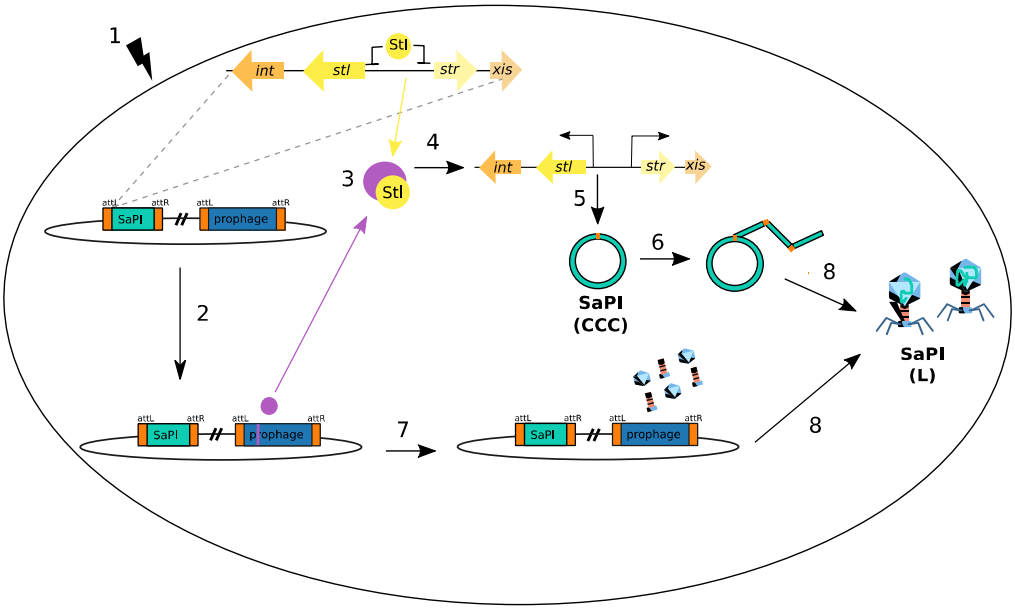

Figure 1.6 SaPI Excision-Replication-Packaging (ERP) cycle. (1) Activation of SOS response after mitomycin C addition. (2) The prophage is induced and the gene transcription starts, while the SaPI passively resides in the genome under the control of the global repressor Stl. (3) The interaction between the Stl and a specific phage derepressor protein induces the SaPI cycle. (4) The SaPI gene transcription starts. (5) SaPI is excised from the chromosome in a covalently closed circular (CCC) state and (6) starts its replication. (7) The phage synthetises the procapsids and tails. (8) The phage is ready to lyse the bacterium (SaPI L indicates linear monomers released from the phage heads). (Cervera-Alamar *et al.*, 2018).

1.3.3.1 Role of *stl* gene

All SaPIs codify the transcriptional regulator *stl* that is the responsible for repressing the replication of SaPIs. It binds to the intergenic region between the divergent SaPI regulatory genes *stl* and *str* blocking SaPI genes' expression (Úbeda *et al.*, 2008) (Papp-Kádár *et al.*, 2016). Stl has functional homology with the CI

repressor of λ phage but, unlike the phage, it is not degraded by auto-proteolysis following SOS response since it lacks the necessary conserved Ala-Gly motif.

SaPI induction is initiated by the inactivation of the Stl as a repressor through the protein-protein interaction between Stl and a specific phage protein. This second protein acts as the SaPI derepressor since its interaction disrupts the Stl-SaPI DNA complex and begins the excision-replication-packaging (ERP) cycle of the island (M. Á. Tormo-Más *et al*., 2010).

Stl, like CI repressor from λ phage, is mostly α-helical and is constituted by two domains with separate functions. The N-terminal domain binds to the SaPI DNA via its helix-turn-helix (HTH) motif while the C-terminal domain seems to bind to the specific phage derepressor protein though its function is still unknown.

Stl proteins are widely divergent between SaPIs and because the interaction with the phage protein is highly specific, the ability of a particular phage to derepress a SaPI depends on whether it encodes for the phage-encoded derepressor able to bind specifically to the Stl. For example, the 80α phage is able to induce almost all the islands tested so far, including SaPI1, SaPI2, SaPIn1, SaPIbov2 and SaPIbov5, while the Φ11 can only induce SaPIbov1 and SaPIbov5. Therefore, some phages can induce several SaPIs as they encode the specific repressor genes in the genome, and there are SaPIs that can be induced by various phage repressors since the Stl can interact with more than one phage derepressor protein.

1.3.3.2 Role of phage de repressor gene

All the derepressor proteins identified so far are moonlighting proteins, meaning that they can perform more than one function. They have a primary function vital for the phage biology and a secondary function related to SaPI induction.

Bacteriophages encode all the essential genes needed to complete their life cycle and they continuously evolve to acquire new functions that are beneficial to them. As a consequence of the restrictions imposed by the limited amount of genomic content that the phage capsids can package, the new functions are acquired by codified proteins. Thus, bacteriophage genomes have plenty of proteins with

more than one function. Due to the difficulty of predicting moonlighting function for a protein, few examples of bacteriophage moonlighting proteins have been extensively studied regarding their secondary function as well as their impact on the pathogenicity of bacteria.

At present, only four phage derepressors have been identified as inducers for a few SaPIs (M. Á. Tormo-Más *et al*., 2010) (J. Bowring *et al*., 2017): (i) the *dut* gene codes for a dUTPase that induces SaPIbov1, SaPIbov5 or SaPIov1 all sharing the same Stl, ii) *sri* gene (phage 80α ORF22) is a derepressor for SaPI1 and also inhibits bacterial replication. Sri protein binds to the putative host helicase loader DnaI (J. Liu *et al*., 2004). (iii) 80α GP15 induces SaPIbov2 but its primary function related to the phage biology is still unknown and (iv) phage recombinase family (Sak, Sak4, Erf and Redβ) has been demonstrated to play an essential role in phage replication and is also involved in SaPI2 induction.

The dUTPase (Dut) is the most studied phage derepressor. This enzyme is responsible for regulating the cellular dUTP level by breaking dUTP down into dUMP and inorganic pyrophosphate. In addition, Dut plays a key role in preventing uracil misincorporation into DNA what could result in cell death (Vértessy & Tóth, 2009). However, Dut protein is also implicated in other cellular processes regarding regulation of the immune system, autoimmunity and apoptosis. The Dut family has recently been defined as signalling molecules that acts analogously to eukaryotic G-proteins (M. Á. Tormo-Más *et al*., 2013).

It was initially demonstrated that the trimeric Duts from 80α and Φ11 phages are the derepressors of some SaPIs that share the same *stl,* including SaPIbov1, SaPIbov5 and SaPIov1. The SaPI derepression depends on the presence of dUTP and involves the conserved motifs III, IV and V and the variable motif VI. The diversity in the motif VI is the determining factor for the existence of trimeric Duts with different affinities for the Stl repressor (M. Á Tormo-Más *et al.*, 2013) (Maiques *et al*., 2016).

Similarly, Hill and Dockland have shown that dimeric dUTPase from ΦNM1 phage is equally effective at derepressing SaPIbov1 as the structurally distinct trimeric dUTPase of 80α and Φ11 (Hill & Dokland, 2016). A recent study thoroughly investigates the SaPIbov1 induction by dimeric Duts and found a close parallelism

with the trimeric Dut induction mechanism. Both Duts present five conserved catalytic domains in combination with a variable domain (domain IV) which is directly linked to Stl interaction. In addition, this interaction is regulated by the presence of dUTP and the Stl-Dut interaction inhibits the Dut catalytic activity (Donderis *et al*., 2017).

The fact that dimeric and trimeric Duts induce the same SaPI, while being completely unrelated regarding both sequence and structure, raises the possibility of the Stl to interact with different proteins, improving the induction and transference ratio (J. Bowring *et al*., 2017).

Another recent study showed that the single-strand annealing proteins (SSAP or recombinases) were involved in SaPI2 induction. First, the Sak protein from the 80α phage, which belongs to the SSAP family, was demonstrated to be essential to phage replication and was implicated in SaPI2 induction (Neamah *et al*., 2017). Since both dimeric and trimeric Duts are able to interact with the same Stl, the study was extended to other proteins from the SSAP family with different sequences and structures. Sak4 from Φ52A, ORF17 from ΦSLT and SA1794 from ΦN315 were studied and all were able to induce SaPI2 (J. Bowring *et al*., 2017).

From all these studies, it can be concluded that SaPIs have acquired the ability to interact with structurally different phage proteins which represents an extraordinary evolutionary adaptation.

1.3.4 SaPI-helper phage interactions

SaPI needs the presence of a helper phage in order to be transferred into other cells. This utilisation of the phage for the SaPI's benefit has a cost that is translated into a reduction of phage proliferation (Frígols *et al*., 2015). The Stl-derepressor phage protein interaction is one of the mechanisms adopted by the SaPI in which the phage is involved and the phage transmission is reduced. To avoid this event, bacteriophages mutate the derepressor protein to decrease affinity for the Stl (M. Á. Tormo-Más *et al*., 2010). SaPIs precisely exploit and disrupts the life cycle of the helper phage in order to enable their own propagation. So far, there are four SaPIs strategies identified, which are summarised in the Figure 1.7.

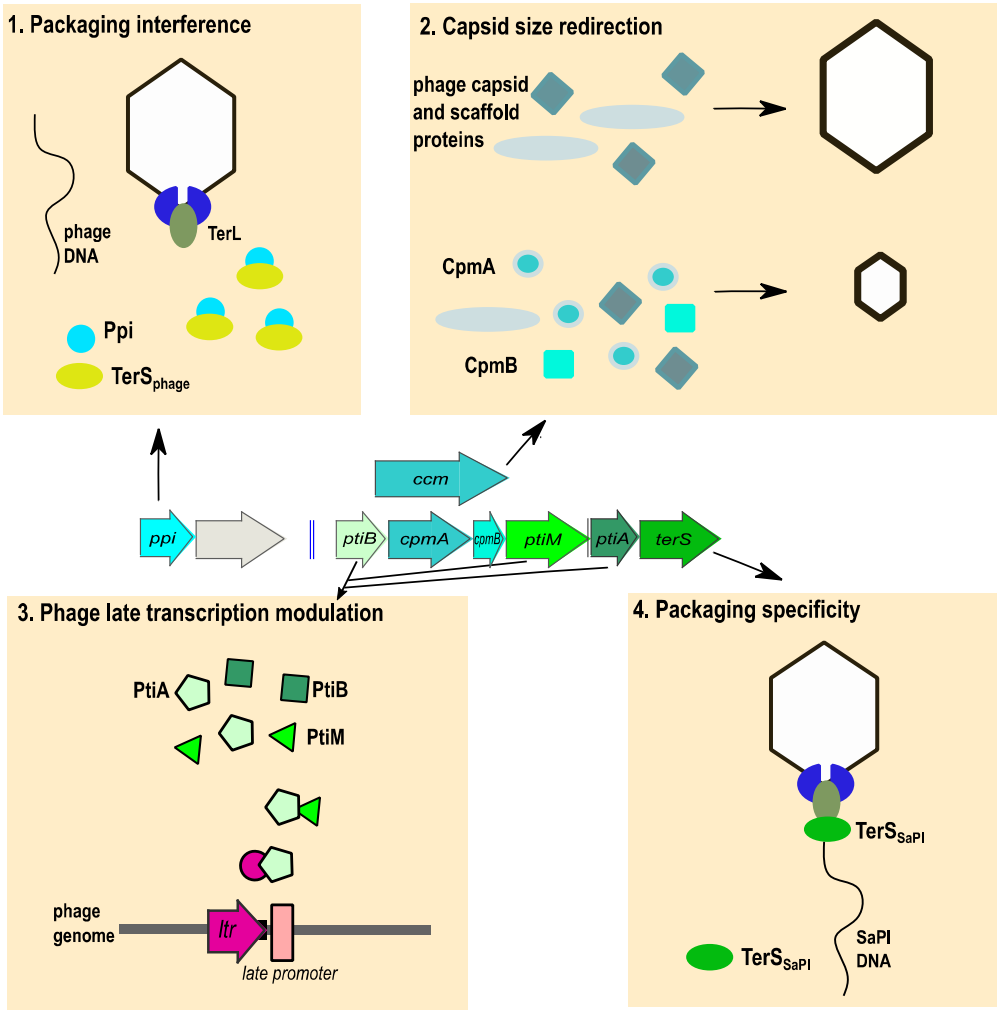

Figure 1.7 SaPI interference mechanisms. 1. Packaging interference: Ppi forms complexes with $TerS_{phage}$ to block its function avoiding phage DNA encapsidation. 2. Capsid size redirection: CpmA and CpmB or Ccm redirect the capsid assembly to produce small capsids which are filled with SaPI DNA. 3. Phage late transcription modulation: Pti binds to LtrC blocking late phage transcription, PtiM interacts with PtiA and modulates its function while PtiB blocks late phage gene expression by unknown mechanism. 4 Packaging specificity: $TerS_{SaPI}$ binds to $TerL_{phage}$ to generate a holoenzyme that recognizes and packages SaPI DNA instead of phage DNA.

1.3.4.1 Capsid size determination

SaPIs do not encode structural proteins but are instead packaged into capsids from structural proteins entirely encoded by phages (Tallent, Langston, Moran, & Christie, 2007) (M. Á. Tormo *et al.*, 2008). In addition, SaPIs redirect the phage capsid assembly pathway to form smaller capsids than those normally produced by the phage alone. The most studied interference mechanism is the SaPI1 mobilisation by 80α phage. The 80α virion has a 63 nm icosahedral head with T=7 architecture, but SaPI1 can modulate the assembly pathway to form capsids that are 45nm with T=4 architecture. The small capsids are incapable of packaging full-

length phage genomes and they interfere with phage propagation (Poliakov *et al*., 2008) (Spilman *et al*., 2011) (Gail E. Christie & Dokland, 2012).

Phage capsid size redirection is dependent on the SaPI-encoded CpmA and CpmB (capsid morphogenesis) proteins in *pac* SaPIs and Ccm proteins in *cos* SaPIs (*cos* capsid morphogenesis) (Úbeda *et al*., 2007) (Damle *et al*., 2012) (Dearborn *et al*., 2011) (Spilman *et al*., 2012). CpmA and CpmB were first studied and it was determined that both are necessary to archieve efficient redirection. CpmB is an internal scaffolding protein (SP) that competes with the SP phage for the binding site in the phage capsid. Although, less is known about the role of CpmA, but it seems to prevent aberrant small capsid formation and could provide access to the phage capsid for CpmB.

Size remodeling is only possible when CpmA-B are compatible with the helper phage capsid system. Indeed, the efficiency of small capsids' formation varies between the same SaPI depending on the present phage (Dearborn & Dokland, 2012). For example, SaPI2 can form small capsids with 80α but not with phage 80, and SaPIbov1 converts 95% of ΦNM1 and ΦNM2 phage capsids into smaller and only 63% of 80α (G. E. Christie *et al*., 2010) (Dearborn & Dokland, 2012). These differences could suggest that other factors are involved in capsid size redirection.

In some studies, it was observed that SaPI1 and SaPIbov1 mutants in *cpm*A-B were transferred at high frequency indicating that these proteins are not essential for SaPI transfer. In fact, there are some SaPIs that do not encode *cpm*AB. For example, SaPIpt1028 lacks CpmAB and is packaged in full-sized phage capsids. In addition, *cos* SaPIbov5 does not encode *cpmA-B* genes but encodes sequentially unrelated *ccm* gene. Ccm drives the production of small capsids by unresolved mechanisms and is part of the capsid shell (Carpena, Manning, Dokland, Marina, & Penade, 2016).

1.3.4.2 DNA packaging interference

Phage DNA can be packaged into procapsids by two different mechanisms that depend on the phage TerS-TerL complex: the unit length and the headful mechanism. In the headful mechanism by pac phages, TerS binds to the pac_{phage} site and TerL

breaks the dsDNA in this point. Once the phage head is filled, the next downstream DNA cleavage is made by the TerL. However, in *cos* phages, TerS binds to a *cos* site and TerL cleavages occurs at a sequence-specific site (*cos*) and needs a second identical *cos* sequence to terminate the cleavage event (Sippy, Patel, Vahanian, Sippy, & Feiss, 2015). SaPIs can also exploit the packaging mechanism to their benefit using different strategies.

Classical or *pac* SaPIs can be packaged in phage capsids through the headful mechanism. To exploit the phage packaging process, the SaPI encodes its own small terminase ($TerS_{SaPI}$) which forms a complex with the $TerL_{phage}$ and favors SaPI DNA packaging against phage DNA (Ubeda *et al*., 2009). $TerS_{SaPI}$ binds to a *pac* site in the SaPI genome, localised upstream of the Operon I, and interacts with $TerL_{phage}$ translocating the SaPI dsDNA into the capsid (Bento, Lane, Read, Cerca, & Christie, 2014). Although $TerS_{SaPI}$ is conserved among SaPIs and is highly different from $TerS_{phage}$ it can interact with $TerL_{phage}$ (J. Chen, Ram, Penadés, Brown, & Novick, 2015).

The $TerS_{SaPI}$ interference mechanism allows to reach a high frequency of SaPI transference. In fact, the transference of SaPI *ter*S mutants is reduced to generalised transduction levels (Ubeda *et al*., 2009).

There is another interference mechanism that also block phage DNA packaging but in a completely different manner. This mechanism is mediated by the SaPI gene *ppi* (phage packaging interference). Ppi interacts directly and specifically with $TerS_{phage}$ and blocks the phage DNA packaging by a mechanism that has to be determined yet. One possible explanation could be that the Ppi-$TerS_{phage}$ interaction would either prevent the $TerS_{phage}$ binding DNA activity or the interaction with TerL (G. Ram *et al*., 2012).

1.3.4.3 Phage late gene transcription interference

Bacteriophages have developed a variety of mechanisms to ensure tight control of the gene expression in the different stages of their life cycle. The morphogenetic genes responsible for packaging the phage and SaPI molecules are expressed during the late phase of the lytic cycle, and they are encoded in a single

polycistronic operon. This operon is positively regulated by the superfamily of Ltr proteins (late transcription regulator) which includes RinA, ArpU, LtrA, LtrB and LtrC (Ferrer *et al*., 2011) (Nuria Quiles-Puchalt *et al*., 2013).

Two SaPI-coded proteins, PtiA and PtiB (phage transcription inhibitor), block the late phage operon transcription responsible for the generation of virion and lysis proteins. The PtiB mechanism of action remains unclear since overexpression of PtiB is toxic for the bacteria. Instead, it was demonstrated that PtiA directly binds to the Ltr regulator to inhibit the transcription. Consequently, this transcription blockage avoids the production of phage, as well as SaPI particles. To prevent a negative effect regarding SaPI transference, a third SaPI protein, PtiM, binds to PtiA and moderates its action. Complex PtiA-M and PtiB enable a low yet sufficient level of phage reproduction to produce SaPI particles (Geeta Ram *et al*., 2014) (Geeta Ram, Chen, Ross, & Novick, 2015).

It is remarkable the existence of at least four unrelated interference mechanisms and all genes involved are located in a small region of the SaPI genome. This is clear example of convergent evolution and highlights its importance to the SaPI life cycle.

1.3.5 SaPIs-mediated generalised transduction

Generalised transduction (GT) mediated by bacteriophages is one of the mechanisms responsible for the quick spread of antibiotic resistance and bacterial virulence. Usually, *pac* phages are able to package the bacterial host DNA into the procapsids instead of its own genome and they can transfer any gene from one bacterium to another (Lennox, 1955).

It was recently described the mechanism by which SaPIs are also able to mediate generalised transduction beyond their own transfer (J. Chen *et al*., 2015). $TerS_{SaPI}$ recognises sequences that resemble the SaPI *pac* site (pseudo-*pac* sites), which are scattered throughout the genome and preferably package downstream of these pseudo-*pac* site. Resemblance of the pseudo-*pac* site varies with respect to the SaPI *pac* site, resulting in the transduction of some genes at quite high frequency and others at low (Schmieger & Backhaus, 1973) (J. Chen *et al*., 2015). Curiously,

pseudo-*pac* sites are adjacent or close to genes involved in adaptation to different environments.

The $TerS_{SaPI}$ is located in the Operon I which is activated by the SOS response. Therefore, the $TerS_{SaPI}$ can be expressed in the absence of SaPI induction and can mediate the generalised transduction independent of the SaPI ERP cycle.

1.3.6 Non-SaPI PICI elements

SaPI-like elements are widely distributed in the bacterial world and have been identified in other Gram-positive and Gram-negative bacteria. This widespread family of MGEs is generically known as phage-inducible chromosomal islands (PICIs) and has some features in common with SaPIs. All PICIs present a modular structure grouped by function and they encode homologous genes which never appear in other genetic elements. Additionally, they also recognise and integrate at unique and specific attachment sites and present a phage-dependent mechanism of induction (Martínez-Rubio *et al*., 2017) (Fillol-Salom *et al*., 2018) (Richard P. Novick & Ram, 2016).

PICIs have been identified in non-*aureus* staphylococci genome of *Staphylococcus pasteuri, Staphylococcus warneri, Staphylococcus lugdunensis* or *S. epidermidis* (H. J. Chen *et al*., 2011). Similarly, this family has been recognised in other Gram-positive bacteria including *Lactoccus lactis* or *Streptococcus pneumonia.* The PICI element EfCI-V583 from *Enterococcus faecalis* was the first element whose PICI functionality was demonstrated outside of the staphylococci (Matos *et al*., 2013) (Martínez-Rubio *et al*., 2017).

Most recently, PICIs have also been characterised in *E. coli* and *Pasteurella multocida* expanding the PICI family to the Gram-negative bacteria. In this case, PICIs use a different strategy to induce the cycle. Instead of presenting an Stl-like element such as SaPIs, they seem to encode an activator AlpA whose expression depends on the helper phage (Fillol-Salom *et al*., 2018). Overall, PICIs represent a unique class of MGEs that have broadly impacted lateral gene transfer in the bacterial world.

2. Objectives

Based on the topics discussed in the introduction section, it is evident that bacteriophages and SaPIs are essential to horizontal gene transfer and, consequently, to bacterial evolution as well as the dissemination of virulence factors between bacterial populations. Therefore, the main aim of the present work is to better understand these MGEs in order to potentially treat bacterial infections from another perspective. To archive this goal, the following specific objectives were addressed:

1º Characterisation of SaPIs in *S. aureus* clinical strains.

SaPIs are involved in several processes related to virulence including toxin production, biofilm formation and host adaptation. However, the relevance of these MGEs in clinical strains has not deeply been studied. For this reason, the first objective of the thesis is to detect the presence of SaPIs in clinical strains. In addition, SaPIs has a conserved organisation of genes in functional modules, and therefore the modules were determined to mainly focus on integration and regulation modules as well as on virulence gene content.

2º Study of SaPIs mobilisation by endogenous bacteriophages in *S. aureus* clinical strains.

To date, SaPI mobilisation has always been studied using only laboratory strains. However, recent evidence has shown that clinical-use antibiotics (ciprofloxacin or trimethoprim) can induce the expression of prophage gene products and lead to the excision and spread of temperate phages that would otherwise remain inactive in the genome. In addition, bacteriophage activation can produce SaPI induction and consequentially dissemination at high frequencies of virulence genes contained in them. As a result of this process, some antibiotics aimed to cure infections can increase the health risks associated with pathogenic bacteria. Therefore, the second objective of the thesis is to study the mobilisation of SaPIs in the clinical environment, which is fundamental in order to understand *S.*

aureus pathogenesis and to consider the antibiotics' effects on the spread of virulence genes.

3º Study of SaPIs role as regulator of chromosomally-encoded virulence gene expression.

Recent literature reflects the important role of SaPIs beyond mere high-speed carriers of virulence genes. Thus, SaPIs interfere with phage reproduction, which ultimately preserves cells from the phage-mediated killing. SaPIs also play a role in bacterial evolution by mediating the transfer of chromosomally-encoded virulence genes and they are also a key driving force of phage evolution and ecology. In addition, SaPIs could be involved in other processes than those described so far, which makes them more critical to the evolution and adaptation of the organism. Thus, the third objective of the thesis is to determine whether SaPIs can regulate the expression of chromosomal genes.

3. Material and Methods

3.1 Bacterial growth

Bacterial strains and plasmids used in the thesis are listed in Appendices 1 and 2. *S. aureus* strains were routinely grown in tryptic soy broth liquid medium (TSB, Scharlab) and TSB agar at 37°C unless otherwise stated. When required, the antibiotics used were erythromycin (10 μg ml^{-1} or 2.5 μg ml^{-1}), tetracycline (3 μg ml^{-1}) and chloramphenicol (20 μg ml^{-1}). *E coli* strains were grown in Luria-Bertani broth (LB, Scharlab) and LB agar plates at 37°C supplemented with ampicillin (100 μg mL^{-1}) or kanamycin (50 μg mL^{-1}) as needed. All the antibiotics were filter sterilized at the requisite stock concentration.

The 200 *S. aureus* strains isolated from cystic fibrosis patients and 48 *S. epidermidis* strains isolated from catheter-associated infections at Hospital Universitario y Politécnico la Fe de Valencia were included as samples of clinical strains. Ethical clearance was sought and obtained from the Biomedical Research Committee of Research Institute la Fe (references 2104/0563 and 2017/0523). Strains were stored at -80° in TSB medium containing 15% glycerol.

3.2 Molecular biology techniques

3.2.1 Chromosomal and plasmid DNA extraction

Two mL of *S. aureus* overnight culture was pelleted by centrifugation and the supernatant was removed. Genomic DNA was isolated using "GenElute™ Bacterial Genomic DNA" according to the manufacturer´s instructions except for the use of lysostaphin (12.5 μg mL^{-1} Sigma-Aldrich) instead of lysozyme to lyse the bacterial wall (45 min at 37°C).

Plasmids were isolated from *E. coli* and *S. aureus* cultures using "GenElute™ Plasmid Miniprep kit" as outlined by the manufacturer except for the additional lysis with lysostaphin in *S. aureus* strains. Both kits were purchased from Sigma-Aldrich.

3.2.2 Polymerase Chain Reaction (PCR)

The primers were obtained from Invitrogen and are listed in Appendix 3. Lyophilised primers were resuspended in HPLC-grade water to a final concentration of 200 µM. Working primer stock solutions of 10 µM were made by 1:20 dilution. Phusion High-Fidelity DNA polymerase (Thermo Fisher) was used for cloning. Routine PCR for simple detection of product was performed by GoTaq® G2 DNA Polymerase (Promega). PCR amplifications were performed according to manufacturer`s instructions.

PCR products were purified for subsequent analysis using "GenElute™ PCR Clean-Up kit" in accordance with the manufacturer`s protocol. DNA was quantified using NanoDrop spectrophotometer (NanoDrop technologies).

Multilocus sequence genotyping (MLST) was performed on selected isolated according to published protocols (Aanensen & Spratt, 2005). Seven *S. aureus* housekeeping genes (*arcC, aroE, glpF, gmk, pta, tpi* and *yqiL*) were amplified by PCR and DNA sequencing was performed for all of the PCR products. DNA sequences were analysed on the MLST website: http://www.mlst.net/

3.2.3 Agarose gel electrophoresis

Success of PCR amplification was assessed using electrophoresis. Agarose gels were prepared by dissolving agarose powder (Agarose D1 Low EEO pronadisa) 0.7-1% in Tris-Acetate-EDTA buffer and adding Gold View staining. Before being loaded into the gel, 1X loading dye (60% glycerol, Tris-HCl pH 7.6 10mM, EDTA 60mM, bromophenol blue 0.03%, xylene cyanol FF 0.03%) was added into each sample. "Gene Ruler 1Kb Plus" and "Gene Ruler 100 bp DNA ladder" from Thermo Scientific were used as a ladder to determine sample band sizes. The gels were visualized using Molecular Imager® ChemiDoc™ XRS + System with Image Lab™ Software (Bio Rad).

3.2.4 Restriction endonuclease digestion of DNA

DNA samples (plasmids or PCR products) were digested usually using double digestion with two different restrictive enzymes (Thermo Fisher) including EcoRI, BamHI or SalI. A buffer providing the maximal activity for both enzymes was selected according to the manufacturer. Analytical digests were incubated at 37° from 3 hours (h) to overnight depending on the DNA an enzyme concentration.

3.2.5 Ligation of DNA fragments with T4 DNA ligase

For the ligation reaction, vector and insert were generally taken in the molar ratio to 1:3, around 20 ng of vector and 60 ng of insert. Ligation was carried out in 20 µl with 1X ligation buffer and 1 µl of enzyme T4 DNA ligase (Thermo Scientific). Reactions were incubated at 22 °C for 1 h and then, the ligase was inactivated at 65°C for 10 min.

3.2.6 Plasmid construction

Generation of plasmid vectors was achieved using standard procedures (Russell & Sambrook, 2001). DNA fragments (inserts) were amplified by PCR, purified and then digested with appropriate restriction enzymes. The digested products were ligated with the corresponding pre-digested vector. After ligation, the plasmid was transformed into chemically competent *E. coli* DH5α cells. The vectors used in this work were pCN51, pCN41, pCU1 and pMAD (Charpentier *et al*., 2004) (Arnaud, Chastanet, & Débarbouillé, 2004). A cadmium-inducible promoter (Pcad) was cloned in pCU1 vector. Protein purification was performed using pET28a plasmid.

3.2.7 DNA sequencing

Cloning fragments were sequenced by GATC BIOTECH company. Sequence data was verified using SnapGene and https://blast.ncbi.nlm.nih.gov/Blast.cgi. Whole-genome sequencing of strains were performed by Microbes NG Service. https:/microbesng.uk/

3.2.8 Southern blot

Southern blot was performed by standard procedures (Russell & Sambrook, 2001). For sample preparation, *S. aureus* strains were grown to OD_{540}: 0.2–0.3 and were then induced with mitomycin C (2 µg/mL Sigma-Aldrich, from *Streptomyces caespitosus*). The cultures were incubated at 32 °C and 80 rpm. Samples were taken 90 min after phage/SaPI induction. Then, 1 mL of the sample was centrifuged, and the pellet was resuspended in 50 µL lysis buffer (47.5 µL of TES-sucrose: 10 mM Tris, 1 mM EDTA, 100 mM NaCl, 0.5 M sucrose and 2.5 µL of lysostaphin (5 mg/mL, Sigma-Aldrich)). After 30 min at 37 °C, 47.5 µL of 2% SDS in H_2O and 2.5 µL of proteinase K (20 mg/mL) were added, and the mixture was incubated at 55 °C for 30 min. Next, 11 µL of loading buffer dye 10X was added, and the mixture was shaken for at least 20 min. The samples were cold (N_2)/heat (65 °C) treated three times. The prepared samples were resolved on a 0.7% agarose gel at 20 V overnight, transferred to a nylon membrane (Roche) via capillary action and blotted with specific Digoxigenin-labelled probes and anti-Digoxigenin-AP (Anti-Digoxigenin-AP Fab fragments, Roche) antibodies. Specific primers were used to generate DIG-labelled probes (Appendix 3). Detection was carried out according to the Immun Star™ AP substrate pack (BIO-RAD).

3.2.9 Western Blot

The strains were treated with $CdCl_2$ (2 µM) and harvested (1 mL) after 2 h plasmid induction. The pellets were resuspended in 200 µL lysis buffer (50 mM Tris-HCl, 20 mM $MgCl_2$, 30% raffinose and 1 µl of lysostaphin (12.5 µg mL^{-1}) and incubated 1 h at 37 °C. SB buffer (Biorad) with β-mercaptoethanol was added, and the samples were heated 10 min at 95 °C, put on ice for 5 min and centrifuged. Samples were electrophoresed on a 15% SDS-polyacrylamide gel (SDS-PAGE) and transferred to a polyvinyl difluoride membrane (PDVF Thermo Scientific) by electroblotting. Membrane was blocked with 5% non-fat milk in TBS-T (10 mM Tris, pH 8.0, 150 mM NaCl, 0.5% Tween 20) for 60 min to overnight. The membrane was washed with TBS-T, incubated with anti-FLAG antibodies (Monoclonal ANTI-FLAG® M2 Peroxidase produced in mouse, Sigma-Aldrich) and

revealed according to the protocol supplied by the manufacturer (ECL™ Prime Western Blotting Detection Reagent).

3.2.10 RNA extraction

RNA was isolated from 2 mL of *S. aureus* culture after 2 h of $CdCl_2$ treatment using "NZY Total RNA isolation kit (Nzytech)" following the manufacturer`s instructions. Cell pellet was resuspended in lysis buffer and subjected to mechanical disruption with glass beads (0,3 g) for 45 s at level 6 using Fast Prep FP120Cell disrupter (Thermo Savant). The purified RNA was quantified and stored at -80 °C.

3.2.11 cDNA synthesis

The cDNA strand was synthesized using "PrimeScript™ Reagent Kit (TAKARA)". 500ng-1µg RNA was added as a template in a 20 µl reaction. The reaction was performed following manufacturer`s instructions. The cDNA was purified and quantified. To verify the absence of genomic DNA in the RNA samples, a reaction without "reverse transcriptase enzyme" was used as a control.

3.2.12 Design of the tiling microarray

The *S. aureus* custom tiling array (NA-Staph-b520729F) was designed in collaboration with Alejandro Toledo Arana, Íñigo Lasa, Jose Rafael Penadés and Affymetrix (Santa Clara, CA USA) in 2011. Specifically, the microarray (format 49-7875 with 11 µm features) contains a total of 522,406 probes, divided into two parts. The first part corresponds to the tiling which contains a total of 384,932 probes (25 mer), which are divided in eight sets. The first set corresponds to 46310 probes recommended by Affymetrix as controls. The sets used in this analysis covers both strands of (i) the *S. aureus* NCTC8325 genome (2,821,347 bp covered by 363127 probes), (ii) the *S. aureus* pathogenicity islands SaPI1 (15229 bp covered by 1892 probes), (iii) SaPIbov1 (15 887 bp covered by 1978 probes), (iv) the phage 80α (43859 bp covered by 5744 probes).

Each 25-mer probe was tiled each 14-nt across the whole genome, resulting in 11-nt overlaps and a 7-nt offset of the tile between strands. This design enables

a 7-nt resolution for hybridization of double-stranded targets and a 14-nt resolution for strand-specific targets. The second part of the chip corresponds to the *S. aureus* gene expression sub-array, which contains 91 164 probes grouped into 3224 probe sets (each probe set includes 14 perfect match and 14 mismatch probes). In all, 3122 probe-sets are targeting unique *S. aureus* genes, whereas 102 probe-sets are controls recommended by Affymetrix.

3.2.13 cDNA synthesis, fragmentation, labelling and array hybridization and scan

Before cDNA synthesis, RNA integrity from each sample was confirmed on Agilent RNA Nano LabChips (Agilent Technologies). In all, 10 mg RNAs were reverse transcribed using SuperScript II reverse transcriptase (Invitrogen Life Technologies) and processed following the protocol of the Affymetrix GeneChip Expression Analysis Technical Manual (P/N 702232 Rev. 2) in the presence of 6 ng/mL Actinomycin D to avoid spurious second-strand cDNA synthesis during reverse transcription reaction. cDNA was digested by DNase I (PIERCE), and the size of digestion products was analysed in the Agilent Bioanalyser 2100 using RNA Nano LabChips to ensure that the fragmentation resulted in a majority of products in the range of 50–200 bp. The fragmented cDNA was then biotinylated using terminal deoxynucleotidyl transferase (Promega) and the GeneChip DNA labelling reagent (Affymetrix) following the manufacturer's recommendations. Biotinylated cDNA (5 mg per array) were hybridized for 16 h according to the Affymetrix protocol in a total volume of 200 mL per hybridization chamber. Following incubation, the arrays were washed and stained in the Fluidics station 450 (Affymetrix) using the protocol nFS450_0005. Scanning of the arrays was then performed using the GeneChip scanner 3000 (Affymetrix). A first scan of the chip was carried out with gene expression sub-array parameters followed by a second scan with tiling sub-array parameters. Intensity signals of each probe cells were computed by the GeneChip operating software.

3.2.14 Expression and purification of proteins

Overexpression of different proteins (Stl, DUF3113, PtiM) using pET28a vector were performed. The genes were amplified by PCR and cloned into the pET28a vector. The primers were designed to generate N-terminal or C-terminal 6xHis-tag fusion proteins. The plasmids were transformed into *E. Coli* BL21(DE3) cells. *E. coli* cells containing the recombinant plasmids were propagated in 500 mL of LB broth (50 μg/mL kanamycin). Cultures were grown until the OD_{650} reached 0.6–0.8, then isopropyl-β-D-thiogalactopyranoside (IPTG, Thermo Scientific) was added to reach a concentration of 1 mM. After induction, the cultures were grown overnight at 16 °C and 200 rpm. Cells were harvested by centrifugation (5000 g for 10 min at 4 °C) and pellets were kept at -20°C before being resuspended in 50 mL of lysis buffer (20 mM Na-phosphate pH 7,4, 150 mM NaCl, 10% glycerol and 0,1 mM PMSF). The cells were lysed by sonication on ice (Amplitude 70% 3 min: 3 s ON and 10 s OFF) and centrifuged (16000 g for 30 min). Supernatants were applied to a metal affinity resin (TALON® Metal Affinity Resin (TAKARA) or Ni-NTA His Bind® Resin (Novagen)), and incubated for 1 hour at RT. The 6xHis-tagged proteins were eluted with 100–500 mM imidazole gradient buffers. Elution fractions were analysed by SDS/PAGE and by size-exclusion chromatography.

3.2.15 Size-exclusion chromatography

A Superdex™ 75 26/60 column running on a ÄKTA™ pure25 system (GE Healthcare Life Sciences) was used to study the interaction of the selected proteins. The column was equilibrated in 20 mM phosphate buffer containing 300 mM NaCl at pH 7.4, and the elution was performed at 2.6 mL/min at 25 °C. Calibration of the column was performed with aldolase (158 KDa), covalbumin (75 KDa), BSA (66 KDa), ovalbumin (44 KDa), dimer ribonuclease (27.4 KDa), dimer ubiquitin (17 KDa), ribonuclease (13.5 KDa) and ubiquitin (8.5 KDa). Peak protein fractions were analysed by SDS-PAGE.

3.2.16 Pull-down assay

The primers were designed to generate N-terminal 6xHis-tag fusion Stl proteins and untagged DUF3113 proteins. The supernatant after sonication and centrifugation of 6xHis-tag proteins were incubated for 1 h with metal affinity resin (TALON, Clontech) and for 1 h with supernatant of untagged proteins. Pull-down elution was performed with 250 mM imidazole buffer. Elution fractions were analysed by SDS-PAGE and by size-exclusion chromatography.

3.3 Bacterial transformation

3.3.1 Preparation of chemically competent E. coli cells

E. coli cells (DH5α and BL21) were inoculated in 10 mL LB medium and grown overnight. From the overnight culture, the cells were diluted 1/100 in 200 mL LB to reach 0.3-0.4 OD_{600}. Then, the culture was put on ice 10 min to stop bacterial growth. The cells were centrifuged at 4000 g 4°C 10 min. The supernatant was discarded and the pellet was resuspended in 1/8 of the original volume (25 mL) of ice cold 0.1M $CaCl_2$ and incubated on ice for 15 min. The cells were again centrifuged and discarded the supernatant and resuspended in 1/25 of the initial volume (8 mL) with 0.1M $CaCl_2$ and incubated on ice for 15 min. The bacteria were then re-pelleted using the same conditions and the pellet were resuspended in 4 mL of ice cold $CaCl_2$ + 15% glycerol, aliquoted and stored at -80°C.

3.3.2 Transformation of *E. coli* cells

The aliquoted competent cells (50µl) were thawed on ice. 10µl of ligation mix or purified plasmid DNA was added (10pg-100ng) to the cells and they were incubated on ice for 30 min. Cells were subjected to heat shock at 42°C for 45s and then placed on ice 2 min. 700 µl of pre-warmed LB was added and the cells were incubated at 37°C with constant shaking (200 rpm) for 1 h. Transformed cells were plated out on appropriate selective LB plates and the plates were incubated at 37°C overnight.

3.3.3 Preparation of electrocompetent *S. aureus* cells

S. aureus (RN4220 and derivatives) cells were prepared from overnight culture grown at 37°C in TSB. 100 mL of LB was inoculated with 1 mL of the overnight culture and incubated at 37°C with constant shaking until 0.2-0.3 OD_{540}. Then, the culture was put on ice for 15 min to stop bacterial growth. Cells were harvested by centrifugation (4000g, 10 min 4°C), the resultant supernatant was discarded and the pellet resuspended in 10 mL of sterile and chilled 0.5M sucrose. Cells were washed twice in the same conditions. Finally, bacteria were re-pelleted and resuspended in 0.5M sucrose + 15% glycerol. The electrocompetent cells obtained were either used immediately for electroporation or frozen in 200 µl aliquots at -80 °C.

3.3.4 Electroporation of *S. aureus* cells

Approximately 1-5 µg of plasmid DNA was added to 200 µl of electrocompetent *S. aureus* cells, mixed and transferred into a 0.2 cm cuvette (Gene Pulser® Cuvette, Bio-Rad). Cells were pulsed at 2.5 kV (25 µF, 100 Ω) with Gene Pulser Xcell™ electroporation system (Bio-Rad) followed immediately by addition of 600 µl TSB. The mixture was transferred to 1.5 mL microcentrifuge tube and incubated at 37°C with 180 rpm for 1 h. Cells were harvested by centrifugation (3000 g, 5 min RT), resuspended in 200 µl of supernatant, and plated out on selective TSA plates. Plates were incubated at 32 or 37 °C for 24 h.

3.4 Biochemical methods: Enzyme assay to quantify β-lactamase activity

β-lactamase assays were performed using nitrocefin (OXOID) as substrate (Tillotson, 2008). Overnight cultures of strains carrying pCN41 and pCU1Pcad derivative plasmids were diluted 1:50 in 10 mL of TSB. pCN41 contains β-lactamase reporter gene. Derivative strains containing the plasmids were treated with MC or $CdCl_2$, as appropriate (OD_{540}= 0,2-0,3). Aliquot of 200 µL were taken after 2 h for β-lactamase activity. Samples were transferred to a flat bottom polystyrene 96 plate. 50 µL of nitrocefin was added to each well. Changes of the

OD at 490 nm were registered during 15 min (every 15 s) in the Synergy H1 microplate reader. β-lactamase activity units were calculated using the formula:

$$\beta - \text{lactamase activity} = \frac{\frac{\Delta Vmax\ (OD_{490}) * 1000}{min}}{OD_{650}}$$

3.5 Bacteriophage and SaPI methodology

3.5.1 Prophage induction

Bacteriophages can be induced from the lysogenic state to the lytic state by stressing the bacteria. Stress can be applied through the addition of Mitomycin C (2 μg mL^{-1} MC). MC causes DNA damage, which triggers the SOS response pathway and this initiates the phage lytic cycle. Lysogenised *S. aureus* strains were grown at 37 °C in TSB (from an overnight culture) until 0.2 OD_{540}. Once OD was reached, MC was added and the induced culture was grown in an orbital shaker at 32 °C with 80 rpm for 4 h or until lysis occurred. The resultant culture was filtered through a 0.22 μM membrane. The potential phage solution was stored at 4 °C.

3.5.2 Strain lysogenisation

MW2 was used as phage donor strain and RN4220 as recipient strain. The recipient strain was grown and adjusted to $1x10^6$ CFU/mL, mixed with soft agar and overlaid onto the TSA plate. MW2 was subjected to prophage induction and the obtained lysate was diluted and spotted on TSA plate with a RN4220 layer and incubated at 37 °C overnight. Colonies from the center of a plaque were picked and their DNA were extracted. PCR of the phage integrase was performed to check the lysogeny.

3.5.3 Bacteriophage infection

Infection of susceptible *S. aureus* strains was performed mainly with laboratory 80α and Φ11 phages. Phage enter the bacteria and start the lytic cycle, amplifying the phage population leading to bacteria lysis. The susceptible infected *S. aureus* strains were diluted 1:50 in TSB (from overnight culture) and grown to 0.15 OD_{650}. Once this OD has been reached, the culture was centrifuged (4000 g,

10 min, 4 °C) and the pellet was resuspended in 1:1 TSB-phage buffer (1mM NaCl, 0.05 M Tris pH 7.8, 1mM $MgSO_4$, 4 mM $CaCl_2$). The resuspended culture was infected with the appropriate amount of phage (100 µL). The culture was incubated at 32 °C with reduced shaking (80 rpm) until lysis occurred. Following lysis, un-lysed bacterial cells were removed by filtration with sterile 0.22 µM filters and the lysate was stored at 4 °C.

3.5.4 Bacteriophage Titering Assay

Bacteriophage transduction was used to mobilize SaPIs or to transfer plasmids. Transduction requires a phage titer of approximately 10^{10} pfu/mL. The strain RN4220 was used as recipient to perform infecting studies in *S. aureus*. To quantify the number of phage particles contained in an interest lysate, recipient strain (RN4220) was diluted 1:50 from the overnight culture and grew to 0.3-0.4 (OD_{540}). 50 µL of the recipient strain were mixed with 100 µL of serial dilutions of the lysate during 10 min at RT. The mixture of culture-phage was plated out on phage base agar (25g of Nutrient Broth No 2 (Oxoid), 7 g agar in 900 mL H_2O) supplemented with 10mM $CaCl_2$. Plates were incubated at 37 °C and the plaques were counted.

Each plaque represents the lysis of a phage-infected bacterial culture and can be designated as a plaque-forming unit (PFU)

$$\frac{PFU}{mL} = \frac{number\ of\ plaques}{(dilution\ of\ plate\ counted\ x\ volume\ of\ lysate)}$$

3.5.5 Transduction tittering Assay

Transduction tittering assay was used to calculate SaPImw2*entC::cat* (chloramphenicol cassette) mobilisation by different bacteriophages. Once the lysates have been obtained, they were then serially diluted with phage buffer. 100 µl of each serial dilution was incubated with 1 mL of receptor strain (RN4220) at OD_{540}= 1.4 for 15 min at 37 °C with the addition of 4.4 mM $CaCl_2$. 3 mL of Top Agar was then added before pouring onto plates with the appropriate antibiotic

(chloramphenicol), cooled and incubated at 37 °C overnight. The colonies were counted. Transduction protocol was also used to transfer derivative plasmids between *S. aureus* strains.

3.5.6 Generation of *S. aureus* mutants by homologous recombination

All mutant strains were generated by allelic replacement using the temperature-sensitive plasmid pMAD as described previously (Arnaud et al., 2004). For deletion of genes from the bacterial chromosome, the flanking region of genes to be deleted were cloned into the pMAD vector. To do that, a 2-piece overlap assembly PCR was used. Two PCRs to amplify ~1.5 kb fragment comprising 750 bp upstream and 750 downstream of the target gene. For the creation of point mutation within the bacterial chromosome, a similar double PCR method was used. However, in this case, the primers incorporate the nucleotide mutation. To insert a new gene in the chromosome, a 3-piece PCR fragment was constructed. Firs the flanking region of gene to be replaced and then the new gene between the flanking regions.

The PCR fragment was cloned into pMAD, transformed first in *E. coli* and then into RN4220 as an intermediary and then transduced to the selected *S. aureus* strains. Allelic replacement was carried out by a two-step procedure. Using the β-gal as indicator, first, the pMAD plasmid was integrated into the chromosome by homologous recombination under non-permissive conditions (44 °C), and single blue colonies were selected. In the second step, the vector backbone was deleted under a permissive temperature (30 °C) as described previously (Arnaud *et al*., 2004). Mutation was confirmed by DNA sequencing.

3.5.7 Staphyloxanthin extraction and quantification

Staphyloxanthin was extracted and quantified according to Rosado, Doyle, Hinds, & Taylor, (2010) with minor modifications. The bacterial cultures (10 mL) were incubated for 16 h at 37 °C. Bacterial cells were measured (OD_{540}) and normalized. The cells were harvested at 4000 g at 4°C for 10 min and the pellets were photographed. Then, the pellets were washed with PBS. After centrifugation, pellets were suspended in 2 mL of 99% methanol and incubated in the dark for 20

min at 55 °C and at -80 °C until bleached. Samples were centrifuged at 10,000 g at 4°C for 15 min, and the absorbance of the supernatant was measured at 465 nm using a spectrophotometer and quartz cuvette. Each experiment was performed at least in triplicate.

3.6 Data analysis

Data was organized using Microsoft® and Excel® and figures were designed using Power Point®, Inkscape and GIMP.

The NCBI Blast server program was used for homology comparison of DNA and protein sequences against the GenBank database. SnapGene Viewer was a tool to visualize DNA sequences and the PRALINE server program was used for protein multiple alignment. Venny 2.1.0 was used to create a Venn diagram for comparing lists of elements (Oliveros, 2016).

After whole-genome sequencing, the bacteriophage and SaPI detection and assembly was performed mainly using Artemis Comparison Tool (ACT) and Mauve aligner. Phaster web server was also used to search for prophage sequences within the main chromosome

Statistical analyses were performed using R software (version 3.6.0 https://www.r-project.org) in collaboration with the biostatistical service of "Instituto de Investigación Sanitaria la Fe".

4. Results

4.1 Mobilisation mechanism of pathogenicity islands by endogenous phages in *S. aureus* clinical strains

4.1.1 Background

The high pathogenicity of *S. aureus* depends on the number of virulence factors presents in the genome, including a variety of surface proteins, proteases, toxins or antigens. Remarkably, many of these virulence genes are part of MGEs, which includes the SaPIs, among others.

The quick adaptation and evolution of *S. aureus* is mainly due to the acquisition of MGEs by horizontal gene transfer (HGT). Bacteriophage transduction is likely the most important HGT mechanism as they are able to package host bacterial DNA and spread it to neighbouring cells. This phenomenon increases the plasticity of *S. aureus* and facilitates its adaptation to different conditions during infection. Some bacteriophages can become hijacked by SaPIs and transfer them at a much higher frequency than other part of the bacterial genome.

Normally, SaPIs remain integrated into the host chromosome. The normal repressed state of SaPIs is due to the master regulator Stl, which acts as a repressor by binding itself to a region between two divergent promoters and inhibiting the transcription of SaPI genes (Úbeda *et al*., 2008) (Papp-Kádár et al., 2016) (Surányi *et al*., 2018). The SaPI life cycle begins with the infection of a helper phage or with the SOS induction of an endogenous helper prophage. Once the prophage is activated or the cell is infected by the phage, a bacteriophage protein binds to the Stl, triggering the release of the Stl repressor from the SaPI DNA. Disruption of the Stl-SaPI DNA complex allows the expression of SaPI proteins as well as the initiation of the ERP cycle of the SaPI (M. Á. Tormo-Más *et al*., 2010).

Different SaPIs encode different Stl repressors whose sequences are not conserved. As the interaction is highly specific, each different Stl interact with different bacteriophage proteins. Therefore, bacteriophages have different capacities to induce SaPIs depending on whether they present derepressor proteins, which are able to bind to specific Stls.

Four different phage derepressor proteins have been identified so far: dimeric and trimeric dUTPases derepress SaPIbov1 and SaPIbov5, Sri protein is a derepressor of SaPI1, ORF15 of 80α induces SaPIbov2 and recombinase family proteins (Sak, Sak4, Redβ and Erf) induce SaPI2. All previously studied phage derepressor proteins are moonlighting proteins since they have at least one more function beneficial to the phage besides their activity as SaPI inducers.

The aim of the research is to study the resident prophage's ability to induce the coresident SaPIs in clinical strains. SaPI induction has always been studied in laboratory lysogenic strains or infected strains with laboratory phages, but there has been no demonstration of SaPI induction by endogenous phages in natural isolates. Thus, we focused on studying SaPI induction in 100 clinical isolates from patients with cystic fibrosis (CF), then we studied pathogenicity island induction in *S. epidermidis* strains isolated from catheter infection, and extended the study to 14 clinical strains from different clinical origins.

4.1.2 SaPI induction by endogenous phages in clinical strains isolated from cystic fibrosis patients

SaPI induction and mobilisation by prophages have been widely studied in laboratory strains. In previous studies, SaPIs and bacteriophage are usually transduced and lysogenised in the phage-cured laboratory *S. aureus* strains RN450 and RN4220. To determine whether this process can naturally occur, a *S. aureus* strain collection of clinical origin was used.

A total of 200 *S. aureus* strains were isolated from sputum samples from 117 cystic fibrosis (CF) patients. Strains were collected from the Hospital Universitario y Politécnico La Fe de Valencia from 2013-2015 and genotyped by MLST (Multilocus Sequence Typing) analysis. The frequency of ST types is shown in Figure 4.1. From the 200 isolates, 29 strains (14%) were identified as a novel STs and can be designated as single locus variants (SLV) or double locus variants (DLV) of the already described Clonal Complexes (CC). The remaining strains belonged to 19 different CC, the most frequent CC being CC5 (22.5%), CC30 (15.5%), CC45 (10.5%) and CC398 (9.5%).

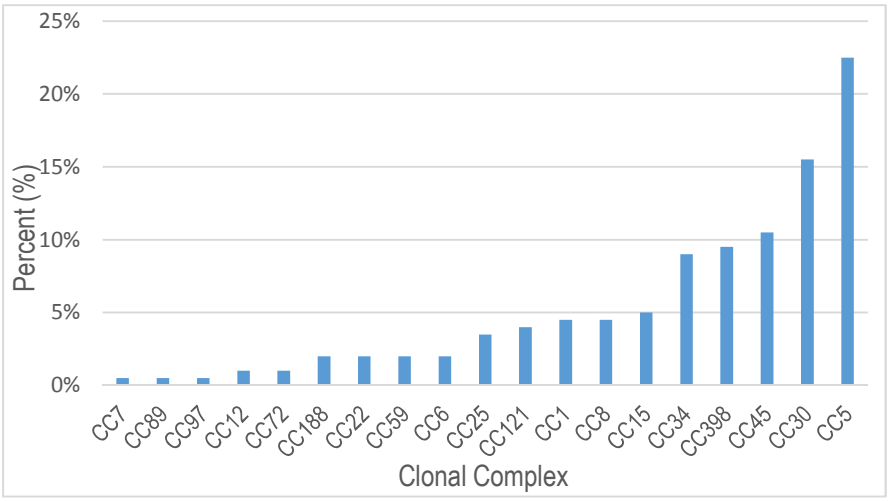

Figure 4.1 Frequency distribution (%) of *S. aureus* Clonal Complexes (CC) isolated from CF patients. A total of 200 *S. aureus* strains were differentiated into 19 genotypes by the seven-gene standard MLST genotyping method.

On the other hand, bacteriophages and SaPIs content was detected and categorised according to their *int* gene, which dictates the insertion site (paper in preparation) (Aanensen & Spratt, 2005) (Subedi, Ubeda, Adhikari, Penadés, & Novick, 2007). SaPIs can be classified by integrase homology and insertion site into five specific groups. Alternatively, bacteriophages can be clustered into seven groups according to integrase nucleotide sequence (Goerke *et al.*, 2009). Notably, no strain appears to have more than one phage or SaPI of each family type, probably due to competition over insertion site (Lindsay & Holden, 2004).

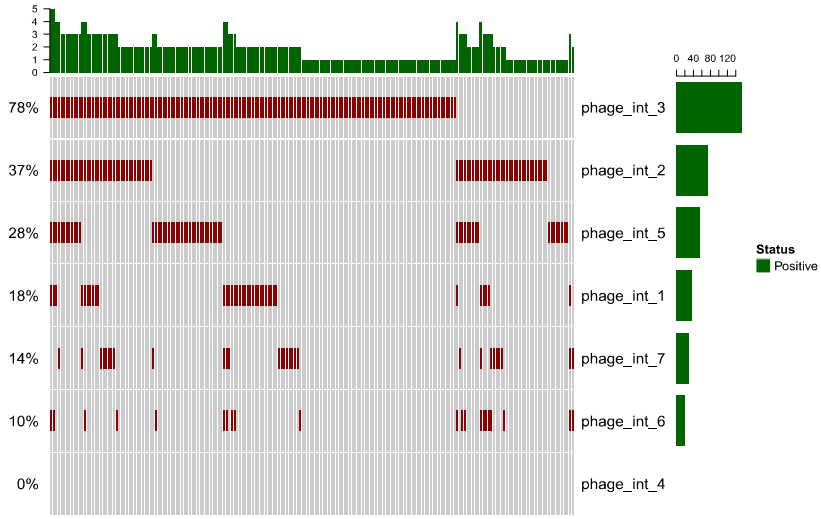

Figure 4.2 Distribution of bacteriophages in *S. aureus* clinical strains. The oncoprint provides an overview of bacteriophage presence in *S. aureus* strains. Each column corresponds to one of the 200 *S. aureus* strains. The seven rows represent the presence of each type of bacteriophage. Each cell is coloured in red or grey according to the presence or absence of the different types of bacteriophages.

The seven different *int* genes from bacteriophages were PCR amplified using the 200 *S. aureus* DNA as a template. All *S. aureus* strains carried at least one bacteriophage and 58.3% contained more than one integrated into their genome. While prophages of the group Sa3*int* were the most frequent (78%), Sa2*int* could be detected in 37%, Sa5*int* in 28%, Sa1*int* in 18.5%, Sa7*int* in 14.5% and Sa6*int* in only 10.5% of the isolates. Sa4*int* was no detected in any *S. aureus* strain (Figure 4.2).

To assess associations between the 7 different *int* type phages distribution and the clonal complex assignation of the 200 *S. aureus* clinical strains, an association plot was performed. Figure 4.3 shows a very high significant association between the presence of specific *int* type phages and particular clonal complexes (p-value =< $2.22e^{-16}$). It should be noted that Sa*int*1 phages were associated with CC12, CC121 and CC5 and Sa*int*2 with CC15. Instead, Sa*int*5 phages were infrequent in the widely distributed CC5 and Sa*int*3 in CC15. Finally, Sa6*int* phages were the least extended in this strain collection but showed association with CC398 and CC5.

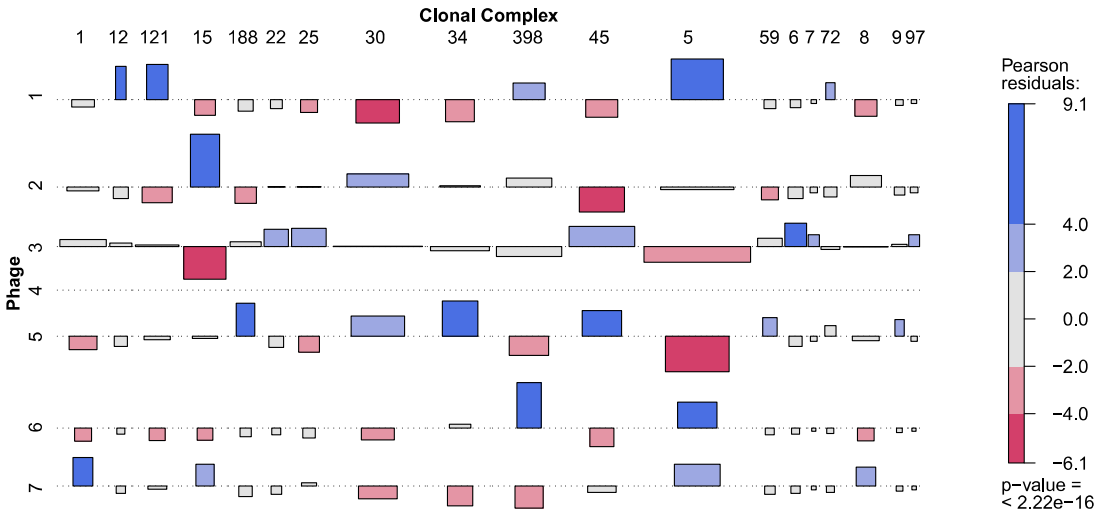

Figure 4.3 Two-way association plot for "phage type" and "Clonal Complex (CC)" in *S. aureus* CF clinical strains. Representation of the contingency table frequencies as a matrix of bars, the dimensions of which are proportional to the observed frequencies of each cross-classification (phage presence type and CC). The color code (pink and blue) and the height of the bars denote the significance level and magnitude of the Pearson's residuals, and their widths are proportional to the sample size. The rectangles also indicate the polarity of the differences between observed and expected frequencies.

In order to identify SaPIs in the *S. aureus* collection, PCR and Southern blot analysis were performed. As a result, 100 strains (50%) contained at least one SaPI integrated in the genome and 23 were positive for more than one integrase type.

SaPI*int*I, SaPI*int*V and SaPI*int*III were the most abundant type with 42%, 36% and 35%, respectively. On the other hand, SaPI*int*2 was detected in 12% and SaPI*int*4 in only 4% of the strains from SaPI positive strains (Figure 4.4).

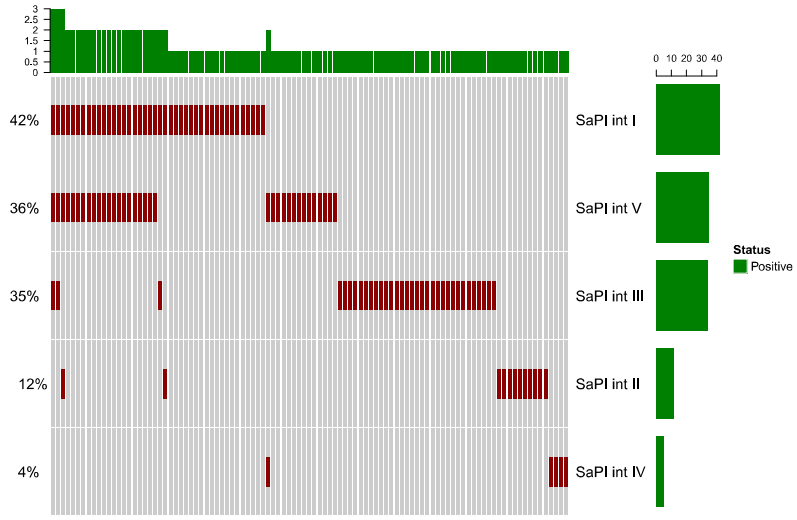

Figure 4.4 Distribution of SaPIs in *S. aureus* clinical strains. The oncoprint provides an overview of SaPI presence in *S. aureus* strains. Each column corresponds to one of the 100 SaPI-positive *S. aureus* strains. The top row shows the number of SaPIs integrated in the genome for each sample (green). The other five rows show the presence (red) or absence (grey) of each type of SaPI.

To study whether there are an association between SaPIs group and *S. aureus* clonal complexes, the 200 *S. aureus* clinical strains were also examined and represented in an association plot (Figure 4.5). A very highly significant association between the presence of specific SaPIs and particular clonal complexes (p-value=<2.22e-16) was also observed. It is worth highlighting that *int*1 SaPIs were associated with CC30 and CC34 and SaPIs *int*3 with CC45 and CC5. In return, although *int*2 and *int*4 SaPIs were scarcely detected, they showed a statistically significance association with CC398 for *int*2 SaPIs and CC59, CC8, CC89 and CC1 for *int*4 SaPIs.

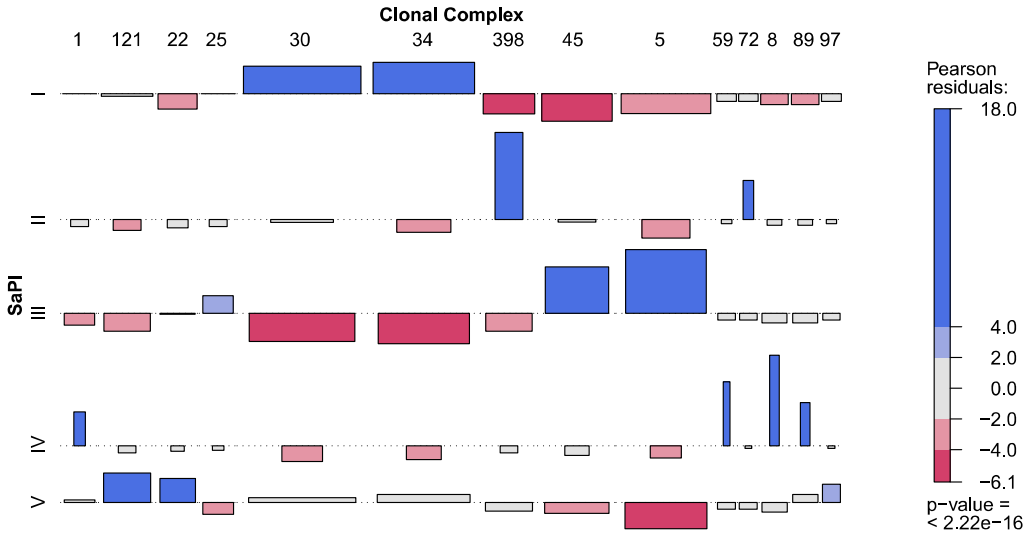

Figure 4.5 Two-way association plot for "SaPI presence" and "Clonal Complex" in *S. aureus* CF clinical strains. Representation of the contingency table frequencies as a matrix of bars, the dimensions of which are proportional to the observed frequencies of each cross-classification (SaPI type and CC). The color code and the height of the bars denote the significance level and magnitude of the Pearson's residuals, and their widths are proportional to the sample size. The rectangles also indicate the polarity of the differences between observed and expected frequencies.

Once the strains were classified according to the SaPI and bacteriophage types, SaPI mobilisation was tested using Southern blot. Strains were treated with mitomycin C (MC) to activate the SOS response and to induce the endogenous phages. The objective was to determine whether these prophages were able to activate and mobilise the resident SaPIs in "real world" clinical strains isolated from CF patients. Southern blot allows the examination of SaPI induction. Previous research demonstrated that the presence of one band corresponded to SaPI being integrated in the chromosome, while the presence of two bands indicated that the SaPI was integrated (upper band) and excised (lower band) from the chromosome. After MC treatment, the samples were processed and electrophoresed on an agarose gel. DNA was transferred to a membrane and hybridised with a probe. The chosen probes were DNA fragments from the different *int* SaPI genes.

Among the 100 SaPI-positive strains, 22 contained SaPIs induced by endogenous phages. Of these 22 strains, only 15 were finally considered and the remaining seven strains were discarded since they were considered duplicates in the collection. This exclusion process was defined as strains isolated from the same patient that had identical ST and the same SaPIs and bacteriophage types.

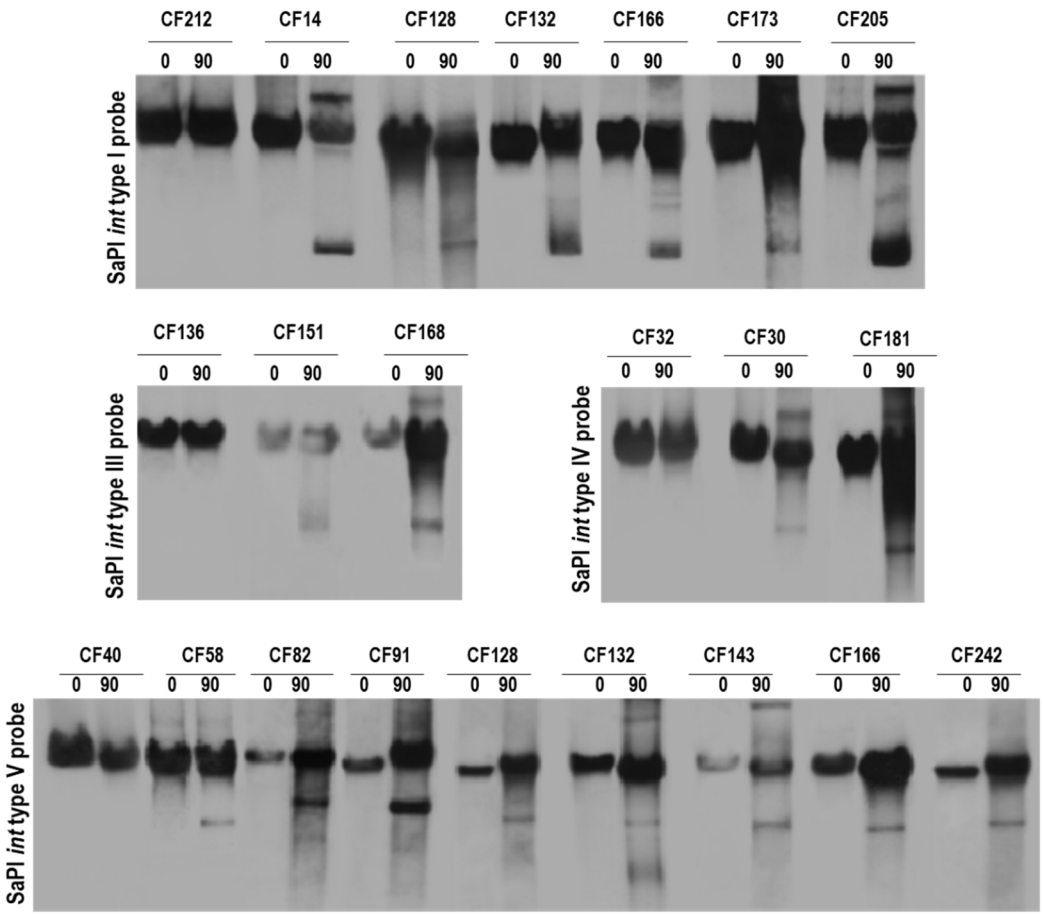

Figure 4.6 SaPI induction of *S. aureus* strains isolated from CF patients. 100 strains were treated with MC to activate the SOS response and induce the resident prophages. Samples were isolated at 0 and 90 min, separated in agarose gel and blotted with SaPI integrase-specific probe. This figure only shows SaPI-inducible positive strains. Upper band corresponds to chromosomal bulk DNA including SaPI and the lower band corresponds to SaPI excised from the chromosome.

Figure 4.6 shows the strains that contained SaPIs induced by endogenous prophages. The first strain of each image is a negative control and corresponds to non-induced SaPIs. From the 22 initial strains, 15 contained SaPIs induced by endogenous phages after MC treatment. Among them, CF14 (CC30), CF128 (CC30), CF132 (CC34), CF166 (CC34), CF173 (CC30) and CF205 (CC30) contained SaPI*int*1 type inducible by endogenous phages. Only two strains contained SaPI*int*III type (CF151 and CF168) and SaPI*int*IV (CF32 and CF181) inducible by prophages. SaPI*int*V type is inducible by resident bacteriophages in 8 strains (CF58, CF82, CF91, CF128, CF132, CF143, CF166 and CF242). Three strains (CF128, CF132 and CF166) contained two SaPIs (SaPI*int* type I and V) and both were induced by prophages. Finally, SaPI*int*II type induction was not detected in any strain.

Based on the results, it can be concluded that some clinical strains contain SaPIs induced and mobilised by endogenous phages under conditions that lead to prophage induction. However, MC was used to induce SOS response in laboratory conditions, prophages can be induced by a number of commonly clinically used antibiotics.

Nine SaPI inducible *S. aureus* strains (CF32, CF128, CF166, CF151, CF168, CF181, CF82, CF91 and CF143) were completely sequenced and SaPIs were identified in the genome. First, the insertion sites and *int* gene were detected in the genome, and then the SaPIs were localised. Predicted ORFs from SaPI sequences were annotated using Protein Blast against the non-redundant protein database maintained by the National Center for Biotechnology Information. The bioinformatic analysis showed that some strains carried identical SaPIs such as SaPI*int*I and SaPI*int*V from CF91, CF128 and CF166 or SaPI*int*III from CF151 and CF168.

The SaPI genomes have two components: genes involved in the ERP cycle organised into functional modules and those contributing to virulence or other phenotypic functions. *int* gene was at the extreme left end and classified the SaPIs according to the sequence in five types (Figure 4.7). All the sequenced SaPIs were *pac* types with the classical Operon I and the *terS* gene. The virulence genes were located in one of three regions: mainly at the extreme right end but also at the extreme left end or immediately to the right of *int* gene. The function of some genes from the virulence module was unknown. Contrary, other virulence genes were annotated and codified for *tst*, exfoliative toxin (*eta*), ferrichrome-binding protein (*fhuD*) or enterotoxins C and L (*entC* and *entL*).

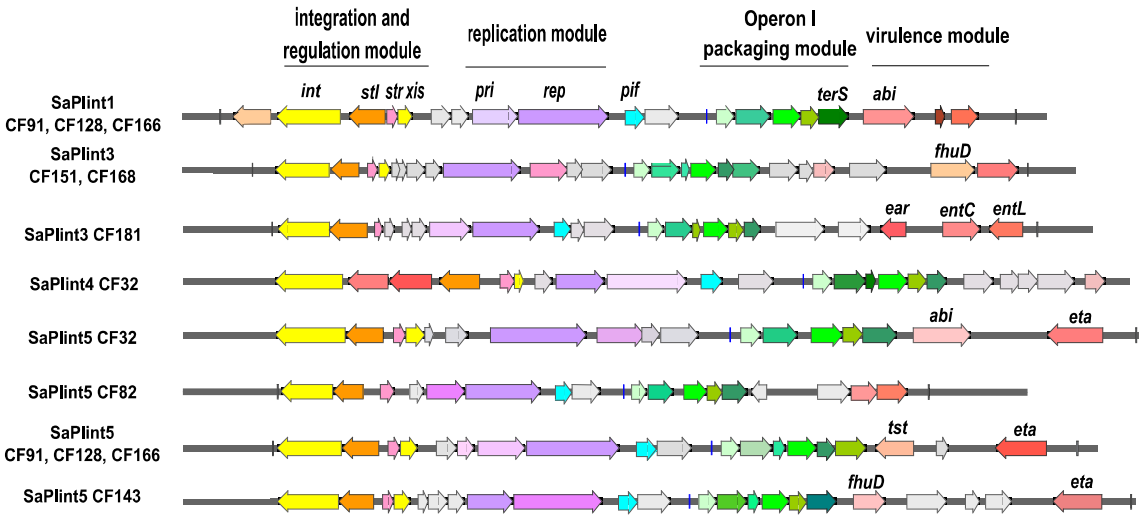

Figure 4.7 Comparison of SaPI genomes from *S. aureus* clinical strains. Genes are coloured according to their sequence and function. *int* (integrase) and *xis* (excisionase) are yellow. Transcription regulator *stl* is orange and *str* is pink. The replication genes including the primase (*pri*) and replication initiator (*rep*) are purple. *pif* (phage interference gene) is coloured in light blue. Packaging-interference genes are green and the accessory-virulence genes are red. Genes with unknown functions are grey.

4.1.3 SePI induction by endogenous phages in clinical strains isolated from catheter infections.

Once pathogenicity island mobilisation was confirmed in *S. aureus* strains isolated from CF patients, we wished to test the mobilisation in a different clinical environment. To this end, 48 *S. epidermidis* strains were collected from patients with catheter infections at the Hospital Universitario y Politécnico La Fe from 2015 to 2016.

S. epidermidis is generally less virulent than *S. aureus*. Although it forms part of normal skin flora. it is now emerging as an opportunistic pathogen causing nosocomial infections in hospital settings, primarily catheter and prosthesis bacteremia (Otto, 2009) (Qin *et al.*, 2017). *Staphylococcus epidermidis* Pathogenicity island (SePI) was first identified in 2011 (Madhusoodanan *et al.*, 2011) and, since then, few articles have focused on the study of SePIs (H. J. Chen *et al.*, 2011) (Argemi *et al.*, 2018).

As previously mentioned, SaPIs can be classified according to the integrase gene into five types as it is highly conserved and allows SaPI to be integrated at five different specific sites. Regarding SePIs, int_{SePI} genes are not described in literature,

therefore, an *in-silico* analysis was performed in order to identify *int*$_{SePI}$ genes. A search was conducted with the available *S. epidermidis* sequences found in the GenBank non-redundant database using different *int*$_{SaPI}$ sequences as probes. The *in-silico* analysis enabled the classification of SePIs according to *int* gene based on sequence homology with *int*$_{SaPIs}$ and to design specific primers to detect five types of SePIs (*int*$_{SePI}$ gene).

PCR and Southern blot were performed to identify the presence of SePIs in 48 *S. epidermidis* isolates. 83.3% of the isolates (n=40) contained at least one SePI, while 43.75% (n=21) encoded more than one SePI. SePI*int*V type was the most abundant (85%) followed by SePI*int*II (32%), SePI*int*I (28%) and SePI*int*III (15%). On the other hand, *int* typeIV was not detected (Figure 4.8).

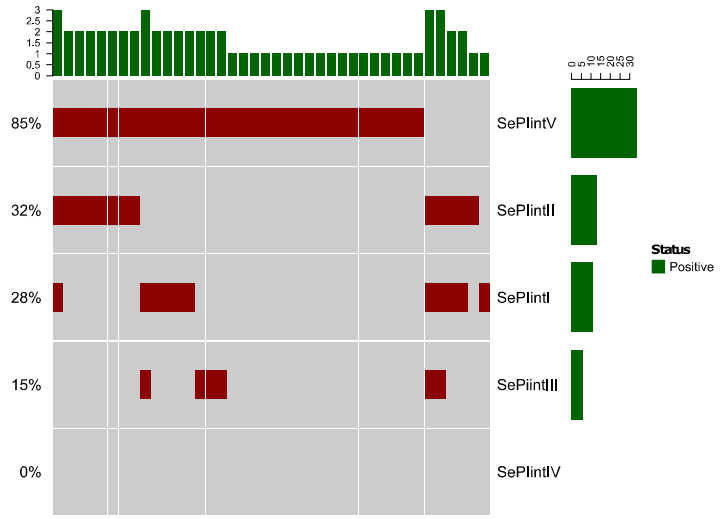

Figure 4.8 Distribution of SePIs in *S. epidermidis* clinical strains. The oncoprint provides an overview of SePI presence in *S. epidermidis* strains. Each column corresponds to one of the 40 SePI-positive *S. epidermidis* strains. The top row shows the number of SePIs integrated in the genome for each sample (green). The other five rows represent the presence of each type of SePI.

Although SePI mobilisation was already studied with no success (Madhusoodanan *et al*., 2011), we were interested in detecting SePI mobilisation by endogenous phages in *S. epidermidis* clinical strains. To this end, 40 SePI positive strains were exposed to MC for 120 min. Cultures were harvested and the pellets were processed (see the methods section). Southern blot was performed using specific probes from five identified *int*$_{SePI}$ genes. Only one strain, namely SCN45, contained SePIs mobilized by endogenous phages (Figure 4.9). SCN45 contained

three SePIs (*int*SePI types 1, 3 and 5) and two of them (*int*SePI types 3 and 5) were mobilised by their endogenous phages.

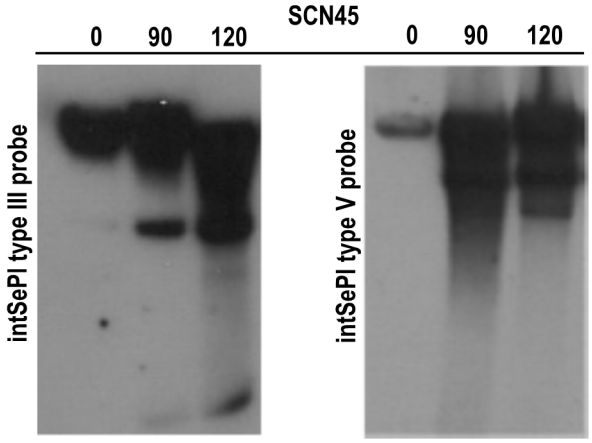

Figure 4.9 SePI induction by endogenous phages in *S. epidermidis* SCN45 strain. Strain SCN45 was induced with MC for 120 min. Samples were taken at time 0, 90 and 120 min. Samples were processed, electrophoresed and hybridised with specific probes.

SCN45 was sequenced using Illumina HiSeq sequencing technology. Bioinformatic analysis was conducted in order to identify SePIs and bacteriophages in the genome. Three SePIs were identified in SCN45 chromosome and the sequence comparison revealed extensive similarities to SaPIs. They shared modular genetic organisation and ORFs were matched to known SaPIs sequences on the basis of homology. All SePIs contained an integration-excision module (*int* and *xis* genes), a transcriptional regulation module (*stl* and *str* genes), a *terS* gene and other genes related to SePI packaging (Figure 4.10). Phaster web server was used to search for prophage sequences within the main chromosome, revealed two prophages, known as ΦSCN45_a and ΦSCN45_b. Their modular organisation was characteristic of the *Siphoviridae* bacteriophage family.

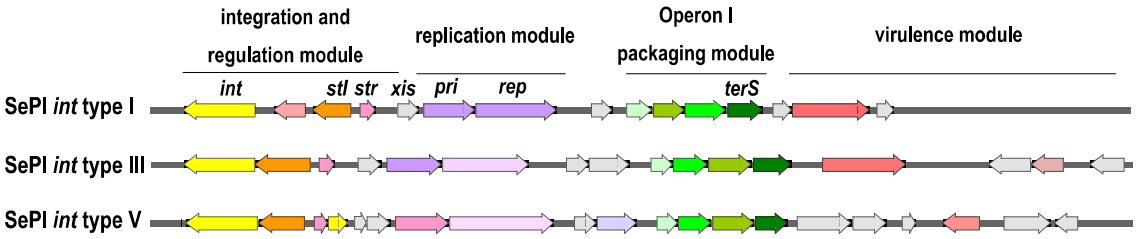

Figure 4.10 Genomic map of SePIs from *S. epidermidis* strain SCN45. Genes are coloured according to their sequence and function. *int* (integrase) is yellow. Transcription regulator *stl* is orange and *str* is pink. The replication genes include the primase (*pri*) and replication initiator (*rep*). Packaging-interference genes are green and the accessory-virulence genes are red. Genes with unknown functions are grey.

4.1.4 SaPI induction by endogenous phages in clinical strains of diverse origins

Once PICIs mobilisation was examined in two different collections (acquired from our hospital), the study was extended to other strains of diverse clinical origins. The selected clinical *S. aureus* strains were MRSA252, E-MRSA16, Mu50, TW20, MN8, USA300, COL, N315, MW2, A, B, C, D and E. These 14 strains were selected due to their broad characterisation, the easy access to their genome sequence and the presence of SaPIs and bacteriophages integrated in their chromosome. Table 4.1 shows some characteristics of SaPIs and bacteriophages integrated in the genome of the selected strains.

Table 4.1 Description of selected clinical strains.

Strain	SaPI (packaging type)	SaPI *int* type	Prophage (packaging type)	SaPI inducible by endogenous phage	Phage derepressor protein
E-MRSA16	SaPI4	I	ΦSa2, ΦSa3	no	-
MRSA252	SaPI4	I	ΦSa2, ΦSa3	no	-
Mu50	SaPIm4, SaPIm1	III, V	ΦSa3	no	-
TW20	SaPI1	IV	ΦSa1, ΦSa3	no	-
MN8	SaPI2	V	ΦSa1, ΦSa2, ΦSa3 ΦSa5, ΦSa6	no	-
USA300	SaPI5 (*cos*)	IV	ΦSa2usa (*pac*), ΦSa3usa (*cos*)	yes	unknown
COL	SaPI3 (*pac*)	IV	ΦCOL (*cos*)	yes	unknown
N315	SaPIn1(*pac*)	V	phiN315 (ΦSa3) (*cos*)	yes	Redβ
MW2	SaPImw2 (*cos*)	III	ΦSa2mw (*cos*), ΦSa3mw (*pac*)	yes	DUF3113
A	SaPIa (*pac*)	I	ΦSa3 (*cos*)	yes	Sri
B	SaPIb	I	ΦSa2, ΦSa3	no	-
C	SaPIc (*cos*)	III	ΦSa2, ΦSa6 (*cos*)	yes	DUF3113
D	SaPId	V	ΦSa1, ΦSa3, ΦSa6	no	-
E	SaPIe (*pac*)	I	ΦSa3 (*cos*)	yes	unknown

SaPI induction of the clinical strains was tested using Southern blot. Following the methodology previously described, the 14 selected strains were treated with MC. The Southern blots showed that 7 strains (50%) contained prophages able to induce and excise the resident SaPIs: MW2, C, USA300, A, E, N315 and COL (Figure 4.11).

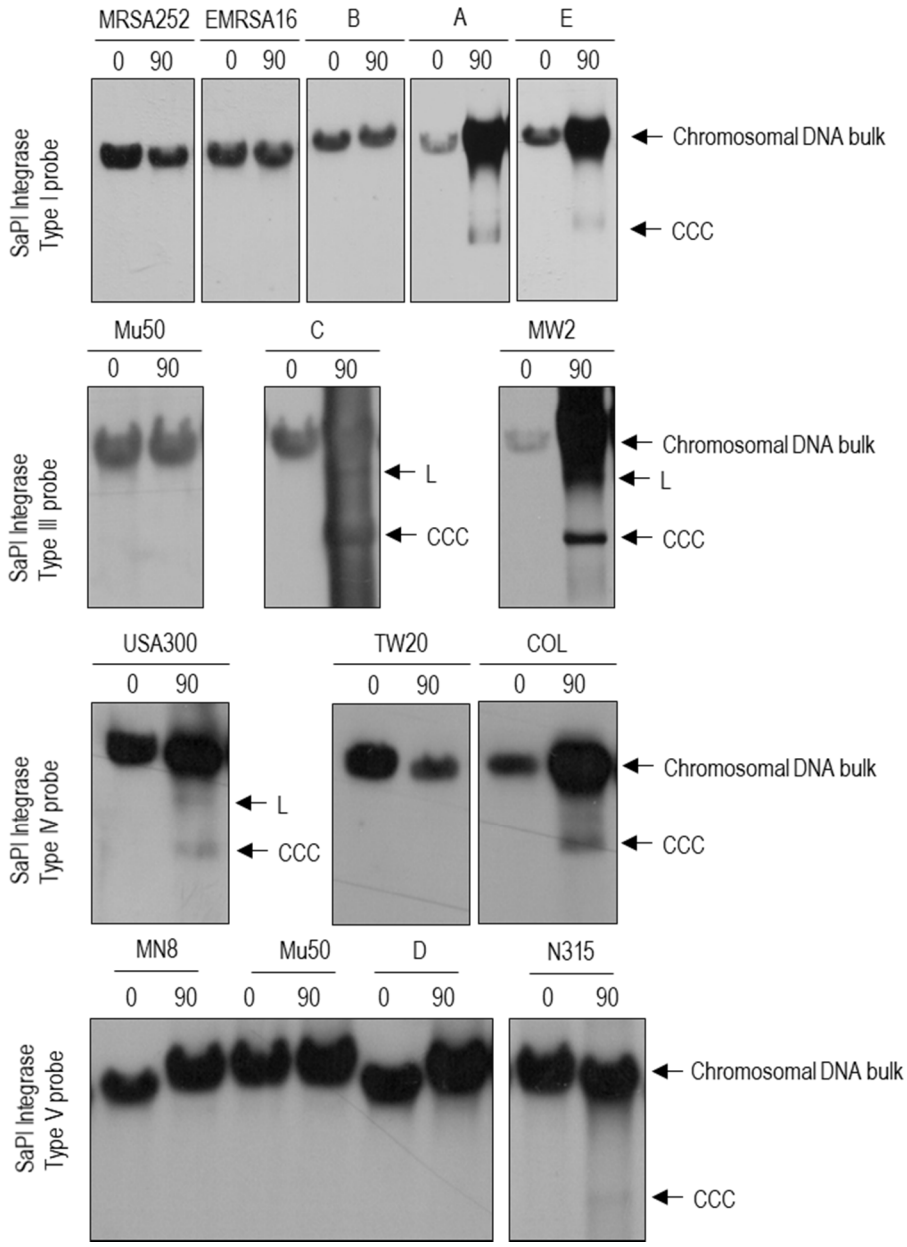

Figure 4.11 SaPI induction of 14 *S. aureus* clinical strains. 14 strains were treated with MC to activate the SOS response and induce the resident prophages. Samples were isolated at 0 and 90 min, separated in agarose gel and blotted with SaPI integrase-specific probe. Upper band corresponds to chromosomal bulk DNA including SaPI and the lower bands correspond to covalently closed circular SaPI (CCC) excised from the genome or SaPI monomers released from the phage capsid (L) if the helper phage mediated induction and packaging

It is important to take into consideration the difference between SaPI induction and SaPI mobilisation in order to correctly understand the results of the Southern blot. The specific interaction between the phage derepressor protein and its corresponding Stl induces the SaPI and initiates the excision and replication. In contrast, in addition to Stl-derepressor interaction, SaPI mobilisation requires compatibility in the packaging mechanism between the SaPI and phage. As

bacteriophages, SaPIs can be packaged using two different mechanisms: *pac* and *cos* (Feiss & Rao, 2012). Generally, *pac* SaPIs can be packaged by *pac* phages and *cos* SaPIs can use the packaging mechanisms of *pac* and *cos* phages (N. Quiles-Puchalt *et al*., 2014).

Regarding the Southern blot results, three different bands could be identified. The top band corresponded to the SaPI integrated in the host chromosome (bulk DNA) and is present in all the strains. The intermediate band corresponded to the SaPI linear monomers (L) released from the phage capsids and could be identified in MW2, C and USA300. This seems to indicate that the phage is able to induce and encapsidate the SaPI. The bottom band comprised covalently closed circular (CCC) molecules, which corresponded to SaPI induced but not packaged and it is present in A, E, N315 and COL strains (Figure 4.11).

To explain the absence of an L SaPI band in A, E, N315 and COL strains, the SaPI and prophage sequences of the clinical strains were analysed in order to identify whether they were *cos* or *pac* types (Table 4.1). The differences in migration and assignment in two types (CCC or L) among the analysed strains were consistent with the presence of *cos* and *pac* SaPIs and phages. MW2, C and USA300 encode *cos* SaPI and *cos* phages with similar *cos* sites, allowing the SaPIs to be induced and mobilized by their endogenous phages. A, E, N315 and COL contain *pac* SaPIs and *cos* phages, which means possible SaPI induction but not packaging.

To clearly confirm SaPI induction and subsequent excision from the chromosome and circularisation, a PCR was performed with inward and outward-directed primers (Figure 4.12). Specific primers were designated to recognise the flanking region of each SaPI type. All strains that produced a band that migrated at a position which suggests that SaPI induction by Southern blot was confirmed as the excised product by PCR.

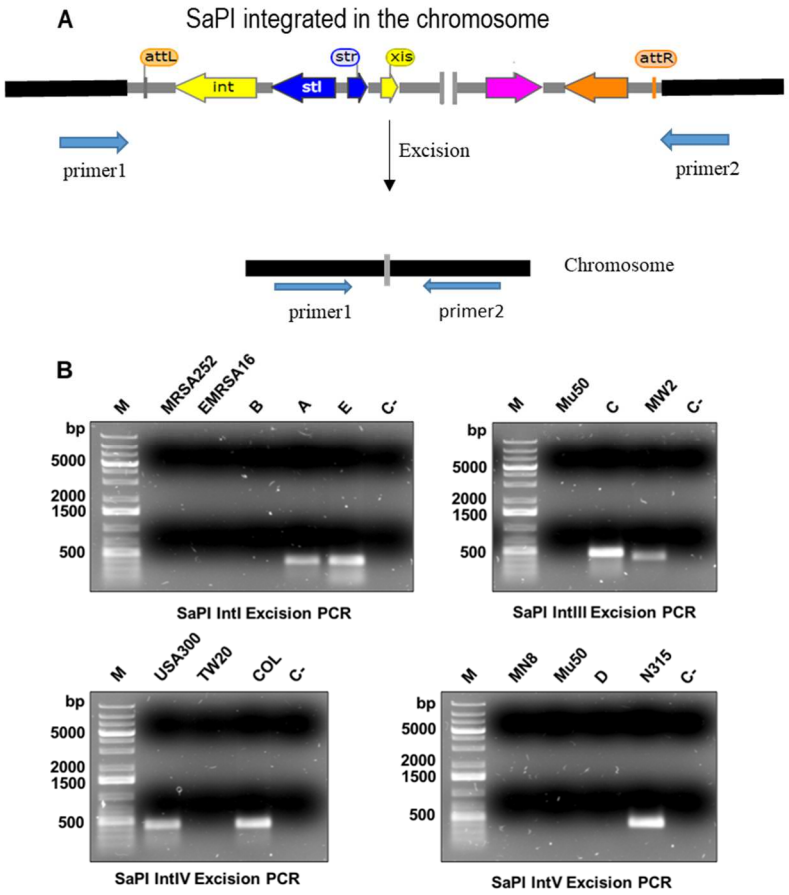

Figure 4.12 SaPI excision after induction in A, E, MW2, USA300, COL and N315 strains. A. Schematic representation of SaPI excision from the chromosome and PCR primers used for the identification. Specific primers recognising the flanking region of each SaPI type were used. **B.** Detection of SaPI excision by PCR amplification.

In addition, SaPI circularisation was tested by PCR using a pair of primers set divergently at both end of each SaPI. Only the positive strains for SaPI induction were tested, and all showed a band indicating SaPI circularisation. Figure 4.13 confirmed that the SaPIs encoded in A, E, C, MW2, USA300, COL and N315 strains were able to excise and circularise after induction.

In conclusion, all these results confirm the existence of SaPI induction after MC treatment in strains from different clinical environments. Some antibiotics can similarly induce SOS response to MC treatment as well as induce and mobilise SaPIs as an undesired effect. Therefore, it is important to thoroughly study the molecular mechanism involved in SaPI mobilization in order to prevent the appearance of unwanted effects.

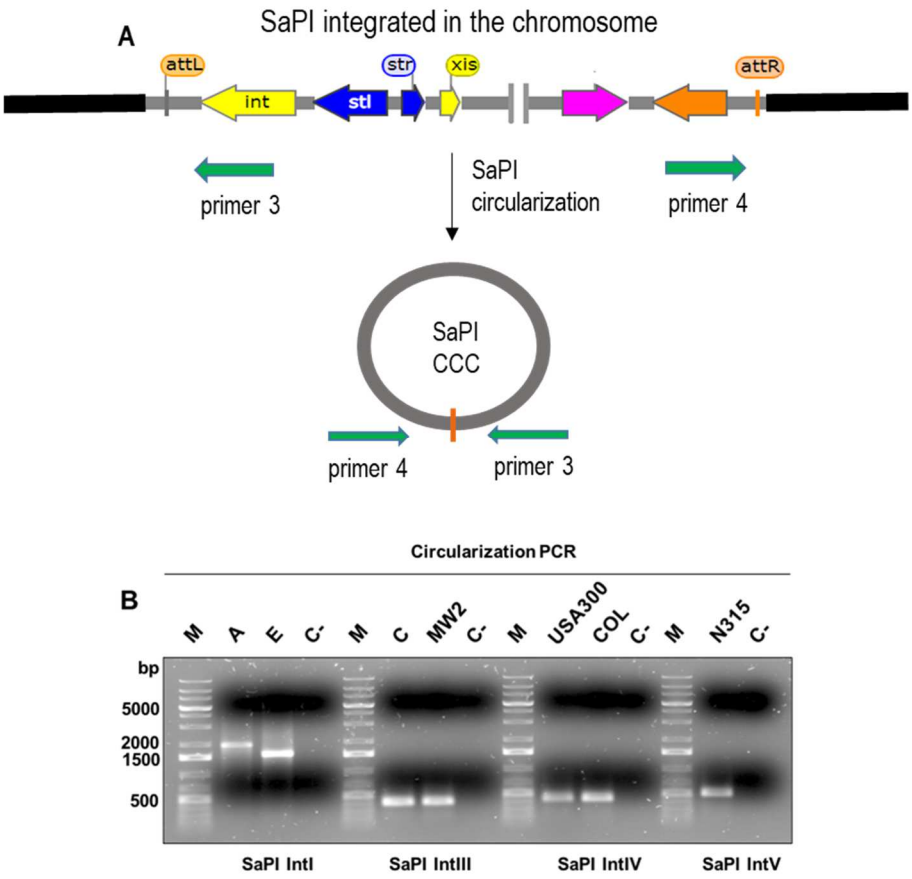

Figure 4.13 SaPI circularisation after excision in A, E, MW2, USA300, COL and N315 strains. A. Schematic representation of SaPI circularisation and PCR primers used. The primers were specifically designed for each SaPI. **B.** Detection of SaPI circularisation by PCR amplification using a pair of primers set divergently at both ends of SaPI. Only excised SaPIs were tested for circularization, all being positive.

The specific interaction between a phage derepressor protein and Stl is essential for SaPI induction. Hence, it is key to search for new phage proteins able to interact with different Stls for studying SaPI mobilisation and the subsequent dissemination of virulence genes contained therein.

Stl sequences were analysed for the 15 strains collected from CF patients, which contained SaPIs inducible by endogenous bacteriophages identified in the previous section. Additionally, 7 strains from the small collection from diverse clinical origins were also included. Analysis of the global repressor Stl sequences revealed its high divergence among SaPIs. In this context, some SaPIs contained identical Stl and could be clustered into seven groups according to the Stl sequence (Figure 4.14 and Table 4.2). Stl such as SaPI1 (strain A and CF32) had already been described and the derepressor was identified as a Sri protein. In addition, Stl like

SaPI2 was previously studied and the derepressors were phage recombinase proteins. The other Stls have not yet been studied and therefore, the phage proteins responsible for their interaction remain unknown.

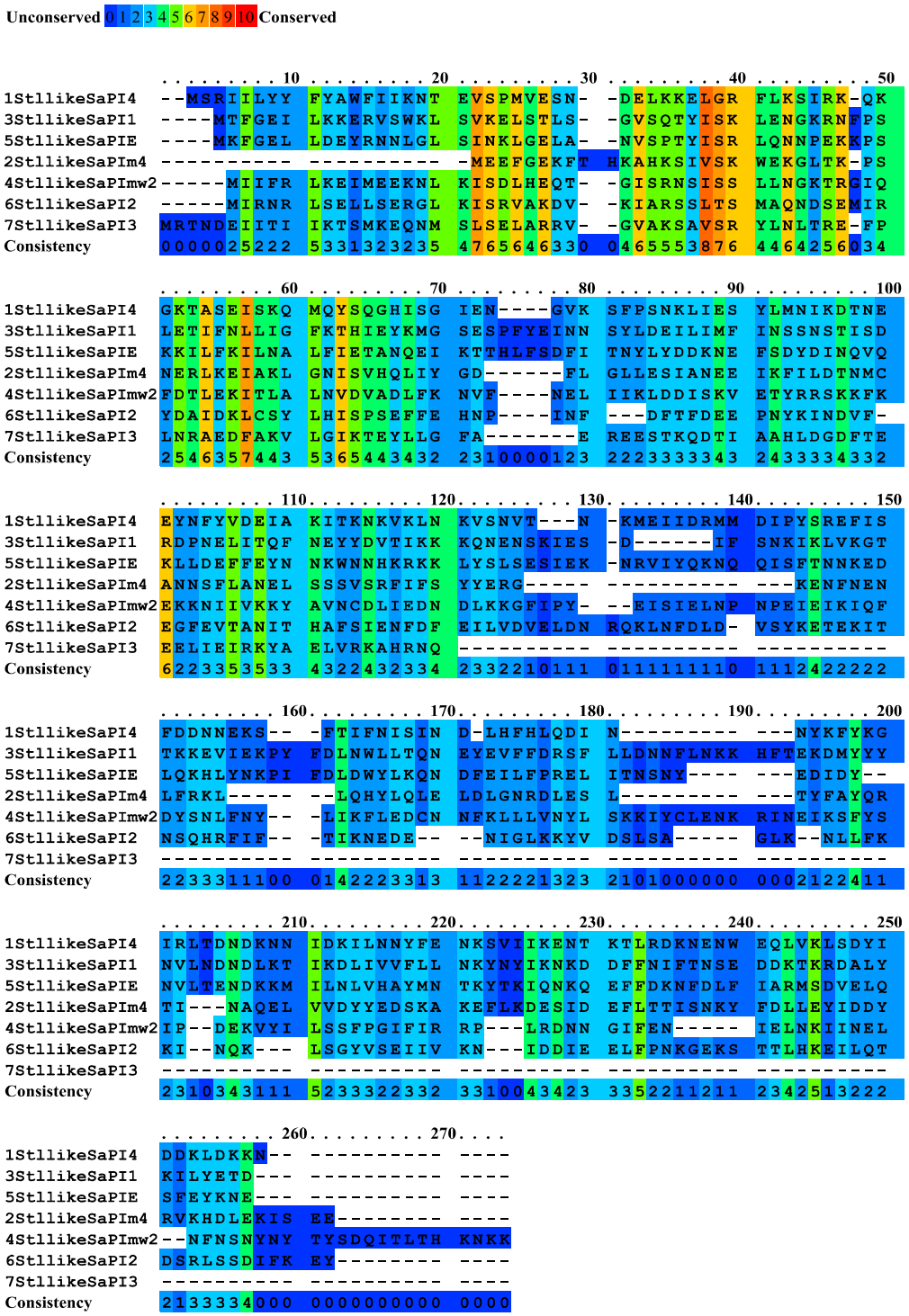

Figure 4.14 Alignment of Stl proteins from induced SaPIs. 1) Stl like SaPI4: CF14, CF128, CF132, CF166, CF173, CF205, 2) Stl like SaPIm4: CF151, CF168. 3) Stl like SaPI1: A, CF32. 4) Stl like SaPImw2: MW2 and C. 5) Stl like SaPIE: E 6) Stl like SaPI2: N315, CF82, CF91, CF128, CF132, CF143, CF166, CF220, CF231, CF242, CF249, CF253. 7)Stl like SaPI3: CF181, COL and USA300.

Table 4.2 Description of Stl and derepressor proteins from SaPI induced strains.

Strains	Stl like	derepressor
CF14, CF128, CF132, CF166, CF173, CF205	SaPI4	unknown
A, CF32, CF181	SaPI1	sri
E	new	unknown
CF151, CF168	SaPIm4	unknown
MW2, C	SaPImw2	unknown
USA300, COL, CF181	SaPI3	unknown
CF58, CF82, CF91, CF128, CF132, CF143, CF166, CF242, N315	SaPI2	recombinase

4.1.5 ΦSa2mw from MW2 induces SaPImw2 replication and transfer

Once SaPI mobilisation was tested in different strain collections, we focused on the identification of the phage protein responsible for the induction of the SaPIs contained in C and MW2 strains, which encode for the same Stl (Stl like SaPImw2). MW2 is a community-acquired, methicillin-resistant strain that carries a wide range of virulence and resistance genes. It was isolated from a child succumbing pneumonia and toxic shock syndrome. Regarding the carriage of MGEs, MW2 strain contains one SaPI (SaPImw2) and two bacteriophages (ϕSa2mw and ϕSa3mw) (Baba *et al*., 2002).

To study the mobilisation of SaPImw2 by endogenous phages, a chloramphenicol cassette was introduced in the island. SaPImw2 is a *cos* SaPI, therefore, to preserve genome size is very important to be transferred at high frequency by *cos* phages (Carpena et al., 2016). The chloramphenicol acetyl transferase gene (*cat*) of 651 bp was inserted to replace the accessory enterotoxin type C gene (*entC*) of 814 bp finally creating SaPImw2*entC::cat*. Given that ϕSa3mw is a *pac* phage and ϕSa2mw is a *cos* phage with the same *cos* site sequence as SaPImw2 (cggcgggggc), we first thought that ϕSa2mw was more likely to be involved in SaPImw2 transference than ϕSa3mw.

To develop this idea, first, a lysogenic strain (RN4220) for ϕSa2mw was generated and SaPImw2*entC::cat* was also transduced into the strain RN4220 ϕSa2mw. SaPImw2*entC::cat* induction was tested by southern blot confirming the ability of ϕSa2mw to mobilize the island (Figure 4.15). SaPI interferes with the helper phage reproduction what finally results in a reduction in the number of

plaques or phage titer (Frígols *et al*., 2015). Indeed, phage titer was reduced by two orders of magnitude in presence of SaPImw2 (Table 4.3) compared to ϕSa2mw lysogenic strain.

Stl represses the expression of the SaPI genes by binding in the intergenic regulatory region between *stl* and *str* genes. Only the interaction of Stl with a specific phage protein can release the repression and start the expression of the SaPI genes (M. Á. Tormo-Más *et al*., 2010). To ensure that ϕSa2mw was the responsible of SaPImw2 induction, a plasmid with a reporter fusion of β-lactamase gene with the intergenic regulatory region of SaPImw2, including *stl* and *str* genes, was constructed (pCN41-*stl-str*$_{SaPImw2}$-β-lactamase). The derivative plasmid was transformed into the lysogenic strain (RN4220 ϕSa2mw), treated with mitomycin C and the β-lactamase expression was measured. The analysis of the results revealed an increase of β-lactamase activity in the derivative strain concluding that ϕSa2mw activated SaPImw2 transcription (Figure 4.15).

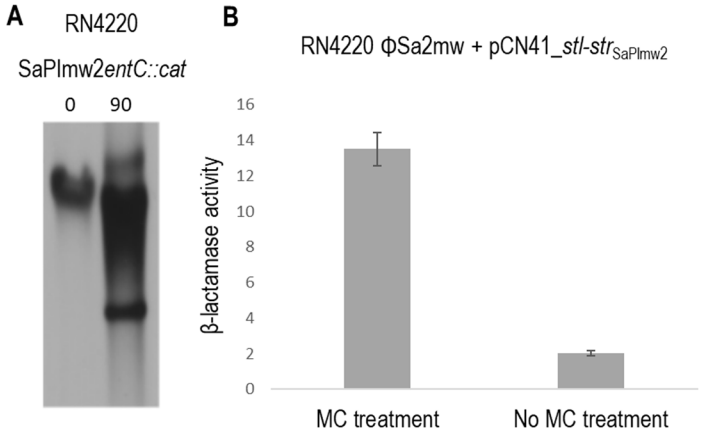

Figure 4.15 Induction of SaPImw2 by ΦSa2mw from MW2. A) Southern blot of ϕSa2mw lysogen in SaPImw2*entC::cat* positive strain (blotted with SaPI integrase type III probe) B) β-lactamase activity of ϕSa2mw lysogen with pCN41*stl-str* SaPImw2.

4.1.6 A hypothetical protein belonging to DUF3113 superfamily is responsible for derepression of SaPImw2

To identify which ϕSa2mw protein is responsible for the interaction with Stl$_{SaPImw2}$ and subsequent SaPImw2 induction, two different approaches were tested. First, seven regions of ϕSa2mw were cloned independently in a plasmid containing a cadmium promoter (pCU1-Pcad-Region$_{ϕSa2mw}$). These regions cover almost all the DNA replication module since all previous phage derepressor genes were located in

this module. Each plasmid was transformed in a derivative strain containing the plasmid pCN41-*stl-str*$_{SaPImw2}$-β-lact. After 2h of induction with $CdCl_2$, β-lactamase activity was measured indicating that the fifth section was the responsible of SaPImw2 induction. In turn, nine ORFs were encoded in the fifth region, and therefore, all were cloned in the plasmid pCU1-Pcad-gene$_{\phi Sa2mw}$. The nine plasmids were independently transformed into RN4220 pCN41-*stl-str*$_{SaPImw2}$ derivative strain and the β-lactamase activity was tested (Figure 4.16). In this case, the ORF responsible for β-lactamase activity corresponded to ORF MW1424, which encodes for a hypothetical protein that belongs to the DUF3113 superfamily.

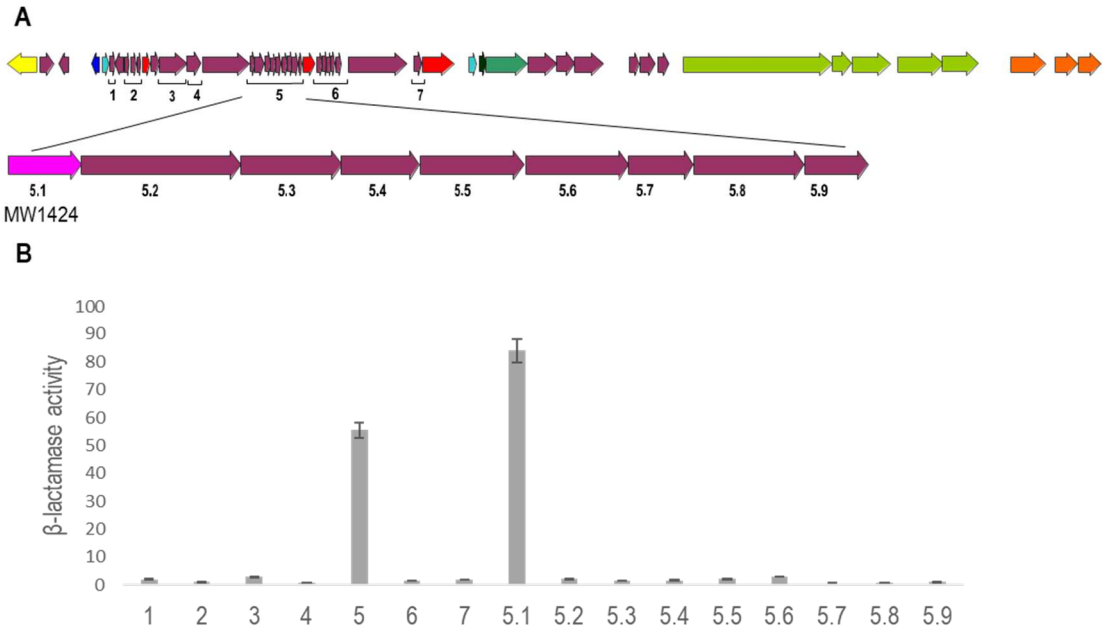

Figure 4.16 MW1424 gene is responsible for SaPImw2 derepression. A) Diagram of Sa2mw marked with the cloned regions. B) β-lactamase activity of derivative strains RN4220 + pCN41$_{stl\text{-}str\ SaPImw2}$ + pCU1P$_{Cad}$-region$_{\phi Sa2mw}$

SaPIs interfere with their helper phage reproduction and block plaque formation. Therefore, phage mutant in the derepressor gene unable to induce the SaPI would form plaque on a SaPI positive strain. As it was expected, ϕSa2mw infection in the SaPImw2 positive strain was obstructed and no plaques were formed. After several attempts, two plaques were obtained and DUF3113 (MW1424) gene was sequenced. Analysis of the sequence revealed changes in the ribosome-binding site and one mutation in the start codon (atg to atc; M1I) supporting the hypothesis that this gene as the inducer of SaPImw2.

DUF3113 (MW1424) gene was overexpressed by pCN51 vector in the SaPImw2 positive (RN4220 SaPImw2) strain and Southern blot test showed that this protein is the unique responsible for the induction. Deletion of DUF3113 ORF in the wt MW2 strain and in the lysogenic strain (RN4220 ϕSa2mw SaPImw2*entC::cat*) blocked SaPI induction and consequently, there was only one band in the Southern blot corresponding to the SaPImw2 integrated in the genome (Figure 4.17).

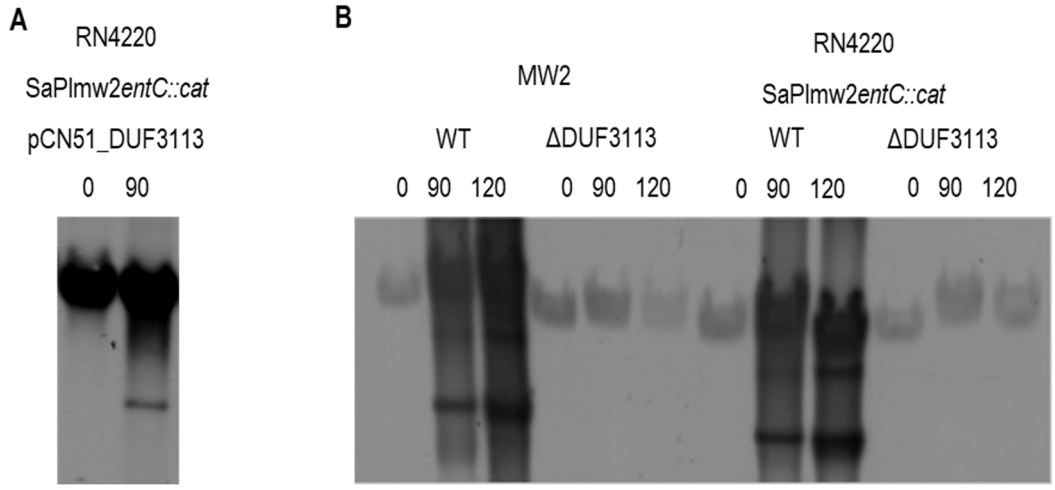

Figure 4.17 SaPImw2*entC*::*cat* induction by DUF3113. A) Southern blot of RN4220 SaPImw2*entC*::*cat* overexpressing DUF3113. B) Southern blot of MW2 and RN4220 SaPImw2*entC::cat* WT and DUF3113.

ϕSa2mw WT and DUF3113 mutant were lysogenised in SaPImw2 positive and negative strain and SaPI transference and phage titers were analysed after MC treatment. As table 4.3 shows, while ϕSa2mw was able to mobilise the SaPImw2, the ΔDUF3113 ϕSa2mw mutant completely abolished the SaPI transference. Complementation of DUF3113 in the mutant strain fully recovered the SaPI titer. Deletion of DUF3113 in ϕSa2mw had no effect in phage titers suggesting that the mutant is as viable as wt.

Table 4.3 Effect of WT and mutant ϕSa2mw in SaPI and phage titers.

Strain	Phage	SaPI	Plasmid	phage titer[T] (pfu)	SaPI titer* (cfu)
MW2	ΦSa2mw, ΦSa3mw	SaPImw2 *entC::cat*	-	4,82E+05	5,50E+07
RN4220	ΦSa2mw	-	-	2,86E+03	-
RN4220	ΦSa2mw	SaPImw2 *entC::cat*	-	9,32E+01	4,69E+04
RN4220	ΦSa2mwΔDUF3113	-	-	4,38E+03	-
RN4220	ΦSa2mwΔDUF3113	SaPImw2 *entC::cat*	-	2,25E+04	0
RN4220	ΦSa2mwΔDUF3113	SaPImw2 *entC::cat*	pCN51 DUF3113	7,00E+01	2,04E+05

The average of results from three independent experiments are shown. [T]Phage titer using RN4220 as an indicator. [*]Transfer of SaPImw2 using RN4220 as a recipient.

4.1.7 DUF3113 protein directly interacts with SaPImw2 Stl

Direct interaction between the Stl repressor and the phage derepressor proteins disrupts the Stl SaPI DNA binding and starts the SaPI induction (M. Á. Tormo-Más *et al.*, 2010) (M. Á. Tormo-Más *et al.*, 2013). To find out whether DUF3113 derepresses SaPImw2 *in vitro* using a similar mechanism, a size exclusion chromatography (SEC) assay was performed and the relevant peaks were analysed by SDS-PAGE. In SEC technique, the solute retention time sensitively depends on the solute's size. In our case DUF3113$_{\Phi Sa2mw}$ is 7.1 kDa and Stl$_{SaPImw2}$ is 32.5 kDa. First, His-tagged Stl$_{SaPImw2}$ and His-tagged DUF3113$_{\phi Sa2mw}$ were purified using a Ni-nta affinity column, and the purified proteins were injected in the SEC column (Figure 4.18). While His-tagged DUF3113$_{\phi Sa2mw}$ eluted in a single peak (PK187), His-tagged Stl$_{SaPImw2}$ eluted in two peaks (PK116 and PK165). The mixture of two proteins was also injected in the SEC column and three peaks were identified (PK116, PK135 and PK187). PK135 was a new peak with different retention volume corresponding with the protein complex as was further confirmed by SDS-PAGE.

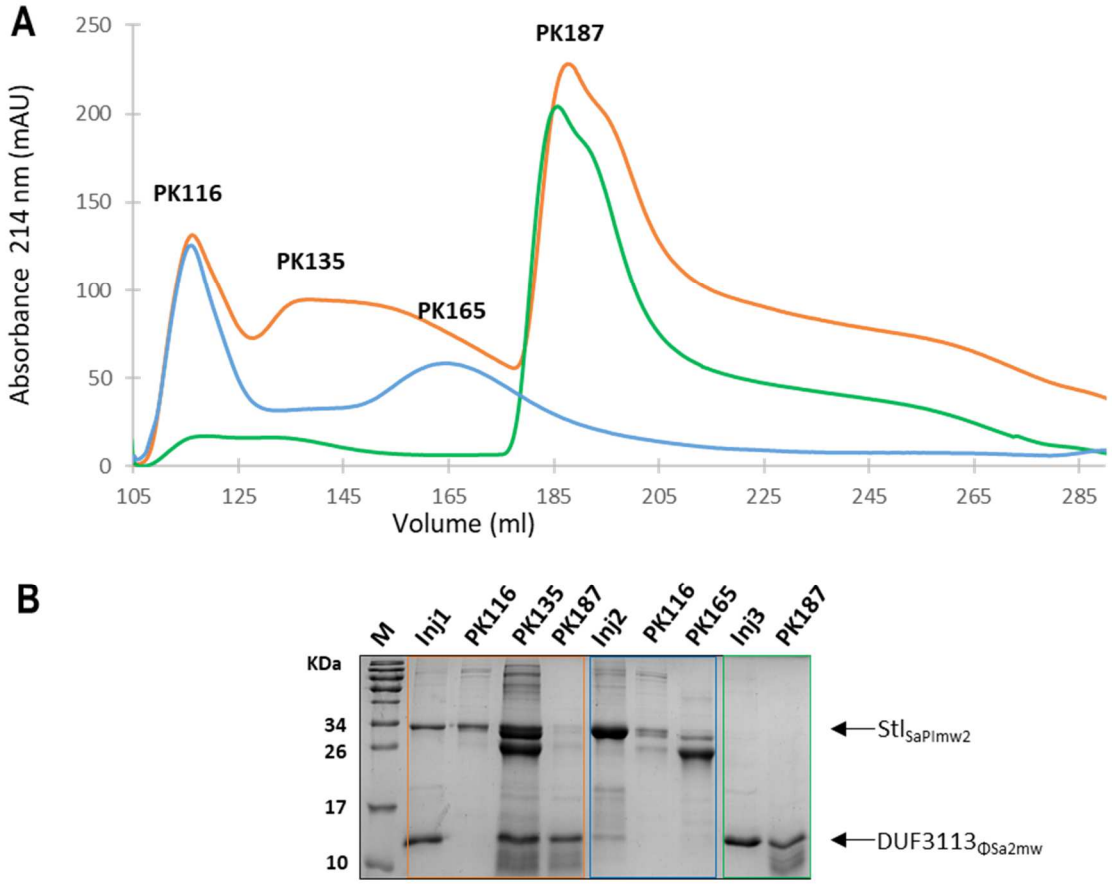

Figure 4.18 Interaction of DUF3113 protein with $Stl_{SaPImw2}$. Size exclusion chromatography assays. Blue: His-tagged-$Stl_{SaPImw2}$ chromatogram (PK116 and PK165). Green: His-tagged-$DUF3113_{\Phi Sa2mw}$ chromatogram (PK187). Orange: interaction between His-tagged-$Stl_{SaPImw2}$ and His-tagged-$DUF3113_{\Phi Sa2mw}$ (PK116, PK135, PK187). Peaks were confirmed by SDS-PAGE.

To verify the interaction, pull down assays were performed (Figure 4.19). His-tagged $Stl_{SaPImw2}$ and untagged $DUF3113_{\phi Sa2mw}$ were expressed and lysed by sonication. The supernatant was mixed, loaded onto a column and co-precipitate what confirmed the interaction. Untagged $DUF3113_{\phi Sa2mw}$ was loaded and was not retained by the resin.

To test the specificity of the interaction, two controls were used. Although His-tagged $Stl_{SaPIbov1}$ was incubated with untagged $DUF3113_{\phi Sa2mw}$, they eluted separately. As it was expected, SaPIbov1 was not induced by ϕSa2mw. Additionally, untagged $DUF3113_{\phi Sa2c,}$ which does not derepress SaPImw2, was incubated with his-tagged $Stl_{SaPImw2}$ and they did not co-elute. All these results support the idea of a direct and specific interaction between $DUF3113_{\phi Sa2mw}$ and $Stl_{SaPImw2}$.

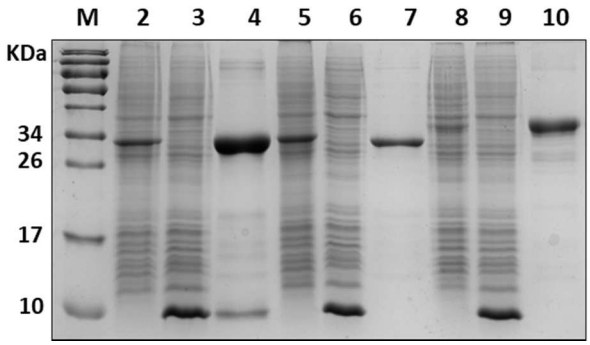

Figure 4.19 Affinity chromatography of DUF3113 for His-tagged Stl. Lane 2, 3, 5, 6, 8 and 9 are proteins induction supernatants after 20h at 16°C of 1mM IPTG treatment. Lane 4 is pull down of His_6-$Stl_{SaPImw2}$ (32.52 kDa) with $DUF3113_{\Phi Sa2mw}$ (7.1 kDa). Lane 7 is pull down of His_6-$Stl_{SaPImw2}$ (32.52 kDa) with $DUF3113_{\Phi 2c}$. Lane 10 is pull down of His_6-$Stl_{SaPIbov1}$ (35 kDa) with $DUF3113_{\Phi Sa2mw}$.

4.1.8 Diverse DUF3113 phage proteins display different capacities to induce SaPImw2

Since there are significant differences in the bacteriophage capacities to induce SaPIs, the ability of 80α, Φ11 and Φ12 to induce SaPImw2 was tested by Southern blot (Maiques *et al.*, 2007). Φ11 cannot induce SaPImw2 (there was no second band), whereas 80α and Φ12 were able to induce it (Figure 4.20). Φ12 is a *cos* phage with the same *cos* site as SaPImw2 and the Southern blot showed a low band corresponding to the monomeric form of SaPImw2 packaged in small capsids. The presence of *ccm* gene in SaPImw2 could be the responsible for Φ12 packaging in small capsids. Contrary, 80α is a *pac* phage and there was no lower band in the Southern blot. This could be explained by the fact that SaPImw2 does not encode CpmA-B proteins responsible for the interference to form small capsids in *pac* phages.

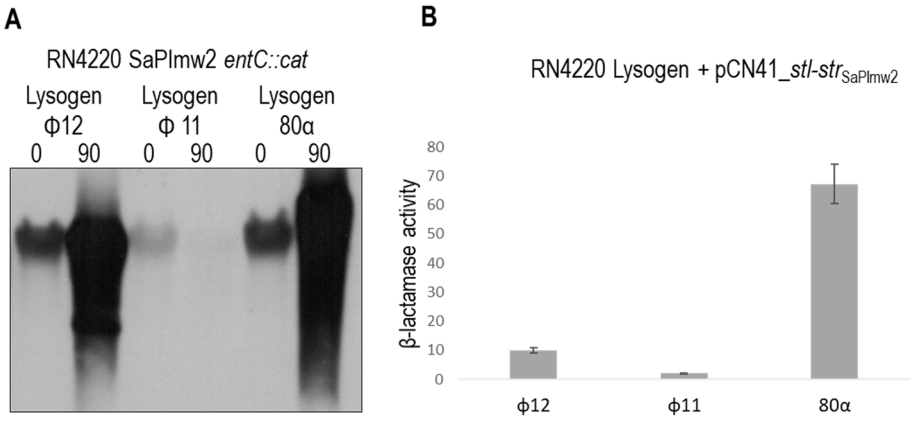

Figure 4.20 SaPImw2 induction by different bacteriophages. A Southern blot of Φ12, Φ11 and Φ80α lysogens in SaPImw2 positive strain blotted with SaPI integrase type III probe. B) β-lactamase activity of ϕ12, ϕ11 and ϕ80α lysogens with $pCN41_{stl\text{-}str\ SaPImw2}$.

SaPImw2*entC*::*cat* was transduced to 80α, Φ11 and Φ12 lysogens in the phage cured RN4220 to test for SaPI transfer. The transduced titers of SaPImw2*entC*::*cat* showed that 80α and Φ12 were able to transduce the SaPI successfully. On the other hand, as it was expected, Φ11 was unable to transfer SaPImw2*entC*::*cat* (Table 4.4).

Table 4.4 Phage (80α, Φ11 and Φ12) and SaPImw2 titers.

Strain	Phage	SaPI	phage titer^{T} (pfu)	SaPI titer* (cfu)
RN451	Φ11	-	2,77E+09	-
RN451	Φ11	SaPImw2 *entC::cat*	9,62E+07	7
RN451	Φ11ΔDUF3113	SaPImw2 *entC::cat*	5,92E+08	0
RN10359	80α	-	1,79E+10	-
RN10359	80α	SaPImw2 *entC::cat*	6,87E+08	3,21E+03
RN4220	Φ12	-	2,05E+05	-
RN4220	Φ12	SaPImw2 *entC::cat*	4,00E+03	1,71E+05
RN4220	Φ12ΔDUF3113	SaPImw2 *entC::cat*	9,23E+04	0

Furthermore, the ability to the lysogens to bind to $Stl_{SaPImw2}$ and activate the transcription was studied. The plasmid pCN41 with β-lactamase reporter gene fused to *stl-str*$_{SaPImw2}$ genes was transformed in the three lysogens. All tested phages showed different abilities to derepress the regulatory region and induce SaPImw2 (Figure 4.20).

It was previously demonstrated that allelic variants of phage derepressor proteins had different affinities for the SaPI-encoded repressors (M. Á Tormo-Más *et al*., 2013). *In silico* analysis of DUF3113 homologues sequences determined that $DUF3113_{Φ12}$ was identical to $DUF3113_{ΦSa2mw}$. Contrary, phage 80α and Φ11 encoded for DUF3113 homologues in ORF25 and ORF19, respectively, and both being identical in 61% of amino acids. Strain C, which was previously mentioned, contains one SaPI mobilized by an endogenous prophage and two prophages in its genome. The phages have been referred in this study as Φ2c and Φ6c and both contain DUF3113 homologues with 98% and 58% identity, respectively (Figure

4.21). Since Stl_{SaPIC} is identical in sequence to $Stl_{SaPImw2}$, $DUF3113_{\Phi 2c}$ or $DUF3113_{\Phi 6c}$ should interact with Stl protein and induce the SaPI.

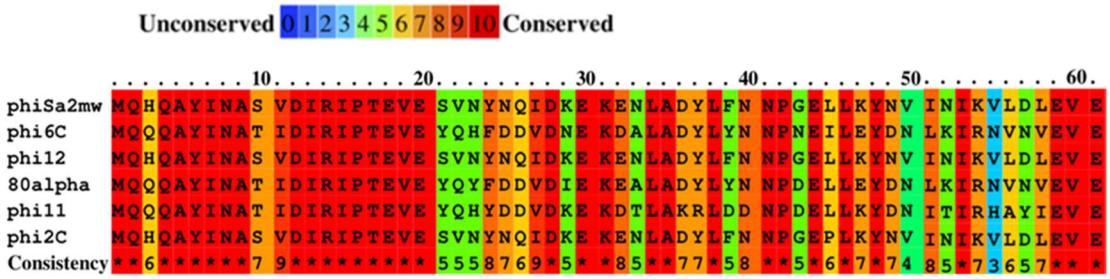

Figure 4.21 Alignment of *S. aureus* phage encoded for DUF3113 protein shows allelic variation. The scoring scheme assigns 0 for the least-conserved alignment position, with values up to 10 for the most-conserved alignment position.

The only function for the phage-encoded DUF3113 protein in the SaPI cycle is to disrupt the binding of the Stl repressor to the target DNA site. Consequently, allelic variants should induce the SaPI transcription in a different manner. This was confirmed by a procedure previously used in this thesis. DUF3113 homologues genes from different phages (80α, Φ11, Φ2c and Φ6c) were cloned into the plasmid pCU1 with cadmiun dependent Pcad promoter. These plasmids were transformed in RN4220 strain which carries a plasmid with β-lactamase reporter gene fused to *stl-str*$_{SaPImw2}$ genes. From this analysis it was observed that the presence of inducing concentration of $CdCl_2$ varied the β-lactamase activity between the tested strains.

In parallel, DUF3113 genes from phages 80α, Φ11, Φ2c and Φ6c were tagged with 3xflag and cloned into pCN51 with the same ribosome binding site (RBS). The plasmids were transformed into strain RN4220 SaPImw2*entC::cat* and treated with $CdCl_2$. Southern blot and Western blot were performed with the derivative strains (Figure 4.22). Southern blot showed differences in SaPImw2 induction when the proteins were equally expressed (as a result of cloning the allelic variants of DUF3113 gene with the same RBS). DUF3113 from 80α and Φ6c were able to induce SaPImw2, whereas, DUF3113 from Φ11 and Φ2c did not seem to induce it. From these results, it could be concluded that allelic variants differ in their capability to derepress SaPImw2 due to differences in their ability to bind to $Stl_{SaPImw2}$.

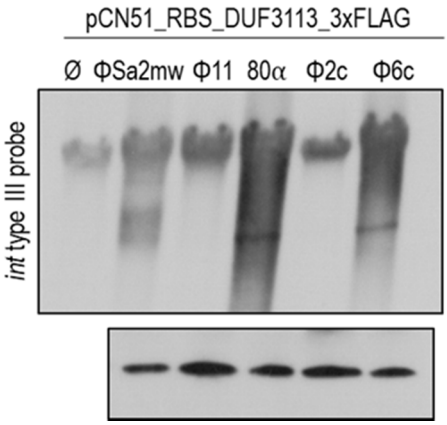

Figure 4.22 Differences in the ability of DUF3113 allelic variants to induce SaPImw2. Southern (top) and Western blots (bottom) of derivative RN4220 SaPImw2*entC::cat* strains carrying pCN51 3xFLAG homologues to the DUF3113 protein from phages ΦSa2mw, Φ11, 80α, Φ2c and Φ6c .

To determine whether DUF3113 was also present in another *Staphylococcus* spp, an *in-silico* study was performed. DUF3113 protein was found in phages from *Staphylococcus argenteus, Staphylococcus warneri, Staphylococcus haemolyticus, Staphylococcus schweitzeri* and *Staphylococcus simiae* (Figure 4.23). In addition, *S. argenteus* clinical strain (BN75) was identified, which contains a SaPI that encodes for a Stl similar to $Stl_{SaPImw2}$ (99% of identity) and a phage with a DUF3113 protein with 61% identity. In the same way, a SaPI that encodes for a Stl and a phage with a DUF3113 was detected in the *S. aureus* TCH70 strain with 100% identity. These results suggest that the mechanism described in this study could be general in *Staphylococcus* spp.

Part of these results have been published in Cervera Alamar *et al.,* 2018.

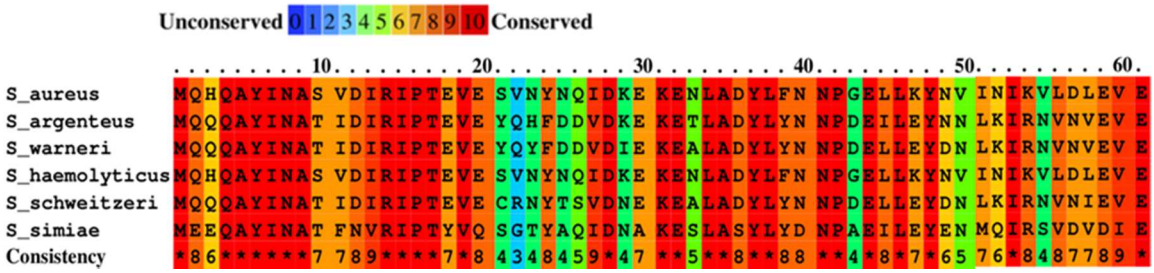

Figure 4.23 Alignment of DUF3113 from different Staphylococcal phage. The scoring scheme assigns 0 as the least-conserved alignment position, with values up to 10 for the most-conserved alignment position.

4.2 SaPIs control of chromosomally-encoded virulence gene expression

4.2.1 Background

SaPIs are clinically relevant due to their important role in the *S. aureus* pathogenicity. They usually carry two or more superantigens and are responsible for most superantigen-related human diseases. SaPIs are efficiently encapsidated into phage particles and can disseminate to other bacteria populations.

SaPIs were serendipitously discovered in 1998 due to their carriage of the gene for toxic shock syndrome toxin-1 (*tst*) of which the SaPIs are the sole source (Lindsay *et al*., 1998). Since then, many virulence genes such as enterotoxins B (*seb*), C (*sec*), K (*sek*) or Q (*seq*) were identified in these MGEs.

SaPIs are not merely virulence gene carriers, they have also developed different strategies that make them critical to the organism's evolution. For example, they mediate their own transfer and direct the horizontal transference of unlinked chromosomal DNA containing virulence genes (J. Chen *et al*., 2015). In addition, SaPIs can interfere with phage reproduction, increase the host survival of phage infections, determine animal host specificity or even force phage evolution (Geeta Ram *et al*., 2015)(Frígols *et al*., 2015) (Viana *et al*., 2015).

Bacteriophages, a type of MGE closely related to SaPIs, are also important contributors to host pathogenicity. They can directly provide virulence factors to the host bacteria by the lysogenic conversion phenomenon, as well as mobilise host genome segments through transduction. In addition, they influence the global regulators of the host, allowing them to adapt to different situations and regulate chromosomal genes (Osmundson, Montero-Diez, Westblade, Hochschild, & Darst, 2012). Some studies have detected small RNAs (sRNAs) in bacteriophages of *S. aureus* that regulate the expression of target chromosomal genes at the posttranscriptional level. Some sRNAs can regulate the expression of virulence factors, suggesting that they play an essential role in staphylococcal virulence (Chabelskaya, Gaillot, & Felden, 2010) (Eyraud, Tattevin, Chabelskaya, & Felden, 2014). A recent study demonstrated that lysogenisation with Φ11 and 80α phages

in *S. aureus* increases the biofilm formation, as well as other transcriptional changes in different regulatory pathways of the host (Fernández *et al*., 2018).

Due to this multifunctionality shown by the SaPIs, it is likely that they can develop other functions that could be important to bacteria. In this study, we explored the hypothesis that SaPIs are involved in the regulation of chromosomal genes expression. For this purpose, we performed a genome-wide transcriptional analysis using RN450 and its derivatives strains.

4.2.2 SaPIs controls the expression of chromosomal genes

To investigate whether SaPIs were able to control the expression of chromosomal genes, a tiling array transcriptome study was conducted in isogenic strains differing in the content of MGEs. The tiling array allows the simultaneous measurement of the expression level of thousands of genes during a particular time. This approach can emphasise whether genes are up- or down- regulated under specific conditions. A significant change in transcript highlights the impact of SaPIs on bacterial physiology. To ensure that all strains used for this study shared the same genetic background, bacteriophages and SaPIs were integrated into the non-lysogenic RN450 strain. In this case, the strains used were:

RN450 WT (Code 4)

RN450 + 80α (Code 1)

RN450 + SaPIbov1 (SB1) + SaPI1 (Code 4S)

RN450 + 80α + SaPIbov1 (SB1) + SaPI1 (Code 1SS)

RN450 SaPIbov1Δ*stl* (Code 4D)

All strains were treated with MC to induce the prophage and SaPIs encoded in their genome. Although, RN450 SaPIbov1Δ*stl* was also treated, it did not need to be induced as this mutant was constitutively induced in the absence of any inducing phage and external stimuli.

Total RNA was isolated from the samples at 0 and 60 min after MC induction, converted to cDNA and the cDNA hybridised with the microarray. In addition, the RNA from RN450 WT was isolated after 60' of MC treatment to be used as a control.

RNA transcript levels were determined for each strain and time point. Pairwise comparisons of gene expression profiles between different strains showed that SaPI induction had a big impact on the transcriptome of RN450. In this study, we focused on four pairwise comparisons that we considered to be the most descriptive and informative in terms of SaPI implication in chromosomal gene expression (Figure 4.24). We decided to only examine transcripts with log fold-changes of >=1 and <=-1. As a result, 33 genes were identified as differentially expressed at one or more time points in at least one of the strains or conditions. Differentially expressed genes (DEGs) are listed in Appendix 4.

Briefly, the expression pattern of RN450 SaPIbov1Δ*stl* (4D) was compared with RN450 + SaPIbov1 + SaPI (4S). This comparison showed differences in gene expression between a strain with uncontrolled SaPI replication due to the mutation in *stl* repressor with respect to the strain with SaPIs integrated and repressed in the genome. Both were untreated with MC, resulting in 1 downregulated and 23 upregulated genes

Comparison of gene expression between RN450 + SaPIbov1 + SaPI1 (4S-60') and RN450 WT (4-60 ') after 60 min of MC treatment revealed statistically significant changes in the expression of 7 genes. RN450 + SaPIbov1 + SaPI1 (4S-60') showed 6 upregulated genes whereas 1 was downregulated. The operon I of SaPIbov1 is regulated by lexA and is, therefore, activated when the SOS response is triggered. Thus, this comparison showed the differences in gene expression related to the activation of the Operon I of SaPIbov1 after induction with MC compared to the control strain induced with island absence.

In addition, RN450 + 80α + SaPIbov1 + SaPI1 (1SS-60) treated with MC reported significant changes in the expression of 12 genes comparing to MC-treated RN450 WT (4-60). 6 genes were negatively and 6 genes positively regulated. DEGs

occurred due to SOS response activation which results in phage induction and as a consequence in SaPI induction and replication.

Finally, the gene expression patterns after 60 min of MC treatment of RN450 + 80α (1S-60) and RN450 + 80α +SaPIbov1 + SaPI1 (1SS-60) were compared. This pattern presented DEGs due to SaPI induction avoiding genes related to phage 80α induction. In this comparison 8 genes were differentially regulated and all were upregulated.

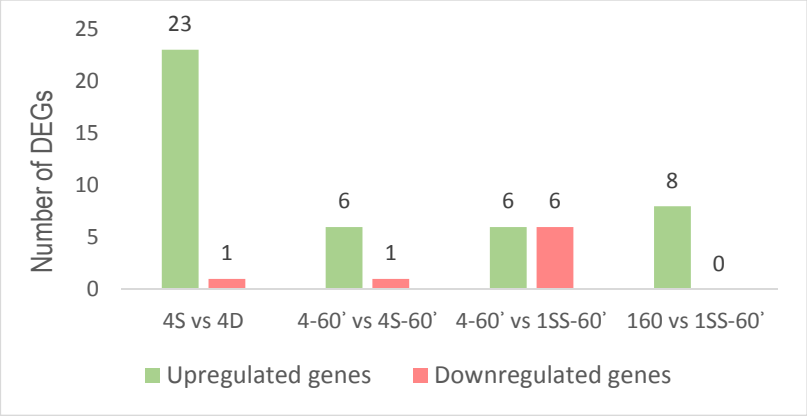

Figure 4.24 Number of DEGs after SaPI induction. Total number of genes upregulated and downregulated in the four pairwise comparisons of derivative RN450 strains.

4.2.3 *crtOPQMN* operon and SAOUHSC_02886 genes are overexpressed by induced SaPIs

The number of unique and overlapping DEGs between all the conditions that were tested are shown in a Venn diagram (Figure 4.25). Most of the DEGs were uniquely identified in a single strain while 6 genes were commonly differentially expressed in the 4 comparisons analysed. 5 of 6 genes corresponded to the *crtOPQMN* operon (SAOUHSC_02882 *crtO*, SAOUHSC_02881 *crtP*, SAOUHSC_02880 *crtQ*, SAOUHSC_02879 *crtM* and SAOUHSC_02877 *crtN*) and the other one was an acetyl transferase gene (SAOUHSC_02886 *gnat*). Interestingly, both loci have previously been linked with virulence.

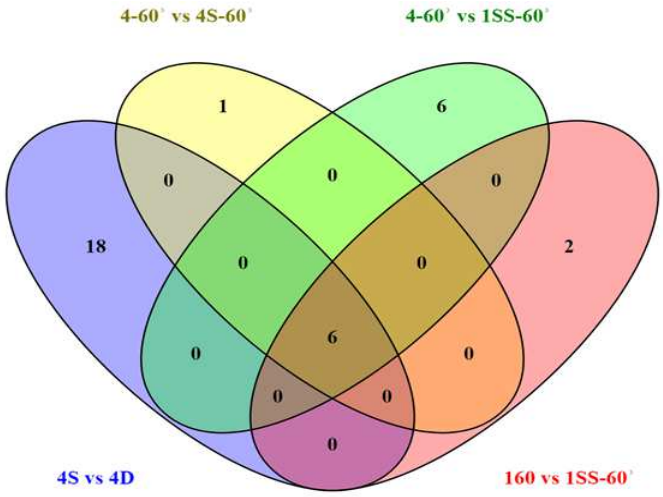

Figure 4.25 Venn diagram of DEGs in array analysis. Four lists of DEGs between the described comparisons were used to construct the Venn diagram. Blue circle indicates core subset of DEGs between 450 SaPIs vs 450 SB1Δ*stl*. Yellow circle, between 450 (t_{60}) vs 450 SaPIs (t_{60}). Green area shows differences between 450 (t_{60}) vs 450 + 80α + SaPIs (t_{60}) and red area indicates differences between 450 80α (t_{60}) vs 450 + SaPIs + 80α (t_{60}). Central area indicates two loci (six genes) that were found to have different expression profile across all tested conditions.

The *crtOPQMN* operon encodes five genes involved in the STX biosynthesis, and is controlled by a sigmaB-dependent promoter. STX is a C_{30} carotenoid pigment responsible for the characteristic golden colour of *S. aureus,* and it is also considered to be an important virulence factor. Additionally, STX plays an important role in resistance to reactive oxygen species (ROS) generated by neutrophils and macrophages, which makes *S. aureus* resistant to innate immune system (G. Y. Liu *et al*., 2005) (Clauditz *et al.,* 2006) (C. I. Liu *et al*., 2008). The first step of STX biosynthesis concerns the condensation of two molecules of farnesyl diphosphate that are catalysed by the dehydrosqualene synthase CrtM resulting in the formation of dehydrosqualene (4,4'-diapophytoene). The dehydrosqualene desaturase CrtN dehydrogenates dehydrosqualene to form the main intermediate yellow called 4,4'-diaponeurosporene. This molecule is then oxidised by CrtP, glycosylated by CrtQ, and esterified by the acyltransferase CrtO to give the STX carotenoid (Marshall & Wilmoth, 1981) (Pelz *et al*., 2005).

Deletion in the *crt* operon resulted a non-pigmented *S. aureus* with significantly reduced virulence in murine models of infections (C. I. Liu *et al*., 2008). Since STX is an important virulence factor, the STX biosynthetic pathway could be one of the targets to combat *S. aureus* infections (F. Chen *et al*., 2016) (Gao, Davies, & Kao, 2017).

SAOUHSC_02886 (*gnat*) codifies for an uncharacterised protein with an acetyltransferase domain (GNAT family). GNAT is a large enzyme superfamily that catalyses the transfer of an acetyl group from acetyl CoA to the primary amine of many substrates. Members of this family can confer resistance to aminoglycosides antibiotics such as kanamycin, gentamycin or tobramycin (Favrot, Blanchard, & Vergnolle, 2016).

To validate the transcriptome results, a qRT-PCR analysis was performed (Figure 4.26). The expression levels of SAOUHSC_02886 and *crtM* genes were compared between RN450 SaPIbov1 (not induced) and RN450 SaPIbov1Δ*stl* strains, in which all SaPI genes were derepressed (uncontrolled constant replication). Both genes were commonly overexpressed in all the conditions tested. Expression of these genes was significantly increased in the Δ*stl* mutant indicating that SaPI induction positively regulates the expression. Thus, the trend observed in the microarray was confirmed by qRT-PCR and the fold change was larger in this experiment than in the microarray.

Given the obtained results, we decided to focus on the study of STX expression due to its deeply demonstrated virulence. In addition, STX is a carotenoid responsible for strain pigmentation and expression changes produce phenotypic differences.

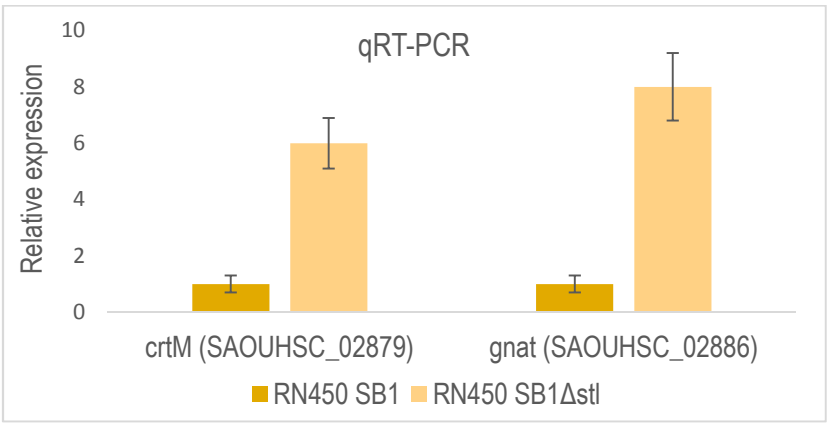

Figure 4.26 qRT-PCR validation of transcriptome data for two selected genes. Expression of *crtM* and SAOUHSC_02886 genes were tested for SB1 wild-type and SB1Δ*stl* mutant. Relative expression was normalized to the abundance of *gyrB* gene.

4.2.4 SaPI regulate STX production by independent Sigma B system

In previous studies, Úbeda and collaborators reported that the repressor *stl* deletion in SaPIbov1 caused orange pigmentation in RN4220 *S. aureus* strain (Úbeda *et al*., 2008). They thought that the stress of unrestricted SaPI replication was responsible for the pigmentation. Using the expression profiling information presented in this work, however, it could be deduced that SaPIs were directly involved in carotenoid expression and consequently in the pigmentation change. STX is a pigment and an important virulence factor due to its antioxidant properties. For this reason, it is important to study the regulatory mechanism that would allow the development of virulence factor-specific drugs.

It is known that the stress response alternative sigma B factor (SigB - σ^B) regulates STX production, since σ^B mutant is white (Lan, Cheng, Dunman, Missiakas, & He, 2010). Although the regulation mechanism is not confirmed, it seems to be direct since a consensus σ^B DNA binding motif was identified in the *crt* operon promoter upstream of *crtO* gene (Kullik, Giachino, & Fuchs, 1998) (Pelz *et al*., 2005). We would like to know whether STX production regulated by SaPIs was linked to SigB regulation. To answer this question, the SaPIbov1Δ*stl* (SB1Δ*stl*) was transduced to a σ^B positive SH1000 strain and to a σ^B mutant strain. Cells were harvested after 16h, the pellet was visualised and the carotenoid pigment was measured (Figure 4.27). The pigmentation was slightly higher in SH1000 SB1Δ*stl* compared to SH1000 WT. The SB1Δ*stl* transduction to the white SH1000 $\Delta\sigma^B$ produced an orange phenotype, which confirms that the regulation mechanism of SB1 occurred independently of σ^B.

To ensure that the SB1 effect on pigmentation was due to changes in *crt* operon expression, *crt*-operon deletion mutant in SH1000 SB1Δ*stl* and SH1000 $\Delta\sigma^B$ SB1Δ*stl* strains were generated. As shown in Figure 4.27, both strains were white despite SB1 derepression.

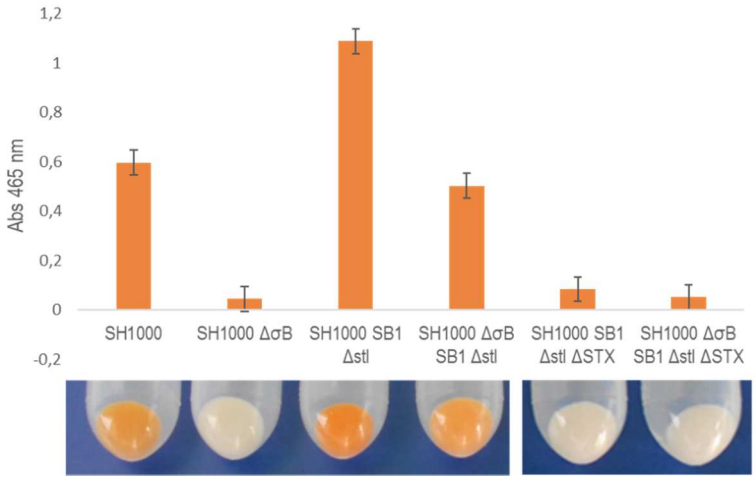

Figure 4.27 Production of staphyloxanthin in *S. aureus* strains SH1000 and its derivatives. The bottom panel shows the colour of the bacterial cell pellet from SH1000 and its derivatives: SB1Δ*stl* and SB1Δ*stl* ΔSTX (WT, $\Delta\sigma^B$). The top graph shows the quantification of carotenoid pigment production. Values indicate the average of three independent experiments.

4.2.5 LexA-controlled SaPI operon I regulates staphyloxanthin production

To further understand the SaPI regulatory mechanism in relation to STX production, we wanted to specify which region or genes from the SaPI could be involved. All SaPIs contain a conserved genomic structure in which genes are organised into functional modules. One of these modules is the Operon I (OpI), which contains genes involved in SaPI encapsidation and transfer. The expression of the Operon I occurs independently of the SaPI cycle but is controlled by a LexA-dependent promoter (Úbeda *et al*., 2007).

The SOS response is a global gene regulatory system used by bacteria to respond to DNA damage (Baharoglu & Mazel, 2014). MC was used to stress the cells and induce SOS response in laboratory conditions. SOS response produces the activation of the protease RecA, which in turn induces autocleavage of the repressor LexA (Butala, Žgur-Bertok, & Busby, 2009). In a strain with a prophage, the active RecA initiates the phage repressor CI autocleavage leading to the expression of the phage genes and the initiation of the lytic cycle (Goerke, Koller, & Wolz, 2006). In a SaPI positive strain, following SOS response, the Lex-A dependent promoter of the Operon I is activated and only promotes the expression of this gene cluster. In a SaPI and helper prophage positive strain, the SOS response triggers prophage induction

and subsequent SaPI induction with the expression of all the SaPI genes (Úbeda *et al*., 2005).

The expression profile showed overexpression of *crtOPQMN* operon in the SaPIs positive strain (4S-60) after 60 min of MC treatment compared to the treated RN450 WT (4-60). Since this strain only contained SaPIs integrated in the genome (prophage negative strain), the SOS response only caused the Lex-A-dependent Operon I activation. Thus, we suspected that Operon I was the only one involved in STX expression. To test this hypothesis, SaPIbov1 Operon I mutant strain (RN450 SB1ΔOpI) was used. The capacity of RN450 SB1ΔOpI to produce STX was analysed after MC induction. As shown in Figure 4.28, the Operon I must be required for STX production since the OpI mutant strain displayed a decreased pigment production compared to the WT.

4.2.6 PtiA and PtiM are involved in STX production

In SaPIbov1, the LexA-mediated Operon I comprises 6 genes (from ORF5 to ORF10) that codify homologues to *terS* (ORF5), *ptiA* (ORF6), *ptiM* (ORF7), *cpmB* (ORF8), *cpmA* (ORF9) and *ptiB* (ORF10). All are involved in SaPI morphogenesis, phage interference and packaging. Thus, to identify the gene or genes responsible for the STX production, SaPIbov1 was transduced to RN450 strain and independent in-frame deletions in each gene of Operon$I_{SaPIbov1}$ were generated. The capacity to produce STX was examined in six derivative strains after MC treatment (RN450 SB1 ΔORF5, RN450 SB1 ΔORF6, RN450 SB1 ΔORF7, RN450 SB1 ΔORF8, RN450 SB1 ΔORF9 and RN450 SB1 ΔORF10). As shown in Figure 4.28, ORFs 6 and 7 seemed to be involved in pigment production as both gene deletions had the greatest effect on pigment reduction.

In order to study STX expression, it is important to take into account the strain selected to perform the genetic manipulations. Basal pigmentation could differ between strains, and this must be considered when interpreting the results. In this instance, RN450 was used as a background, but SH1000 and SH1000 $\Delta\sigma^B$ were used in the previous section. RN450 (also known as 8325-4) is a SigB positive strain with a 11 bp deletion in *rsbU,* which is a positive activator of SigB. Deletion of *rsbU* results

in defective SigB activity, and phenotypically, RN450 results in a poor basal pigmentation. SH1000 derives from RN450 but the *rsbU* was repaired and showed high pigmentation. SH1000 $\Delta\sigma^B$ was generated from SH1000 and displayed a colourless phenotype (white) due to the lack of the positive regulator σ^B (Herbert *et al*., 2010) (Bæk *et al*., 2013).

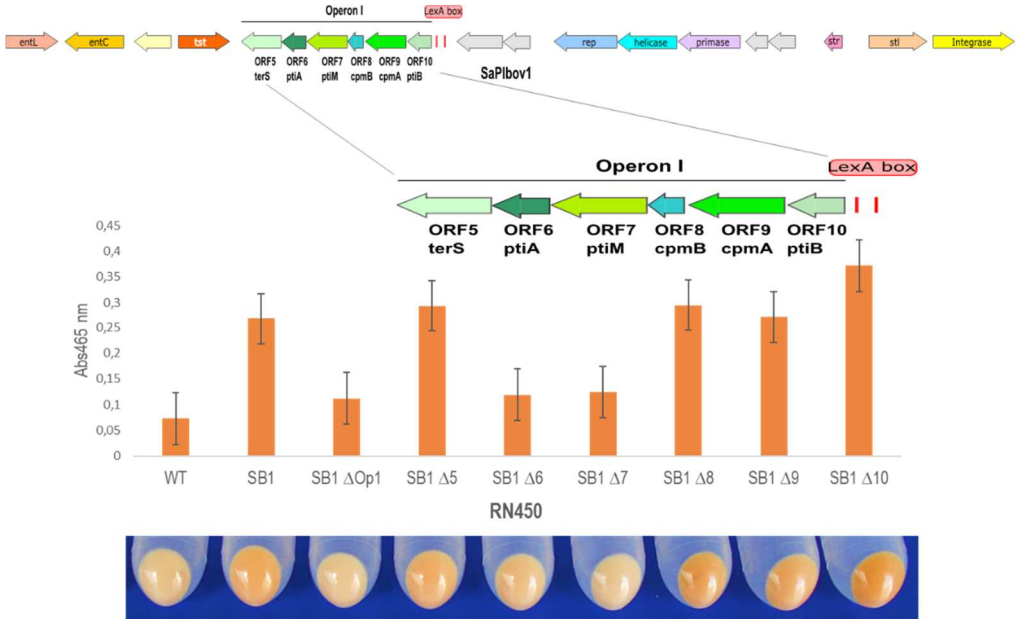

Figure 4.28 Pigment production of RN450 SB1 mutants. The top part shows the quantification of pigment production in RN450 SB1 Operon I mutant and in all genes contained in the OpI: ΔORF5 (*terS),* ΔORF6 (*ptiA),* ΔORF7 (*ptiM)*, ΔORF8 (*cpmB)*, ΔORF9 (*cpmA)* and ΔORF10 (*ptiB*). The bottom panel shows the colour of the bacterial cell pellet.

To confirm whether ORF6 and ORF 7 were involved in strain pigmentation, both genes were cloned separately and together into a pCN51 vector containing a cadmium-inducible promoter (P_{cad}) and then transformed into the laboratory strain RN450. The vector positive strains were induced with $CdCl_2$ and the obtained pellets were analysed to measure pigment production. As shown below, in Figure 4.29, ORF6 overexpression had no effect regarding colour changes, while ORF7 showed a slight increase in pigmentation. Remarkably, when ORFs 6 and 7 were co-expressed, staphyloxanthin pigment production was significantly increased.

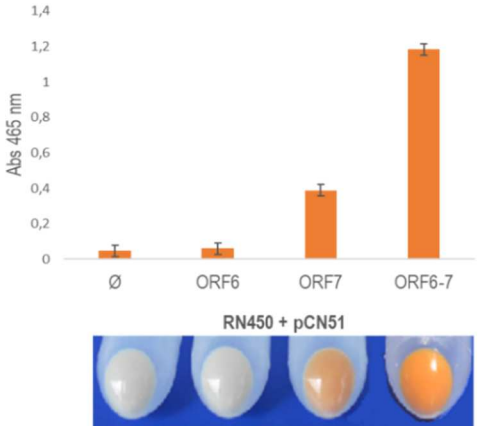

Figure 4.29 Production of STX by ORF6 and ORF7 overexpression. Top: Measurement of STX extracted in methanol of RN450 and its derivatives with pCN51 vector. Bottom: colour of analysed bacterial cells. Data represents the mean of the results of three independent experiments.

The plasmid pCN51 ORF6-7 was also transformed into SH1000 and SH1000Δσ^B strains, and the derivative strains were treated with $CdCl_2$. Quantification of pigment production by carotenoid extraction confirmed that, regardless of the presence or absence of SigB, both strains exhibited strong orange pigmentation. To ensure that differences in pigmentation were related to changes in *crtOPQMN* expression, deletion of this operon was generated in SH1000 and SH1000Δσ^B (Δ*crt* = ΔSTX), then the pCN51 ORF6-7 plasmid was transformed and the derivative strains were finally treated with $CdCl_2$. The absence of colony pigmentation in both strains verified that *crtOPQMN* operon was responsible for strain coloration (Figure 4.30).

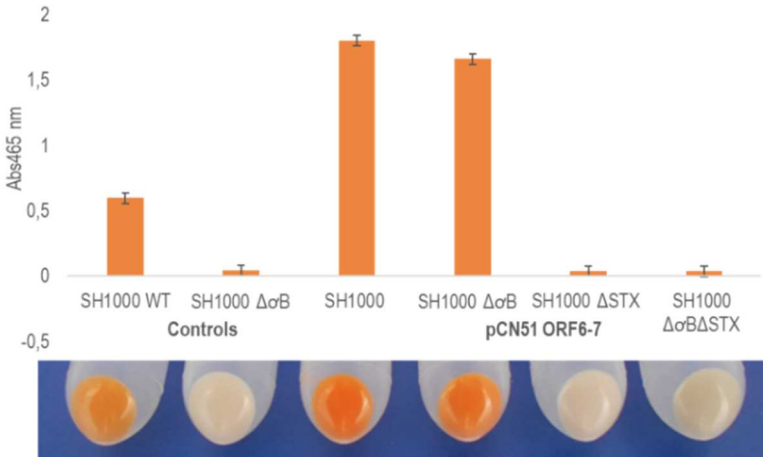

Figure 4.30 Production of STX when ORF6 and ORF7 from SaPIbov1 were overexpressed in SH1000 derivatives strains The bottom panel shows the colour of the bacterial cell pellets from SH1000 and its derivatives: pCN51_ORF6-7 (WT, Δσ^B) and ΔSTX pCN51_ORF6-7 (WT, Δσ^B). The top graph shows the quantification of carotenoid pigment production.

qRT-PCR was performed to confirm the different expression of *crtM* (from *crt*-operon) and SAOUHSC_02886 genes when PtiA and PtiM were overexpressed. RNA from RN450 with pCN51Ø or pCN51_*ORF6-7* was isolated, the cDNA was synthetized and the expression was quantified. Both *crtM* and SAOUHSC_02886 genes were overexpressed with PtiA-M presence (Figure 4.31).

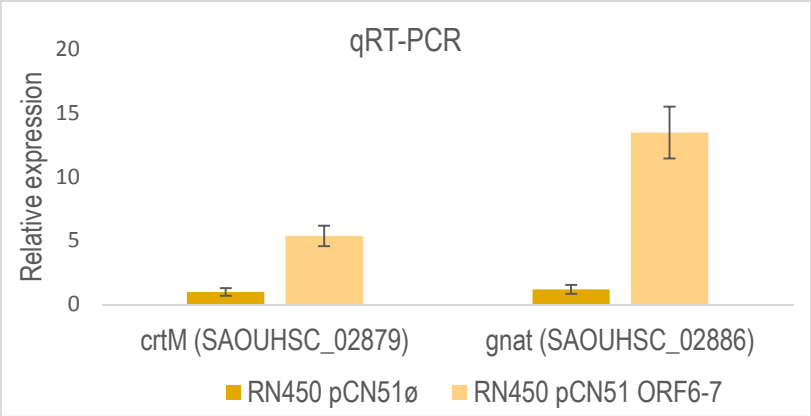

Figure 4.31 qRT-PCR for *crtM* and SAOUHSC_02886 genes. Expression of *crtM* and SAOUHSC_02886 genes were tested for RN450 pCN51Ø and pCN51_ORF6-7. Relative expression was normalized to the abundance of *gyrB* gene.

The ORF6 and ORF7 homologs present in SaPI2 have been previously characterised and named *pti*A and *pti*M, respectively. These two genes are involved in phage interference mechanisms. PtiA binds to the phage transcription activator LtrC and blocks the production of virion and lysis proteins. PtiM acts as a modulator protein, binds to PtiA and precisely modulates their activity to not interfere with SaPI particle formation (Geeta Ram et al., 2014) (Geeta Ram et al., 2015). Therefore, PtiA and PtiM directly interact to form a complex that enhances STX production. They are moonlighting proteins sine they have more than one function, one related to STX production and the other related to phage packaging interference.

To confirm that the STX regulation by PtiA-M (ORF6-7) is not exclusive to SaPIbov1, homologous to PtiA-M in different SaPIs (SaPI2, SaPIeq and SaPIpt1028) were cloned in the pCN51 vector and transformed into the RN450 strain. Sequence alignment of the proteins showed that all were highly similar (more than 90% identity). After $CdCl_2$ treatment, all strains showed an increase in pigmentation, and therefore homologues to PtiA and PtiM could be responsible for STX expression (Figure 4.32 and Figure 4.33)

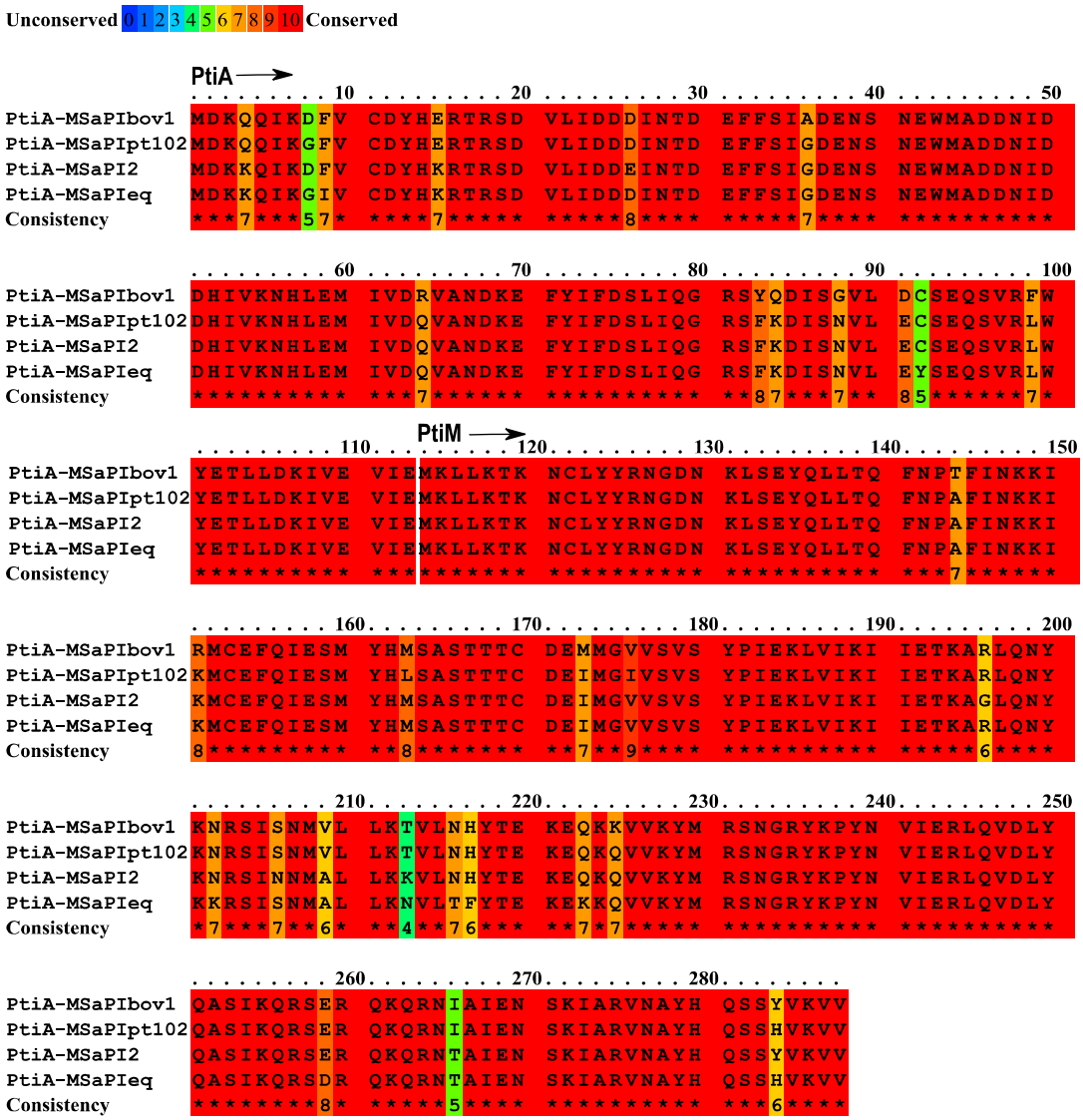

Figure 4.32 Alignment of PtiA and PtiM from different SaPIs. PtiA and PtiM from SaPIbov1, SaPI2, SaPIeq and SaPIpt1028 were aligned using PRALINE multiple sequence alignment program.

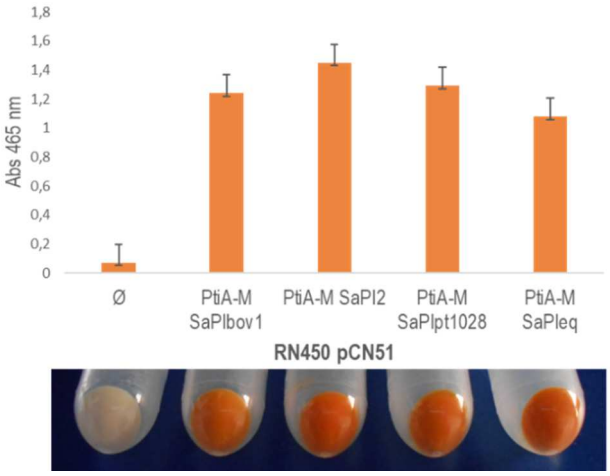

Figure 4.33 STX production by homologous proteins to PtiA and PtiM from SaPIbov1. Top graph: quantification of STX carotenoid of RN450 strains and its derivatives: RN450 + pCN51_*ptiA-M* from SaPIbov1, SaPI2, SaPIpt1028 and SaPIeq. Bottom image: colour of the bacterial cell pellets.

Finally, to rule out the possibility that the observed effect was dependent on a regulatory mRNA contained in *ptiA-M,* four independent nonsense mutations in *ptiA-M* genes were introduced: K2stop and R64stop in *ptiA* and K2stop and K88stop in *ptiM*. The result of these mutations was a shorter and unfinished protein product that did not affect the gene structure. The mutated ORFs were introduced into the plasmid pCN51 and transformed into RN450 strain. Figure 4.34 illustrates the difference in pigmentation of derived RN450 strains containing plasmids after $CdCl_2$ treatment. Both mutations in PtiM (ORF7) abolished STX production whereas mutations in PtiA (ORF6) showed little pigmentation similar to PtiM overexpression shown in Figure 4.29. These results were consistent with those obtained previously by PtiA, PtiM and PtiA-M overexpression. In this context, PtiA alone had no effect in pigmentation, while PtiM slightly increased it (Lorenz *et al*., 2011). This result indicated that SaPI proteins PtiA and PtiM are linked to form a complex, and they overexpress *crtOPQMN* with a positive effect in pigmentation.

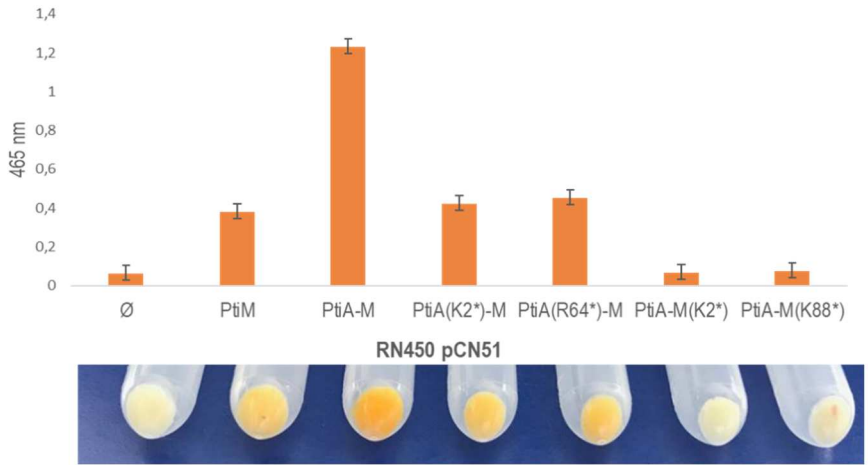

Figure 4.34 Pigment production of RN450 derivatives with shortened proteins PtiA and PtiM. The top graph shows the quantification of carotenoid pigment production. Values indicate the average of three independent experiments. The bottom panel shows the colour of the bacterial cell pellet from RN450 derivatives strains.

4.2.7 PtiA-M complex does not bind to the operon *crtOPQMN* promoter

The simplest hypothesis about STX regulation mechanism by SaPIs was that PtiA-M directly regulate the transcription of *crtOPQMN* by specifically binding to the respective promoter. To probe this idea, the promoter region upstream the operon was deleted in RN450 strain (Δ*crtO*), and pCN51ø / pCN51_*ptiA-M* were independently transduced and induced to quantify strain pigmentation. Figure 4.35

shows that deletion of the promoter appeared to have no negative effect on STX production when *ptiA-M* were overexpressed. These results suggested that PtiA-M did not act as transcriptional regulator by binding to *crtOPQMN* promoter and therefore, an alternative mechanism must exist.

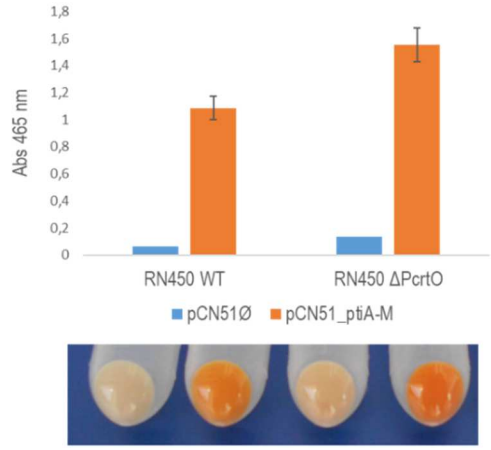

Figure 4.35 STX production by PtiA-M in P*crtO* mutant strains. Top graph: Quantification of STX production of RN450 ΔP*crtO* strain. Bottom panel: Colour of bacterial cell pellets.

4.2.8 Flanking genes to *crtOPQMN* are co-transcribed when PtiA-M are overexpressed

To elucidate the regulation mechanism of *crtOPQMN* expression by PtiA-M, the transcriptional profile in the *crtOPQMN* operon region of RN450 SaPIbov1Δ*stl* (4D) / RN450 +SaPIbov1 + SaPI1 (4S) was carefully examined (Figure 4.36). The first observation was that, in the RN450 SaPIbov1Δ*stl* (4D) strain (uncontrolled constant SaPI replication), high expression signals (red lines) were observed far upstream of the *crtO* promoter, specifically starting downstream of the *isaA* gene (SAOUHSC_02887), continuing until over GNAT (SAOUHSC_02886) and passing over the *crtOPQMN* operon. Interestingly, this high signal continued until reaching the *copA* gene, where signals decreased to basal levels. Note that *copA* gene is encoded at the opposite DNA strand. In contrast, the expression level profile of 4S strain (blue) in that genomic region was as expected, meaning that transcriptomic signals began and ended at the corresponding transcriptional start and stop sites for *gnat*, *oat*, *ssaA* and *crtOPQMN* genes, respectively. The transcriptional behaviour observed in the SaPIbov1Δ*stl* (4D) strain suggested that the genes are co-transcribed in a single transcript which could indicate the presence of a putative processive antitermination mechanism.

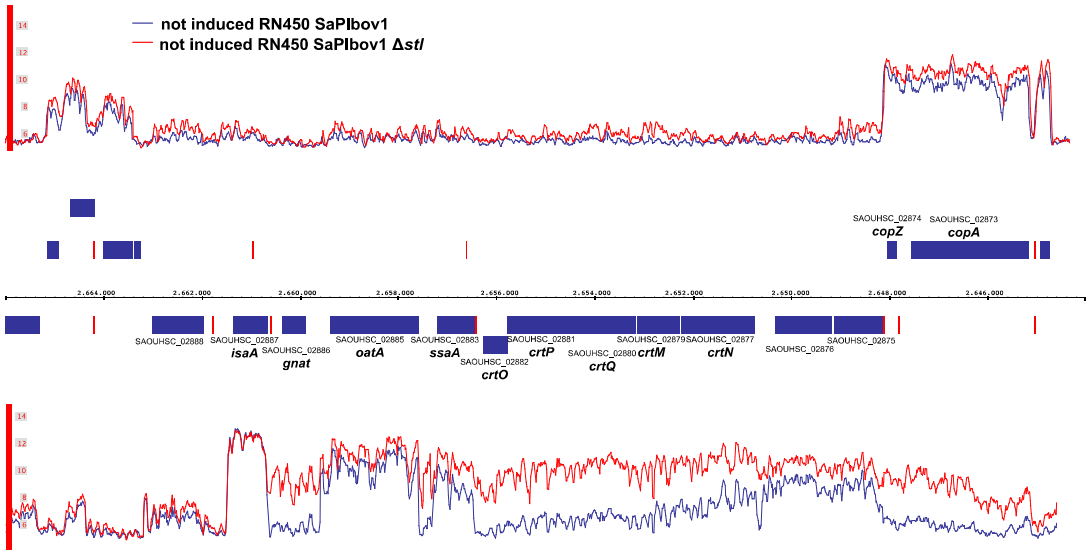

Figure 4.36 Transcriptional activity of the *crt*-operon region. The blue line represents transcription from RN450 SaPIbov1 while the red line represents transcription from RN450 SaPIbov1 Δ*stl*. The analysed region was between SAOUHSC_02888 and SAOUHSC_02873 (20 kb).

RT-PCR analysis was performed to validate the hypothesis regarding the existence of an antitermination mechanism affecting the *crtOPQMN* operon and the flanking zones. RT-PCR represents a valuable tool for identifying co-transcribed genes. Specifically, RT-PCR was used to assess the co-transcription of several gene sets localised between SAOUHSC_02887 (*isaA*) and *crtN* genes. To conduct the RT-PCR analysis, RNA was extracted from two derivative strains: SH1000 $\Delta\sigma^B$ complemented with pCN51_ø or pCN51_*ptiA-M* after 2h of $CdCl_2$ treatment. Whole-cell RNA was first reversely transcribed into a cDNA, and was then amplified via PCR. The primers used, hybridised with intergenic regions of the supposed transcript. Successful PCR amplification was taken as a signal that the genes were co-transcribed into a polycistronic mRNA, whereas non-amplification meant that genes were independently transcribed. Figure 4.37 shows PCR products with all the primers used from *isaA* to *crtN* with the cDNA from derivative strain SH1000 $\Delta\sigma^B$ pCN51_*ptiA-M*. Contrary, there were no PCR products with any primers when the cDNA from strain SH1000 $\Delta\sigma^B$ pCN51_ø was used as a template. The reverse transcription mixture lacking reverse transcriptase was used as a negative control and no PCR products were detected for all the primer pairs in similar reactions.

These results demonstrated that *isaA, gnat, oat, ssaA* and the *crt*-operon were co-transcribed as a single operon when PtiA-M were overexpressed, whereas they were transcribed individually in standard conditions.

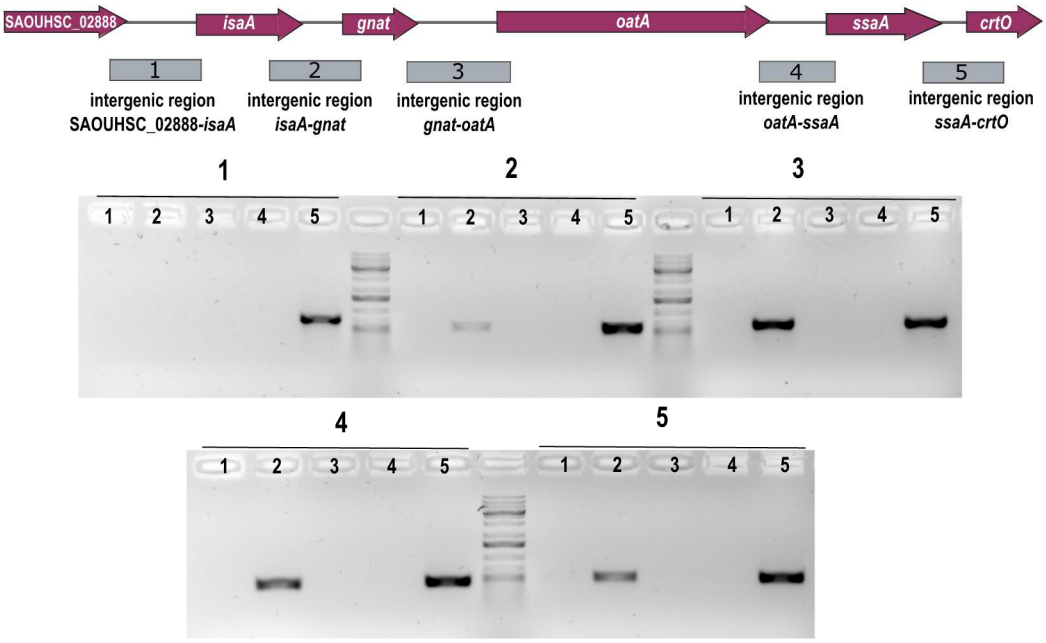

Figure 4.37 RT-PCR analysis of the intergenic regions of the cluster from *isaA* to *crtO* genes. Top: Location of the amplicons were indicated by the numbered boxes. Bottom: Gel electrophoresis of amplicons from RT-PCR experiments: 1=cDNA SH1000 $\Delta\sigma^B$ + pCN51ø, 2=cDNA SH1000 $\Delta\sigma^B$ + pCN51*ptiA-M,* 3 and 4 negatives controls in absence of reverse transcriptase in the reaction, 5=positive control of genomic DNA.

4.2.9 PtiA-M regulates STX production through an antitermination mechanism that begins at *isaA*

In the simplest instance of a bacterial operon, genes are transcribed in a unique polycistronic mRNA, which is initiated at a promoter located upstream of the first gene. However, in many cases, regulation of the operon is more complex and involves internal promoters. To analyse the possible responsibility of the internal promoters included in the gene cluster in relation to STX expression, deletion of some promoters from the hypothetical operon was performed. Promoters of *isaA, gnat, oatA* and *ssaA* genes were chosen since they are upstream of the *crtOPQMN* operon and they could influence its expression. Four independent derivative strains were generated (Δ*PisaA,* Δ*Pgnat*, Δ*PoatA and* Δ*PssaA* in RN450 strains) and complemented with plasmids expressing the PtiA-M complex. The carotenoid production was measured after $CdCl_2$ treatment.

Promoter deletions of P*gnat*, P*oatA* and P*ssaA* did not cause phenotypic changes. However, P*isaA* deletion dramatically decreased pigmentation (Figure 4.38). These findings suggest that PtiA-M complex regulates STX overexpression through an indirect mechanism involving *isaA*.

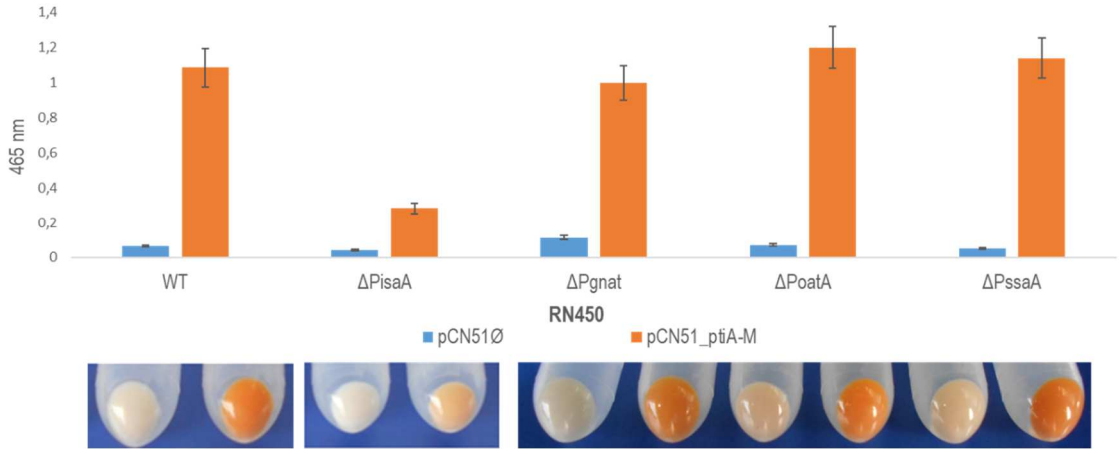

Figure 4.38 Pigment production of RN450 derivatives strains with different promoters deletion. Top graph: STX quantification by methanol extraction of ΔP*isaA,* ΔP*gnat*, ΔP*oatA and* ΔP*ssaA* derivatives strains. Bottom panel: Colour of bacterial pellet. Values indicate the average of three independent experiments.

To understand the role of *isaA* gene in *crtOPQMN* expression, a similar strategy was used. Deletion of *isaA* gene (promoter + gene + terminator) was generated in the derivative RN450 pCN51*_ptiA-M* strain and carotenoid production was measured following $CdCl_2$ treatment. Complete deletion of *isaA* was correlated with the total absence of pigment (Figure 4.39).

The *isaA* gene codifies for a surface protein known as the immunodominant staphylococcal antigen A, which is a soluble lytic transglycosylase. To determine whether *isaA* protein was responsible for this regulatory mechanism, a premature stop codon was introduced into the *isaA* coding sequence (L10stop) causing a shorten protein. Following $CdCl_2$ treatment, STX production was measured in the derivative RN450 *isaA* L10* + pCN51*_ptiA-M* strain. It was observed that this mutation did not cause any effect on STX production and the phenotype was the same as the WT. This experiment excluded the involvement of IsaA protein in STX production. Therefore, PtiA-M seems to exert the regulation by an antitermination mechanism that begins at *isaA* gene.

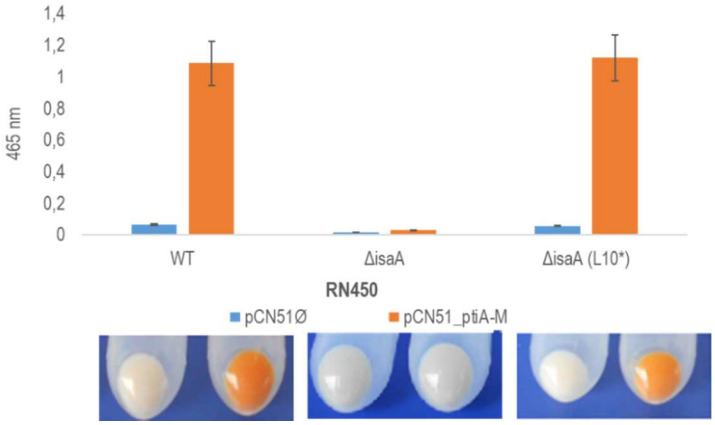

Figure 4.39 STX production of *isaA* mutants in RN450 derivative strains. Top graph: Measurement of STX extracted in methanol of RN450 and its derivatives. Bottom: color of analysed bacterial cells. The data represent the average of three independent experiments.

4.2.10 Reverse orientation of *oatA* gene deletes STX production when PtiA-M are overexpressed

An important feature of operons is that genes are transcribed in the same orientation. In fact, the transcriptional profile in the *crtOPQMN* operon region of RN450 SaPIbov1Δ*stl* (4D) strain showed high expression levels from the *isaA* to the *copA* gene, which is transcribed in the opposite orientation. One method to confirm the co-transcription of genes within an operon is by reversing the orientation of one gene and study how STX production is affected. The gene chosen was *oatA* which is localised upstream of *crtOPQMN* and downstream of the *isaA* gene.

First, a derivative strain with in-frame deletion for the *oatA* gene was created (*ΔoatA*) ensuring that the reading frame remained intact. The gene deletion introduced did not seem to adversely affect the STX operon transcription because there was an increase in pigmentation comparable to the WT when *ptiA-M* was overexpressed. Finally, the derivative Δ*oatA* strain was complemented with the *oatA* gene inserted in the opposite "reverse" orientation. The gene replacement caused operon disruption and the golden STX pigment was absent despite PtiA-M overexpression (Figure 4.40).

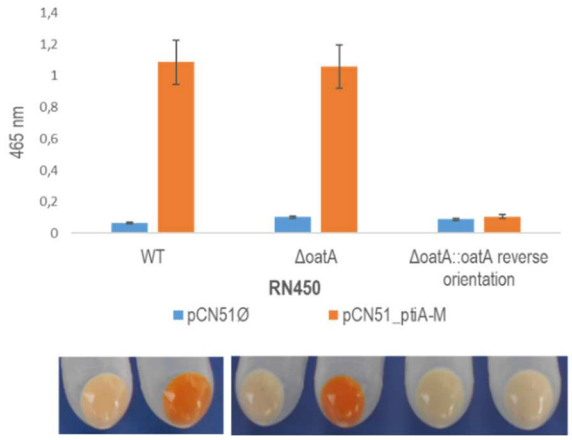

Figure 4.40 STX production of *oatA* gene reverse orientated in RN450 derivative strain. Top part: Quantification of pigment production in RN450 derivative strain (Δ*oatA* and *oatA* reverse). The bottom panel shows the colour of the bacterial cell pellet.

4.2.11 Role of PtiA and PtiM as antitermination proteins

To provide experimental support to the antitermination regulatory mechanism, two plasmid constructions were generated with the *blaZ* reporter gene fused to the *isaA* gene (pCN41_*isaA*) and the region between *isaA* gene and *ssaA* (pCN41_*isaA*-*ssaA*). Both regions contained the intergenic regions (promoter and terminator) as well as the coding genes (Figure 4.41). The fusions plasmids were transduced into derivative SH1000 $\Delta\sigma^B$ and complemented with a second plasmid. This second plasmid encoded for a cadmium-inducible promoter (pCU1P_{cad}) and was used as a control (pCU1P_{cad}ø) or carry *ptiA-M* (pCU1P_{cad} *ptiA-M*). The initial idea was that after $CdCl_2$ treatment, strains containing the pCU1P_{ptiA-M} plasmid, should show an increase in β-lactamase activity. This could be explained as the PtiA-M proteins binds to the transcript preventing the end of transcription. On the other hand, the strains with pCU1P_{cad}ø should remain unchanged. In case there was a terminator, the transcription should stop and no β-galactosidase activity should be observed. As it was expected, both strains containing pCU1P_{ptiA-M} presented β-lactamase activity, which indicated that ptiA-M avoided transcription termination.

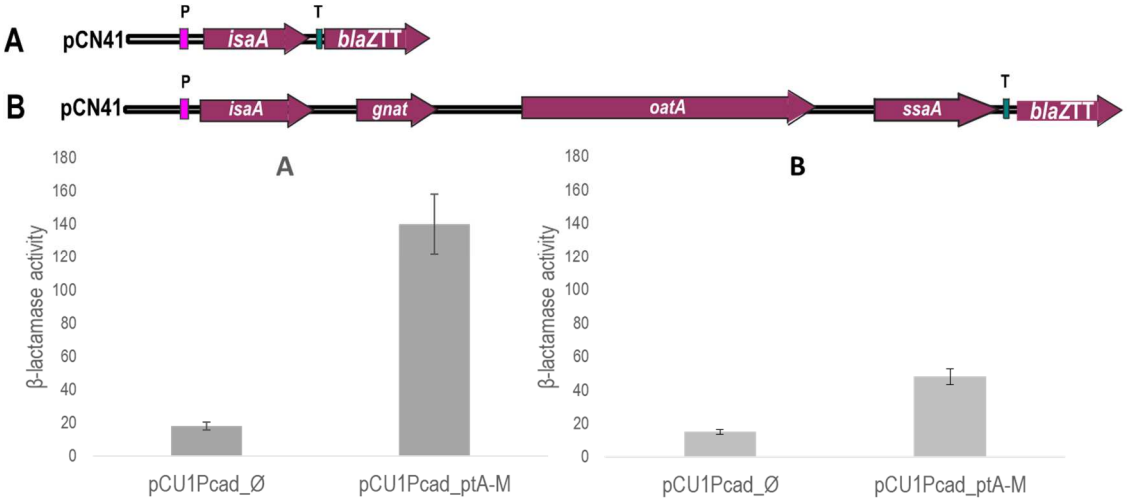

Figure 4.41 PtiA and PtiM avoid *isaA* terminator. Top part: Diagram of derivative vector pCN41: pCN41_*isaA* (A) and pCN41_*isaA-ssaA* fused to *blaZ* reporter gene (B). β-lactamase activity after 2h of $CdCl_2$ treatment of SH1000 ΔσB derivative strain

5. Discussion

5.1 Mobilisation mechanism of pathogenicity islands by endogenous phages in *S. aureus* clinical strains

PICIs are a key factor in pathogenesis of *Staphylococcus* spp. Previous studies have determined and classified MGEs, such as bacteriophages, in clinical strains. Despite their relevance, there are no studies that determine the presence of PICIs in *Staphylococcus* spp. clinical strains (Goerke *et al*., 2009). In this thesis we demonstrated that SaPIs are widely distributed. They were detected in 50% of *S. aureus* strains isolated from patients with CF. Interestingly, the presence of SePIs was higher than SaPIs. SePIs were identified in 83.3% of *S. epidermidis* strains from catheter-associated infections. Since there are no studies to compare with, we do not know whether this higher percentage is characteristic of *S. epidermidis* or is due to the niche where each type of infection is developed. One possible explanation could be that in chronic infections, such as those caused by *S. aureus* (observed in the lungs of CF patients) require less presence of PICIs that in an acute infection, such as catheter-associated infections. Thus, more studies are needed to determine whether the variability found regarding the presence of PICIS is due to the species differences or due to the niche where the infection is developed.

On the other hand, if we analyze the distribution of SaPIs in specific CC, we observed that the distribution of these MGEs varied remarkably between lineages and we demonstrated that there is a highly association of determined MGE with different CC indicating that the spread of SaPIs in the bacterial population is at least partially restricted. In certain CC some SaPIs groups were completely absent and others were significantly less or, on the other hand, significantly more frequent. There are several studies that characterise diverse CCs and determine the presence of SaPIs in these CCs, but there are no studies focused on the association between the different SaPIs and CCs (Kurt et al., 2013). However, Goerke and colleges showed that the distribution of bacteriophages varied remarkably between lineages, and they suggest that a possible explanation the existence of restriction-based barriers (Goerke et al., 2009). Other authors speculate that MGE promote the spreading of determined bacterial clones. This hypothesis was emphasized by the emergence of

highly virulent MRSA strains carrying a phage harboring a new adherence factor (Schulte, Bierbaum, Pohl, Goerke, & Wolz, 2013).

PICIs mobilisation could promote the acquisition of virulence determinants by *S. aureus* strains and transform a benign bacterium into a virulent pathogen. For this reason, SaPI induction and transference has been widely studied. However, the problem is that studies have always been limited to laboratory strains. In this thesis, for the first time, SaPI mobilisation was investigated in clinical strains.

SaPI mobilisation was studied in a collection of *S. aureus* clinical strains from CF patients and the study was then extended to *S. aureus* strains from different clinical origins. SePI mobilisation was also investigated in *S. epidermidis* clinical strains from infected catheters. The obtained results confirmed the existence of SaPI and SePI mobilisation by endogenous phages in "real world" strains. We consider this finding highly important since the antibiotics used in clinical practice, such as β-lactams or fluoroquinolones, can induce the SOS response. SOS activation by antibiotics induce the prophage and promote the spread of SaPIs and staphylococcal virulence genes as an unwanted consequence.

In addition, based on the fact that SaPI interference mechanisms are counterselective for the phage, endogenous phages able to induce coresident SaPIs are considered unlikely to be found (Frígols *et al.*, 2015). However, as already mentioned, this study showed that SaPI induction by endogenous phages naturally occurs in human clinical isolates. One possible reason is the different interference mechanism between *pac* and *cos* types. *pac* SaPIs encode for genes that are responsible for interference with *pac* phage packaging. On the other hand, *cos* SaPIs encode for the *ccm* gene which causes severe interference in *cos* phage packaging. However, there are no described interference mechanisms between *pac* SaPIs and *cos* phages. Strains such as COL, N315, A or E, in which the SaPIs are *pac* and their coresident phages are *cos*, bacteriophages may have been acquired by horizontal transfer regardless of the SaPI presence, since they are not affected by interference mechanisms.

Furthermore, SaPImw2 mobilisation was characterised in depth. As described, MW2 contains the *cos* SaPImw2 which is induced and packaged by the *cos* phage ΦSa2mw. DUF3113 is a small protein and is the responsible for SaPImw2 derepression. All derepressor phage proteins have a dual role: a primary function related to the phage biology and a secondary function related to SaPI mobilisation. The best studied example is the dUTPase protein, which prevents the incorporation of uracil into DNA and is also able to mobilize SaPIbov1 (M. Á Tormo-Más *et al.,* 2013). DUF3113 is a protein with an unknown function for the phage. Deletion of the MW1424 gene was generated but did not change either phage replication or phage titer. So, under laboratory conditions, DUF3113 is not an essential protein. In this study, a new phage derepressor protein was identified. DUF3113 is added to the four described derepressors but there are still many derepressors to identify.

cos SaPIc from clinical strain C, which encodes for the same *stl* gene as SaPImw2, is induced by endogenous *cos* ϕ6c. However, this phage is probably not responsible for SaPIc packaging since ϕ6c *cos* site and its flanking region, which are important to SaPI packaging, are completely different to *cos* site and flanking region of SaPIc. Strain C presents another prophage, called ϕ2c. In this case, ϕ2c *cos* site and its flanking region presents 89% identity with the corresponding sequence in SaPIc. Ccm in *cos* SaPIs drives phage interference by the production of small capsids and the phage major capsid protein (CP) is its target (Carpena et al., 2016). In addition, Ccm and CP present high structural similarity in the C-terminal domain. In this case, Ccm SaPIc presents higher structural similarity with $CP_{\Phi 2c}$ than with $CP_{\Phi 6c}$ (Figure 5.1). So, it is possible that SaPIc only interferes with ϕ2c. Therefore, it could be supposed that SaPIc seems to be able to hijack both prophages to carry out its ERP cycle; ϕ6c is involved in SaPIc induction and ϕ2c in packaging.

This type of mechanisms in which SaPI utilise more than one phage for its induction is identified for the first time in this study. Until now, only one phage had been hijacked by the SaPIs. However, phage cooperation can also be suspected in MW2 strain. SaPImw2 mobilisation from MW2 (ϕSa2mw and ϕSa3mw) was three orders of magnitude higher than that of ϕSa2mw lysogen, which indicates that, though ϕSa2mw is responsible for SaPImw2 induction,

ϕSa3mw increases SaPImw2 mobilisation. In this case, both phages are *cos* type and the *cos* site of ϕSa2mw is identical to that of SaPImw2. Therefore, ϕSa2mw phage could both induce and encapsidate SaPImw2. However, lysis of the MW2 culture after MC treatment was higher than in the ϕSa2mw lysogen. This suggests that ϕSa3mw is involved in lysis and could improve the mobilisation rate.

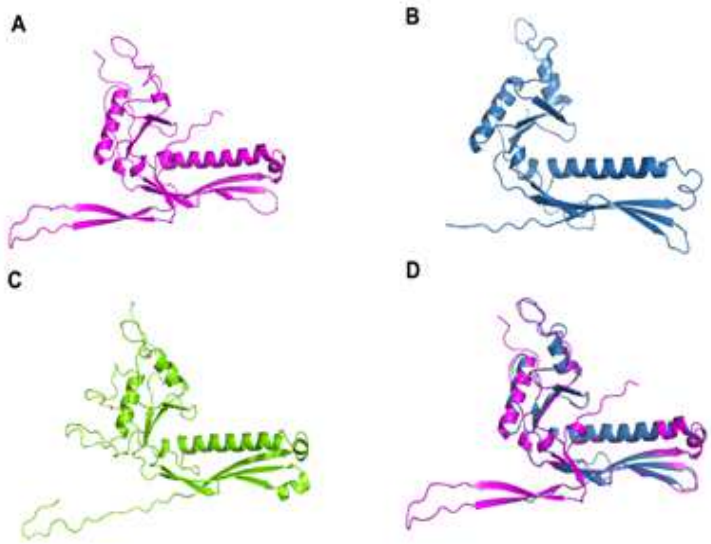

Figure 5.1 C-terminal portion of Ccm is structurally more similar to CPϕ2c than CPϕ6c. Cartoon representation of (A) Ccm SaPIc (residues 156-351), (B) CP ϕ2c (residues 156-401) and (C) CP ϕ6c (residues 156-401). Structural models for these proteins were generated using RaptorX. Superposition of Ccm SaPIc and CP ϕ2c (D), using PyMol, exhibited a root-mean-square-deviation (RMSD) of 0.511. The same comparison involving Ccm SaPIc and CP ϕ6c revealed a significant higher value (14.622).

To sum up, this is the first study where SaPI mobilisation was demonstrated in a collection of *S. aureus* clinical strains. SaPImw2 mobilisation was studied in depth and the gene responsible for its derepression has been described. Additionally, the study shows signs of a new SaPI strategy hijacking the machinery of more than one phage to increase SaPI mobilisation.

Instead, there are still many uncertainties about SaPI induction and further studies are necessary to identify new phage derepressors and to better understand the mechanism of SaPI mobilisation and the consequent transfer of virulence factors.

As future work we propose different research lines:

Determine the phage inductors responsible for the induction of both SaPIs and SePIs described in this thesis, whose repressors have not been previously studied.

All the SaPI inductors described until now are moonlighting proteins, as in addition to their activity as inducers of specific SaPIs, develop a function for the the phage itself, so we consider that it is of interest to define which function these proteins perform in the phage biology.

Investigate in more detail the cooperation among phages in order to mobilize other MGEs such as the SaPIs.

5.2 SaPIs control of chromosomally-encoded virulence gene expression

The importance of SaPIs to the physiology, adaptation and evolution of the microorganism is unquestionable. Embedded in bacterial genomes, these entities can prevent phage infection and interfere with phage reproduction (M. Á. Tormo-Más *et al*., 2010) (Frígols *et al*., 2015). Once SaPIs are induced, they can contribute to the enhancement of their hosts' pathogenicity by transferring their virulence genes and unlinked virulence genes scattered throughout the chromosome (J. Chen et al., 2015). But would it be possible for SaPIs to develop more functions than those described so far?

Recent studies showed the regulation of chromosomal genes by MGEs such as bacteriophages, but what about SaPIs? Are these MGEs able to regulate also chromosomal genes? In this study we resolve that question and demonstrate that SaPIs develop a new function associated with the regulation of chromosomal gene expression.

The tiling array of *S. aureus* NCTC8325 genome was used to assess RN450 WT, lysogenised RN450 and SaPI-positive RN450 strains for differential expression. The data obtained from this analysis revealed significant changes in the bacterial transcriptome resulting from SaPI induction. In all different conditions 33 genes were found differentially expressed after SaPI induction. A cluster of five genes encoded hypothetical proteins were reported. When SaPIbov1 was in constant replication (Δ*stl*), 24 genes were differentially expressed compared to the WT (see Appendix 4). Some of these genes are related to virulence such as *coa*, *cidA* or *trfB* genes while others are involved in nitrate and nitrite metabolism as well as transport. The *coa* gene (SAUHSC_00192) codifies for an extracellular virulence factor that causes the coagulation of plasma (Hemker, Bas, & Muller, 1975). The *cidA* gene (SAOUHSC_02851) encodes for a murein hydrolase and also contributes to DNA release and biofilm development. On the other hand, *trfB* (SAOUHSC_00936) is involved in resistance to two classes of cell-wall-active antibiotics: teicoplanin and vancomycin (Bayles *et al*., 2007) (Renzoni *et al*., 2009).

The most interesting finding of the tiling array analysis concerned the upregulation of two loci in the four considered comparisons. The *crtOPQMN* operon and *gnat* gene were upregulated and also involved in virulence. Thus, these results suggest that SaPIs regulate chromosomal genes, mainly, genes involved in virulence and pathogenicity.

The upregulation of staphyloxanthin gene cluster (*crtOPQMN*) by SaPI induction was studied in depth. The yellowish-orange (golden) pigment produced by *S. aureus* has been linked to virulence owing to its antioxidant property. Previous studies have reported that the pigmentation is controlled by *rsbUVW-sigB* system, but is also affected by purine biosynthesis, the tricarboxylic acid cycle or the oxidative phosphorylation (Kullik et al., 1998) (Lan et al., 2010). In addition, the cold shock protein A (CspA) is a positive regulator in a SigB-dependent manner (Katzif, Lee, Law, Tzeng, & Shafer, 2005) (Caballero *et al*., 2018). This study showed that STX production is also regulated by SaPI induction. In addition, the data further indicates that the SaPI regulatory system is independent of SigB, as the SigB mutant (white phenotype) showed pigmented colonies after SB1 (*Δstl*) transduction.

Specifically, PtiA and PtiM are the two SaPI proteins responsible for pigment production. The deletion of *ptiA-M* genes affected STX production and the colony was not pigmented. However, overexpression of these genes increased colony pigmentation. PtiA and PtiM form a complex and their function has been described in previous researches (Geeta Ram et al., 2014) (Geeta Ram et al., 2015). As it was reported, they interfere with phages by modulating the late-module transcription, which is responsible for the production of virion and lysis proteins. This study uncovered a new function for this protein complex related to STX production. Although plenty of multifunctional proteins have been described in bacteriophages, this represents the first time that multifunctional proteins are identified in SaPIs (M. Á. Tormo-Más *et al*., 2010) (Boyd, 2012) (J. Z. Bowring, Marina, Penadés, & Quiles-Puchalt, 2016).

The regulatory mechanism of the *crtOPQMN* operon by PtiA-M was unexpected and complex. Previous studies have revealed that regulation was

direct and that the SigB or AirR specifically binds to the *crt* promoter region (Pelz *et al*., 2005) (Hall, Yang, Guo, & Ji, 2017). In this study, the deletion of the *crt* promoter resulted in no loss of pigmentation. This demonstrated that PtiA-M neither acted as transcriptional factors nor binded to the operon promoter. This study proposed that PtiA-M are proteins that regulate *crtOPQMN* and *gnat* genes through processive antitermination (PA) mechanism.

In PA mechanisms, antitermination factors are associated with a bacterial RNA polymerase (RNAP) elongation complex that leads to the bypass of terminator sites even over longer genomic distances (Weisberg & Gottesman, 1999) (Santangelo & Artsimovitch, 2011). Although proteins that control transcription initiation have been identified, those affecting transcription elongation have not broadly been investigated yet. The best characterised PA mechanisms are mediated by the phage λ proteins N and Q (Yarnell & Roberts, 1992) (Patterson *et al*., 1994). New studies are emerging that uncover more examples of PA mechanism, however, more information needs to be acquired (A. V. Yakhnin, Yakhnin, & Babitzke, 2008)(Goodson et al., 2017).

In our study, intergenic RT-PCR revealed that the *crtOPQMN* operon and flanking regions were cotranscribed as a single transcript only when PtiA-M were overexpressed. These results confirmed the high gene expression level observed in the transcriptional profile of RN450 SB1 (*Δstl*).

The deletion of the *isaA* promoter resulted in a decrease of pigmentation, and the deletion of the *isaA* gene caused total pigment loss when PtiA-M were overexpressed. IsaA is a lytic transglycosilase protein that was previously reported to be implicated in cleaving peptidoglycan (Stapleton *et al*., 2007). A stop mutation was introduced in the *isaA* gene though it did not negatively affect STX production. This result suggested that IsaA protein was not responsible for STX production. *isaA* gene is around 4400 bp upstream of *crtOPQMN* operon and seemed to be located where transcription begins. When PtiA and PtiM are overexpressed, the RNAP binds to the *isaA* promoter and continues, bypassing all terminator sequences until the *copA* gene, which is transcribed in the opposite direction.

In this study, the *oat* gene, localised within the hypothetical operon upstream of the *crt*-operon, was replaced with the *oat* gene in the opposite direction and the pigment production was abolished. On the other hand, using a *blaZ*-based reporter assay, it was demonstrated that the overexpression of PtiA-M lead to the RNAP to read throughout the *isaA* and *ssaA* terminator sites.

However, the complete basis by which PtiA-M regulates *crtOPQMN* could not be discovered. Further analyses are required in order to understand the mechanism of action of PtiA and PtiM. In addition, transcriptional profiles of different genomic regions are now under study in order to determine the existence of additional PA mechanisms controlled by SaPIs. Until now, five regions were chosen according to the known information about different genetic expression patterns following SaPI induction. Preliminary observations of transcriptional profiles suggest that PA mechanisms are present in other parts of the chromosome, which highlights the impact of SaPIs on the regulation of chromosomal genes (Figure 5.2).

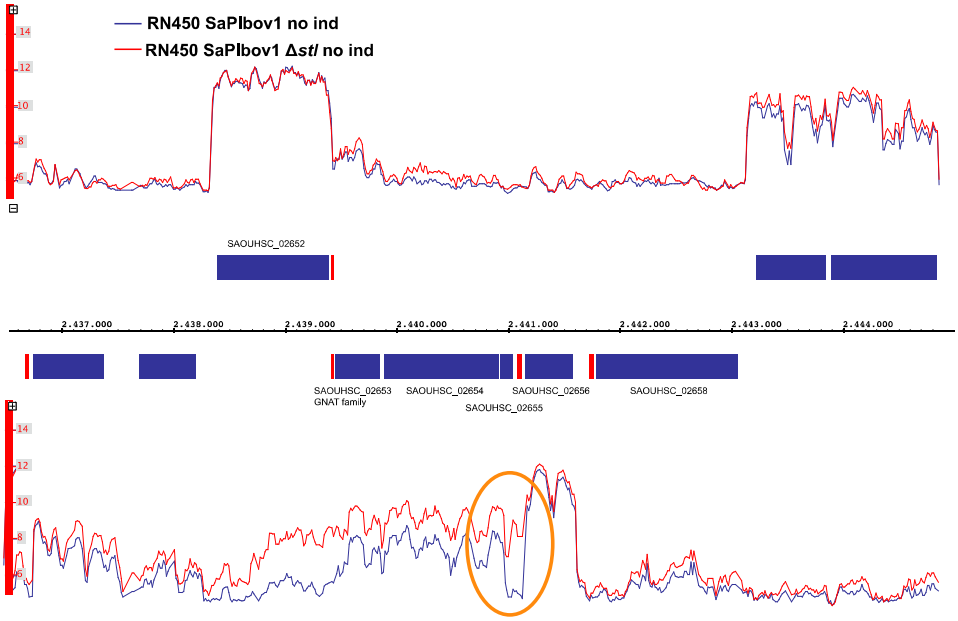

Figure 5.2. Example of transcriptional profile of SAOUHSC_02665 region. The blue line represents transcription from RN450 SaPIbov1 while the red line represents transcription from RN450 SaPIbov1 Δ*stl*. The analysed region was between SAOUHSC_02650 and SAOUHSC_02660. Orange circle marks the region where an antitermination mechanism seems to modulate transcription termination of SAOUHSC_02656.

Future studies should aim to examine in detail the binding region of *isaA* gene by PtiA-M proteins. To do that, different plasmid constructions (pCN41) have been

generated with the *blaZ* reporter gene fused to portions of *isaA* gene (promoter and terminator are always included). The fusions plasmids have been transduced into derivative SH1000 $\Delta\sigma^B$ and complemented with pCU1P_{cad}ø (used as a control) or pCU1P_{cad} *ptiA-M*. The experiments are currently being performed and there are still no conclusive results.

In addition, Western blot needs to be performed in order to study the expression of PtiM in presence and absence of PtiA. It should be noted that PtiA by itself did not increase STX expression, PtiM slightly increased STX expression and PtiA-M coexpression significantly increased pigment production. Moreover, it has been demonstrated that PtiM and PtiA interacts to modulate their activity. Two constructions in pCN51 vector have been obtained to study differences in PtiM expression. *ptiA-M* tagged with 3xflag and *ptiA*(K2*)-*M* tagged with 3xflag have been independently cloned in pCN51 and transformed into RN4220. Preliminary results have showed an increased level of PtiM when is coexpressed with WT *ptiA* compared to *ptiA*(K2*) mutated. To explain this results could be that the PtiA protein stabilises or prevents PtiM degradation, and higher levels of PtiM show a greater increase in STX expression. Nevertheless, more research is required to validate this finding.

Finally, genes which are homologous to *ptiA* and *ptiM* were identified in *S. epidermidis* and *S. chromogenes*. These two species are non-pigmented and lack the *crt*-operon. These genes have also been cloned into pCN51 vector and transformed into *S. aureus*. While *ptiA-M* from *S. epidermidis* has only showed a slight increment in pigmentation, *ptiA-M* from *S. chromogenes* has showed intense pigmentation. Future work will focus on analysing whether PtiA and PtiM from *S. epidermidis* and *S. chromogenes* also act as antitermination factors.

6. Conclusions

In this doctoral thesis, SaPIs mobilisation mechanism and their role in the regulation of chromosomal gene expression have been presented, analysed and studied in detail. The main conclusions can be summarised as follows:

1. PICIs are present in 50% of the *S. aureus* strains collection isolated from patients with CF and in 83.3% of *S. epidermidis* isolates collected from catheter-associated infections.
2. SaPIs could be mobilised by endogenous bacteriophages in *S. aureus* clinical strains. In particular, 18.75% of the analysed clinical isolates from CF patients as well as 50% of clinical isolates of diverse origins showed SaPI induction by endogenous prophages.
3. Two SePIs from one *S. epidermidis* strain (SCN45) isolated from an infected catheter were induced by its endogenous phages.
4. The MW1424 gene from ΦSa2mw codifies for a hypothetical protein with DUF3113 domain. This protein is responsible for SaPImw2 induction and mobilisation in the MW2 strain through direct protein interaction with the SaPI repressor Stl.
5. Allelic variants of the MW1424 gene code for proteins with different ability to induce SaPImw2.
6. SaPIs are able to hijack more than one bacteriophage to conduct their mobilisation. SaPIc from strain C seems to be induced by the endogenous Φ6c and encapsidated by Φ2c. SaPImw2 is present in MW2 and is induced and mobilised by ΦSa2mw but the presence of ΦSa3mw increases transfer frequency.
7. SaPIs are able to regulate the expression of chromosomal genes. Specifically, they are responsible for the overexpression of the *crtOPQMN* operon and the SAOUHSC_02886 genes.
8. PtiM from SaPIbov1 is the main protein responsible for the overexpression of the *crtOPQMN* operon and, consequently, for STX production. PtiA interacts with PtiM contributing to this effect. Homologous proteins to PtiA and PtiM identified in other SaPIs are also able to positively regulate STX production.

9. PtiA and PtiM are not transcriptional factors that bind to the *crt*-operon promoter to regulate the expression. PtiA-M complex regulates the *crt*-operon expression by a PA mechanism regulating the gene cluster from the *isaA* gene to the *copA* gene

7. Bibliography

Aanensen, D. M., & Spratt, B. G. (2005). The multilocus sequence typing network: mlst.net. *Nucleic Acids Research*, *33*(SUPPL. 2). https://doi.org/10.1093/nar/gki415

Ackermann, H. W. (1975). Classification of the bacteriophages of Gram-positive cocci: *Micrococcus, Staphylococcus,* and *Streptococcus* La classification des bacteriophages des cocci Gram-positifs: *Micrococcus*, *Staphylococcus* et *Streptococcus. Pathol.Biol.Paris*, *23*(3), 247–253.

Alibayov, B., Baba-Moussa, L., Sina, H., Zdeňková, K., & Demnerová, K. (2014). *Staphylococcus aureus* mobile genetic elements. *Molecular Biology Reports*, *41*(8), 5005–5018. https://doi.org/10.1007/s11033-014-3367-3

Argemi, X., Nanoukon, C., Affolabi, D., Keller, D., Hansmann, Y., Riegel, P., … Prévost, G. (2018). Comparative genomics and identification of an enterotoxin-bearing pathogenicity island, SEPI-1/SECI-1, in staphylococcus epidermidis pathogenic strains. *Toxins*, *10*(3). https://doi.org/10.3390/toxins10030093

Arnaud, M., Chastanet, A., & Débarbouillé, M. (2004). New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Applied and Environmental Microbiology*, *70*(11), 6887–6891. https://doi.org/10.1128/AEM.70.11.6887-6891.2004

Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K. I., Oguchi, A., … Hiramatsu, K. (2002). Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet*, *359*(9320), 1819–1827. https://doi.org/10.1016/S0140-6736(02)08713-5

Bæk, K. T., Frees, D., Renzoni, A., Barras, C., Rodriguez, N., Manzano, C., & Kelley, W. L. (2013). Genetic Variation in the *Staphylococcus aureus* 8325 Strain Lineage Revealed by Whole-Genome Sequencing. *PLoS ONE*, *8*(9). https://doi.org/10.1371/journal.pone.0077122

Baharoglu, Z., & Mazel, D. (2014). SOS, the formidable strategy of bacteria against aggressions. *FEMS Microbiology Reviews*. https://doi.org/10.1111/1574-6976.12077

Bayles, K. W., Weiss, E. C., Rice, K. C., Cassat, J. E., Mann, E. E., Smeltzer, M. S., & Endres, J. L. (2007). The *cidA* murein hydrolase regulator contributes to DNA release and biofilm development in *Staphylococcus aureus*. *Proceedings of the National Academy of Sciences*, *104*(19), 8113–8118. https://doi.org/10.1073/pnas.0610226104

Benson, N., & Youderian, P. (1989). Phage cro protein and *cI* repressor use two different patterns of specific protein-DNA interactions to achieve sequence specificity in vivo. *Genetics*, *121*(1), 5–12.

Bento, J. C., Lane, K. D., Read, E. K., Cerca, N., & Christie, G. E. (2014). Sequence determinants for DNA packaging specificity in the S. aureus pathogenicity island SaPI1. *Plasmid*, *71*(1), 8–15. https://doi.org/10.1016/j.plasmid.2013.12.001

Bisognano, C., Kelley, W. L., Estoppey, T., Francois, P., Schrenzel, J., Li, D., … Vaudaux, P. (2004). A RecA-LexA-dependent Pathway Mediates Ciprofloxacin-induced Fibronectin Binding in *Staphylococcus aureus*. *Journal of Biological Chemistry*, *279*(10), 9064–9071. https://doi.org/10.1074/jbc.M309836200

Bowring, J., Neamah, M. M., Donderis, J., Mir-Sanchis, I., Alite, C., Ciges-Tomas, J. R., … Penades, J. R. (2017). Pirating conserved phage mechanisms promotes promiscuous staphylococcal pathogenicity island transfer. *eLife*, *6*, e26487. https://doi.org/10.7554/eLife.26487

Bowring, J. Z., Marina, A., Penadés, J. R., & Quiles-Puchalt, N. (2016). Bacteriophage Moonlighting Proteins in the Control of Bacterial Pathogenicity. In *Moonlighting Proteins: Novel Virulence Factors in Bacterial Infections* (pp. 387–412). https://doi.org/10.1002/9781118951149.ch22

Boyd, E. F. (2012). Bacteriophage-Encoded Bacterial Virulence Factors and Phage-Pathogenicity Island Interactions. *Advances in Virus Research*, *82*, 91–118. https://doi.org/10.1016/B978-0-12-394621-8.00014-5

Brammar, W. J. (2013). Specialized Transduction. In *Brenner's Encyclopedia of Genetics: Second Edition* (pp. 501–502). https://doi.org/10.1016/B978-0-12-374984-0.01450-9

Brussow, H., Canchaya, C., & Hardt, W.-D. (2004). Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. *Microbiology and Molecular Biology Reviews*, *68*(3), 560–602. https://doi.org/10.1128/MMBR.68.3.560-602.2004

Brüssow, H., & Hendrix, R. W. (2002). Phage Genomics: Small is beautiful. *Cell*. https://doi.org/10.1016/S0092-8674(01)00637-7

Butala, M., Žgur-Bertok, D., & Busby, S. J. W. (2009). The bacterial LexA transcriptional repressor. *Cellular and Molecular Life Sciences*. https://doi.org/10.1007/s00018-008-8378-6

Caballero, C. J., Menendez-Gil, P., Catalan-Moreno, A., Vergara-Irigaray, M., García, B., Segura, V., … Toledo-Arana, A. (2018). The regulon of the RNA chaperone CspA and its auto-regulation in *Staphylococcus aureus*. *Nucleic Acids Research*, *46*(3), 1345–1361. https://doi.org/10.1093/nar/gkx1284

Canchaya, C., Fournous, G., & Brüssow, H. (2004). The impact of prophages on bacterial chromosomes. *Molecular Microbiology*. https://doi.org/10.1111/j.1365-2958.2004.04113.x

Carpena, N., Manning, K. A., Dokland, T., Marina, A., & Penade, R. (2016). Convergent evolution of pathogenicity islands in helper *cos* phage interference. https://doi.org/10.1098/rstb.2015.0505

Casjens, S. R., & Hendrix, R. W. (2015). Bacteriophage lambda: Early pioneer and still relevant. *Virology*. https://doi.org/10.1016/j.virol.2015.02.010

Cervera-Alamar, M., Guzmán-Markevitch, K., Žiemytė, M., Ortí, L., Bernabé-Quispe, P., Pineda-Lucena, A., … Tormo-Mas, M. Á. (2018). Mobilisation Mechanism of Pathogenicity Islands by Endogenous Phages in *Staphylococcus aureus* clinical strains. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-34918-2

Chabelskaya, S., Gaillot, O., & Felden, B. (2010). A *Staphylococcus aureus* small RNA is required for bacterial virulence and regulates the expression of an immune-evasion molecule. *PLoS Pathogens*, *6*(6). https://doi.org/10.1371/journal.ppat.1000927

Chambers, H. F., & DeLeo, F. R. (2009). Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nature Reviews Microbiology*. https://doi.org/10.1038/nrmicro2200

Charpentier, E., Anton, A. I., Barry, P., Alfonso, B., Fang, Y., & Novick, R. P. (2004). Novel cassette-based shuttle vector system for gram-positive bacteria. *Applied and Environmental Microbiology*, *70*(10), 6076–6085. https://doi.org/10.1128/AEM.70.10.6076-6085.2004

Chen, F., Di, H., Wang, Y., Cao, Q., Xu, B., Zhang, X., … Lan, L. (2016). Small-molecule targeting of a diapophytoene desaturase inhibits S. aureus virulence. *Nature Chemical Biology*,

12(3), 174–179. https://doi.org/10.1038/nchembio.2003

Chen, H. J., Tsai, J. C., Hung, W. C., Tseng, S. P., Hsueh, P. R., & Teng, L. J. (2011). Identification of *fusB*-mediated fusidic acid resistance islands in Staphylococcus epidermidis isolates. *Antimicrobial Agents and Chemotherapy*, *55*(12), 5842–5849. https://doi.org/10.1128/AAC.00592-11

Chen, J., Quiles-Puchalt, N., Chiang, Y. N., Bacigalupe, R., Fillol-Salom, A., Chee, M. S. J., … Penadés, J. R. (2018). Genome hypermobility by lateral transduction. *Science*, *362*(6411), 207–212. https://doi.org/10.1126/science.aat5867

Chen, J., Ram, G., Penadés, J. R., Brown, S., & Novick, R. P. (2015). Pathogenicity island-directed transfer of unlinked chromosomal virulence genes. *Molecular Cell*, *57*(1), 138–149. https://doi.org/10.1016/j.molcel.2014.11.011

Cho, E. H., Gumport, R. I., & Gardner, J. F. (2002). Interactions between integrase and excisionase in the phage lambda excisive nucleoprotein complex. *Journal of Bacteriology*, *184*(18), 5200–5203. https://doi.org/10.1128/JB.184.18.5200-5203.2002

Christie, G. E., & Dokland, T. (2012). Pirates of the Caudovirales. *Virology*. https://doi.org/10.1016/j.virol.2012.10.028

Christie, G. E., Matthews, A. M., King, D. G., Lane, K. D., Olivarez, N. P., Tallent, S. M., … Novick, R. P. (2010). The complete genomes of *Staphylococcus aureus* bacteriophages 80 and 80α-Implications for the specificity of SaPI mobilization. *Virology*, *407*(2), 381–390. https://doi.org/10.1016/j.virol.2010.08.036

Cirz, R. T., Jones, M. B., Gingles, N. A., Minogue, T. D., Jarrahi, B., Peterson, S. N., & Romesberg, F. E. (2007). Complete and SOS-mediated response of *Staphylococcus aureus* to the antibiotic ciprofloxacin. *Journal of Bacteriology*, *189*(2), 531–539. https://doi.org/10.1128/JB.01464-06

Clauditz, A., Resch, A., Wieland, K. P., Peschel, A., & Götz, F. (2006). Staphyloxanthin plays a role in the fitness of *Staphylococcus aureus* and its ability to cope with oxidative stress. *Infection and Immunity*, *74*(8), 4950–4953. https://doi.org/10.1128/IAI.00204-06

Clokie, M. R. J., Millard, A. D., Letarov, A. V., & Heaphy, S. (2011). Phages in nature. *Bacteriophage*, *1*(1), 31–45. https://doi.org/10.4161/bact.1.1.14942

Cramton, S. E., Gerke, C., Schnell, N. F., Nichols, W. W., & Götz, F. (1999). The intercellular adhesion (ica) locus is present in *Staphylococcus aureus* and is required for biofilm formation. *Infection and Immunity*, *67*(10), 5427–5433.

Cucarella, C., Tormo, M. Á., Úbeda, C., Trotonda, M. P., Monzón, M., Peris, C., … Penadés, J. R. (2004). Role of Biofilm-Associated Protein Bap in the Pathogenesis of Bovine *Staphylococcus aureus*. *Infection and Immunity*, *72*(4), 2177–2185. https://doi.org/10.1128/IAI.72.4.2177-2185.2004

Damle, P. K., Wall, E. A., Spilman, M. S., Dearborn, A. D., Ram, G., Novick, R. P., … Christie, G. E. (2012). The roles of SaPI1 proteins gp7 (CpmA) and gp6 (CpmB) in capsid size determination and helper phage interference. *Virology*, *432*(2), 277–282. https://doi.org/10.1016/j.virol.2012.05.026

Dearborn, A. D., & Dokland, T. (2012). Mobilization of pathogenicity islands by *Staphylococcus aureus* strain Newman bacteriophages. *Bacteriophage*, *2*(2), 70–78. https://doi.org/10.4161/bact.20632

Dearborn, A. D., Spilman, M. S., Damle, P. K., Chang, J. R., Monroe, E. B., Saad, J. S., … Dokland, T. (2011). The *Staphylococcus aureus* pathogenicity island 1 protein gp6 functions as an internal scaffold during capsid size determination. *Journal of Molecular Biology*, *412*(4), 710–722. https://doi.org/10.1016/j.jmb.2011.07.036

Deghorain, M., & Van Melderen, L. (2012). The staphylococci phages family: An overview. *Viruses*. https://doi.org/10.3390/v4123316

Donderis, J., Bowring, J., Maiques, E., Ciges-Tomas, J. R., Alite, C., Mehmedov, I., … Marina, A. (2017). Convergent evolution involving dimeric and trimeric dUTPases in pathogenicity island mobilization. *PLOS Pathogens*, *13*(9), e1006581. https://doi.org/10.1371/journal.ppat.1006581

Epshtein, V., Dutta, D., Wade, J., & Nudler, E. (2010). An allosteric mechanism of Rho-dependent transcription termination. *Nature*, *463*(7278), 245–249. https://doi.org/10.1038/nature08669

Eyraud, A., Tattevin, P., Chabelskaya, S., & Felden, B. (2014). A small RNA controls a protein regulator involved in antibiotic resistance in *Staphylococcus aureus*. *Nucleic Acids Research*, *42*(8), 4892–4905. https://doi.org/10.1093/nar/gku149

Favrot, L., Blanchard, J. S., & Vergnolle, O. (2016). Bacterial GCN5-Related N-Acetyltransferases: From Resistance to Regulation. *Biochemistry*, *55*(7), 989–1002. https://doi.org/10.1021/acs.biochem.5b01269

Fechter, P., Caldelari, I., Lioliou, E., & Romby, P. (2014). Novel aspects of RNA regulation in *Staphylococcus aureus*. *FEBS Letters*. https://doi.org/10.1016/j.febslet.2014.05.037

Feiss, M., & Rao, V. B. (2012). The bacteriophage DNA packaging machine. *Advances in Experimental Medicine and Biology*, *726*, 489–509. https://doi.org/10.1007/978-1-4614-0980-9_22

Fernández, L., González, S., Quiles-Puchalt, N., Gutiérrez, D., Penadés, J. R., García, P., & Rodríguez, A. (2018). Lysogenization of *Staphylococcus aureus* RN450 by phages ϕ11 and ϕ80α leads to the activation of the SigB regulon. *Scientific Reports*, *8*(1), 12662. https://doi.org/10.1038/s41598-018-31107-z

Ferrer, M. D., Quiles-Puchalt, N., Harwich, M. D., Tormo-Más, M. Á., Campoy, S., Barbé, J., … Penadés, J. R. (2011). RinA controls phage-mediated packaging and transfer of virulence genes in Gram-positive bacteria. *Nucleic Acids Research*, *39*(14), 5866–5878. https://doi.org/10.1093/nar/gkr158

Fillol-Salom, A., Martínez-Rubio, R., Abdulrahman, R. F., Chen, J., Davies, R., & Penadés, J. R. (2018). Phage-inducible chromosomal islands are ubiquitous within the bacterial universe. *The ISME Journal*, 1. https://doi.org/10.1038/s41396-018-0156-3

Fokine, A., & Rossmann, M. G. (2014). Molecular architecture of tailed double-stranded DNA phages. *Bacteriophage*, *4*(2), e28281. https://doi.org/10.4161/bact.28281

Foster, T. J. (2017). Antibiotic resistance in *Staphylococcus aureus*. Current status and future prospects. *FEMS Microbiology Reviews*. https://doi.org/10.1093/femsre/fux007

Frígols, B., Quiles-Puchalt, N., Mir-Sanchis, I., Donderis, J., Elena, S. F., Buckling, A., … Penadés, J. R. (2015). Virus Satellites Drive Viral Evolution and Ecology. *PLoS Genetics*, *11*(10). https://doi.org/10.1371/journal.pgen.1005609

Frost, L. S., Leplae, R., Summers, A. O., & Toussaint, A. (2005). Mobile genetic elements: The agents of open source evolution. *Nature Reviews Microbiology*. https://doi.org/10.1038/nrmicro1235

Gao, P., Davies, J., & Kao, R. Y. T. (2017). Dehydrosqualene desaturase as a novel target for anti-virulence therapy against *Staphylococcus aureus*. *mBio*, *8*(5). https://doi.org/10.1128/mBio.01224-17

Goerke, C., Koller, J., & Wolz, C. (2006). Ciprofloxacin and trimethoprim cause phage induction and virulence modulation in *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, *50*(1), 171–177. https://doi.org/10.1128/AAC.50.1.171-177.2006

Goerke, C., Pantucek, R., Holtfreter, S., Schulte, B., Zink, M., Grumann, D., … Wolz, C. (2009). Diversity of prophages in dominant *Staphylococcus aureus* clonal lineages. *Journal of Bacteriology*, *191*(11), 3462–3468. https://doi.org/10.1128/JB.01804-08

Goodson, J. R., Klupt, S., Zhang, C., Straight, P., & Winkler, W. C. (2017). LoaP is a broadly conserved antiterminator protein that regulates antibiotic gene clusters in *Bacillus amyloliquefaciens.* *Nature Microbiology*, *2*. https://doi.org/10.1038/nmicrobiol.2017.3

Greenblatt, J., McLimont, M., & Hanly, S. (1981). Termination of transcription by *nusA* gene protein of Escherichia coli. *Nature*, *292*(5820), 215–220. https://doi.org/10.1038/292215a0

Grundmann, H., Aires-de-Sousa, M., Boyce, J., & Tiemersma, E. (2006). Emergence and resurgence of meticillin-resistant *Staphylococcus aureus* as a public-health threat. *Lancet*. https://doi.org/10.1016/S0140-6736(06)68853-3

Guttman, B., Raya, R., & Kutter, E. (2004). Basic Phage Biology. In *Bacteriophages*. https://doi.org/10.1201/9780203491751.ch3

Hall, J. W., Yang, J., Guo, H., & Ji, Y. (2017). The *Staphylococcus aureus* AirSR two-component system mediates reactive oxygen species resistance via transcriptional regulation of staphyloxanthin production. *Infection and Immunity*, *85*(2). https://doi.org/10.1128/IAI.00838-16

Hatfull, G. F. (2008). Bacteriophage genomics. *Current Opinion in Microbiology*. https://doi.org/10.1016/j.mib.2008.09.004

Hemker, H. C., Bas, B. M., & Muller, A. D. (1975). Activation of a pro-enzyme by a stoichiometric reaction with another protein. The reaction between prothrombin and staphylocoagulase. *BBA - Protein Structure*, *379*(1), 180–188.

https://doi.org/10.1016/0005-2795(75)90020-3

Herbert, S., Ziebandt, A. K., Ohlsen, K., Schäfer, T., Hecker, M., Albrecht, D., … Götz, F. (2010). Repair of global regulators in *Staphylococcus aureus* 8325 and comparative analysis with other clinical isolates. *Infection and Immunity*, *78*(6), 2877–2889. https://doi.org/10.1128/IAI.00088-10

Hill, R. L. L., & Dokland, T. (2016). The Type 2 dUTPase of Bacteriophage PdblNM1 Initiates Mobilization of *Staphylococcus aureus* Bovine Pathogenicity Island 1. *Journal of Molecular Biology*, *428*(1), 142–152. https://doi.org/10.1016/j.jmb.2015.11.009

Hiramatsu, K., Hanaki, H., Ino, T., Yabuta, K., Oguri, T., & Tenover, F. C. (1997). Methicillin-resistant *Staphylococcus aureus* clinical strain with reduced vancomycin susceptibility [1]. *Journal of Antimicrobial Chemotherapy*. https://doi.org/10.1093/jac/40.1.135

Hu, C., Xiong, N., Zhang, Y., Rayner, S., & Chen, S. (2012). Functional characterization of lipase in the pathogenesis of *Staphylococcus aureus*. *Biochemical and Biophysical Research Communications*, *419*(4), 617–620. https://doi.org/10.1016/j.bbrc.2012.02.057

Ibarra, J. A., Pérez-Rueda, E., Carroll, R. K., & Shaw, L. N. (2013). Global analysis of transcriptional regulators in *Staphylococcus aureus. BMC Genomics*, *14*(1). https://doi.org/10.1186/1471-2164-14-126

Katzif, S., Lee, E. H., Law, A. B., Tzeng, Y. L., & Shafer, W. M. (2005). CspA regulates pigment production in *Staphylococcus aureus* through a SigB-dependent mechanism. *Journal of Bacteriology*, *187*(23), 8181–8184. https://doi.org/10.1128/JB.187.23.8181-8184.2005

Keen, E. C., & Dantas, G. (2018). Close Encounters of Three Kinds: Bacteriophages, Commensal Bacteria, and Host Immunity. *Trends in Microbiology*. https://doi.org/10.1016/j.tim.2018.05.009

Kullik, I., Giachino, P., & Fuchs, T. (1998). Deletion of the Alternative Sigma Factor in *Staphylococcus aureus* Reveals Its Function as a Global Regulator of Virulence Genes. *Journal of Bacteriology*, *180*(18), 4814–4820.

Kurt, K., Rasigade, J. P., Laurent, F., Goering, R. V., Žemličková, H., Machova, I., … Nübel, U. (2013). Subpopulations of *Staphylococcus aureus* Clonal Complex 121 Are Associated with Distinct Clinical Entities. *PLoS ONE*, *8*(3).

https://doi.org/10.1371/journal.pone.0058155

Kutter, E., & Sulakvelidze, A. (2005). *Bacteriophages: biology and applications*. *New York*. https://doi.org/10.1201/9780203491751

Lan, L., Cheng, A., Dunman, P. M., Missiakas, D., & He, C. (2010). Golden pigment production and virulence gene expression are affected by metabolisms in *Staphylococcus aureus. Journal of Bacteriology*, *192*(12), 3068–3077. https://doi.org/10.1128/JB.00928-09

Le Scornet, A., & Redder, P. (2018). Post-transcriptional control of virulence gene expression in *Staphylococcus aureus*. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*. https://doi.org/10.1016/j.bbagrm.2018.04.004

Lennox, E. S. (1955). Transduction of linked genetic characters of the host by bacteriophage P1. *Virology*, *1*(2), 190–206. https://doi.org/10.1016/0042-6822(55)90016-7

Lindsay, J. A., & Holden, M. T. G. (2004). *Staphylococcus aureus:* Superbug, super genome? *Trends in Microbiology*. https://doi.org/10.1016/j.tim.2004.06.004

Lindsay, J. A., & Holden, M. T. G. (2006). Understanding the rise of the superbug: Investigation of the evolution and genomic variation of *Staphylococcus aureus. Functional and Integrative Genomics*. https://doi.org/10.1007/s10142-005-0019-7

Lindsay, J. A., Ruzin, A., Ross, H. F., Kurepina, N., & Novick, R. P. (1998). The gene for toxic shock toxin is carried by a family of mobile pathogenicity islands in *Staphylococcus aureus. Molecular Microbiology*, *29*(2), 527–543. https://doi.org/10.1046/j.1365-2958.1998.00947.x

Liu, C. I., Liu, G. Y., Song, Y., Yin, F., Hensler, M. E., Jeng, W. Y., … Oldfield, E. (2008). A cholesterol biosynthesis inhibitor blocks *Staphylococcus aureus* virulence. *Science*, *319*(5868), 1391–1394. https://doi.org/10.1126/science.1153018

Liu, G. Y., Essex, A., Buchanan, J. T., Datta, V., Hoffman, H. M., Bastian, J. F., … Nizet, V. (2005). *Staphylococcus aureus* golden pigment impairs neutrophil killing and promotes virulence through its antioxidant activity. *The Journal of Experimental Medicine*, *202*(2), 209–215. https://doi.org/10.1084/jem.20050846

Liu, J., Dehbi, M., Moeck, G., Arhin, F., Banda, P., Bergeron, D., … DuBow, M. (2004). Antimicrobial drug discovery through bacteriophage genomics. *Nature Biotechnology*,

22(2), 185–191. https://doi.org/10.1038/nbt932

Lorenz, U., Lorenz, B., Schmitter, T., Streker, K., Erck, C., Wehland, J., … Ohlsen, K. (2011). Functional antibodies targeting IsaA of *Staphylococcus aureus* augment host immune response and open new perspectives for antibacterial therapy. *Antimicrobial Agents and Chemotherapy*. https://doi.org/10.1128/AAC.01144-10

Madhusoodanan, J., Seo, K. S., Remortel, B., Park, J. Y., Hwang, S. Y., Fox, L. K., … Gill, S. R. (2011). An enterotoxin-bearing pathogenicity island in *Staphylococcus epidermidis*. *Journal of Bacteriology*, *193*(8), 1854–1862. https://doi.org/10.1128/JB.00162-10

Maiques, E., Quiles-Puchalt, N., Donderis, J., Ciges-Tomas, J. R., Alite, C., Bowring, J. Z., … Marina, A. (2016). Another look at the mechanism involving trimeric dUTPases in *Staphylococcus aureus* pathogenicity island induction involves novel players in the party. *Nucleic Acids Research*, *44*(11), 5457–5469. https://doi.org/10.1093/nar/gkw317

Maiques, E., Úbeda, C., Tormo, M. Á., Ferrer, M. D., Lasa, Í., Novick, R. P., & Penadés, J. R. (2007). Role of staphylococcal phage and SaPI integrase in intra- and interspecies SaPI transfer. *Journal of Bacteriology*, *189*(15), 5608–5616. https://doi.org/10.1128/JB.00619-07

Malachowa, N., & Deleo, F. R. (2010). Mobile genetic elements of *Staphylococcus aureus*. *Cellular and Molecular Life Sciences*. https://doi.org/10.1007/s00018-010-0389-4

Maniloff, J., & Ackermann, H. W. (1998). Taxonomy of bacterial viruses: Establishment of tailed virus genera and the order Caudovirales. *Archives of Virology*, *143*(10), 2051–2063. https://doi.org/10.1007/s007050050442

Marshall, J. H., & Wilmoth, G. J. (1981). Proposed pathway of triterpenoid carotenoid biosynthesis in *Staphylococcus aureus*: Evidence from a study of mutants. *Journal of Bacteriology*, *147*(3), 914–919.

Martínez-Rubio, R., Quiles-Puchalt, N., Martí, M., Humphrey, S., Ram, G., Smyth, D., … Penadés, J. R. (2017). Phage-inducible islands in the Gram-positive cocci. *ISME Journal*, *11*(4), 1029–1042. https://doi.org/10.1038/ismej.2016.163

Matos, R. C., Lapaque, N., Rigottier-Gois, L., Debarbieux, L., Meylheuc, T., Gonzalez-Zorn, B., … Serror, P. (2013). *Enterococcus faecalis* Prophage Dynamics and Contributions to

Pathogenic Traits. *PLoS Genetics*, *9*(6). https://doi.org/10.1371/journal.pgen.1003539

Mcnamara, P. J., Milligan-Monroe, K. C., Khalili, S., & Proctor, R. A. (2000). Identification, cloning, and initial characterization of rot, a locus encoding a regulator of virulence factor expression in *Staphylococcus aureus*. *Journal of Bacteriology*, *182*(11), 3197–3203. https://doi.org/10.1128/JB.182.11.3197-3203.2000

Merino, E., & Yanofsky, C. (2005). Transcription attenuation: A highly conserved regulatory strategy used by bacteria. *Trends in Genetics*. https://doi.org/10.1016/j.tig.2005.03.002

Mir-Sanchis, I., Martínez-Rubio, R., Martí, M., Chen, J., Lasa, Í., Novick, R. P., … Penadés, J. R. (2012). Control of *Staphylococcus aureus* pathogenicity island excision. *Molecular Microbiology*, *85*(5), 833–845. https://doi.org/10.1111/j.1365-2958.2012.08145.x

Mondal, S., Yakhnin, A. V., Sebastian, A., Albert, I., & Babitzke, P. (2016). NusA-dependent transcription termination prevents misregulation of global gene expression. *Nature Microbiology*, *1*(1). https://doi.org/10.1038/nmicrobiol.2015.7

Moreillon, P., Entenza, J. M., Francioli, P., McDevitt, D., Foster, T. J., Francois, P., & Vaudaux, P. (1995). Role of *Staphylococcus aureus* coagulase and clumping factor in pathogenesis of experimental endocarditis. *Infection and Immunity*, *63*(12), 4738–4743.

Neamah, M. M., Mir-Sanchis, I., Lopez-Sanz, M., Acosta, S., Baquedano, I., Haag, A. F., … Penadés, J. R. (2017). Sak and Sak4 recombinases are required for bacteriophage replication in *Staphylococcus aureus*. *Nucleic Acids Research*, *45*(11), 6507–6519. https://doi.org/10.1093/nar/gkx308

Novick, R. P., Christie, G. E., & Penadés, J. R. (2010). The phage-related chromosomal islands of Gram-positive bacteria. *Nature Reviews Microbiology*. https://doi.org/10.1038/nrmicro2393

Novick, R. P., & Ram, G. (2016). The Floating (Pathogenicity) Island: A Genomic Dessert. *Trends in Genetics*. https://doi.org/10.1016/j.tig.2015.11.005

Novick, R. P., Ross, H. F., Projan, S. J., Kornblum, J., Kreiswirth, B., & Moghazeh, S. (1993). Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA

molecule. *The EMBO Journal*, *12*(10), 3967–75. https://doi.org/10.1016/j.jnutbio.2011.09.003

Oliveros, J. C. (2016). Venny. An interactive tool for comparing lists with Venn﷽s diagrams. 2007--2015. h ttp. *Bioinfogp. Cnb. Csic. Es/tools/venny/index. Html. Accessed*, *1*.

Oppenheim, A. B., Kobiler, O., Stavans, J., Court, D. L., & Adhya, S. (2005). Switches in Bacteriophage Lambda Development. *Annual Review of Genetics*, *39*(1), 409–429. https://doi.org/10.1146/annurev.genet.39.073003.113656

Osmundson, J., Montero-Diez, C., Westblade, L. F., Hochschild, A., & Darst, S. A. (2012). Promoter-specific transcription inhibition in *Staphylococcus aureus* by a phage protein. *Cell*, *151*(5), 1005–1016. https://doi.org/10.1016/j.cell.2012.10.034

Otto, M. (2009). Staphylococcus epidermidis - The "accidental" pathogen. *Nature Reviews Microbiology*. https://doi.org/10.1038/nrmicro2182

Otto, M. (2014). *Staphylococcus aureus* toxins. *Current Opinion in Microbiology*. https://doi.org/10.1016/j.mib.2013.11.004

Papp-Kádár, V., Szabó, J. E., Nyíri, K., & Vertessy, B. G. (2016). In Vitro analysis of predicted DNA-binding sites for the Stl repressor of the *Staphylococcus aureus* SaPIBov1 pathogenicity island. *PLoS ONE*, *11*(7). https://doi.org/10.1371/journal.pone.0158793

Patterson, T. A., Zhang, Z., Baker, T., Johnson, L. L., Friedman, D. I., & Court, D. L. (1994). Bacteriophage lambda N-dependent transcription antitermination: Competition for an RNA site may regulate antitermination. *Journal of Molecular Biology*, *236*(1), 217–228. https://doi.org/10.1006/jmbi.1994.1131

Peacock, S. J., De Silva, I., & Lowy, F. D. (2001). What determines nasal carriage of *Staphylococcus aureus*? *Trends in Microbiology*. https://doi.org/10.1016/S0966-842X(01)02254-5

Pelz, A., Wieland, K. P., Putzbach, K., Hentschel, P., Albert, K., & Götz, F. (2005). Structure and biosynthesis of staphyloxanthin from *Staphylococcus aureus*. *Journal of Biological Chemistry*, *280*(37), 32493–32498. https://doi.org/10.1074/jbc.M505070200

Poliakov, A., Chang, J. R., Spilman, M. S., Damle, P. K., Christie, G. E., Mobley, J. A., & Dokland,

T. (2008). Capsid Size Determination by *Staphylococcus aureus* Pathogenicity Island SaPI1 Involves Specific Incorporation of SaPI1 Proteins into Procapsids. *Journal of Molecular Biology*, *380*(3), 465–475. https://doi.org/10.1016/j.jmb.2008.04.065

Qin, L., Da, F., Fisher, E. L., Tan, D. C. S., Nguyen, T. H., Fu, C. L., … Otto, M. (2017). Toxin Mediates Sepsis Caused by Methicillin-Resistant *Staphylococcus epidermidis*. *PLoS Pathogens*, *13*(2). https://doi.org/10.1371/journal.ppat.1006153

Quiles-Puchalt, N., Carpena, N., Alonso, J. C., Novick, R. P., Marina, A., & Penades, J. R. (2014). Staphylococcal pathogenicity island DNA packaging system involving cos-site packaging and phage-encoded HNH endonucleases. *Proceedings of the National Academy of Sciences*, *111*(16), 6016–6021. https://doi.org/10.1073/pnas.1320538111

Quiles-Puchalt, N., Tormo-Más, M. Á., Campoy, S., Toledo-Arana, A., Monedero, V., Lasa, Í., … Penadés, J. R. (2013). A super-family of transcriptional activators regulates bacteriophage packaging and lysis in Gram-positive bacteria. *Nucleic Acids Research*, *41*(15), 7260–7275. https://doi.org/10.1093/nar/gkt508

Ram, G., Chen, J., Kumar, K., Ross, H. F., Ubeda, C., Damle, P. K., … Novick, R. P. (2012). Staphylococcal pathogenicity island interference with helper phage reproduction is a paradigm of molecular parasitism. *Proceedings of the National Academy of Sciences*, *109*(40), 16300–16305. https://doi.org/10.1073/pnas.1204615109

Ram, G., Chen, J., Ross, H. F., & Novick, R. P. (2014). Precisely modulated pathogenicity island interference with late phage gene transcription. *Proceedings of the National Academy of Sciences*, *111*(40), 14536–14541. https://doi.org/10.1073/pnas.1406749111

Ram, G., Chen, J., Ross, H. F., & Novick, R. P. (2015). An insight into staphylococcal pathogenicity island-mediated interference with phage late gene transcription. *Bacteriophage*, *5*(2), e1028608. https://doi.org/10.1080/21597081.2015.1028608

Rao, V. B., & Feiss, M. (2015). Mechanisms of DNA Packaging by Large Double-Stranded DNA Viruses. *Annual Review of Virology*, *2*(1), 351–378. https://doi.org/10.1146/annurev-virology-100114-055212

Renzoni, A., Kelley, W. L., Barras, C., Monod, A., Huggler, E., Frangois, P., … Lew, D. P. (2009). Identification by genomic and genetic analysis of two new genes playing a key role in

intermediate glycopeptide resistance in *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, *53*(3), 903–911. https://doi.org/10.1128/AAC.01287-08

Rosado, H., Doyle, M., Hinds, J., & Taylor, P. W. (2010). Low-shear modelled microgravity alters expression of virulence determinants of *Staphylococcus aureus*. *Acta Astronautica*, *66*(3–4), 408–413. https://doi.org/10.1016/j.actaastro.2009.06.007

Russell, D. W., & Sambrook, J. (2001). *Molecular cloning: A laboratory manual*. *Cold Spring Harbour*. Retrieved from www.molecularcloning.com

Ruzin, A., Lindsay, J., & Novick, R. P. (2001). Molecular genetics of SaPI1--a mobile pathogenicity island in *Staphylococcus aureus*. *Molecular Microbiology*, *41*(2), 365–77. https://doi.org/mmi2488 [pii]

Santangelo, T. J., & Artsimovitch, I. (2011). Termination and antitermination: RNA polymerase runs a stop sign. *Nature Reviews Microbiology*, *9*(5), 319–329. https://doi.org/10.1038/nrmicro2560

Schindler, D., & Echols, H. (1981). Retroregulation of the int gene of bacteriophage A: Control of translation completion (bacteriophage A integration-excision/regulation by recombination/regulation by RNase II). *Genetics*, *78*(7), 4475–4479. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC319814/pdf/pnas00658-0516.pdf

Schmieger, H., & Backhaus, H. (1973). The origin of DNA in transducing particles in P22-mutants with increased transduction-frequencies (HT-mutants). *MGG Molecular & General Genetics*, *120*(2), 181–190. https://doi.org/10.1007/BF00267246

Schulte, B., Bierbaum, G., Pohl, K., Goerke, C., & Wolz, C. (2013). Diversification of clonal complex 5 methicillin-resistant *Staphylococcus aureus* strains (Rhine-Hesse clone) within Germany. *Journal of Clinical Microbiology*, *51*(1), 212–216. https://doi.org/10.1128/JCM.01967-12

Sevostyanova, A., Belogurov, G. A., Mooney, R. A., Landick, R., & Artsimovitch, I. (2011). The β Subunit Gate Loop Is Required for RNA Polymerase Modification by RfaH and NusG. *Molecular Cell*, *43*(2), 253–262. https://doi.org/10.1016/j.molcel.2011.05.026

Sippy, J., Patel, P., Vahanian, N., Sippy, R., & Feiss, M. (2015). Genetics of critical contacts and clashes in the DNA packaging specificities of bacteriophages λ and 21. *Virology*, *476*, 115–123. https://doi.org/10.1016/j.virol.2014.11.028

Skalka, A., Poonian, M., & Bartl, P. (1972). Concatemers in DNA replication: Electron microscopic studies of partially denatured intracellular lambda DNA. *Journal of Molecular Biology*, *64*(3). https://doi.org/10.1016/0022-2836(72)90081-2

Spilman, M. S., Damle, P. K., Dearborn, A. D., Rodenburg, C. M., Chang, J. R., Wall, E. A., … Dokland, T. (2012). Assembly of bacteriophage 80α capsids in a *Staphylococcus aureus* expression system. *Virology*, *434*(2), 242–250. https://doi.org/10.1016/j.virol.2012.08.031

Spilman, M. S., Dearborn, A. D., Chang, J. R., Damle, P. K., Christie, G. E., & Dokland, T. (2011). A conformational switch involved in maturation of *Staphylococcus aureus* bacteriophage 80α capsids. *Journal of Molecular Biology*, *405*(3), 863–876. https://doi.org/10.1016/j.jmb.2010.11.047

Srivastava, P., Khandokar, Y. B., Swarbrick, C. M. D., Roman, N., Himiari, Z., Sarker, S., … Forwood, J. K. (2014). Structural characterization of a Gcn5-related N-Acetyltransferase from *Staphylococcus aureus*. *PLoS ONE*, *9*(8). https://doi.org/10.1371/journal.pone.0102348

Stapleton, M. R., Horsburgh, M. J., Hayhurst, E. J., Wright, L., Jonsson, I. M., Tarkowski, A., … Foster, S. J. (2007). Characterization of IsaA and SceD, two putative lytic transglycosylases of *Staphylococcus aureus*. *Journal of Bacteriology*, *189*(20), 7316–7325. https://doi.org/10.1128/JB.00734-07

Subedi, A., Ubeda, C., Adhikari, R. P., Penadés, J. R., & Novick, R. P. (2007). Sequence analysis reveals genetic exchanges and intraspecific spread of SaPl2, pathogenicity island involved in menstrual toxic shock. *Microbiology*, *153*(10), 3235–3245. https://doi.org/10.1099/mic.0.2007/006932-0

Surányi, É. V., Hírmondó, R., Nyíri, K., Tarjányi, S., Kőhegyi, B., Tóth, J., & Vértessy, B. G. (2018). Exploiting a phage-bacterium interaction system as a molecular switch to decipher macromolecular interactions in the living cell. *Viruses*, *10*(4). https://doi.org/10.3390/v10040168

Tallent, S. M., Langston, T. B., Moran, R. G., & Christie, G. E. (2007). Transducing particles of *Staphylococcus aureus* pathogenicity island SaPI1 are comprised of helper phage-encoded proteins. *Journal of Bacteriology*, *189*(20), 7520–7524. https://doi.org/10.1128/JB.00738-07

Thierauf, A., & Maloy, S. (2011). Generalized Transduction. In *Prokaryotic Genomics* (pp. 50–64). https://doi.org/10.1007/978-3-0348-8963-6_6

Tillotson, G. S. (2008). *Staphylococcus* : Molecular Genetics . *Expert Review of Anti-Infective Therapy*, *6*(6), 849–850. https://doi.org/10.1586/14787210.6.6.849

Toledo-Arana, A., Repoila, F., & Cossart, P. (2007). Small noncoding RNAs controlling pathogenesis. *Current Opinion in Microbiology*. https://doi.org/10.1016/j.mib.2007.03.004

Tong, S. Y. C., Davis, J. S., Eichenberger, E., Holland, T. L., & Fowler, V. G. (2015). *Staphylococcus aureus* infections: Epidemiology, pathophysiology, clinical manifestations, and management. *Clinical Microbiology Reviews*, *28*(3), 603–661. https://doi.org/10.1128/CMR.00134-14

Tormo-Más, M. Á., Mir, I., Shrestha, A., Tallent, S. M., Campoy, S., Lasa, Í., … Penadés, J. R. (2010). Moonlighting bacteriophage proteins derepress staphylococcal pathogenicity islands. *Nature*, *465*(7299), 779–782. https://doi.org/10.1038/nature09065

Tormo-Más, M. ángeles, Donderis, J., García-Caballer, M., Alt, A., Mir-Sanchis, I., Marina, A., & Penadés, J. R. (2013). Phage dUTPases Control Transfer of Virulence Genes by a Proto-Oncogenic G Protein-like Mechanism. *Molecular Cell*, *49*(5), 947–958. https://doi.org/10.1016/j.molcel.2012.12.013

Tormo, M. Á., Ferrer, M. D., Maiques, E., Úbeda, C., Selva, L., Lasa, Í., … Penadés, J. R. (2008). *Staphylococcus aureus* pathogenicity island DNA is packaged in particles composed of phage proteins. *Journal of Bacteriology*, *190*(7), 2434–2440. https://doi.org/10.1128/JB.01349-07

Tormo, M. Á., Úbeda, C., Martí, M., Maiques, E., Cucarella, C., Valle, J., … Penadés, J. R. (2007). Phase-variable expression of the biofilm-associated protein (Bap) in *Staphylococcus aureus*. *Microbiology*, *153*(6), 1702–1710. https://doi.org/10.1099/mic.0.2006/005744-0

Touchon, M., Moura De Sousa, J. A., & PC Rocha, E. (2017). Embracing the enemy: the diversification of microbial gene repertoires by phage-mediated horizontal gene transfer. *Current Opinion in Microbiology*, *38*, 66–73. https://doi.org/10.1016/J.MIB.2017.04.010

Úbeda, C., Maiques, E., Barry, P., Matthews, A., Tormo, M. Á., Lasa, Í., … Penadés, J. R. (2008). SaPI mutations affecting replication and transfer and enabling autonomous replication in the absence of helper phage. *Molecular Microbiology*. https://doi.org/10.1111/j.1365-2958.2007.06027.x

Úbeda, C., Maiques, E., Knecht, E., Lasa, Í., Novick, R. P., & Penadés, J. R. (2005). Antibiotic-induced SOS response promotes horizontal dissemination of pathogenicity island-encoded virulence factors in staphylococci. *Molecular Microbiology*, *56*(3), 836–844. https://doi.org/10.1111/j.1365-2958.2005.04584.x

Úbeda, C., Maiques, E., Tormo, M. Á., Campoy, S., Lasa, Í., Barbé, J., … Penadés, J. R. (2007). SaPI operon I is required for SaPI packaging and is controlled by LexA. *Molecular Microbiology*, *65*(1), 41–50. https://doi.org/10.1111/j.1365-2958.2007.05758.x

Ubeda, C., Olivarez, N. P., Barry, P., Wang, H., Kong, X., Matthews, A., … Novick, R. P. (2009). Specificity of staphylococcal phage and SaPI DNA packaging as revealed by integrase and terminase mutations. *Molecular Microbiology*, *72*(1), 98–108. https://doi.org/10.1111/j.1365-2958.2009.06634.x

Ubeda, C., Tormo-Más, M. ángeles, Penadés, J. R., & Novick, R. P. (2012). Structure-function analysis of the SaPIbov1 replication origin in *Staphylococcus aureus*. *Plasmid*, *67*(2), 183–190. https://doi.org/10.1016/j.plasmid.2012.01.006

Valle, J., Vergara-Irigaray, M., Merino, N., Penadés, J. R., & Lasa, I. (2007). σB regulates IS256-mediated *Staphylococcus aureus* biofilm phenotypic variation. *Journal of Bacteriology*, *189*(7), 2886–2896. https://doi.org/10.1128/JB.01767-06

Vértessy, B. G., & Tóth, J. (2009). Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases. *Accounts of Chemical Research*, *42*(1), 97–106. https://doi.org/10.1021/ar800114w

Viana, D., Comos, M., McAdam, P. R., Ward, M. J., Selva, L., Guinane, C. M., … Penadés, J. R. (2015). A single natural nucleotide mutation alters bacterial pathogen host tropism. *Nature Genetics*, *47*(4), 361–6. https://doi.org/10.1038/ng.3219

Weisberg, R. A., & Gottesman, M. E. (1999). Processive antitermination. *Journal of Bacteriology*.

Wertheim, H. F., Melles, D. C., Vos, M. C., van Leeuwen, W., van Belkum, A., Verbrugh, H. a, &

Nouwen, J. L. (2005). The role of nasal carriage in *Staphylococcus aureus* infections. *The Lancet Infectious Diseases*, *5*(12), 751–762. https://doi.org/10.1016/S1473-3099(05)70295-4

Xia, G., & Wolz, C. (2014). Phages of *Staphylococcus aureus* and their impact on host evolution. *Infection, Genetics and Evolution*, *21*, 593–601. https://doi.org/10.1016/j.meegid.2013.04.022

Yakhnin, A. V., Murakami, K. S., & Babitzke, P. (2016). NusG is a sequence-specific RNA polymerase pause factor that binds to the non-template DNA within the paused transcription bubble. *Journal of Biological Chemistry*, *291*(10), 5299–5308. https://doi.org/10.1074/jbc.M115.704189

Yakhnin, A. V., Yakhnin, H., & Babitzke, P. (2008). Function of the *Bacillus subtilis* transcription elongation factor NusG in hairpin-dependent RNA polymerase pausing in the trp leader. *Proceedings of the National Academy of Sciences*, *105*(42), 16131–16136. https://doi.org/10.1073/pnas.0808842105

Yarnell, W. S., & Roberts, J. W. (1992). The phage λ gene Q transcription antiterminator binds DNA in the late gene promoter as it modifies RNA polymerase. *Cell*, *69*(7), 1181–1189. https://doi.org/10.1016/0092-8674(92)90639-T

Young, R. (1992). Bacteriophage lysis: mechanism and regulation. *Microbiological Reviews*, *56*(3), 430–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1406491%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC372879

Young, R. (2014). Phage lysis: Three steps, three choices, one outcome. *Journal of Microbiology*. https://doi.org/10.1007/s12275-014-4087-z

Zylicz, M., Liberek, K., Wawrzynow, A., & Georgopoulos, C. (1998). Formation of the preprimosome protects lambda O from RNA transcription-dependent proteolysis by ClpP/ClpX. *Proc Natl Acad Sci U S A*, *95*(26), 15259–15263. https://doi.org/10.1073/pnas.95.26.15259

8. Appendices

Appendix 1: Bacterial strains used in this study

1.1 List of strains used in the section "Mobilisation of SaPIs by endogenous phages in clinical strains"

Strain	Description	Reference
RN4220	Restriction-defective derivate of RN450	Lab strain
DH5α	Recombination-deficient cloning strain	Lab strain
BL21(DE3)	Expression strain	Stratagene
RN10359	RN450 lysogenic for 80α	Lab strain
RN451	RN450 lysogenic for Φ11	Lab strain
A	Isolated in France (1981)	(Viana *et al*., 2015)
B	Isolated in USA (2001)	
C	Isolated in Spain (2002)	
D	Isolated in Belgium (2003)	
E	Isolated in Lesser Antilles (2005)	
COL	Isolated in England (1960)	(Gill *et al*., 2005)
E-MRSA 16	Isolated in UK	
MRSA252	Isolated in 1997	(Holden *et al*., 2004)
MW2	Isolated in USA (1998)	(Baba *et al*., 2002)
Mu50	Isolated in Japan (1997)	(Kuroda *et al*., 2001)
MN8	Isolated in (1980)	(Schlievert & Blomster, 1983)
N315	Isolated in Japan (1982)	(Kuroda *et al*., 2001)
TW20	Isolated in England (2010)	(Holden *et al*., 2010)
USA300	Isolated in San Francisco (2005)	(Diep *et al*., 2006)
GTM750	MW2 SaPImw2 *entC::cat*	This study
GTM751	RN4220 SaPImw2 *entC::cat*	This study
GTM752	RN4220 lysogenic for ϕSa2mw	This study
GTM753	GTM752 SaPImw2 *entC::cat*	This study
GTM754	GTM752 pGTM2	This study
GTM755	RN4220 pGTM2	This study
GTM756	MW2 ΔDUF3113 SaPImw2 *entC::cat*	This study
GTM757	RN4220 ΦSa2mw ΔDUF3113 SaPImw2 *entC::cat*	This study
GTM758	RN4220 ΦSa2mw ΔDUF3113	This study
GTM759	GTM757 pGTM20	This study
GTM760	GTM755 pGTM4	This study
GTM761	GTM755 pGTM5	This study
GTM762	GTM755 pGTM6	This study
GTM763	GTM755 pGTM7	This study
GTM764	GTM755 pGTM8	This study
GTM765	GTM755 pGTM9	This study
GTM766	GTM755 pGTM10	This study
GTM767	GTM755 pGTM11	This study
GTM768	GTM755 pGTM12	This study

GTM769	GTM755 pGTM13	This study
GTM770	GTM755 pGTM14	This study
GTM771	GTM755 pGTM15	This study
GTM772	GTM755 pGTM16	This study
GTM773	GTM755 pGTM17	This study
GTM774	GTM755 pGTM18	This study
GTM775	GTM755 pGTM19	This study
GTM776	BL21 pGTM22	This study
GTM777	BL21 pGTM23	This study
GTM778	BL21 pGTM24	This study
GTM779	BL21 pGTM25	This study
GTM780	BL21 pGTM26	This study
GTM781	RN4220 80α pGTM2	This study
GTM782	RN4220phi11 pGTM2	This study
GTM783	RN4220 phi12 pGTM2	This study
GTM784	RN4220 80α SaPImw2 *entC::cat*	This study
GTM785	RN4220 phi11 SaPImw2 *entC::cat*	This study
GTM786	RN4220 phi12 SaPImw2 *entC::cat*	This study
GTM787	RN4220 pGTM27 SaPImw2 *entC::cat*	This study
GTM788	RN4220 pGTM28 SaPImw2 *entC::cat*	This study
GTM789	RN4220 pGTM29 SaPImw2 *entC::cat*	This study
GTM790	RN4220 pGTM30 SaPImw2 *entC::cat*	This study
GTM791	RN4220 pGTM31 SaPImw2 *entC::cat*	This study
GTM792	RN4220 pGTM2 pGTM32	This study
GTM793	RN4220 pGTM2 pGTM33	This study
GTM794	RN4220 pGTM2 pGTM34	This study
GTM795	RN4220 pGTM2 pGTM35	This study
GTM796	RN4220 pGTM2 pGTM36	This study
GTM797	RN4220 pGTM37 SaPImw2 *entC::cat*	This study
GTM798	RN4220 pGTM38 SaPImw2 *entC::cat*	This study
GTM799	RN4220 pGTM39 SaPImw2 *entC::cat*	This study
GTM800	RN4220 pGTM40 SaPImw2 *entC::cat*	This study
GTM801	RN4220 pGTM2 pGTM41	This study
GTM802	RN4220 pGTM2 pGTM42	This study
GTM803	RN4220 pGTM2 pGTM43	This study
GTM804	RN4220 pGTM2 pGTM44	This study
GTM805	RN4220 SaPImw2 *entC::cat* pGTM20	This study
GTM806	RN4220 SaPImw2 *entC::cat* phi11 ΔDUF3113	This study
GTM807	RN4220 SaPImw2 *entC::cat* phi12 ΔDUF3113	This study

1.2 List of strains used in the section "SaPIs control of chromosomally-encoded virulence gene expression"

Strain	Description	Reference
RN450		Lab strain
SH1000	8325-4 (*rsbU*$^+$)	Lab strain
SH1000 Δσ^B		Lab strain
RN10359	RN450 (80α)	Lab strain
GTM887	RN450 SaPIbov1 SaPI1	This study
GTM888	RN450 80α SaPIbov1 SaPI1	This study
JP2015	RN450 SaPIbov1Δ*stl*	(Úbeda *et al.,* 2008)
GTM889	SH1000 SB1Δ*stl*	This study
GTM890	SH1000 σ^B SB1Δ*stl*	This study
GTM891	SH1000 SB1Δ*stl* Δ*crtOPQMN*	This study
GTM892	SH1000 σ^B SB1Δ*stl ΔcrtOPQMN*	This study
GTM893	RN450 SB1 ΔOpI (transduced from JP1997)	(Úbeda *et al*., 2007)
GTM894	RN450 SB1 ΔORF5 (transduced from JP1998)	(Úbeda *et al.,* 2008)
GTM895	RN450 SB1 ΔORF6 (transduced from JP1999)	
GTM896	RN450 SB1 ΔORF7 (transduced from JP2000)	
GTM897	RN450 SB1 ΔORF8 (transduced from JP2001)	
GTM898	RN450 SB1 ΔORF9(transduced from JP2002)	
GTM899	RN450 SB1 ΔORF10 (transduced from JP2003)	
GTM900	RN450 pCN51_*ptiA*	This study
GTM901	RN450 pCN51_*ptiM*	This study
GTM902	RN450 pCN51_*ptiA-M*	This study
GTM903	SH1000 pCN51_*ptiA-M*	This study
GTM904	SH1000 σ^B pCN51_*ptiA-M*	This study
GTM905	SH1000 Δ*crtOPQMN* pCN51_*ptiA-M*	This study
GTM906	SH1000 σ^B *ΔcrtOPQMN* pCN51_*ptiA-M*	This study
GTM907	RN450 pCN51_*ptiA-M*$_{SaPI2}$	This study
GTM908	RN450 pCN51_*ptiA-M*$_{SaPIeq}$	This study
GTM909	RN450 pCN51_*ptiA-M*$_{SaPIpt1028}$	This study
GTM910	RN450 pCN51_*ptiA*K2*-*M*	This study
GTM911	RN450 pCN51_*ptiA*R64*-*M*	This study
GTM912	RN450 pCN51_*ptiA-M*K2*	This study
GTM913	RN450 pCN51_*ptiA-M*K88*	This study
GTM914	RN450 Δ*crtO*	This study
GTM915	RN450 Δ*crtO* pCN51ø	This study

GTM916	RN450 ΔP*crtO* pCN51_*ptiA-M*	This study
GTM917	RN450 Δ*PisaA*	This study
GTM918	RN450 ΔPi*saA* pCN51*ø*	This study
GTM919	RN450 Δ*PisaA* pCN51_*ptiA-M*	This study
GTM920	RN450 Δ*Pgnat*	This study
GTM921	RN450 Δ*Pgnat* pCN51*ø*	This study
GTM922	RN450 Δ*Pgnat* pCN51_*ptiA-M*	This study
GTM923	RN450 ΔPo*atA*	This study
GTM924	RN450 ΔPo*atA* pCN51*ø*	This study
GTM925	RN450 ΔPo*atA* pCN51_*ptiA-M*	This study
GTM926	RN450 Δ*PssaA*	This study
GTM927	RN450 Δ*PssaA* pCN51*ø*	This study
GTM928	RN450 Δ*PssaA* pCN51_*ptiA-M*	This study
GTM929	RN450 Δ*isaA*	This study
GTM930	RN450 Δi*saA* pCN51*ø*	This study
GTM931	RN450 Δ*isaA* pCN51_*ptiA-M*	This study
GTM932	RN450 *isaA*L10*	This study
GTM933	RN450 i*saA*L10* pCN51*ø*	This study
GTM934	RN450 *isaA* L10* pCN51_*ptiA-M*	This study
GTM935	RN450 Δo*atA*	This study
GTM936	RN450 Δo*atA* pCN51*ø*	This study
GTM937	RN450 Δo*atA* pCN51_*ptiA-M*	This study
GTM938	RN450 Δo*atA::oatA*(reverse)	This study
GTM939	RN450 Δo*atA::oatA* pCN51*ø*	This study
GTM940	RN450 Δo*atA::oatA* pCN51_*ptiA-M*	This study
GTM941	SH1000 σB pCN41_*isaA* pCUP$_{Cad}$*ø*	This study
GTM942	SH1000 σB pCN41_*isaA* pCUP$_{Cad}$ ptiA-M	This study
GTM943	SH1000 σB pCN41_*isaA-ssaA* pCUP$_{Cad}$*ø*	This study
GTM944	SH1000 σB pCN41_*isaA-ssaA* pCUP$_{Cad}$ ptiA-M	This study

Appendix 2: List of plasmids used in this study

2.1 List of plasmids used in the section "Mobilisation of SaPIs by endogenous phages in clinical strains

Plasmid	Description	Reference
pCN51	expression vector	(Charpentier *et al*., 2004)
pCN41	expression vector	(Charpentier *et al*., 2004)
pCU1	expression vector	
pET28a	expression vector	Novagen
pGTM1	pMAD *entC::cat*	This study
pGTM2	pCN41 *stlstr*$_{SaPImw2}$ (regulatory region SaPImw2)	This study
pGTM3	pCU1 cadmium- inducible promoter P_{cad}	This study
pGTM4	pGTM3 MW_RS07680 (MW1434)	This study
pGTM5	pGTM3 region from MW1432 to MW1430	This study
pGTM6	pGTM3 region from MW1429 to MW1428	This study
pGTM7	pGTM3 MW1426	This study
pGTM8	pGTM3 region from MW1424 to MW1416	This study
pGTM9	pGTM3 region from MW1414 to MW1410	This study
pGTM10	pGTM3 MW1406	This study
pGTM11	pCU1 MW1424 (DUF3113)	This study
pGTM12	pGTM3 MW1423	This study
pGTM13	pGTM3 MW1422	This study
pGTM14	pGTM3 MW1421	This study
pGTM15	pGTM3 MW1420	This study
pGTM16	pGTM3 MW1419	This study
pGTM17	pGTM3 MW1418	This study
pGTM18	pGTM3 MW1417	This study
pGTM19	pGTM3 MW1416	This study
pGTM20	pCN51 MW1424 (DUF3113)	This study
pGTM21	pMAD ΔDUF3113 ϕSa2mw	This study
pGTM22	pET28a-Stl$_{SaPImw2}$ (C-terminal His6 tag)	This study
pGTM23	pET28a DUF3113$_{\phi Sa2mw}$ (C-terminal His6 tag)	This study
pGTM24	pET28a DUF3113$_{\phi Sa2mw}$	This study
pGTM25	pET28a DUF3113$_{\phi Sa2c}$	This study
pGTM26	pET28a-Stl$_{SaPIbov1}$ (C-terminal His6 tag)	This study
pGTM27	pCN51_ DUF3113$_{\phi Sa2mw}$ _3xflag	This study
pGTM28	pCN51_DUF3113$_{80\alpha}$_3xflag	This study
pGTM29	pCN51_DUF3113$_{\phi 11}$_3xflag	This study
pGTM30	pCN51_DUF3113$_{\phi 2c}$_3xflag	This study
pGTM31	pCN51_DUF3113$_{\phi 6c}$_3xflag	This study
pGTM32	pGTM3_DUF3113$_{\phi Sa2mw}$_3xflag	This study
pGTM33	pGTM3_DUF3113$_{80\alpha}$_3xflag	This study
pGTM34	pGTM3_DUF3113$_{\phi 11}$_3xflag	This study

pGTM35	pGTM3_ DUF3113$_{\phi 2c}$_3xflag	This study
pGTM36	pGTM3_ DUF3113$_{\phi 6c}$_3xflag	This study
pGTM37	pCN51_RBS_ DUF3113$_{80\alpha}$_3xflag	This study
pGTM38	pCN51_RBS_ DUF3113$_{\phi 11}$_3xflag	This study
pGTM39	pCN51_RBS_ DUF3113$_{\phi 2c}$_3xflag	This study
pGTM40	pCN51_RBS_ DUF3113$_{\phi 6c}$_3xflag	This study
pGTM41	pGTM3_RBS_ DUF3113$_{80\alpha}$_3xflag	This study
pGTM42	pGTM3_RBS_ DUF3113$_{\phi 11}$_3xflag	This study
pGTM43	pGTM3_RBS_ DUF3113$_{\phi 2c}$_3xflag	This study
pGTM44	pGTM3_RBS_ DUF3113$_{\phi 6c}$_3xflag	This study

2.2 List of plasmids used in the section "SaPIs control of chromosomally-encoded virulence gene expression"

Plasmid	Description	References
	pCN51ø	(Charpentier *et al.*, 2004)
pGTM50	pCN51_*ptiA-M*	This study
pGTM51	pCN51_*ptiA-M*$_{SaPI2}$	This study
pGTM52	pCN51_*ptiA-M*$_{SaPIeq}$	This study
pGTM53	pCN51_*ptiA-M*$_{SaPIpt1028}$	This study
pGTM54	pCN51_*ptiA*K2*-*M*	This study
pGTM55	pCN51_*ptiA*R64*-*M*	This study
pGTM56	pCN51_*ptiA-M*K2*	This study
pGTM57	pCN51_*ptiA-M*K88*	This study
pGTM58	pCN41_*isaA*	This study
pGTM59	pCN41_*isaA-ssaA*	This study
pGTM60	pCU1P$_{cad}$ø	This study
pGTM61	pCUP$_{cad}$ *ptiA-M*	This study

Appendix 3: List of primers used in this study

3.1 List of primers used in the section "Mobilisation of SaPIs by endogenous phages in clinical strains

Plasmid/ construction	oligonucleotides	Sequence
Oligonucleotides for mutant constructions		
pMAD *entC::cat* (Introduction of chloramphenicol cassette) pGTM1	SaPImw2_1mB	CGCGGATCCGTTTCAGCAGCTTTTACAACG
	SaPImw2_2c	CTTTCTTATCTTGATAATAAGGAGAAACAGAGGATTTCTAAGC
	SaPImw2_3m	CCCCGTTAGTTGAAGAAGGGGATAATGTTAATCCGATTTTG
	SaPImw2_4cE	CCGGAATTCAGTCTTATCTAACGGCGATG
	cat194_1m	CTTATTATCAAGATAAGAAAG
	cat194_2c	CCTTCTTCAACTAACGGGGG
pMAD ΔDUF3113 mutation in DUF3113 gene pGTM21	duf3113_mw2_5mB	CGCGGATCCAGATGTGTTCAACACACACGG
	duf3113_mw2_6c	CATTCAATCACACTCCTAATCC
	duf3113_mw2_7m	GGATTAGGAGTGTGATTGAATGGGTGAACTATTAAAATATAACG
	phiI_mw2_21cE	CCGGAATTCGTTGCCATCGTTACCCCTC
Oligonucleotides for phage derepressor identification		
pGTM2	SaPImw2_5mS	ACGCGTCGACTAATTCGTCCATCATTTCGTTG
	SaPImw2_6cB	CGCGGATCCATGCCGTACACACTAAAAGC
pGTM4	phiI_mw2_1mB	CGCGGATCCTTACAGAGTTTCCAAAATGTC
	phiI_mw2_2cE	CCGGAATTCTGTTGTTTACTTTGAAAATGAG
pGTM5	phiI_mw2_3mB	CGCGGATCCTCATCAGATATCAAGGGCATG
	phiI_mw2_4cE	CCGGAATTCTTGCACTGTTACTTGCTCC
pGTM6	phiI_mw2_5mB	CGCGGATCCAAGTAACAGTGCAAGATGAG
	phiI_mw2_6cE	CCGGAATTCTGCATGAGCTCTATTTGAGTG
pGTM7	phiI_mw2_7mB	CGCGGATCCTTAACACTTGCTACCGAGTCG
	phiI_mw2_8cE	CCGGAATTCAATTAAAGTGGGGCTAAAACC
pGTM8	phiI_mw2_9mS	ACGCGTCGACGATTGAATTTGAATAGTGACGG
	phiI_mw2_10cB	CGCGGATCCAAGTTCTTTTAATTGATCTACTG
pGTM9	phiI_mw2_11mS	ACGCGTCGACGACGCAGGAAAAGGATACGTG
	phiI_mw2_12cB	CGCGGATCCAATCCTGGAGAGGGCTTTAGC
pGTM10	phiI_mw2_13mB	CGCGGATCCTAGTGTTAGGGAGTAAAGAGG
	phiI_mw2_14cB	CCGGAATTCGATAGCTATGTGGTTTGAAGTC
pGTM11/pGTM20	phiI_mw2_9mS	ACGCGTCGACGATTGAATTTGAATAGTGACGG
	phiI_mw2_17cB	CGCGGATCCTCTTCTAGCCATCATTCCACC

pGTM12	phiI_mw2_18mB	CGCGGATCCAGGTGAACTATTAAAATATAAC
	phiI_mw2_19cE	CCGGAATTCATGCTCATTATGCTTCACTCC
pGTM13	phiI_mw2_20mB	CGCGGATCCTCCGTACTGGTTCGATGTCAC
	phiI_mw2_21cE	CCGGAATTCGTTGCCATCGTTACCCCTC
pGTM14	phiI_mw2_22mB	CGCGGATCCATGAGGATTTAGCAAAGGCG
	phiI_mw2_23cE	CCGGAATTCCTTTTAACAATTAGGCAGTCC
pGTM15	phiI_mw2_24mB	CGCGGATCCAGACAAGCATAAGGAATGAAG
	phiI_mw2_25cE	CCGGAATTCCTCTAAAATAAAGTTAGTTGC
pGTM16	phiI_mw2_26mB	CGCGGATCCAGCAACATGGTTTGGAATACG
	phiI_mw2_27cE	CCGGAATTCACTAATACTCATTTTCCTGC
pGTM17	phiI_mw2_28mB	CGCGGATCCAGGTATTGAACTTGATGAAGC
	phiI_mw2_29cE	CCGGAATTCTTCCATTCCACTCACTCGTCC
pGTM18	phiI_mw2_30mB	CGCGGATCCTGTCTGATGATTCAGTTATTAG
	phiI_mw2_31cE	CCGGAATTCAAGCATTTACTCGTCCTCC
pGTM19	phiI_mw2_32mS	ACGCGTCGACTTGCTAACGAGCATGAGATTG
	phiI_mw2_10cB	CGCGGATCCAAGTTCTTTTAATTGATCTACTG
Oligonucleotides for MLST genotyping		
MLST	arc_1m	TTGATTCACCAGCGCGTATTGTC
	acr_2c	AGGTATCTGCTTCAATCAGCG
	tpi_1m	TCGTTCATTCTGAACGTCGTGAA
	tpi_1m	TTTGCACCTTCTAACAATTGTAC
	glp_1m	CTAGGAACTGCAATCTTAATC
	glp_2c	TGGTAAAATCGCATGTCCAATTC
	gmk_1m	ATCGTTTTATCGGGACCATC
	gmk_2c	TCATTAACTACAACGTAATCGTA
	pta_1m	GTTAAAATCGTATTACCTGAA
	pta_2c	GACCCTTTTGTTGAAAAG CTT
	aro_1m	ATCGGAAATCCTATTTCACATTC
	aro_2c	GGTGTTGTATTAATAACGATATC
	yqi_1m	CAGCATACAGGACACCTATTGGC
	yqi_2c	CGTTGAGGAATCGATACTGGAAC
Oligonucleotides for PICIs and bacteriophages identification		
SaPIs *int*	IntI_F	GATAGCAAAGTACTCATCATCAG
	IntI_R	CATTGTCTGTCTTGATTG
	IntII_F	GACTGCCTATATGTTTG
	IntII_R	CGTTGATATGTGCTTTCTC
	IntIII_F	CTCGTAAGTACATAACAC
	IntIII_R	CTTCATTGCATGATTCAG

	IntIV_F	AGGATTCAAAATGGAATGATGG
	IntIV_R	CCTTATCCATCTGTTCAGTAACG
	IntV_F	GTCTGCTATAGACTTTGAG
	IntV_R	CAATGAGTATGACGAAATG
SePIs *int*	SePI_intI_1m	ACAAATGGGGACGTGTTAGC
	SePI_intI_2c	AGCAAGTAGCTCGCCGATAC
	SePI_intII_1m	GCTGCAAACAGAAGTACAAGC
	SePI_intII_2c	CCATGTTAAAGCCATTAATTCGCC
	SePI_intIII_1m	GTACCTAAAGAAAATGCACGTG
	SePI_intIII_2c	TGTCACTTCACCGTGATAGTC
	SePI_intIV_1m	ATCAAGCATGTGACGAATGG
	SePI_intIV_2c	TGCGTTACCTTCATTACGCC
	SePI_intV_1m	AATGGTCTGATATAGACTTTGAG
	SePI_intV_2c	GTGGTATTTATATCTTTGTGGCG
Phage *int*	Sa1int_1m	AAGCTAAGTTCGGGCACA
	Sa1int_2c	GTAATGTTTGGGAGCCAT
	Sa2int_1m	TCAAGTAACCCGTCAACTC
	Sa2int_2c	ATGTCTAAATGTGTGCGTG
	Sa3int_1m	TTATTGACTCTACAGGCTGA
	Sa3int_2c	GAAAAACAAACGGTGCTAT
	Sa4int_1m	ATTGATATTAACGGAACTC
	Sa4int_2c	TAAACTTATATGCGTGTGT
	Sa5int_1m	AAAGATGCCAAACTAGCTG
	Sa5int_2c	CTTGTGGTTTTGTTCTGG
	Sa6int_1m	GCCATCAATTCAAGGATAG
	Sa6int_2c	TCTGCAGCTGAGGACAAT
	Sa7int_1m	GTCCGGTAGCTAGAGGTC
	Sa7int_2c	GGCGTATGCTTGACTGTGT
Oligonucleotides for SaPI excision and circularization		
SaPI*int*I excision	SaPI_Int1exc_1m	GCAGGTGGACCAAGAAGAGG
	SaPI_Int1exc_2c	CTTTGTGCAAATCCCAAACAC
SaPI*int*I circularization	SaPI_Int1circ_1c	ATCTACACTTGCGCAATTAGG
	SaPI_Int1circ_2m	AGGCGTAATTGTGAAACTAGG
SaPI*int*III excision	SaPI_Int3exc_1m	TGCTGTTAACGCGATTCAAC

	SaPI_Int3exc_2c	CTTTTTTGACACTTATTTGACAC
SaPI*int*III circularization	SaPI_Int3circ_1c	ATGACACCAGCTTTTTGGGT
	SaPI_Int3circ_2m	GAGGTGATGTTGGGTGCAC
SaPI*int*IV excision	SaPI_Int4exc_1m	TCGAACTTTGCAATCGAACA
	SaPI_Int4exc_2c	CTTGAGGGAGTGGGACTGAA
SaPI*int*IV circularization	SaPI_Int4circ_1c	ATGGATCGTGTAGGCCATTC
	SaPI_Int4circ_2m	GTGTGGTTGTTTTCAAGGGC
SaPI*int*V excision	SaPI_Int5exc_1m	GCTGCTACAAACGAGTGGGT
	SaPI_Int5exc_2c	ATACCATAACTCCCGCACCA
SaPI*int*V circularization	SaPI_Int5circ_1c	TACAAGAACGCTTGCGACAC
	SaPI_Int5circ_2m	TATGAGGTGATGTTGGGTGC
Oligonucleotides for Stl-DUF3113 interaction		
pET28a-Stl$_{SaPImw2}$ (C-terminal His6 tag) pGTM22	Stl_SaPImw2_1B	CGCGGATCCATGATAATTTTTCGATTAAAAGAAAT
	Stl_SaPImw2_2cE	CCGGAATTCGGCAATCACCTATTTCATTTC
pET28a_ DUF3113$_{\phi Sa2mw}$ pGTM24	duf3113_mw2_3mS	ACGCGTCGACTATAATAATAATTTGTTTAACTTTAA GAAGGAGATATACCATGCAACATCAAGCTTATATC
	duf3113_4cN	ATAAGAATGCGGCCGCTCTTCTAGCCATCATTCCACC
pET28a_ DUF3113 $_{\phi Sa2mw}$ (C-terminal His6 tag) pGTM23	duf3113_8mB	CGCGGATCCATGCAACATCAAGCTTATATC
	duf3113_9cS	ACGCGTCGACTTATAACTTTTCTTCTAGCC
pET28_ DUF3113$_{\phi Sa2c}$ pGTM25	duf3113_phi11_1mS	ACGCGTCGACTATAATAATAATTTGTTTAACTTTAA GAAGGAGATATACCATGCAACAACAAGCATATATAA ACG
	duf3113_phi11_2cN	ATAAGAATGCGGCCGCGTTCAACTTCATTTATTCCA C
pET28_ Stl$_{SaPIbov1}$ (C-terminal His6 tag) pGTM26	Stl_SB1_1mB	CGCGGATCCATCGAAGGAGCTGGTCAAATG
	Stl_SB1_2cE	CCGGAATTCTACCTTGTGGCGTGATGAT
	phiII_mw2_9mS	ACGCGTCGACGATTGAATTTGAATAGTGACGG

pCN51/pGTM3_duf3113 $_{\phi Sa2mw}$ _3xflagC pGTM27 pGTM32	phiII_mw2_3xflagC_1cB	CGCGGATCCTTATTTATCATCATCATCTTTATAATC GATATCGTGATCTTTATAATCATCGTGATCTTTATA ATTTCCACCTCTAAATCTAAAACC
pCN51/pGTM3_ duf3113φ11 3xflagC pGTM28 pGTM33	phiV_FQ166_37mS	AGCCGTCGACAAGAATGGATGGTATATGAGG
	phi11_3xflagC_1cB	CGCGGATCCTTATTTATCATCATCATCTTTATAATC GATATCGTGATCTTATAATCGCCATGCTGATCTTTA TAATCTTCCACCTCTATATATGCATG
pCN51/pGTM3_ duf311380α 3xflagC pGTM29 pGTM34	duf3113_phi80α_1mS	ACGCGTCGACTAGCTCAAAGCGTTATGG
	phi80α_3xflagC_1cB	CGCGGATCCTTATTTATCATCATCATCATCTTTATA ATCGATATCGTGATCTTTATAATCGCCATCGTGATC TTTATAATCTTCCACCTCTACATTTACATTT
pCN51/pGTM3_ duf3113φ2.C 3xflagC pGTM30 pGTM35	phiII_mw2_9mS	ACGCGTCGACGATTGAATTTGAATAGTGACGG
	phiII_mw2_3xflagC_1cB	CGCGGATCCTTATTTATCATCATCATCTTTATAATC GATATCGTGATCTTTATAATCATCGTGATCTTTATA ATTTCCACCTCTAAATCTAAAACC
pCN51/pGTM3_ duf3113φ6.C 3xflagC pGTM31 pGTM36	duf3113_phi80α_1mS	ACGCGTCGACTAGCTCAAAGCGTTATGG
	duf3113_3xflagC_phiVI_1cB	CGCGGATCCTTATTTATCATCATCATCTTTATAATC GATATCGTGATCTTTATAATCGCCATCGTGATCTTT ATAATTTTTCACCTCTACATTTACGTTT
pCN51/pGTM3_ duf3113φ11 3xflagC pGTM38 pGTM42	duf3113_phi11_5m	GAAGGATTAGGAGTGTGATTGAATGCAACAACAAGC ATATATAAACG
	phi11_3xflagC_1cB	CGCGGATCCTTATTTATCATCATCATCTTTATAATC GATATCGTGATCTTATAATCGCCATGCTGATCTTTA TAATCTTCCACCTCTATATATGCATG
	duf3113_mw2_13mS	ACGCGTCGACGATTGAATTTGAATAGTGACGGATTT ACTTCTCCGTTTTATATGAAGGATTAGGAGTGTGAT TGAATGC
	phi11_3xflagC_1cB	CGCGGATCCTTATTTATCATCATCATCTTTATAATC GATATCGTGATCTTATAATCGCCATGCTGATCTTTA TAATCTTCCACCTCTATATATGCATG
pCN51/pGTM3_ duf311380α 3xflagC pGTM37 pGTM41	duf3113_80α_3m	GAAGGATTAGGAGTGTGATTGAATGCAACAGCAAGC ATATATAAACG
	phi80α_3xflagC_1cB	CGCGGATCCTTATTTATCATCATCATCTTTATAATC GATATCGTGATCTTTATAATCCCCATCGTGGATCTT TATAATCTTCCACCTCTACATTTACATTT
	duf3113_mw2_13mS	ACGCGTCGACGATTGAATTTGAATAGTGACGGATTT ACTTCTCCGTTTTATATGAAGGATTAGGAGTGTGAT TGAATGC
	phi80α_3xflagC_1cB	CGCGGATCCTTATTTATCATCATCATCTTTATAATC GATATCGTGATCTTTATAATCCCCATCGTGGATCTT TATAATCTTCCACCTCTACATTTACATTT
pCN51/pGTM3_duf3113φ2.C 3xflagC pGTM39 pGTM43	duf3113_phiII_cepaC_2m	GAAGGATTAGGAGTGTGATTGAATGCAACATCAAGC TTATATC
	phiII_mw2_3xflagC_1cB	CGCGGATCCTTATTTATCATCATCATCTTTATAATC GATATCGTGATCTTTATAATCATCGTGATCTTTATA ATTTCCACCTCTAAATCTAAAACC

	duf3113_mw2_13mS	ACGCGTCGACGATTGAATTTGAATAGTGACGGATTTACTTCTCCGTTTTATATGAAGGATTAGGAGTGTGATTGAATGC
	phiII_mw2_3xflagC_1cB	CGCGGATCCTTATTTATCATCATCATCTTTATAATCGATATCGTGATCTTTATAATCATCGTGATCTTTATAATTTCCACCTCTAAATCTAAAACC
pCN51/pGTM3_duf3113φ6.C 3xflagC pGTM40 pGTM44	duf3113_phi11_5m	GAAGGATTAGGAGTGTGATTGAATGCAACAACAAGCATATATAAACG
	duf3113_3xflagC_phiVI_1cB	CGCGGATCCTTATTTATCATCATCATCTTTATAATCGATATCGTGATCTTTATAATCGCCATCGTGATCTTTATAATTTTTCACCTCTACATTTACGTTT
	duf3113_mw2_13mS	ACGCGTCGACGATTGAATTTGAATAGTGACGGATTTACTTCTCCGTTTTATATGAAGGATTAGGAGTGTGATTGAATGC
	duf3113_3xflagC_phiVI_1cB	CGCGGATCCTTATTTATCATCATCATCTTTATAATCGATATCGTGATCTTTATAATCGCCATCGTGATCTTTATAATTTTTCACCTCTACATTTACGTTT

3.2 List of primers used in the section "SaPIs control of chromosomally-encoded virulence gene expression"

Plasmid/ Contruction	Oligonucleotides	Sequence
Oligonucleotides for mutant constructions		
Δ*PcrtO*	STX_87mB	CGCGGATCCGTATACAATTGATGCACAAGG
	STX_88c	AATAATGGCATAGTAAGCTCAAAAACGCCAATACTCTCATG
	STX_89m	GTAATATCGGAGGCTTTATTACTATGTCAGTATGATTGAGAG
	STX_90cS	ACGCGTCGACAATCAAAGCCAAAGCCATCTG
Δ*PgnaT*	STX_111mB	CGCGGATCCCCACTATTAATTGGTGACTCG
	STX_103c	TGTTCCTAATAAAGCTTCATTG
	STX_104m	CAATGAAGCTTTATTAGGAACAGTCGCAAAACTATAAAAGTTATAATTG
	STX_105cS	ACGCGTCGACCACTGTATAGCCAGATGATGC
Δ*PisaA*	STX_98mB	CGCGGATCCGTTACCAATATTAACGACAGC
	STX_99c	GTCTTTTATTTCTTTCAACGGACAATATCAAGTGTTATTTGTAATG
	STX_100m	CCGTTGAAAGAAATAAAAGAC
	STX_109cS	ACGCGTCGACTTGACCATTTGATTCACGAGC
Δ*isaA*	STX_98mB	CGCGGATCCGTTACCAATATTAACGACAGC
	STX_126c	TAATGTGAAATATGTGCCCTTTAG
	STX_127m	AAGGGCACATATTTCACATTAGTGGATATTGAAGTTAGGGAA
	STX_128cE	CCGGAATTCATATTTAAACGCTATGTGTCAC

isaA L10*	STX_98mB	CGCGGATCCGTTACCAATATTAACGACAGC
	STX_139c	CTTCtGCAGCGTGTGCTTaATGTCCTGTACCTGCTGCG
	STX_122cE	CCGGAATTCAATTAATAAAAAAACTCTCCCC
	STX_138m	CATtAAGCACACGCTGCaGAAGTAAACGTTGATCAAGC
Δ*oatA*	STX_148mB	CGCGGATCCTCTTTATTAAGTGCCAAGTGG
	STX_159c	ACCAAACGATATATCTCGTCAG
	STX_160m	CTGACGAGATATATCGTTTGGTCAGAAAAAGTGTTGTAAACATTC
	STX_153cE	CCGGAATTCGTTACTTCCTGATGCATAACC
Δ*gnat*	STX_154mB	CGCGGATCCGATGGTGCATATGACATCCAC
	STX_119c	CAATATCCACTTTTTAAAAAATG
	STX_155m	CATTTTTTAAAAAGTGGATATTGTACTGAGTACACTAGATAC
	STX_128cE	CCGGAATTCATATTTAAACGCTATGTGTCAC
Δ*oatA*	STX_148mB	CGCGGATCCTCTTTATTAAGTGCCAAGTGG
	STX_156c	ACCAAACGATATATCTCGTCAG
	STX_157c	CTGACGAGATATATCGTTTGGTCTACAAAATAGAGAAAAATAAATG
	STX_158cE	CCGGAATTCAATGCTAATATACAACCAAGC
oatA reverse	STX_148mB	CGCGGATCCTCTTTATTAAGTGCCAAGTGG
	STX_149c	TCAAATCCCATAGTAACGGTGACCAAACGATATATCTCGTCAG
	STX_150m	CTGTTACATTAGAGGATAGGCTGC
	STX_151c	GGTCACCGTTACTATGGGATTTGA
	STX_152m	GCAGCCTATCCTCTAATGTAACAGAAAAAGTGTTGTAAACATTC
	STX_153cE	CCGGAATTCGTTACTTCCTGATGCATAACC
Δ*crtOPQMN*	STX_1mB	CATCTAAATTGAATCACTCTC
	STX_2c	CATCTAAATTGAATCACTCTC
	STX_12m	CCTATTGTCTTAACGAGTGCG
	STX_13cE	GATAACACATGTTGGTACTGC
oligonucleotides for intergenic PCRs		
region between *isaA-gnat*	STX_140	CGTACTGGTGTTTCAGCT
	STX_141	GATGTTCTCCAAATCAGGTTG
region between *gnat-oatA*	STX_142m	AAGCGAGGAGATTTACCAAGT
	STX_143c	ATTCCTATAACTGCGAATGCC
region between *oatA-ssaA*	STX_144m	CAGGTAGTAAAGCGCTGACTG
	STX_145c	TAGTGGCAAGTCCTGCTGTAG
	STX_146m	TACAATTTCAGCAAACCAAGC

region between *ssaA-crtO*	STX_147c	AATGAATCAACTCTGCCCTCT
oligonucleotides for *ptiA-M* constructions		
pCN51_ *ptiA*	SB1_230mB	CGCGGATCCATGGATAAACAGCAAATAAAAGAC
	SB1_223mK	CAAAACGCGCTTGTTTTGC
pCN51_ *ptiM*	SB1_229mB	CGCGGATCCAATAGCAATGATTGCCTATCC
	SB1_228mK	GTTACACCACTTTTACATATG
pCN51_ *ptiA-M*	SB1_223mK	CAAAACGCGCTTGTTTTGC
	SB1_229mB	CGCGGATCCAATAGCAATGATTGCCTATCC
pCN51_ *ptiA-M*$_{SaPI2}$	ORF16SaPI_1mB	CTCACTCATTCTATCACCTCC
	ORF17_SaPI_2cK	CTCACTCATTCTATCACCTCC
pCN51 _*ptiA-M*$_{SaPIeq}$	pt1028_17mB	ATAACGAAGTGTTCCGTAACG
	SaPIeq_orf6_1cK	GTTAACTCACTCATTCTATCACC
pCN51 _*ptiA-M*$_{pt1028}$	pt1028_17mB	ATAACGAAGTGTTCCGTAACG
	pt1028_18cK	GTTAACTCACTCATTCTATAACC
pCN51 _*ptiA*-K2*-*M*	229mB	CGCGGATCCAATAGCAATGATTGCCTATCC
	Orf6_SB12k_2c	CGAAGTCTTTTATTTGCTGTTAATCCATTGTTACACCAC
	Orf6_k2stop_1m	GTGGTGTAACAATGGTTAACAGCAAATAAAAGACTTCG
	223mK	CAAAACGCGCTTGTTTTGC
pCN51_*ptiA-R64*-M*	229mB	CGCGGATCCAATAGCAATGATTGCCTATCC
	Orf6_R64*_2c	CTCTTTATCATTAGCTACTCAGTCAACAATCATTTCTAAGTG
	Orf6_R64*_1m	CACTTAAGAAATGATTGTTGACTGAGTAGCTAAAGAG
	223mK	CAAAACGCGCTTGTTTTGC
pCN51*ptiA-M* K2*	Orf7_stop_1m	GTTTATAGGGGTGAACAAATGTAACTGCTTAAAACGAAGAATTG
	SB1_223mK	CAAAACGCGCTTGTTTTGC
	Orf7_stop_2mB	CGCGGATCCAATAGCAATGATTGCCTATCCAATTCGGGTAGGCTCTGTTTATAGGGGTGA
	SB1_223mK	CAAAACGCGCTTGTTTTGC
pCN51 *ptiA-M* K88*	SB1_229mB	CGCGGATCCAATAGCAATGATTGCCTATCC
	Orf7_stop_4c	ATAGTTTTGTAATCTTGCCTTTG
	Orf7_stop_3m	CAAAGGCAAGATTACAAAACTATTAAAATCTATAAGTAATATG
	SB1_223mK	CAAAACGCGCTTGTTTTGC
pCN41_*isaA*	STX_106m	GGCTTGCACATATCATTTATCG
	STX_122cE	CCGGAATTCAATTAATAAAAAAACTCTCCCC
pCN41_*isaA-ssaA*	STX_106m	GGCTTGCACATATCATTTATCG
	STX_108m	CCGGAATTCTTGGCGCGCCAACAATTAACCAATACATACTAC

oligonucleotides for qPCR		
crtM	crtM_1m	TAGAATCATGATGGCGCTTCAG
	crtM_2c	TACAGCGTTCTTCTGAACCACT
gnat	STX_155m	TGTACTGAGTACACTAGATAC
	STX_110c	CCAGAATATTCCTTCTCGAT

Appendix 4: List of genes differentially expressed resulting from tiling array

List of genes differentially expressed between RN450 SaPIbov1Δ*stl* (4D) and RN450 + SaPIbov1 + SaPI (4S).

Gene Identifier	Description	logFC	adj.P.Val	B
SAOUHSC_02882	conserved hypothetical protein	**-4,3**	7,769E-06	8,42
SAOUHSC_02880	glycosyl transferase, group 2 family protein	**-4,1**	1,452E-07	12,89
SAOUHSC_02881	phytoene dehydrogenase	**-4,1**	1,682E-07	12,67
SAOUHSC_02886	acetyltransferase, gnat family	**-4,0**	2,459E-05	7,12
SAOUHSC_02879	squalene desaturase	**-3,5**	6,555E-05	6,12
SAOUHSC_00192	staphylocoagulase	**-3,2**	1,464E-06	10,17
SAOUHSC_01386	phosphate ABC transporter, permease protein	**-2,6**	1,294E-06	10,40
SAOUHSC_02877	amine oxidase:FAD oxidoreductase	**-2,5**	5,208E-04	3,98
SAOUHSC_01387	phosphate ABC transporter, permease protein	**-2,4**	2,932E-06	9,44
SAOUHSC_01385	phosphate ABCtransporter ATP-binding	**-2,4**	1,462E-06	10,22
SAOUHSC_00347	conserved hypothetical protein	**-2,0**	6,568E-05	6,08
SAOUHSC_02679	nitrate reductase molybdenum cofactor	**-2,0**	9,031E-03	0,83
SAOUHSC_02655	conserved hypothetical protein	**-2,0**	2,060E-03	2,56
SAOUHSC_02685	transcriptional regulator NirR	**-1,8**	6,147E-03	1,27
SAOUHSC_02680	nitrate reductase, beta subunit	**-1,8**	1,334E-02	0,43
SAOUHSC_01384	phosphate transport system regulatory protein	**-1,8**	9,917E-06	8,10
SAOUHSC_00936	transcription factor	**-1,7**	1,482E-02	0,30
SAOUHSC_02144	conserved hypothetical protein	**-1,6**	3,038E-03	2,12
SAOUHSC_02675	regulatory protein, LuxR:Response regulator	**-1,4**	8,240E-03	0,94
SAOUHSC_02851	LrgA family	**-1,2**	7,058E-03	1,11
SAOUHSC_02397	ABC transporter, ATP-binding protein	**-1,1**	3,658E-03	1,86
SAOUHSC_03010	imidazole glycerol phosphate synthase	**-1,1**	4,605E-03	1,61
SAOUHSC_02654	pyridine nucleotide-disulphide oxidoreductase	**-1,0**	2,706E-03	2,27
SAOUHSC_02407	conserved hypothetical protein	**1,2**	1,077E-03	3,23

List of genes differentially expressed between RN450 + SaPIbov1 + SaPI1 (4S-60') and RN450 WT (4-60).

Gene Identifier	Description	logFC	adj.P.Val	B
SAOUHSC_02882	conserved hypothetical protein	**-4,7**	2,19E-06	10,335
SAOUHSC_02880	glycosyl transferase, group 2 family protein	**-4,4**	3,22E-08	14,633
SAOUHSC_02881	phytoene dehydrogenase	**-4,2**	7,96E-08	13,654
SAOUHSC_02879	squalene desaturase	**-3,9**	2,09E-05	8,0449
SAOUHSC_02886	acetyltransferase, gnat family	**-3,5**	0,00048	4,8259
SAOUHSC_02877	amine oxidase:FAD dependent oxidoreductase	**-2,4**	0,00216	3,169
SAOUHSC_01849	acetoin utilization protein AcuC	**1,1**	0,03973	0,2902

List of genes differentially expressed between RN450 + 80α + SaPIbov1 + SaPI1 (1SS-60) and RN450 WT (4-60)

Gene Identifier	Description	logFC	adj.P.Val	B
SAOUHSC_02882	conserved hypothetical protein	**-4,9**	1,10E-06	10,997
SAOUHSC_02880	glycosyl transferase, group 2 family protein	**-4,6**	8,57E-09	15,936
SAOUHSC_02881	phytoene dehydrogenase	**-4,5**	2,55E-08	14,771
SAOUHSC_02879	squalene desaturase	**-4,2**	5,90E-06	9,2589
SAOUHSC_02886	acetyltransferase, gnat family	**-3,8**	0,00013	6,0918
SAOUHSC_02877	amine oxidase:FAD dependent oxidoreductase	**-2,3**	0,00231	2,9978
SAOUHSC_02609	metallothiol transferase FosB	**1,0**	0,02096	0,4951
SAOUHSC_00045	conserved hypothetical protein	**1,1**	0,01576	0,8903
SAOUHSC_01917	lipoprotein, putative	**1,1**	0,02096	0,5249
SAOUHSC_00408	conserved hypothetical protein	**1,3**	0,0234	0,317
SAOUHSC_01030	glutaredoxin	**1,4**	0,02096	0,5138
SAOUHSC_02795	conserved hypothetical protein	**1,6**	0,02455	0,2365

List of genes differentially expressed between RN450 + 80α (1S-60) and RN450 + 80α +SaPIbov1 + SaPI1 (1SS-60)

Gene Identifier	Description	logFC	adj.P.Val	B
SAOUHSC_02882	conserved hypothetical protein	**-4,0**	4,88E-05	7,1929
SAOUHSC_02880	glycosyl transferase, group 2 family protein	**-3,9**	5,14E-07	11,94
SAOUHSC_02881	phytoene dehydrogenase	**-3,8**	8,92E-07	11,298
SAOUHSC_02886	acetyltransferase, gnat family	**-3,8**	0,00013	6,0788
SAOUHSC_02879	squalene desaturase	**-3,6**	7,25E-05	6,7257
SAOUHSC_02877	amine oxidase:FAD dependent oxidoreductase	**-2,5**	0,00082	4,2024
SAOUHSC_02681	nitrate reductase, alpha subunit	**-2,2**	0,00158	3,3342
SAOUHSC_02883	staphylococcal secretory antigen SsaA	**-1,5**	0,02902	0,2772