



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**FACULTAD DE ESTUDIOS SUPERIORES IZTACALA**

**CARRERA DE BIOLOGÍA**



**“ANÁLISIS DE LA HISTORIA EVOLUTIVA DEL ELEMENTO GENÉTICO MÓVIL  
*Staphylococcal Cassette Chromosome*”**

**TESIS PROFESIONAL**

**PARA OBTENER EL TÍTULO DE:**

**BIÓLOGO**

**PRESENTA**

David Vargas Peralta

**DIRECTOR DE TESIS**

Dr. Santiago Castillo Ramírez



**CCG**

Centro de Ciencias Genómicas

Los Reyes, Iztacala, Estado de México

2016



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Esta Tesis fue realizada con mucho amor, esfuerzo, trabajo, etc.... y es dedicada:

A mi madre Ocotlán Adelina

Y a mi madrina Teresa, que ha sido como una segunda madre

A ambas, muchas gracias

## Agradecimientos

A mi tutor, Santiago Castillo Ramírez, muchas gracias por permitirme desarrollar este proyecto en su laboratorio, y por la dedicación y la paciencia que me brindó a lo largo de dicho proyecto.

A mi madre Ocotlán Adelina, gracias por sacarme adelante y porque nunca te rendiste; así mismo te doy las gracias por tu amor, comprensión y apoyo incondicional que me has brindado todo este tiempo.

A mi madrina Teresa Sánchez, eres como una segunda madre para mí, gracias por tu amor y apoyo incondicional que también me has brindado durante este tiempo.

A mi tío Gaudencio Peralta, que lamentablemente ya no se encuentra físicamente con nosotros, pero siempre seguirá vivo en nuestros recuerdos, muchas gracias por todo.

A toda mi familia muchas gracias, en especial a mi tía Ninfa Peralta, por el amor, las pláticas, los consejos y soporte que me ha dado durante todos estos años.

A Claudia Pineda por estos 4 años que hemos compartido, donde nos hemos visto crecer como personas y como profesionistas, te agradezco por contar siempre con tu apoyo en las buenas y en las malas.

Muchas gracias a todos mis amigos, en especial a Francisco Pedraza que sin conocerme lo suficiente creyó en mí y por esa razón tuve la oportunidad de acercarme a los análisis *in silico* y tener una nueva visión de la ciencia.

# ÍNDICE

<b>RESUMEN</b> .....	1
<b>INTRODUCCIÓN</b> .....	2
<i>Staphylococcus aureus</i> y su importancia médica .....	4
El elemento <i>SCCmec</i> y sus componentes .....	5
Clasificación del <i>SCCmec</i> .....	8
<i>Staphylococcus aureus</i> resistente a meticilina ( <i>MRSA</i> ).....	9
<b>ANTECEDENTES</b> .....	9
<b>OBJETIVOS</b> .....	12
Objetivo general .....	12
Objetivo particular .....	12
<b>MÉTODOS</b> .....	13
Homólogos, alineamientos y filogenias de máxima verosimilitud.....	13
Análisis de recombinación y selección.....	13
Diagramas eBURST.....	14
<b>RESULTADOS</b> .....	15
El <i>SCCmec</i> no se encuentra presente en todas las cepas .....	15
Patrones de descendencia evolutiva (vertical) de <i>S. aureus</i> a partir del <i>MLST</i> .....	17
Los genes del <i>SCCmec</i> presentaron diferentes historias evolutivas.....	20

Los genes del <i>SCCmec</i> muestran señales de recombinación y diferencias en variación genética .....	23
El papel de la selección en los genes <i>ccrA</i> , <i>ccrB</i> y <i>mecA</i> .....	28
<b>DISCUSIÓN</b> .....	30
Las relaciones evolutivas verticales y las limitantes de eBURST .....	30
Los patrones de TGH y el papel de la selección .....	33
<b>CONCLUSIONES</b> .....	38
<b>PERSPECTIVA</b> .....	38
<b>REFERENCIAS</b> .....	39
<b>GLOSARIO</b> .....	50
Glosario de abreviaciones .....	50
Glosario de términos .....	51
<b>ANEXOS</b> .....	54
Anexo 1. Alineamientos con TRANALIGN .....	54
Anexo 2. Evaluación de la selección natural y la deriva génica .....	54
Modelos de sitio y el Test de máxima verosimilitud (LRT) para su comparación ...	54
Anexo 3. Breve definición de los diagramas eBURST .....	58
Diagramas eBURST de <i>S. aureus</i> .....	58
Anexo 4. Análisis de recombinación con RDP .....	64

## RESUMEN

*Staphylococcus aureus* es un patógeno oportunista y es la principal causa de infecciones nosocomiales. Esto se debe a factores de virulencia y/o resistencia a múltiples agentes antimicrobianos que son codificados por elementos genéticos móviles. Dentro de estos elementos, el *Staphylococcal Cassette Chromosome (SCC)* es uno de los más característicos de *S. aureus* ya que frecuentemente presenta el gen *mecA*; el cual, codifica para la resistencia a varios antibióticos. Con el fin de entender mejor la historia evolutiva de este elemento, se buscaron los homólogos de tres genes de este elemento en 49 genomas; se construyeron sus filogenias, y se realizaron análisis de recombinación. Por último se evaluó el papel de la selección y la deriva génica, calculando el cociente de sustituciones sinónimas (dS) y no sinónimas (dN).

Nuestros resultados sugieren que este elemento presenta una historia evolutiva compleja; con frecuentes eventos de transferencia horizontal y pérdida del elemento completo, o partes de él. También se encontraron diferencias entre los niveles de variación genética, para los tres genes, donde los genes *ccrA* y *ccrB* presentaron diversidades nucleotídicas altas. Finalmente los tres genes parecen estar bajo un proceso de selección purificadora ya que presentaron mayores mutaciones sinónimas que mutaciones no sinónimas.

## INTRODUCCIÓN

Los elementos genéticos móviles (*MGEs*, por sus siglas en inglés) fueron descritos por primera vez en el genoma del maíz a finales de 1940 (McClintock, 1950; McClintock, 1951) y son un medio importante para la transferencia de información genética. El traspaso de *MGEs* entre las células se conoce como transferencia genética horizontal (TGH) y ésta se produce entre procariontas, entre procariontas y eucariotas, e incluso entre eucariotas (Jain *et al.*, 1999; Keeling y Palmer, 2008). Para la TGH en bacterias se han descrito al menos 3 mecanismos: (1) la transformación (captación de genes desde el entorno), (2) transducción bacteriófaga, y (3) la conjugación (el contacto directo entre las células bacterianas) (Malachowa y DeLeo, 2010) (Ver figura 1).

Los *MGEs* son fragmentos de ADN extracromosómicos, que en bacterias, codifican una variedad de factores de virulencia y/o resistencia, así como enzimas que median su propia transferencia e integración en el ADN del hospedero (Frost *et al.*, 2005). En bacterias, los *MGEs* pertenecen al genoma accesorio; el genoma bacteriano se compone de un genoma principal, presente en todas las cepas de una especie, y un genoma accesorio que sólo se encuentra en algunas cepas de una especie bacteriana. El genoma principal contiene todos los genes vitales para la supervivencia celular (metabolismo, síntesis de ADN y ARN, y la replicación). Mientras que los genes accesorios pueden codificar proteínas que favorezcan la adaptación de las bacterias a diferentes nichos ecológicos (resistencias, factores de virulencia, etc.). Estos genes accesorios se obtienen a partir de otras especies bacterianas, (Hacker y Kaper, 2000; Lindsay y Holden, 2004) por medio de la TGH, que es una de las principales fuerzas que generan un “salto cuántico evolutivo” en procariontes (Hacker y Kaper, 2000). La presencia de los *MGEs*

juega un papel crucial en la plasticidad de las bacterias, ya que permite que se adapten a nuevos entornos. La presión selectiva del ambiente impulsa el enriquecimiento de genes específicos que promueven la aptitud y la supervivencia. Un ejemplo de presión selectiva, es el uso inadecuado de antibióticos en ambientes hospitalarios, esto ocasiona que bacterias como *Staphylococcus aureus*, se vean obligadas a adquirir *MGEs* con genes que ofrecen resistencia a múltiples agentes antimicrobianos o factores de virulencia (Malachowa y DeLeo, 2010).

### Mecanismos de adquisición de *MGEs*

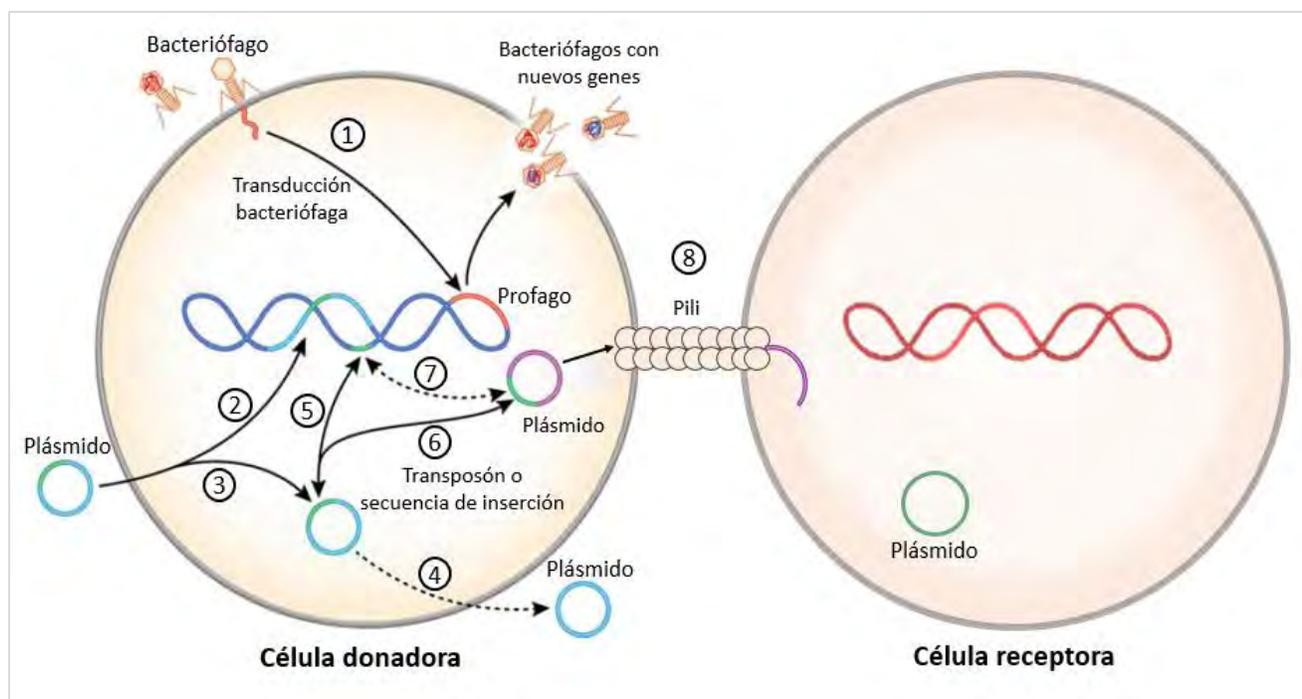


Fig. 1 | **Mecanismos de adquisición de *MGEs*** | 1 Transducción bacteriófaga. 2 Incorporación de plásmidos en el ADN genómico. 3 Plásmido libre. 4 Plásmido suicida. 5 Transferencia de un transposón o una secuencia de inserción entre el plásmido y ADN genómico. 6 Transferencia entre plásmidos. 7 Transferencia entre el ADN genómico y un plásmido. 8 Pili sexual (Modificado de Malachowa y DeLeo, 2010).

## ***Staphylococcus aureus* y su importancia médica**

*Staphylococcus aureus* es un microorganismo de gran importancia médica, ya que desde hace muchos años se le ha reconocido como uno de los principales agentes patógenos para el humano. Esta especie forma parte del género *Staphylococcus*, –el cual consta de bacterias Gram positivas –, que alberga más de 30 especies diferentes. *S. aureus* se puede encontrar habitando la piel y las membranas mucosas del ser humano como un organismo comensal (Howe *et al.*, 1996; Velázquez-Meza, 2005; Kanafani y Fowler, 2006); sin embargo, también es considerado un patógeno oportunista y es responsable de una amplia gama de enfermedades. Por ejemplo, puede causar intoxicación alimentaria, shock tóxico, septicemia, infecciones en el sistema nervioso central, osteomielitis, endocarditis, neumonía y es la principal causa de infecciones nosocomiales (Bustos-Martínez *et al.*, 2006). Entre otros factores, esto es debido a su capacidad para expresar una extensa gama de genes que codifican factores de virulencia y/o resistencia a múltiples agentes antimicrobianos; dichos genes a menudo se localizan en *MGEs*. *S. aureus* contiene diversos tipos de *MGEs* tales como: secuencias de inserción, transposones, fagos, plásmidos, islas de patogenicidad, y el *Staphylococcal Cassette Chromosome* (*SCC*) (Malachowa y DeLeo, 2010), éste último incluye al *SCCmec*; que alberga los genes –denominados *mec* –que codifican las proteínas que confieren la resistencia a la meticilina y a muchos antibióticos  $\beta$ -lactámicos (Ito *et al.*, 2003). Este fenotipo de resistencia a antibióticos –es decir; cepas resistentes a meticilina, que en inglés es denominado *Methicillin-resistant Staphylococcus aureus* (*MRSA*) –es uno de los mayores problemas para tratar infecciones causadas por *S. aureus*. Actualmente, se conocen dos tipos de cepas *MRSA* asociadas a seres humanos el *MRSA* asociado a

hospitales (*HA-MRSA*, del inglés *Hospital-Associated MRSA*), siendo este el fenotipo típico y el *MRSA* asociado a comunidad (*CA-MRSA*, del inglés *Community-Associated MRSA*) (Lindsay, 2010).

### **El elemento *SCCmec* y sus componentes**

El elemento *SCCmec* puede llegar a medir entre 21 - 67 Kpb (Hanssen y Sollid, 2006), este elemento, se inserta en el cromosoma de *S. aureus* en un sitio específico –attB o la secuencia del sitio de integración ISS –dentro del extremo 3' del gen *orfX*, que codifica una metiltransferasa ribosomal (Boundy *et al.*, 2013). El *SCCmec* tiene tres componentes génicos: los complejos de genes *mec* (genes que confieren la resistencia a antibióticos), los genes *ccr* (que codifican para recombinasas sitio-especificas) y una región conocida como J (*Junkyard*) (Chambers y DeLeo, 2009; Shore y Coleman, 2013).

Los genes del complejo *mec* incluyen el gen *mecA* y sus genes reguladores *mecR1-mecI-mecR2* (Arêde *et al.*, 2012), (ver figura 2). El gen *mecA* codifica para una proteína de unión a la penicilina (PBP) de 78 kDa, llamada PBP2a; la cual, presenta baja afinidad para la meticilina y todos los antibióticos  $\beta$ -lactámicos. La proteína PBP2a continúa sintetizando peptidoglicano para la pared celular aun cuando las PBP normales estén inhibidas por los antibióticos (Velázquez-Meza, 2005; Foster, 2004; Hiramatsu *et al.*, 2001). Por otro lado, los genes *ccr* son: *ccrA*, y *ccrB*, los cuales codifican para recombinasas responsables de la movilización del *SCCmec*; ellas median su integración y escisión (Carroll *et al.*, 1995).

## Modelo de la inducción de *mecA* por *mecR1-mecI-mecR2*

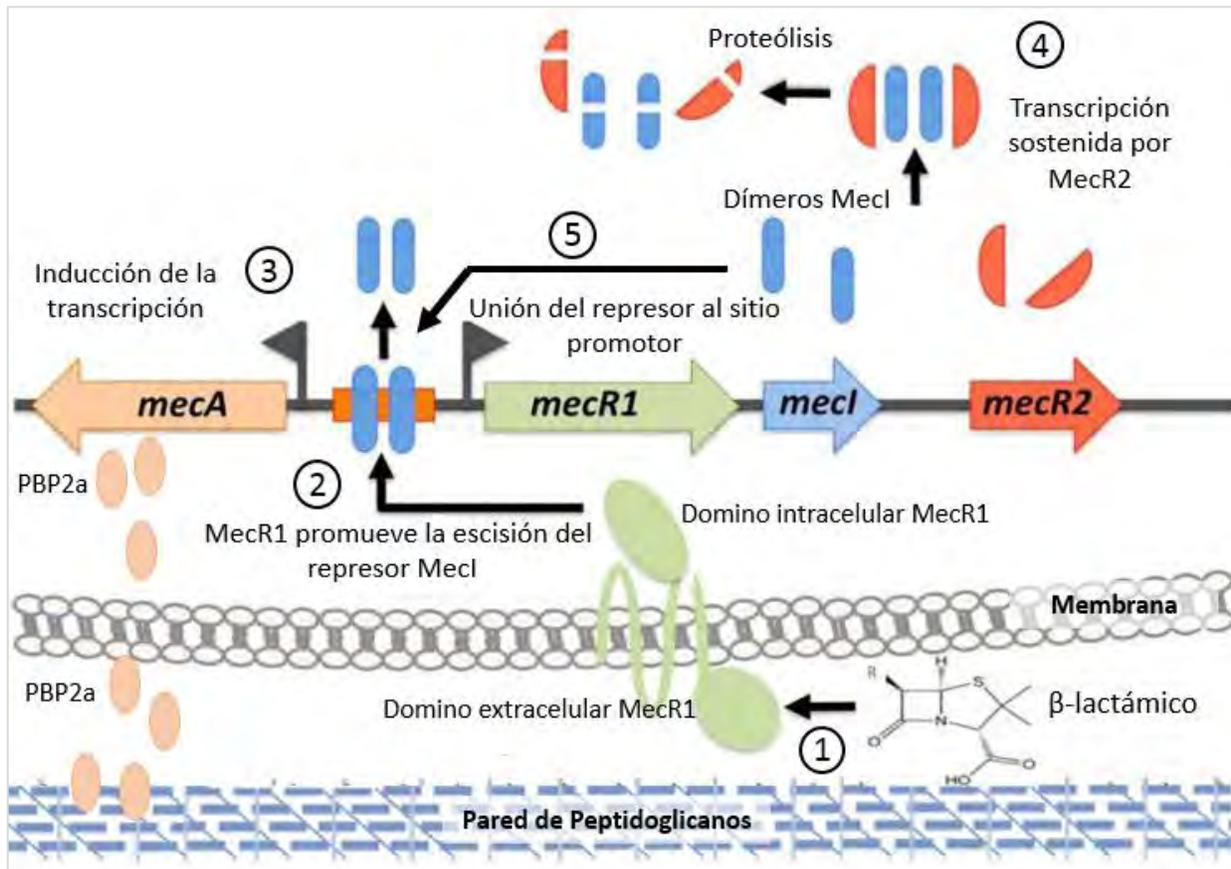


Fig. 2 | **Modelo de la inducción de *mecA* por *mecR1-mecI-mecR2*** | (1) En presencia de un antibiótico  $\beta$ -lactámico, el dominio intracelular MecR1 se activa y (2) promueve la escisión del represor MecI, (3) esto induce la expresión del gen *mecA* y *mecR1-mecI-mecR2*. (4) MecR2 sostiene la transcripción ya que promueve la proteólisis de los dímeros MecI, también MecR2 puede destabilizar la unión de estos con el sitio promotor. (5) En ausencia de  $\beta$ -lactámicos, MecR1 ya no está activo, por lo que ya no hay señal de transducción y los dímeros MecI que se encuentren estables se unen nuevamente con el sitio promotor. (Modificado de Arède *et al.*, 2012).

El resto del cassette cromosómico contiene secuencias de la región J o *Junkyard*, que comprende tres fragmentos denominados J1, J2 y J3; que constituyen componentes no esenciales del cassette. Estas regiones pueden portar determinantes de resistencia

adicionales como: genes de resistencia a antibióticos no  $\beta$ -lactámicos y a metales pesados (Zhang *et al.*, 2005). La región J1 se localiza entre el complejo *ccr* y el extremo cromosómico flanqueante derecho; J2 se encuentra entre los complejos *mec* y *ccr*; y J3 está entre *orfX* y el complejo *mec*. Aunque los fragmentos J no son específicos de cada tipo de *SCCmec*, se observa que algunos de ellos se encuentran más comúnmente en ciertos tipos de *SCCmec* (Kondo *et al.*, 2007) (Ver figura 3).

### Elementos del *SCCmec* tipo II

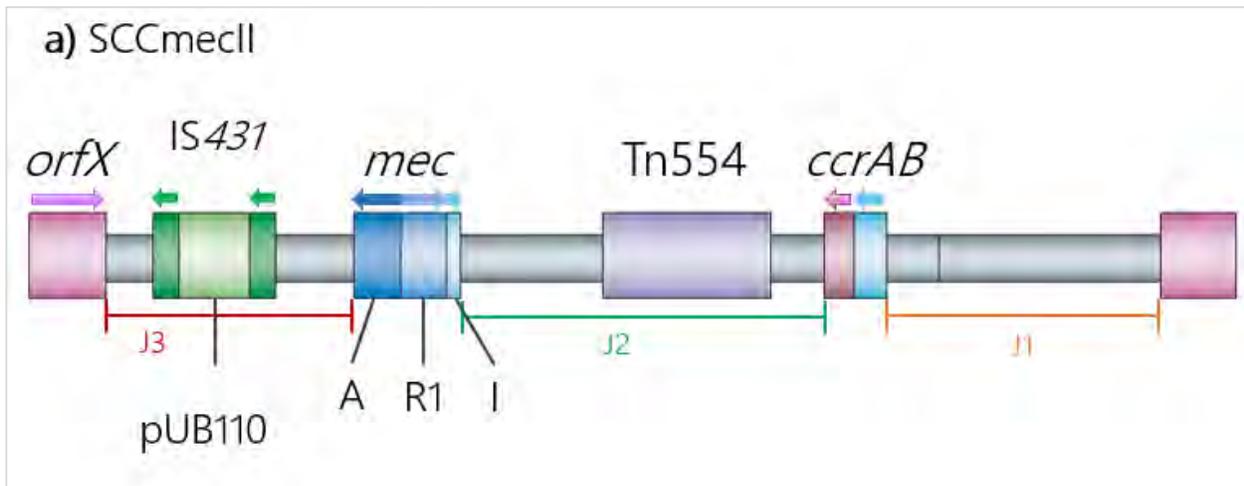


Fig. 3 | **Elementos del *SCCmec* tipo II** | (De izquierda a derecha). Tenemos el gen *orfX*, seguido de la región J3 que alberga la secuencia de inserción 431 (IS431) interrumpida por el plásmido puB110, continuo a esta, se encuentra el complejo *mec*, posteriormente se localiza la región J2, que porta el transposón 554 (Tn554), finalmente tenemos el complejo *ccr* y la región J1 que no lleva ningún componente. Las flechas indican el sentido de la transcripción. (Modificado de Chambers y DeLeo, 2009).

## Clasificación del SCCmec

El SCCmec se clasifica con base en los polimorfismos de sus diferentes elementos: las diferentes combinaciones del complejo *mec* (A, B, C y D) y los distintos alotipos del complejo *ccr* (1, 2, 3 y 4) generan los diferentes tipos del *cassette* SCCmec. Las variaciones en la región J –aun teniendo la misma combinación de los complejos *mec* y *ccr* –definen los subtipos del SCCmec (Anderson *et al.*, 2009; Deurenberg y Stobberingh, 2008). A la fecha se han descrito 8 tipos de SCCmec designados del I al VIII (Lowy, 2003; Deurenberg y Stobberingh, 2008) (ver figura 4).

### Comparación de dos tipos de SCCmec

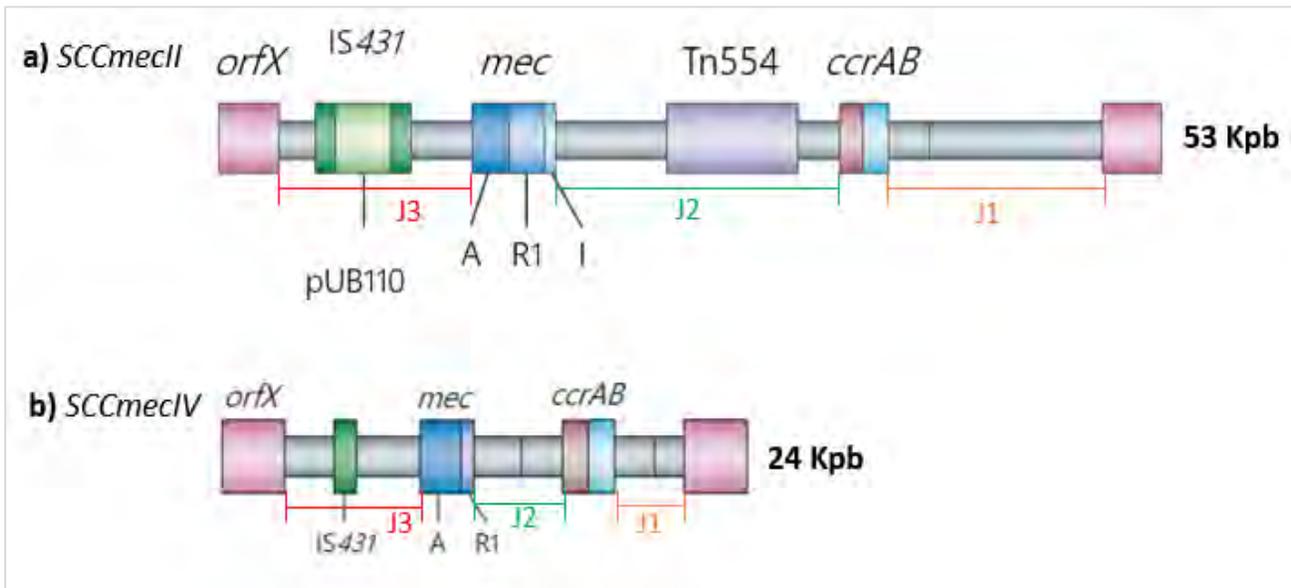


Fig. 4 | **Comparación de dos tipos de SCCmec** | El SCCmec tipo IV es diferente al SCCmec tipo II, ya que el primero, carece del plásmido pUB110 en la región J3 y del transposón 554 (Tn554) en la región J2, e inclusive su región J1 es menor (Modificado de Chambers y DeLeo, 2009).

## ***Staphylococcus aureus* resistente a meticilina (MRSA)**

En 1961 se publicó el primer registro de *MRSA*, apenas 2 años después de la introducción de la meticilina para el tratamiento de las infecciones por *S. aureus* resistentes a la penicilina (Chambers y DeLeo, 2009). Todas las cepas de *MRSA* contienen *SCCmec*. Los elementos *SCCmec* han estado sujetos a eventos de transferencia horizontal (Moore y Lindsay, 2001; Goerke *et al.*, 2006); incluso, se ha descrito que los diferentes componentes de estos elementos tienen diferentes historias evolutivas (Couto *et al.*, 1996; Tsubakishita *et al.*, 2010a, 2010b; Wu *et al.*, 1998). Aunque algunos trabajos han estudiado la relación del *SCCmec* y las cepas que lo presentan (Castillo-Ramírez *et al.*, 2012; Holden *et al.*, 2013), estos estudios solo se han limitado a algunas cuantas clonas. A la fecha no ha habido un estudio reciente que se haya centrado en la historia evolutiva de este elemento incluyendo diversos linajes de *S. aureus*.

## **ANTECEDENTES**

Couto y colaboradores en 1996, en un intento de investigar el origen y/o reservorio del o de los determinantes genéticos de la resistencia a la meticilina en *Staphylococcus aureus*, examinaron la presencia del gen *mecA* en más de 200 cepas que representaban 13 especies distintas del género *Staphylococcus*, construyeron una sonda de ADN usando como referencia una cepa de *Staphylococcus aureus* resistente a meticilina (*MRSA*). Donde 134 cepas de *Staphylococcus sciuri* –una de las especies consideradas taxonómicamente la más primitivas entre *Staphylococcus*– dieron positivo a la sonda, dos tercios de estas cepas, eran de la subespecie *sciuri*, 9 de 11 eran subespecies *carnaticum* estas cepas mostraron resistencia a la meticilina (concentración mínima inhibitoria de

0.75 -6.0 microgramos/ml), mientras que la subespecie *rodentius* (13 aislamientos en total) mostró resistencia heterogénea a los antibióticos muy similar al fenotipo *MRSA*. Finalmente, la presencia uniforme de *mecA* en las cepas de *S. sciuri* pertenecientes a distintos ribotipos y patrones de macrorrestricción, recuperados en un período de 20 años a partir de una amplia variedad de fuentes animales y sitios geográficos sugiere que *mecA* puede ser un elemento genético nativo con una función fisiológica aún no identificada en esta especie de *Staphylococcus*.

Por otro lado, Wu y colaboradores, en 1998, buscaron secuencias homólogas del gen *mecA* y las regiones que lo flanquean usando como referencia a *Staphylococcus aureus* resistente a meticilina (*MRSA*), en 4 subespecies de *Staphylococcus sciuri*: *Staphylococcus sciuri* subsp. *sciuri* (cepa K1), cepas K3 y K8 que representan a *S. sciuri* subsp. *rodentius* y un representante de *S. sciuri* subsp. *carnaticum* (cepa K11). Ellos encontraron que las cepas K1 y K11 eran susceptibles mientras que las cepas K3 y K8 mostraban una resistencia heterogénea a meticilina, debido a que (ambas pertenecientes a la subespecie *rodentius*) contenían genes *mecA* idénticos a *S. aureus*; además una de las cepas (K3) contenía dos copias del gen *mecA*: una (*mecA2*) prácticamente idéntica al determinante PBP2a de *S. aureus*, y la otra (*mecA1*) prácticamente idéntica al homólogo *mecA* de K1 (subsp. *sciuri*) y K11 (subsp. *Carnaticum*).

Las cepas que llevan esta forma de *mecA* son susceptibles a la meticilina, lo que sugiere que esta variante alélica no confiere resistencia a los antibióticos. La segunda cepa, K8, perdió el homólogo nativo y sólo lleva el *mecA* tipo *MRSA*. También identificaron un marco de lectura abierto que define un polipéptido truncado N-terminal, NTORF101, con un alto grado de homología con un segmento de ADN en la región hipervariable de *MRSA*

–similar al gen *ugpQ* de *Escherichia coli* –fue también identificado río abajo del homólogo *mecA* de K11. El gen *ugpQ* no está presente en cepas de *S. aureus* susceptibles a meticilina. La presencia del gen *ugpQ* junto con el homólogo de *mecA* en la cepa K11 apoya la especulación de que estos elementos genéticos pueden ser familiares y/o precursores evolutivos del determinante genético de *MRSA*.

Más recientemente, Tsubakishita y colaboradores en 2010a, analizaron los genomas de 3 cepas (JCSC5402, JCSC7096 y JCSC7528) de *Micrococcus caseolyticus* –una bacteria comensal de animales productores de alimentos –, ellos identificaron el plásmido *pMCCL2*, que lleva un transposón (*Tn6045*) que contiene una estructura peculiar, *mecl-mecR1-mecA-blaZ* (designado *mecl<sub>m</sub>-mecR1<sub>m</sub>-mecA<sub>m</sub>-blaZ<sub>m</sub>*, por *Micrococcus*), que es diferente de *mecl-mecR1-mecA* presente en *SCCmec*. Este transposón fue encontrado río abajo del sitio *orfX* del cromosoma, flanqueado por secuencias de repetición directa *DR* (por sus siglas en inglés Direct Repeats), reconocidas por las recombinasas (*ccr*). El elemento *Staphylococcal Cassette Chromosome* (*SCC*), designado *SCC7096* (por la cepa JCSC7096), se integra a un lado del transposón *mec*, separado de este último por un *DR*. Mediante experimentos por PCR anidado, mostraron que el transposón *Tn6045* –que contiene el complejo *mec*–, no sólo se escinde por separado, sino también se co-escinde con *SCC7096*. Los elementos co-escindidos forman el elemento *SCCmec*, un ADN circular cerrado extracromosómico. Lo anterior reveló un mecanismo potencial para la generación de nuevos elementos *SCCmec*.

El mismo grupo (Tsubakishita y *et al.*, 2010b) investigaron el origen/depósito del gen *mecA*. Ellos mencionan a *S. fleuretti*, una especie relacionada con animales, como un origen altamente probable del gen *mecA*: este gen se encontró en el cromosoma

vinculado con los genes esenciales para el crecimiento de *Staphylococcus* y no se asoció con *SCCmec*. El locus *mecA* del cromosoma *S. fleurettii* tiene una secuencia prácticamente idéntica a la de la región que contiene *mecA* (~ 12 Kpb de largo) de *SCCmec*. Ellos realizaron un análisis de correspondencia a las cepas *S. sciuri*, *S. vitulinus* y *S. fleurettii* y pudieron determinar que homólogo *mecA* de *S. fleurettii*, desciende de su antepasado y que no fue adquirido recientemente. Así ellos sugieren que el *SCCmec* fue generado en las células de *Staphylococcus* que viven en los animales mediante la adquisición de la región *mecA* de *S. fleurettii*, ya que este por sí mismo no puede formar el *SCCmec*, ya que es intrínsecamente resistente a los antibióticos  $\beta$ -lactámicos.

## OBJETIVOS

### Objetivo general

- Analizar la historia evolutiva del *SCCmec*.

### Objetivo particular

- Encontrar secuencias homólogas de los genes *mecA*, *ccrA* y *ccrB* en los 49 genomas descargados de la página de la NCBI, usando como referencia la cepa *TW20*.
- Construir filogenias de máxima verosimilitud.
- Detectar eventos de recombinación mediante los programas Recombination Detection Program (RDP) versión 2 y Splitstree4.
- Evaluar el papel de la selección y la deriva génica con el cociente de sustituciones sinónimas (dS) y no sinónimas (dN).

## MÉTODOS

### Homólogos, alineamientos y filogenias de máxima verosimilitud

De la página NCBI se descargaron los 49 genomas. Se buscaron las secuencias homólogas de los genes *ccrA*, *ccrB* y *mecA* usando el programa BLAST (Altschul *et al.*, 1990), utilizando como referencia las secuencias de la cepa *TW20*. Para obtener alineamientos de ADN conservando el marco de lectura, se procedió de la siguiente manera: se utilizó el programa TRANALING, que es parte del software de EMBOSS (Rice *et al.*, 2000), para alinear las secuencias de ADN usando como referencia secuencias proteicas –es decir de amino ácidos– (ver anexo 1). Una vez que se obtuvieron los alineamientos de ADN conservando el marco de lectura, se utilizó el “Alignment Explorer” de MEGA6 (Tamura *et al.*, 2013) para editar el alineamiento y quitar aquellas regiones con deleciones. Con “Model Selection” de MEGA6, se corrió un test para cada uno de los alineamientos –un alineamiento por cada gen–, para seleccionar el modelo de árbol que se ajuste mejor a nuestros datos. Usando el alineamiento y con el modelo adecuado, se construyeron Filogenias de Máxima Verosimilitud por medio de PhyML (Guindon y Gascuel, 2003), a través del programa Seaview (Galtier *et al.*, 1996).

### Análisis de recombinación y selección

Con el fin de detectar los posibles eventos de recombinación, se utilizaron los programas Recombination Detection Program (RDP) versión 2 (Martin *et al.*, 2010) y SplitsTree4. Con este último programa, se construyeron redes filogenéticas para cada gen, se obtuvo la diversidad filogenética y se corrió el test phi (Huson y Bryant, 2006). Posteriormente, se evaluó el papel de la selección y la deriva génica con el cociente dN/dS, referido como

radio  $\omega$  –el cual compara las tasas de sustituciones sinónimas (dS) y no sinónimas (dN), donde  $dN/dS > 1$  implica selección positiva,  $dN/dS < 1$  significa selección purificadora y  $dN/dS = 1$  sugiere evolución neutral–. Del mismo modo, se corrieron modelos de sitios (ver tabla 1) y se calculó el radio test de máxima verosimilitud (*Likelihood Ratio Test*, por sus sigla en inglés *LRT*), –que permite comparar modelos evolutivos de codones para tratar de inferir el tipo de selección–, (ver anexo 2). Esto se realizó por medio del programa Codeml que es parte del software Phylogenetic Analysis by Maximum Likelihood (PAML) (Yang, 2007).

**Tabla 1. Modelos de sitio**

<b>M0</b>	Asume la misma presión de selección para todos los sitios (una solo $\omega$ ).
<b>M0 vs M3</b>	Determina si hay variación a lo largo del gen.
<b>M1a vs M2a</b>	Permite determinar si hay sitios bajo selección positiva.

Tabla 1 | **Modelos de Sitio** | Estos modelos al compararse permiten detectar la presión de selección entre los sitios (por codón).

## Diagramas eBURST

Sé corrió eBURST a través del sitio web <http://www.mlst.net/>, mediante el uso *Sequence Types (STs)* y perfiles alélicos disponibles en la base de datos *MLST Multi Locus Sequence Type*, que comprende 2,812 *STs* para *S. aureus* (<http://saureus.mlst.net/>); a la cual, se le añadió nuestra propia base de datos que abarca 49 *STs* –un *STs* por cepa–. Finalmente, se usó la definición más estricta (>6/7 alelos compartidos), para la asignación de los grupos (ver anexos 3) (Feil *et al.*, 2004; Spratt *et al.*, 2004).

## RESULTADOS

### El *SCCmec* no se encuentra presente en todas las cepas

Se examinó la presencia de *SCCmec* y de sus genes individuales en el conjunto de 49 cepas que consideramos. Más de un quinto de las cepas (10 de 49) no parecen tener el elemento *SCCmec* ya que no presentaron ninguno de los 3 genes (ver figura 5). De las diez cepas, dos de ellas –55/2053 y 6850–, no se pudieron determinar a qué complejo pertenecían, de las ocho cepas restantes, cuatro están asociadas a seres humanos: la cepa *ECT-R2* que pertenece al complejo clonal 5 (CC5), las cepas *NCTC 8325*, *Newman* y *VC40* del complejo clonal 8 (CC8). Respecto a las cuatro cepas que faltan, solo dos se encuentran asociadas únicamente a animales: la cepa *ED133* asociada a ovinos, del complejo clonal 133 (CC133) y la cepa *ED98*, asociada a aves de corral del CC5, y las últimas 2 cepas se asocian a humanos y animales: la cepa *RF122* asociada a ganado y humanos, del complejo clonal 151 (CC151) y la cepa *71193* asociada a porcinos y humanos del complejo clonal 398 (CC398). Dado que estas diez cepas no presentan el *SCCmec*, es muy probable que estas cepas sean sensibles a meticilina y otros antibióticos  $\beta$ -lactámicos.

A pesar de esto, se encuentran formando complejos clonales con cepas resistentes. Por otra lado, las cepas que presentaron sólo el gen *mecA* fueron *08BA02176* (CC398), *M013* del complejo clonal 59 (CC59), clasificada como *CA-MRSA* (Community–Acquired *MRSA*, cepas *MRSA* asociadas a comunidad) (Huang *et al.*, 2012) y *SA957* (CC59), notablemente todas ellas son cepas *MRSA* y una de ellas, la cepa *08BA02176*, ha sido clasificada como *LA-MRSA* (*Livestock associated MRSA*, cepas *MRSA* asociadas a ganado) (Chua *et al.*, 2014; Chen *et al.*, 2013). Tanto la cepa *COL* (CC8) como *ST398*

(CC398) presentaron los genes *ccrA* y *mecA*; ambas cepas son *MRSA*, aunque el ST398 es *LA-MRSA*. En contra parte, la cepa *MSSA476* del complejo clonal 1 (CC1), careció del gen *mecA*; sin embargo, presentó los genes *ccrA* y *ccrB* y ha sido clasificada como *MSSA* (*Methicillin-Susceptible S. aureus*, cepas sensibles a meticilina) (Chua *et al.*, 2014). A su vez, en el CC5 las cepas 10388, 10497, 15532, 16035, 16125, 18341, 18412 y 18583 presentaron 2 homólogos para el gen *ccrB*. Dentro del CC8, la cepa que destacó fue M1 ya que tuvo 3 homólogos para el gen *ccrA* y 2 homólogos para el gen *ccrB*. Todas las cepas anteriores han sido descritas como *MRSA*, sin embargo esta última se asocia a la comunidad (*CA-MRSA*) (Molina *et al.*, 2008; Larner-Svensson *et al.*, 2013; Chambers y DeLeo, 2009).

**Distribución de los homólogos de *ccrA*, *ccrB* y *mecA***

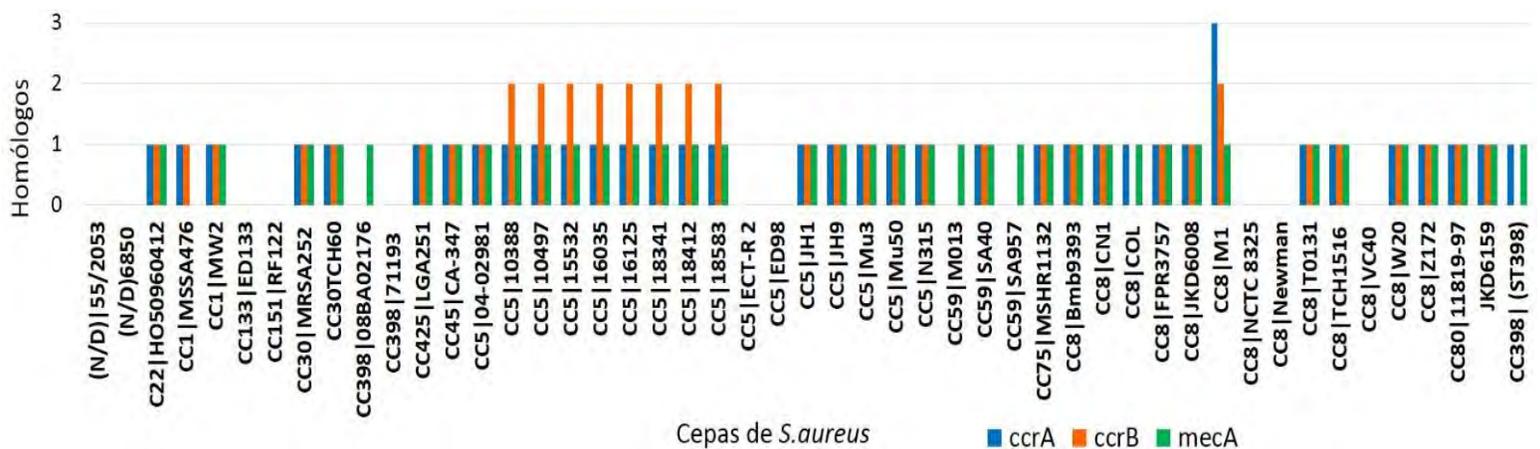


Fig. 5 | **Distribución de los homólogos de *ccrA*, *ccrB* y *mecA* en las 49 cepas analizadas** | Debajo del gráfico se encuentran los nombres de las cepas de *S. aureus*, de acuerdo al NCBI (GenBank), seguido del Complejo Clonal (CC) al que pertenecen. Las cepas con N/D (No fue posible determinar el complejo al que corresponden).

## **Patrones de descendencia evolutiva (vertical) de *S. aureus* a partir del MLST**

Para poder comprender los eventos de recombinación del elemento *SCCmec*, se determinaron las relaciones filogenéticas (verticales) de 47 de nuestras 49 cepas, debido a que las cepas 55/2053 y 6850 fueron omitidas ya que no se encontraron en la base de datos del MLST. El análisis eBURST arrojó 83 diagramas distintos, de los cuales, en 11 de ellos se hallaron 47 de nuestros STs. La cepa JKD6159 con el ST93 fue la única que no formó parte de algún diagrama. A continuación se presenta el diagrama eBURST que albergó el mayor número de nuestros STs –31 de 47 STs –(ver figura 6), en este diagrama, se encontró al CC5, conformado por los STs 228, 225, 105, 5 –siendo este último un ST fundador–, el CC8 presentó los STs 250, 239, 72, 8 y el CC1, solo albergó al ST1. Cabe mencionar que los STs que conforman al complejo clonal, están muy relacionados entre sí.

Encontrar diferentes complejos clonales, nos sugiere que del genotipo fundador –es decir del ST5–, han surgido diferentes variantes de locus simple (del inglés *Single Locus Variant, SLV*) –diferentes STs–, y algunas de estas variantes han sido tan exitosas que han diversificado formando sus propios SLV, estableciendo así sus propios complejos clonales, tal como sucede con el CC8 y el CC1.

## Diagrama eBURST del CC5, CC8 y CC1

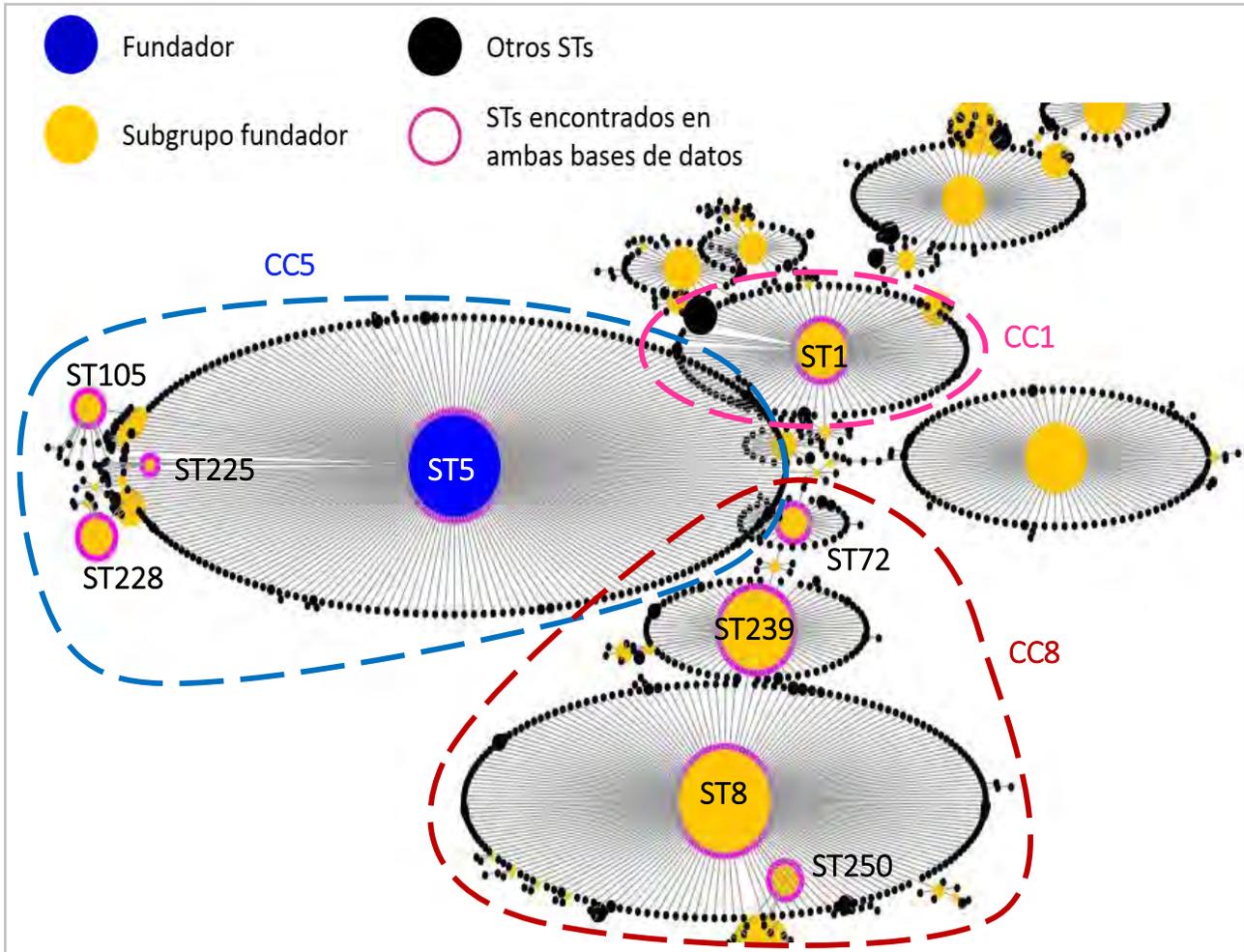


Fig. 6 | **Diagrama eBURST del CC5, CC8 y CC1** | El CC5 –alberga los STs 228, 225, 105, 5 – señalado por la línea punteada azul, el CC8 –agrupa los STs 250, 239, 72, 8 –por la línea punteada roja y finalmente la línea punteada rosa, indica al CC1 –presenta ST1–. Los STs que conforman cada uno de los complejos clonales están relacionados.

El siguiente diagrama eBURST presentó únicamente el complejo clonal 30 (CC30), –este alojó 2 de nuestros 47 STs–, este complejo albergó al ST30 –ST fundador–, y al ST36, siendo este último un SLV del fundador (ver figura 7). También cabe mencionar que el ST30 es menos exitoso con respecto al ST5, –ambos fundadores–. Además, el CC30 es un grupo aislado y distante al CC5, CC8 y CC1 ya que no hay variantes de locus que puedan cohesionar dichos complejos. Este mismo caso ocurre para los diagramas restantes, que podrán ser consultados en el Anexo 3.

### Diagrama eBURST del CC30

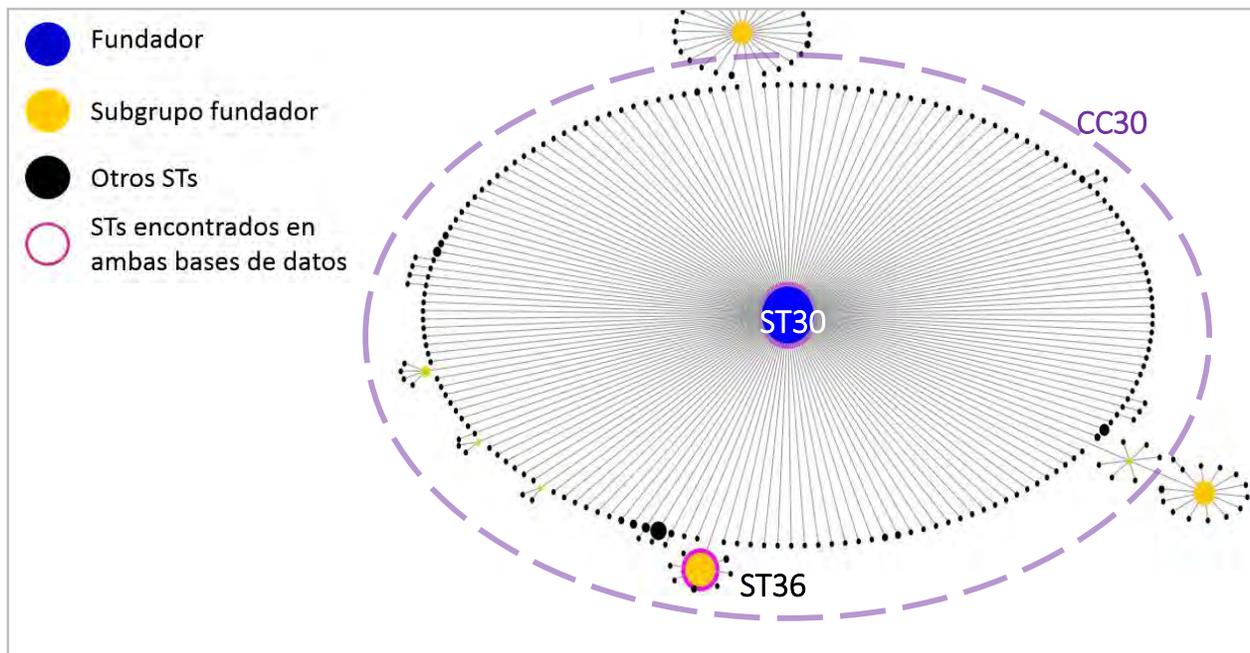


Fig. 7 | **Diagrama eBURST del CC30** | El CC30 alberga al ST30 siendo este un ST fundador y tenemos al ST36 que es un SLV del ST30. Ambas cepas están muy relacionadas.

## Los genes del *SCCmec* presentaron diferentes historias evolutivas

Para evaluar qué tan similares son las historias evolutivas de los genes de este elemento, se construyeron filogenias de máxima verosimilitud. Aunque los genes *ccrA* y *ccrB* pertenecen al mismo complejo de genes, presentan distintas historias evolutivas ya que la topología de los árboles es diferente (ver figura 8). Claramente la disposición de las cepas es diferente en estos genes; por ejemplo, las cepas *TCH60* y *CN1* (grupo I-A) presentan una posición más basal en la filogenia para el gen *ccrA*, ambas cepas son *MRSA* asociadas a humano (Seong *et al.*, 2012; Chen *et al.*, 2013), y para el gen *ccrB* la cepa con posición más basal es *N315*, (grupo I-B) también *MRSA* asociada a humanos (Holden *et al.*, 2004). En la filogenia del gen *ccrA*, el grupo II-A formó un grupo hermano con la cepa *LGA251* (cepa *LA-MRSA*), mientras que esto no es el caso para el gen *ccrB*; donde el grupo II-B formó un grupo hermano con la cepa *CA-347*, cepa asociada a humanos (Stegger *et al.*, 2013).

Por último, ambos árboles mostraron que ha habido diferentes eventos de transferencia genética horizontal (TGH), dentro de los diferentes complejos clonales (CC); esto se puede apreciar claramente cuando se observa que cepas que pertenecen a un mismo complejo clonal no se agrupan todas juntas. Por ejemplo, nótese que las cepas *MSSA476* y *MW1*, aunque las dos pertenecen al *ST1*, están localizadas en diferentes partes del árbol. Lo mismo ocurre con las cepas que presentan el *ST8*. Finalmente, la filogenia de máxima verosimilitud para el gen *mecA* (ver figura 9) presentó en su posición más basal el grupo I-C, que agrupa a las cepas *CN1* y *LGA251* –ambas *MRSA*, la primera asociada a humanos y la segunda ha ganado–, aunque *mecA* y *ccrA* son elementos con diferentes historias evolutivas, en ambos la cepa *CN1* presentó una posición basal.

## Filogenias de máxima verosimilitud del gen *ccrA* y *ccrB*

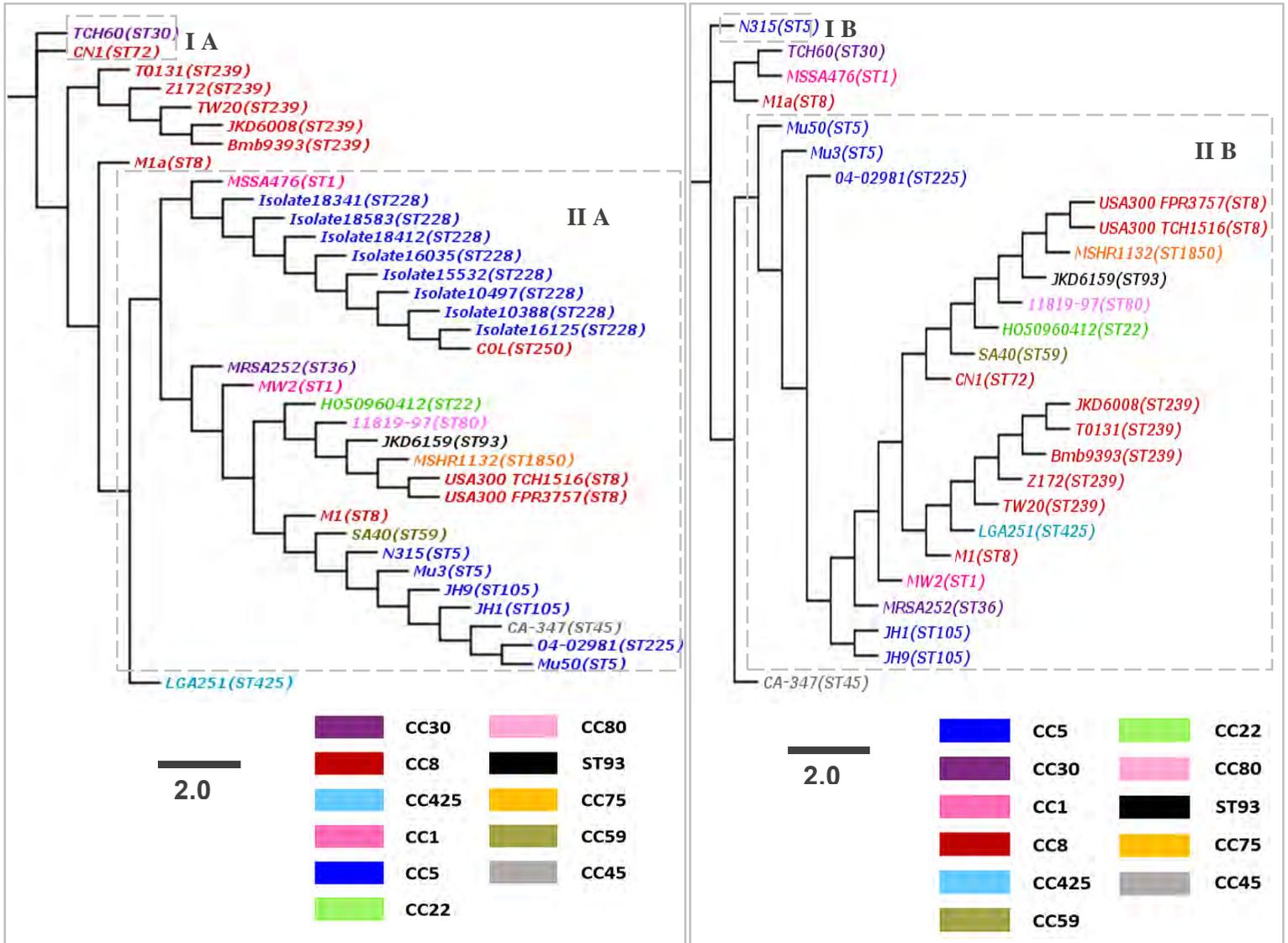


Fig. 8 | Filogenias de máxima verosimilitud del gen *ccrA* y *ccrB*, en cepas de *S. aureus* | Los nombres corresponden a las cepas de *S. aureus*, seguido del *Sequence Type* (ST) entre paréntesis. Las cepas que presentaron al menos dos secuencias homólogas, se les colocó la letra “a” a una de ellas para diferenciarla. Los colores representan el complejo clonal (CC) al que pertenecen y los rectángulos punteados los clados que formaron: I A y II A para el gen *ccrA* y I B y II B para *ccrB*.

## Filogenia de máxima verosimilitud del gen *mecA*

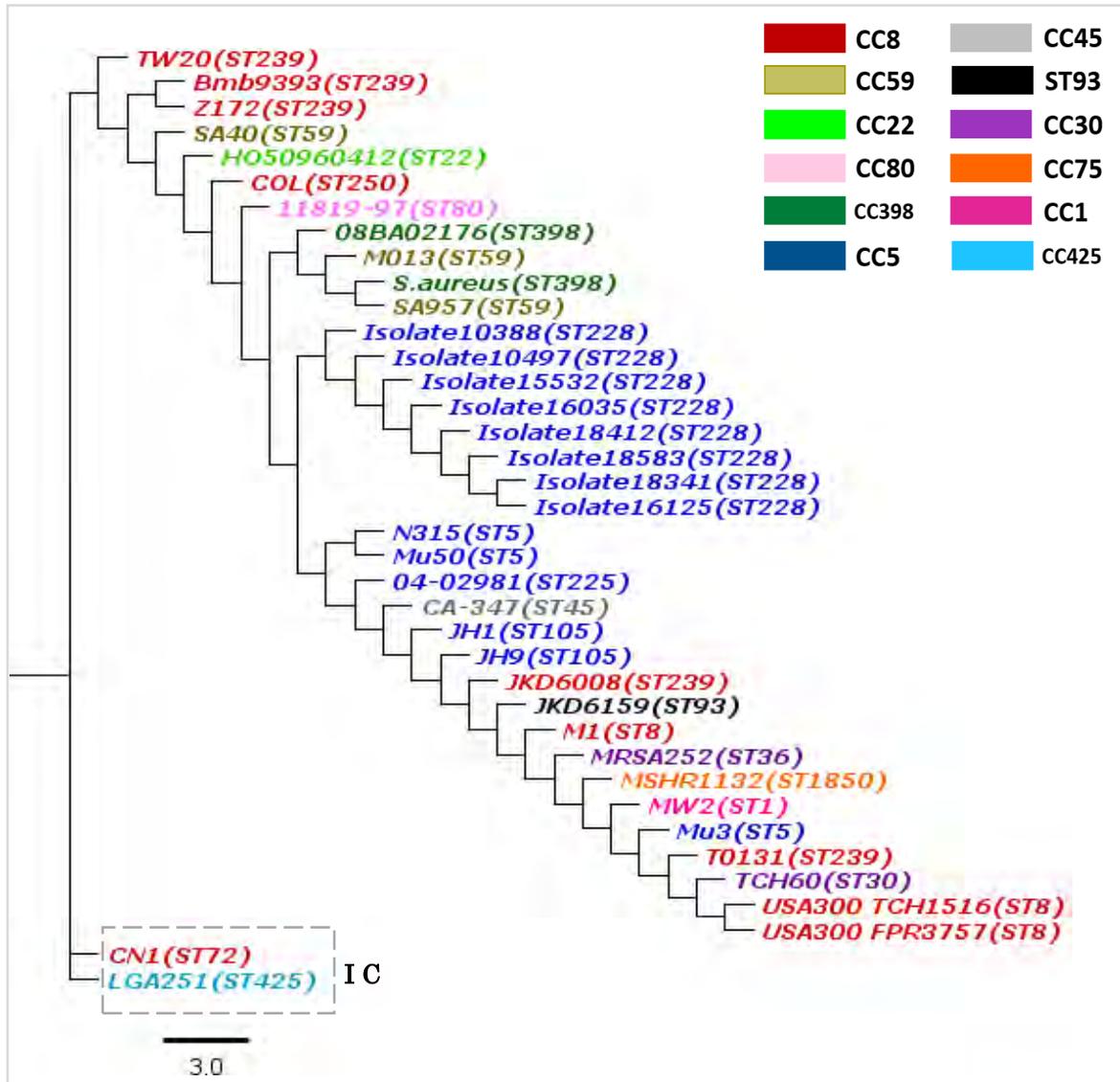


Fig. 9 | Filogenia de máxima verosimilitud del gen *mecA*, en cepas de *S. aureus* | Los nombres corresponde a las cepas de *S. aureus*, seguido del *Sequence Type* (ST) entre paréntesis. Los colores representan el complejo clonal (CC) al que pertenecen y los rectángulos punteados los clados que formaron: I C.

## **Los genes del SCCmec muestran señales de recombinación y diferencias en variación genética**

Dado que varios eventos de transferencia horizontal parecen haber intervenido en la evolución de estos genes, se decidió usar redes “Neighbor” ya que permiten visualizar relaciones genealógicas cuando las partes de un gen tienen distinta historia evolutiva. La red del gen *ccrA* (ver figura 10), formó tres grupos (óvalos de colores), donde el grupo señalado por el óvalo verde presentó cepas *HA-MRSA* y *CA-MRSA*, todas ellas asociadas a seres humanos. Al menos en este grupo, el gen *ccrA*, ha estado sujeto a eventos de TGH en un mayor número de complejos clonales. El grupo señalado por el óvalo azul fue un grupo estrecho a comparación del grupo anterior, ya que está conformado en su mayor parte por cepas pertenecientes al CC8; este grupo presenta cepas del tipo *HA-MRSA*, todas asociadas a seres humanos. El último grupo señalado por el óvalo amarillo, es el único grupo que presentó una cepa asociada a animales (*LA-MRSA*), la cepa *LGA251*, mientras que el resto de las cepas que forman parte de este grupo son asociadas a seres humanos. La cepa *M1*, presentó dos secuencias *M1* y *M1a*, ambas están sujetas a diferentes eventos de recombinación, ya que se encontraron en grupos diferentes. Para el gen *ccrB* (ver figura 11), la red filogenética también formó tres grupos (óvalos de colores), donde el grupo con el óvalo verde presentó cepas *HA-MRSA* y *CA-MRSA*, todas ellas asociadas a seres humanos. Al menos en este grupo, el gen *ccrB*, ha estado sujeto a eventos de TGH en un mayor número de complejos clonales. El grupo indicado por el óvalo azul, es un grupo muy estrecho conformado por cepas del CC8, estas cepas son *HA-MRSA*, asociadas a humanos, además en este grupo se encontró la cepa *LGA251* que es clasificada como *LA-MRSA*. El último grupo (óvalo naranja), se encontró la cepa *MSSA476* –que es sensible a meticilina –formando un grupo con cepas

*MRSA*, a pesar de esto todas están asociadas a seres humanos. Por otra parte la cepa *M1*, presentó dos secuencias *M1* y *M1a*, ambas se encuentran en diferentes grupos, sin embargo la cepa *M1a* presentó eventos de recombinación con la cepa *MSSA476* y *M1* no.

En la red filogenética para el gen *mecA* (ver figura 12), las cepas *CN1* (asociado a humanos) y *LGA251* (*LA-MRSA*) formaron un grupo muy estrecho con respecto a las demás cepas. Debido a que los tres genes *ccrA*, *ccrB* y *mecA* presentaron varios puntos de reticulación, implicando señales de recombinación, para corroborar esto, se corrió el Test phi, implementado en SplitsTree; el cual es comúnmente usado para detectar dichas señales. El Test phi arrojó que los genes *ccrA* y *ccrB* presentan señales de recombinación (con una  $P= 0.0$ ), mientras que el gen *mecA* dio negativo al test ( $P= 0.4115$ ). Después para tratar de ver que partes de los genes se vieron involucrados en los eventos de recombinación, utilizó el programa "RDP". Los resultados de éste se muestran en el anexo 4. Por otra parte la diversidad genética –relacionada con el número mutaciones sinónimas (dS) –para los genes *ccrA* y *ccrB* fue cercana a 1.5, mientras que para el gen *mecA* fue de 1.0154, el más bajo entre los tres genes.

## Red filogenética del gen *ccrA*

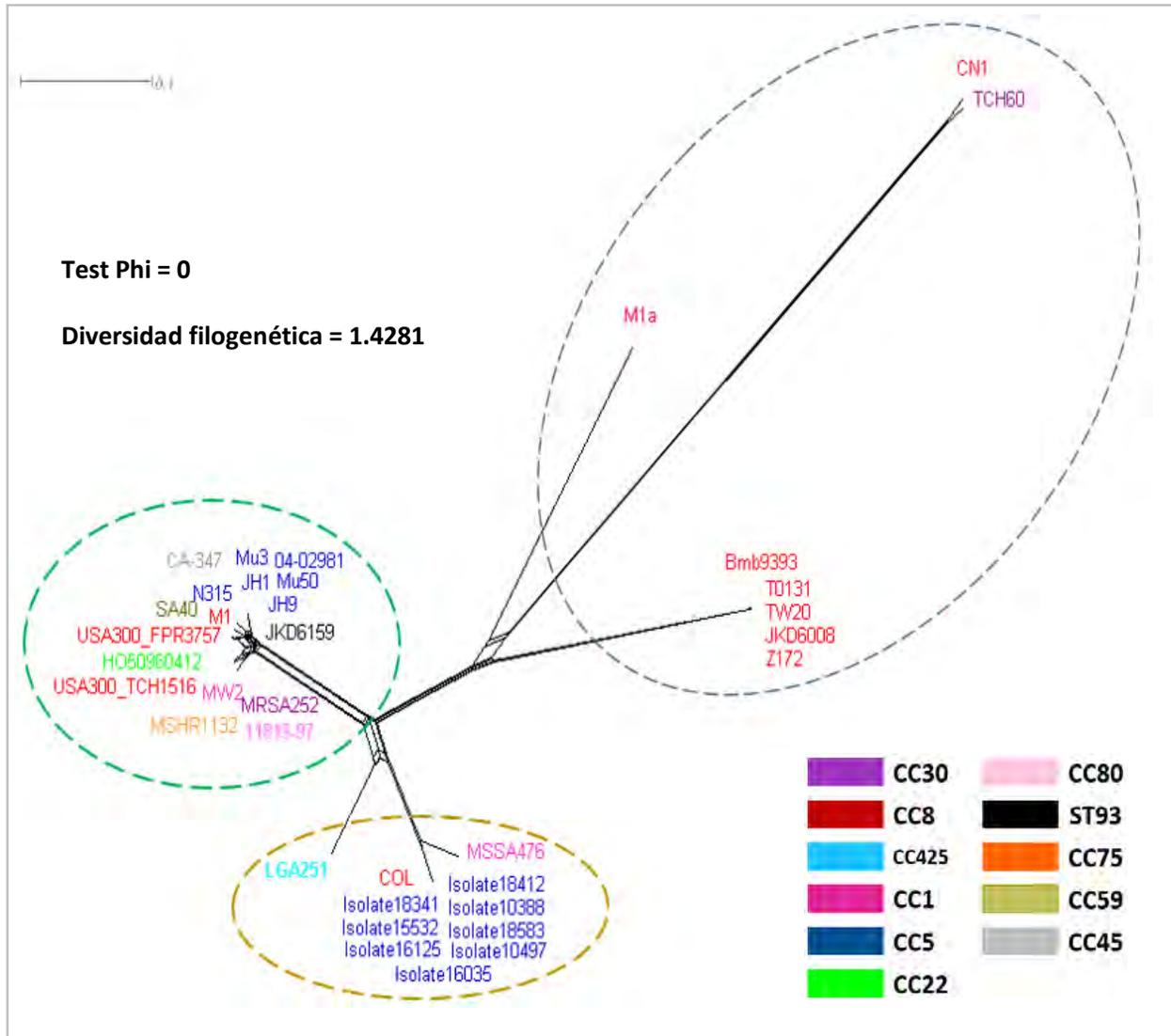


Fig. 10 | **Red filogenética del gen *ccrA*** | Los nombres corresponden a las cepas de *S. aureus* y el color representa el complejo clonal (CC) al que pertenecen. Los óvalos punteados de colores indican los diferentes grupos formados. El grupo del óvalo verde, está conformado por cepas *HA-MRSA* y *CA-MRSA*, mientras que el grupo del óvalo azul está conformado en su mayoría por cepas *HA-MRSA*, en ambos grupos las cepas están asociadas a seres humanos. En el grupo del óvalo amarilla, fue el único que albergó una cepa *LA-MRSA*, la cepa *LGA251*, el resto de las cepas que forman este grupo son asociadas a humanos. Finalmente al correr el Test Phi, dio positivo para recombinación ( $P=0$ ), mientras que la diversidad filogenética fue de 1.4281.

## Red filogenética del gen *ccrB*

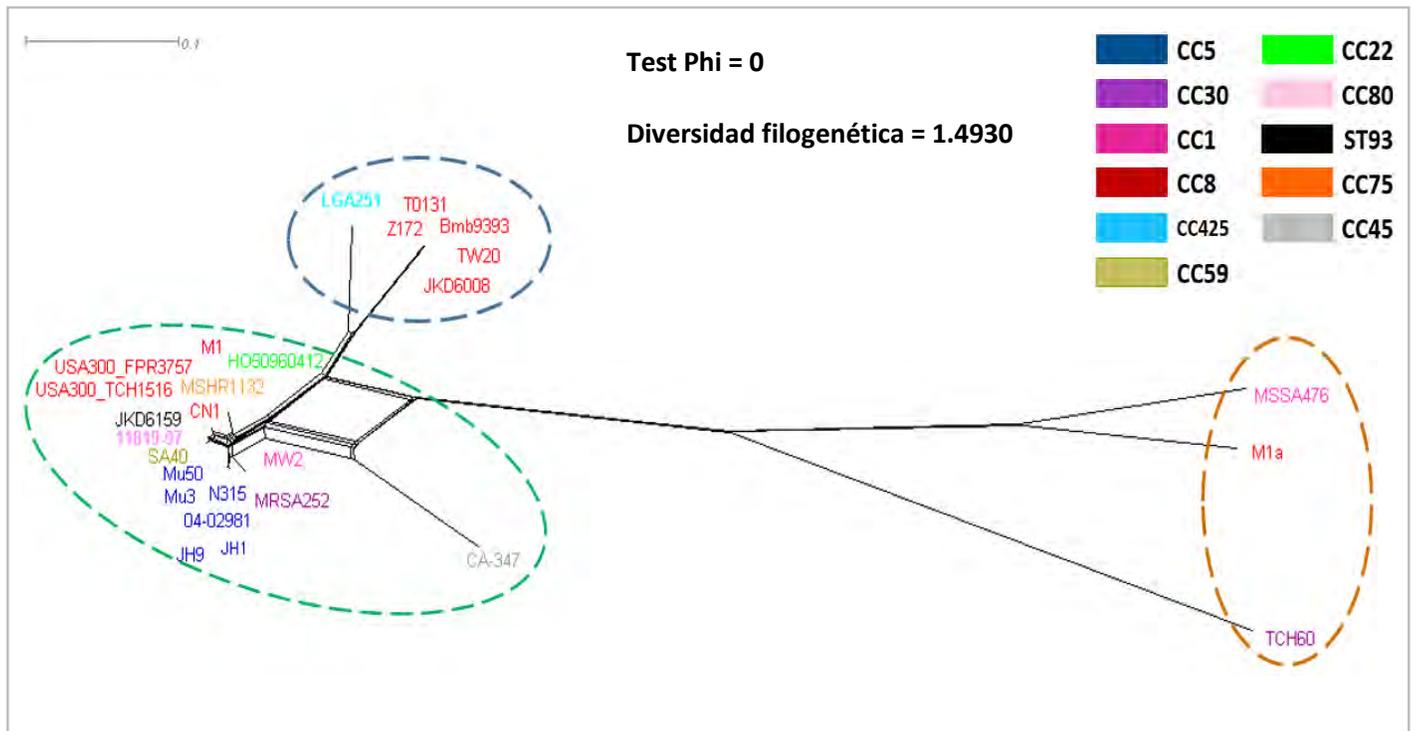


Fig. 11 | **Red filogenética del gen *ccrB*** | Los nombres corresponden a las cepas de *S. aureus* y el color representa el complejo clonal (CC) al que pertenecen. Los óvalos punteados de colores indican los diferentes grupos formados. El grupo del óvalo azul, fue el único que albergó a una cepa *LA-MRSA*, la cepa *LGA251*, el resto de las cepas son *HA-MRSA* asociadas a humanos. El grupo del óvalo verde presentó cepas *HA-MRSA* y *CA-MRSA*, también asociadas a humanos. Por otra el grupo del óvalo amarillo, se encontró la cepa *MSSA476* (reportada como sensible a metilina), formando un grupo con cepas *MRSA*. Finalmente al correr el Test Phi, dio positivo para recombinación ( $P=0$ ), mientras que la diversidad filogenética fue de 1.4930.

## Red filogenética del gen *mecA*

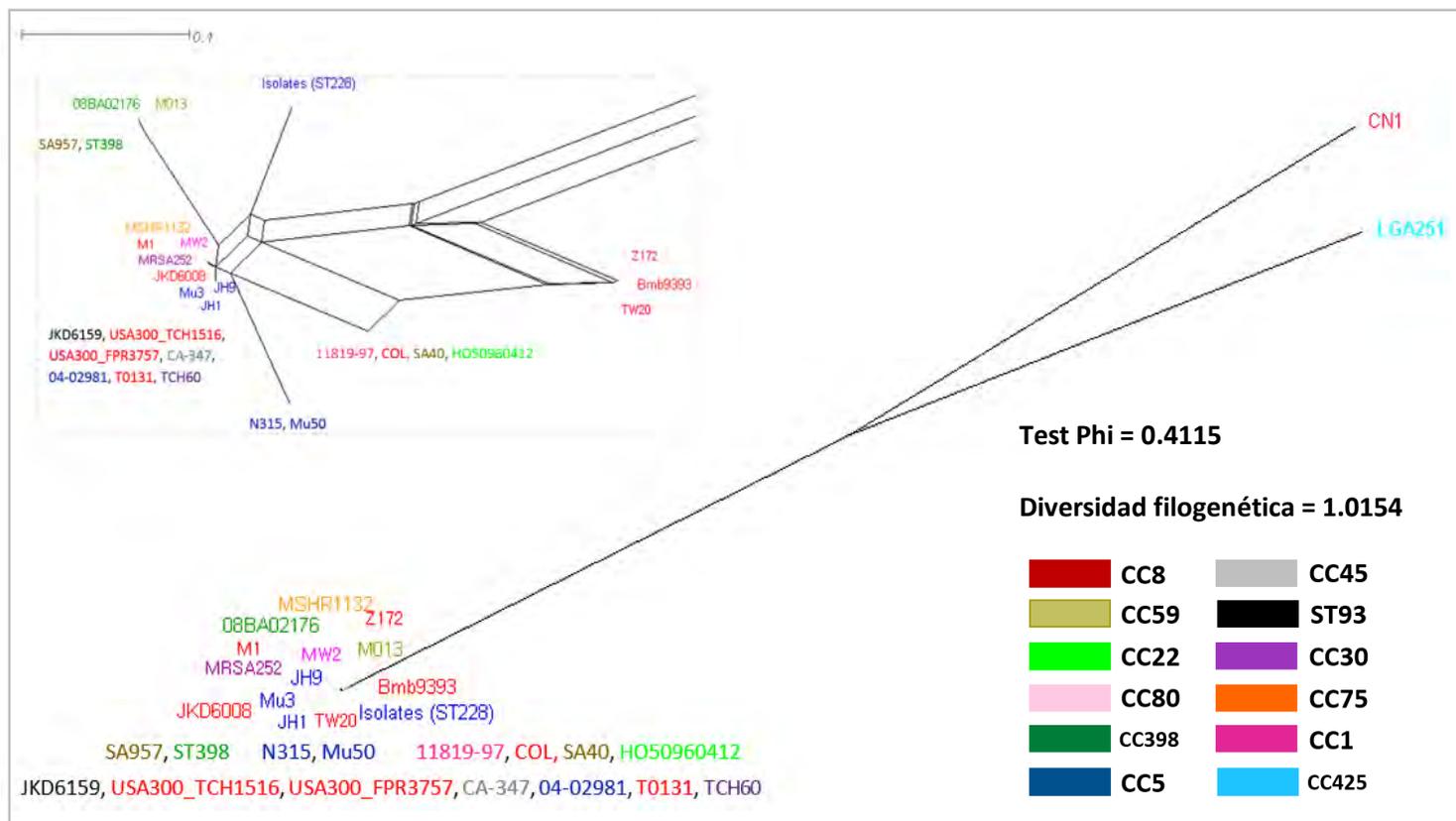


Fig. 12 | **Red filogenética del gen *mecA*** | Los nombres corresponden a las cepas de *S. aureus* y el color representa el complejo clonal (CC) al que pertenecen. Los óvalos punteados de colores indican los diferentes grupos formados. El Test Phi, dio negativo para recombinación ( $P=0$ ), por lo que no se formaron grupos. Por otra parte la diversidad filogenética fue de 1.0154, la más baja entre los tres genes analizados.

## El papel de la selección en los genes *ccrA*, *ccrB* y *mecA*

Finalmente, analizamos el tipo de selección que ha estado actuando sobre los genes que componen el *SCCmec*; como una medida del tipo de selección natural, usamos el cociente dN/dS, para esto empleamos los modelos implementados en el programa PAML (ver métodos). Primero, determinamos el promedio de la selección del gen –mediante el Modelo 0 –para cada uno de los tres genes (ver tabla 2). Claramente, los tres genes parecen estar bajo presión selectiva purificadora pues sus valores de dN/dS son menores a uno. Sin embargo, la magnitud de la selección purificadora ha variado considerablemente; mientras que los genes *ccrA* y *ccrB* están bajo una fuerte selección purificadora con valores dN/dS de 0.1040 y 0.0815 respectivamente, el gen *mecA* presenta un valor casi de un orden de magnitud más alta (dN/dS = 0.56470); lo cual, sugiere que en este caso la selección purificadora no ha sido tan intensa.

**Tabla 2. Promedio de la selección a lo largo de todo el gen, mediante el Modelo M0**

	Mutaciones no sinónimas (dN)	Mutaciones sinónimas (dS)	(dN/dS) $\omega$
<i>ccrA</i>	6.8327	65.6680	0.1040
<i>ccrB</i>	6.7122	82.2749	0.0815
<i>mecA</i>	17.1642	30.3950	0.5647

Tabla 2 | **Promedio de la selección a lo largo de todo el gen, mediante el Modelo M0** | El Modelo M0 se calculó para cada uno de los tres genes. Tanto el gen *ccrA* y *ccrB* presentaron menos mutaciones no sinónimas que el gen *mecA*, sin embargo este último gen presentó menos mutaciones sinónimas que *ccrA* y *ccrB*. Finalmente el cociente dN/dS ( $\omega$ ), sugiere que los tres genes están bajo selección purificadora.

Después usamos la comparación del Modelo 0 (M0) versus el Modelo 3 (M3) para determinar si hay variación a lo largo del gen; es decir, si el cociente dN/dS es diferente para cada sitio (codón). La comparación de modelos se lleva a cabo por medio de “*Likelihood Ratio Tests*” (LRT), ver métodos. Para la comparación de M0 vs M3, los genes *ccrA* y *ccrB* tuvieron LRTs significativos ( $P > 0.001$ ), con  $LRT= 257.4438$  y  $LRT=212.8695$  (respectivamente), por lo que ambos genes dieron positivo para la variación a lo largo del gen. Sin embargo, para el caso del gen *mecA*, no encontramos evidencia de variación a lo largo del gen ya que su valor LRT no fue significativo ( $P < 0.001$ ) ver tabla 3.

Dado que *ccrB* y *ccrA* presentaron variación a lo largo del gen –es decir los codones tenía diferentes dN/dS–, exploramos la posibilidad de que algunos sitios de estos genes hayan estado sujetos a selección positiva. Para esto, empleamos la comparación del Modelo 1a (M1a) versus Modelo 2a (M2a), ver métodos; tanto en *ccrB* como en *ccrA* el LRT no fue significativo (ver tabla 3), lo que indica que no parece haber sitios bajo selección positiva en estos dos genes.

**Tabla 3. Presión de selección por sitio (codones), a través de la comparación de modelos de sitio, para cada uno de los tres genes**

	Variación a lo largo del gen (M0 vs M3)	Sitios bajo selección positiva (M1a vs M2a)
<i>ccrA</i>	257.4438	-0.0013
<i>ccrB</i>	212.8695	0
<i>mecA</i>	0.0004	-

Tabla 3 | **Presión de selección por sitio (codones), para cada uno de los tres genes** | Los genes *ccrA* y *ccrB* presentaron variación en cada uno de los sitios, sin embargo ninguno de los sitios se encuentra bajo selección positiva. El gen *mecA* dio negativo a variación de sitios.

## DISCUSIÓN

### Las relaciones evolutivas verticales y las limitantes de eBURST

Para tratar de discernir la relación genéticas verticales entre los aislados, se usó eBURST. Al correr éste, obtuvimos 11 diagramas diferentes (>6/7 alelos compartidos), y no solo un diagrama, en donde estén representadas todas las relaciones de los *STs*, esto se debe a que aún faltan muchos aislados por tipificar, porque, aunque la base de datos de *MLST* (<http://www.MLST.net/>) para *S. aureus* en Enero de 2016 contiene 5,889 aislados y 2,812 *STs* para *S. aureus*, estos no son suficientes para mostrarnos todos los patrones de descendencia evolutiva vertical presentes para esta especie. Y esta misma problemática la vemos en estudios como el de Feil y colaboradores en 2004, donde la base *MLST* para ese entonces presentaba 1,072 aislados y 191 *STs*, ellos elaboran un diagrama eBURST Snapshot (0/7 alelos compartidos); sin embargo, su diagrama no representa todos los patrones de descendencia e inclusive tienen problemas para definir algunos *ST* fundadores. No obstante a pesar de esta falta de aislados por tipificar, los diagramas eBURST, al menos para *S. aureus* son lo bastante fiables debido a que es una especie que presenta bajas y moderadas tasas de recombinaciones y mutaciones para estos alelos (ver figura 13) (Enright et al., 2000).

## Comparación de las tasas de recombinación entre especies de tres estudios

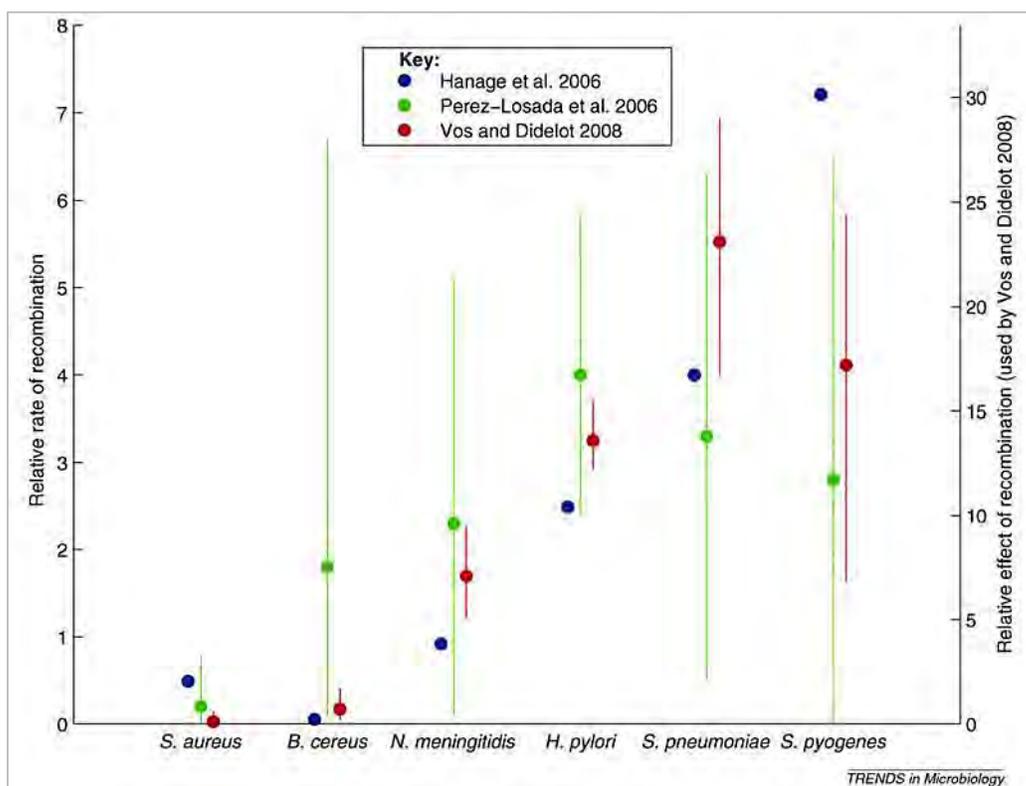


Fig. 13|Comparación de las tasas de recombinación entre especies de tres estudios| Los resultados de Hanage *et al.*, Pérez-Losada *et al.* y Vos y Didelot, son mostrados para las seis especies consideradas por los tres estudios, estas son: *S. aureus*, *B. cereus*, *N. meningitidis*, *H. pylori*, *S. pneumoniae* y *S. pyogenes*.

Dentro de nuestras 49 cepas, solo fue posible trabajar con 47 de sus STs debido a que dos cepas (55/2053 y 6850) no están tipificadas porque no están presentes en la base MLST. Por otra parte, la cepa JKD6159 con el ST93, fue denominada como “única”, debido a que no formó parte de ninguno de los 83 diagramas arrojados por eBURST, así como otros 383 STs pertenecientes a la base MLST. Entre otros factores se debe a que la cepa JKD6159, presenta un genoma lejanamente relacionado con otros genomas de

*S. aureus* previamente secuenciados (Chua *et al.*, 2010), abordando de nuevo la problemática ya descrita. Por lo que secuenciación de las diversas cepas de *S. aureus* facilitará la descripción de toda la gama de genotipos presentes por esta especie y proporcionara una base más robusta, para así tener un mejor entendimiento de los patrones de descendencia evolutiva.

El diagrama eBURST en el que está presente el CC5, CC8 y CC1 (ver figura 6) fue el que mayor número *STs* albergó tanto de nuestra base de datos como de la base *MLST*. Los *STs* de estos complejos en su mayoría albergaban cepas *MRSA* de dos tipos: *HA-MRSA* y *CA-MRSA*. El primer tipo presenta los *SCCmec* tipo II y III, mientras que el segundo tipo presenta el *SCCmec* tipo IV, que es pequeño –al resto de los demás *SCCmec*– y potencialmente más móvil. Es probable que la presencia de estos 3 tipos de *SCCmec* en los complejos clonales ya mencionados haya contribuido a su rápida expansión, ya que las cepas de *MRSA* tempranas de principios de 1960, que portaban *SCCmec* tipo I fueron sustituidas por linajes más exitosos pertenecientes al CC5 y CC8 que portaban *SCCmec* II y *SCCmec* III, debido a que las recombinasas del *SCCmec* I eran defectuosas limitando su TGH. Mientras que el *SCCmec* tipo IV parece conferir poco o ningún coste en el organismo (Huang *et al.*, 2006; Chambres y DeLeo, 2009).

Por otra parte, también dentro de los complejos como: CC5, CC133, CC151, CC425 y CC398, se encontraron cepas *LA-MRSA*. Entre los máximos representantes de este tipo de *MRSA* es el ST398 (CC398), que se ha extendido rápidamente en todo el mundo y se asocia principalmente con la colonización de ganado demostrando altas tasas de prevalencia alrededor del 49% en cerdos y terneros (de Neeling *et al.*, 2007). Sin embargo, también se ha reportado que este *ST* puede causar varias infecciones en seres

humanos como: infecciones de tejidos blandos, en las vías respiratorias, bacteriemia y endocarditis (Huijsdens *et al.*, 2006; Ekkelenkamp *et al.*, 2006; Declercq *et al.*, 2008). Esto demuestra que la adaptación de las cepas a distintos huéspedes mamíferos genera un amplio reservorio de genes que pueden ayudar a la propagación de *S. aureus* (Lindsay, 2010).

### **Los patrones de TGH y el papel de la selección**

La evolución y éxito de muchas clonas de *S. aureus* se ha debido, entre otras cosas a la transferencia horizontal de *MGEs*. Idealmente, el nuevo *MGE* adquirido debe ser estable y no causar ningún coste significativo y proporcionar alguna ventaja selectiva para permitir la rápida expansión y diseminación de la cepa. El presente trabajo se centró en la historia evolutiva de los genes que componen al *MGE*, un elemento de gran importancia para muchas cepas de *S. aureus*.

En este estudio, los datos sugieren una historia evolutiva compleja para el *SCCmec*. A continuación, se describen los tres patrones más relevantes que observamos.

Primero, no todas las cepas presentaron el *SCCmec*, de hecho más de un quinto de las cepas (10 de las 49 cepas) no lo tuvieron, estas fueron las cepas 55/2053, 6850, ED133 (CC133), RF122 (CC151), 71193 (CC398), ED98, ECT-R2 (CC5) y NCTC 8325, Newman, VC40 (CC8), probablemente se deba a que el sitio de unión del *SCCmec* – denominado *attB*–, presenta una composición genética diferente, por lo que la inserción mediada por las recombinasas no sea efectiva o bien esta región puede estar ocupada por otro elemento genético móvil, ya que es un punto de inserción común (Noto *et al.*, 2008). Un buen ejemplo de esto es la cepa ED133. En esta cepa, el elemento *SCCmec*

fue sustituido por genes que codifican un elemento de inserción, una transposasa, y una DNA helicasa (Saov0027), estos fueron encontrados en *attB*. También la cepa *ED133* presentó otros *MGEs* –distintos al *SCCmec*, como una isla de patogenicidad denominada, SaPlov2, que en vez de otorgar una resistencia a antibióticos, le confiere la capacidad para coagular plasma de rumiantes–, que la han permitido dispersarse en ovinos, dejando en claro el papel central que juegan los *MGEs* en la adaptación de las bacterias a diferentes nichos (Guinane *et al.*, 2010).

Además, ciertas cepas solo presentaron algunos de los genes pero no todos. Por ejemplo, las cepas *08BA02176*, *M013*, *SA957*, solo tiene el gen *mecA* y las cepas COL y *S. aureus ST398*, albergan el gen *mecA* y el gen *ccrA*, todas ellas *MRSA* y solo dos de ellas (*08BA02176* y *S. aureus ST398*) son *LA-MRSA*. Es muy probable que en estas cepas el *SCCmec* no se pueda transferir horizontalmente ya que carecen del gen *ccrB* o de todo el complejo *ccr*, encargado de codificar para recombinasas sitio específico y cuya presencia de ambos genes –*ccrA* y *ccrB* –son necesarios para el proceso de escisión e integración del *SCCmec* (Katayama *et al.*, 2000; Wang y Archer, 2010; Misiura *et al.*, 2013). En contraparte, se encontró que la cepa *MSSA476* presentó los genes *ccr* pero no *mecA*; la cual, ha sido descrita como sensible a meticilina. Esto sugiere –para los casos en los que no hay gen *mecA*–, que el fenotipo resistente a meticilina no es esencial en todos los nichos de *S. aureus*, pues una característica común de las bacterias sometidas a la adaptación de nicho es la pérdida de la función de los genes superfluos o perjudiciales para las bacterias en el nuevo hábitat (Eppinger *et al.*, 2006; Herron-Olson *et al.*, 2007; Stinear *et al.*, 2007).

Finalmente, las cepas 10388, 10497, 15532, 16035, 16125, 18341, 18412 y 18583, que pertenecen al CC5 –todas *MRSA*–, presentaron 2 secuencias homólogas para el gen *ccrB*, la primera de ellas mide ~630 pb y la segunda mide ~210 pb, ninguna de ellas funcional, por lo que se descartaron en los posteriores análisis debido a que introducían demasiados gaps (deleciones) generando filogenias menos fiables. Sin embargo presentaron valores altos para ser considerados como homólogos. A su vez, la cepa M1 (CC8) (reportada como *CA-MRSA*), albergó 3 secuencias homólogas para el gen *ccrA*, denotadas como M1, M1a y M1b, esta última no es funcional y tan solo mide 276 pb, por lo que fue descartada, debido a que introducía bastantes gaps. Esta cepa también presentó 2 secuencias homólogas para el gen *ccrB*, de manera que tiene dos juegos de complejos *ccr*, perfectamente funcionales. Las cuatro secuencias homólogas han estado sujetas a diferentes eventos de TGH (ver figuras 8,10 y 11) e inclusive tienen diferentes tiempos de aparición (ver figura 8). La presencia de dos complejos clonales sugiere, que puede haber una sobre expresión de proteínas recombinasas *ccrA* y *ccrB*, lo que provocaría la escisión del *SCCmec* (Noto y Archer, 2006). Se ha descrito que la cepa M1 presenta el *SCCmec* tipo IV subtipo a (Larner-Svensson et al., 2013), este *SCCmec* ha sido responsable de la última epidemia (alrededor del 2000) a nivel mundial causada por *CA-MRSA* y se encuentra ampliamente distribuido entre los aislados de *S. aureus* (Lina et al., 2006), por lo que se le ha considerado como uno de los *SCCmec* más exitosos, quizá parte de este éxito se deba a la capacidad de escindir e integrarse debido a los 2 juegos de complejos *ccr*.

Segundo, cada uno de los genes tiene una historia evolutiva particular, ya que las filogenias presentan diferentes topologías y diferentes niveles de variación genética (ver tabla 2 y 3). Varios estudios han demostrado que el SCC y *mecA* han existido como componentes genéticos individuales, que se originaron y evolucionaron en diferentes comensales animales (Couto *et al.*, 1996; Tsubakishita *et al.*, 2010a, 2010b; Wu *et al.*, 1998); lo cual, claramente concuerda con nuestros resultados. En las filogenias se aprecia que las secuencias que pertenecen al mismo complejo clonal se encuentran en diferentes partes de los arboles filogenéticos y que no forman grupos monofiléticos (ver figura 8 y 9); estos son claros eventos de transferencia horizontal. Notablemente, en dos de estos genes (*ccrA* y *ccrB*) se encontraron claras señales de recombinación lo cual provee evidencia adicional de casos de transferencia horizontal. Claramente hay una amplia diferencia en los niveles de variación genética (ver valores de diversidad filogenética de las figuras 10,11 y 12, así como la tabla 2). En este sentido el gen *mecA* es el que ha acumulado menos cambios, con una diversidad filogenética apenas arriba de uno. Mientras que los genes *ccrB* y *ccrA* tuvieron diversidades muy parecidas acercándose a 1.5.

Finalmente los tres genes parecen estar bajo un proceso de selección purificadora ya que presentaron mayores mutaciones sinónimas que mutaciones no sinónimas (ver tabla 2): esto concuerda con lo ya reportado por Castillo-Ramírez y colaboradores en 2011 y 2012. Ellos mencionan que *S. aureus* presenta una alta tasa de recombinación y un enriquecimiento de mutaciones sinónimas –así como bajos radios dN/dS–, en sus *MGES* (ver figura 14). Tanto *ccrB* como *ccrA*, presentaron señales de recombinación, estos dos genes presentaron radios más bajos de dN/dS implicando una selección purificadora más

intensa o que han estado expuesto durante un mayor tiempo al efecto de la selección. En ese sentido, mientras que para estos dos genes se pudo detectar variación de dN/dS a lo largo de la secuencia del gen, no se detectó dicha variación en el caso de *mecA*. De manera interesante, en ningún de los tres genes fue posible detectar sitios que se encuentran bajo selección positiva. Lo anterior implicaría que, cuando los genes están presentes en algunas cepas, estos parecen estar confiriendo alguna ventaja –y que estos genes presentan restricciones funcionales importantes–, ya que la mayoría de las mutaciones no sinónimas están siendo eliminadas.

### Funciones afectadas por la recombinación

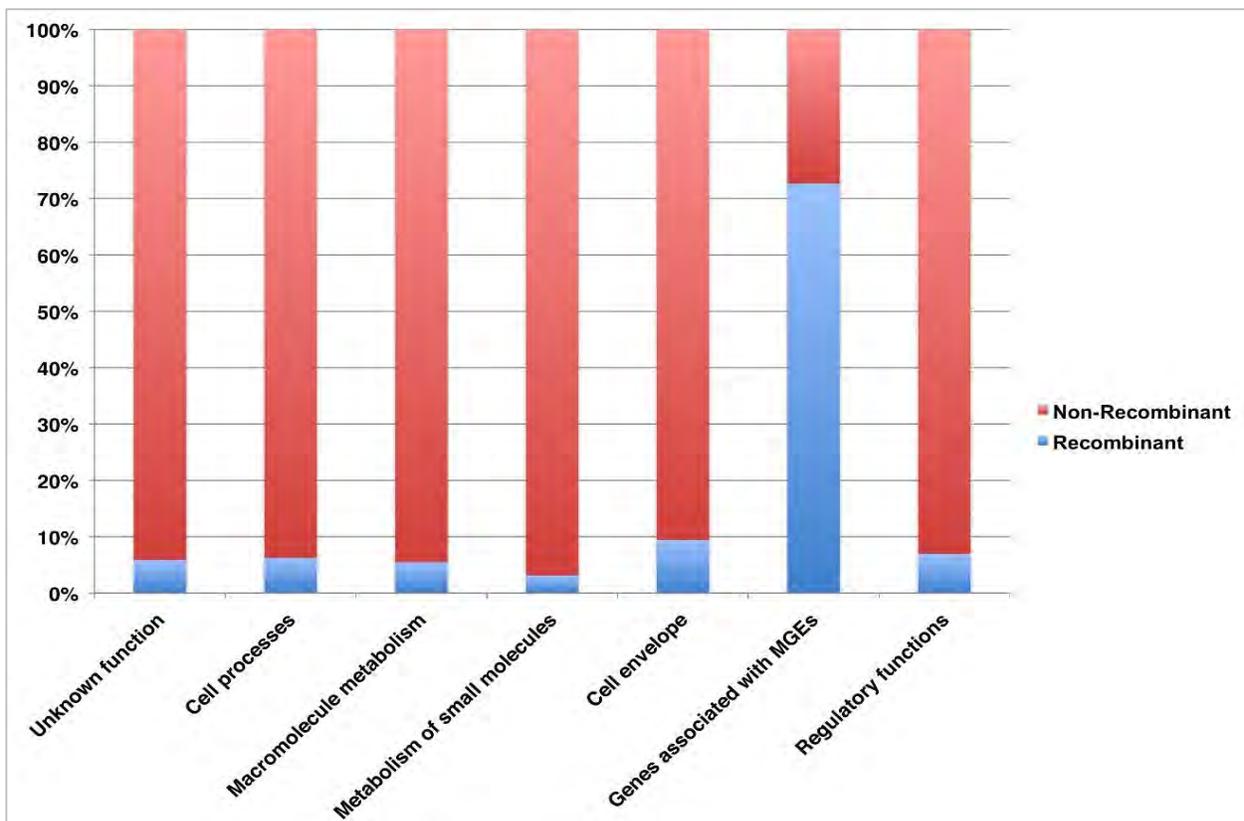


Fig. 14 | **Funciones afectadas por la recombinación** | La sección azul de las barras representa el porcentaje de genes de cada categoría funcional afectada por recombinación. La categoría “extracromosómico” aquí se renombró *MGE* para evitar confusiones ya que estos genes se encuentran físicamente en el cromosoma (Castillo-Ramírez *et al.*, 2012).

## CONCLUSIONES

La pérdida de genes o incluso de todo el elemento *SCCmec*, parece ser un evento frecuente, y esto, muy probablemente se asocia a los requerimientos del hábitat. En ese sentido, la transferencia y pérdida de genes han jugado un papel importante en la historia evolutiva de los diferentes elementos *SCCmec*, permitiendo que cada gen tenga una historia evolutiva particular. Aunque los tres genes se encuentran bajo selección purificadora, esta parece ser más intensa y variar a lo largo del gen para los genes *ccrA* y *ccrB*. A la vez, estos son los dos genes afectados por la recombinación.

## PERSPECTIVA

Para un estudio más fino de los procesos evolutivos del *SCCmec*, deben de considerarse varias cosas: primero se necesita un mayor número de cepas *MSSA* secuenciadas, segundo es fundamental conocer el tipo *SCCmec* –presente en la cepa –y si este se asocia a hospitales, a la comunidad o ganado. Tercero considerar trabajar con las regiones *Junkyard*, ya que en estas regiones, se pueden encontrar otros tipos de elementos genéticos que participen en la adaptación de la cepa. Cuarto, establecer los sitios en la proteína donde hubo eventos de mutaciones sinónimas (dS) y no sinónimas (dN) y cómo éstas afecta la conformación de la proteína o los sitios activos. Y por último, sería interesante corroborar que en la cepa M1 los dos juegos del complejo *ccr* participan en la escisión del *SCCmec*.

## REFERENCIAS

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J Mol Biol*, 215:403–410.
- Anderson, D. J., Kaye, K. S., Chen, L. F., Schmader, K. E., Choi, Y., Sloane, R., & Sexton, D. J. (2009). Clinical and Financial Outcomes Due to Methicillin Resistant *Staphylococcus aureus* Surgical Site Infection: A Multi-Center Matched Outcomes Study. *PLoS ONE*, 4(12), e8305. <http://doi.org/10.1371/journal.pone.0008305>.
- Arêde, P., Milheiriço, C., de Lencastre, H., & Oliveira, D. C. (2012). The Anti-Repressor MecR2 Promotes the Proteolysis of the *mecA* Repressor and Enables Optimal Expression of  $\beta$ -lactam Resistance in MRSA. *PLoS Pathogens*, 8(7), e1002816. <http://doi.org/10.1371/journal.ppat.1002816>.
- Boundy, S., Safo, M. K., Wang, L., Musayev, F. N., O'Farrell, H. C., Rife, J. P., & Archer, G. L. (2013). Characterization of the *Staphylococcus aureus* rRNA Methyltransferase Encoded by *orfX*, the Gene Containing the Staphylococcal Chromosome Cassette *mec* (SCC*mec*) Insertion Site. *The Journal of Biological Chemistry*, 288(1), 132–140. <http://doi.org/10.1074/jbc.M112.385138>.
- Bustos-Martínez, J.H., Hamdan-Partida, A., Gutiérrez-Cárdenas, M. (2006). *Staphylococcus aureus*: la reemergencia de un patógeno en la comunidad. *Rev Biomed*, 17:287–305.
- Carroll, D., Kehoe, M.A., Cavanagh, D., Coleman, D.C. (1995). Novel organization of the site-specific integration and excision recombination functions of the

*Staphylococcus aureus* serotype F virulence-converting phages phi 13 and phi 42. *Mol Microbiol*, 16: 877–893.

Castillo-Ramírez S, Harris SR, Holden MTG, He M, Parkhill J, et al. (2011) The Impact of Recombination on dN/dS within Recently Emerged Bacterial Clones. *PLoS Pathog* 7(7): e1002129. doi: 10.1371/journal.ppat.1002129.

Castillo-Ramírez, S., Corander, J., Marttinen, P., Aldeljawi, M., Hanage, W. P., Westh, H., ... Feil, E. J. (2012). Phylogeographic variation in recombination rates within a global clone of methicillin-resistant *Staphylococcus aureus*. *Genome Biology*, 13(12), R126. <http://doi.org/10.1186/gb-2012-13-12-r126>.

Chambers, H. F., & DeLeo, F. R. (2009). Waves of Resistance: *Staphylococcus aureus* in the Antibiotic Era. *Nature Reviews. Microbiology*, 7(9), 629–641. <http://doi.org/10.1038/nrmicro2200>.

Chen, C.-J., Unger, C., Hoffmann, W., Lindsay, J. A., Huang, Y.-C., & Götz, F. (2013). Characterization and Comparison of 2 Distinct Epidemic Community-Associated Methicillin-Resistant *Staphylococcus aureus* Clones of ST59 Lineage. *PLoS ONE*, 8(9), e63210. <http://doi.org/10.1371/journal.pone.0063210>.

Chua, K., Seemann, T., Harrison, P. F., Davies, J. K., Coutts, S. J., Chen, H., ... Stinear, T. P. (2010). Complete Genome Sequence of *Staphylococcus aureus* Strain *JKD6159*, a Unique Australian Clone of ST93-IV Community Methicillin-Resistant *Staphylococcus aureus*. *Journal of Bacteriology*, 192(20), 5556–5557. <http://doi.org/10.1128/JB.00878-10>.

- Chua, L.Y.K., Howden, P.J., Jiang, Hang-Jih., Stinear, T., Peleg, Y.A. (2014). Population genetics and the evolution of virulence in *Staphylococcus aureus*. *Infection, Genetics and Evolution*, 21, 554–562.
- Couto, I., de Lencastre, H. E. R. M. I. N. I. A., Severina, E., Kloos, W., Webster, J. A., Hubner, R. J., ... & Tomasz, A. (1996). Ubiquitous presence of a mecA homologue in natural isolates of *Staphylococcus sciuri*. *Microbial Drug Resistance*, 2(4), 377–391.
- de Neeling, A.J., van den Broek, M.J.M., Spalburg, E.C., van Santen-Verheuver, M.G., Dam-Deisz, W.D.C., Boshuizen, H.C., van de Giessen A.W., van Duijkeren E., Huijsdens, X.W. (2007). High prevalence of methicillin resistant *Staphylococcus aureus* in pigs. *Vet Microbiol*, 122, 366–372. doi: 10.1016/j.vetmic.2007.01.027.
- Declercq, P., Petre, D., Gordts, B., Voss, A. (2008). Complicated community-acquired soft tissue infection by *MRSA* from porcine origin. *Infection*, 36, 590–592.
- Deurenberg, R. H., & Stobberingh, E. E. (2008). The evolution of *Staphylococcus aureus*. *Infection genetics and evolution*, 8(6), 747–763.
- Ekkelenkamp, M.B., Sekkat, M., Carpaij, N., Troelstra, A., Bonten, M.J. (2006). [Endocarditis due to methicillin-resistant *Staphylococcus aureus* originating from pigs] *Ned Tijdschr Geneesk*, 150, 2442–2447.
- Enright, M. C., Day, N. P. J., Davies, C. E., Peacock, S. J., & Spratt, B. G. (2000). Multilocus Sequence Typing for Characterization of Methicillin-Resistant and Methicillin-Susceptible Clones of *Staphylococcus aureus*. *Journal of Clinical Microbiology*, 38(3), 1008–1015.
- Eppinger, M, Baar, C, Linz, B, Raddatz, G, Lanz, C, *et al.*, (2006) Who Ate Whom? Adaptive *Helicobacter* Genomic Changes That Accompanied a Host Jump from Early

Humans to Large Felines. *PLoS Genet*, 2(7): e120. doi: 10.1371/journal.pgen.0020120.

Feil, E. J., Li, B. C., Aanensen, D. M., Hanage, W. P., & Spratt, B. G. (2004). eBURST: Inferring Patterns of Evolutionary Descent among Clusters of Related Bacterial Genotypes from Multilocus Sequence Typing Data. *Journal of Bacteriology*, 186(5), 1518–1530. <http://doi.org/10.1128/JB.186.5.1518-1530.2004>.

Foster, T. J. (2004). The *Staphylococcus aureus* “superbug.” *Journal of Clinical Investigation*, 114(12), 1693–1696. <http://doi.org/10.1172/JCI200423825>.

Frost, L.S, Leplae, R., Summers, A.O., Toussaint, A. (2005). Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol*, 3, 722–732.

Galtier, N., Gouy, M. and Gautier, C. (1996). SeaView and Phylo\_win, two graphic tools for sequence alignment and molecular phylogeny. *Comput Applic Biosci*, 12, 543–548.

Goerke, C., Wirtz, C., Flückiger, U., Wolz, C. (2006). Extensive phage dynamics in *Staphylococcus aureus* contributes to adaptation to the human host during infection. *Mol Microbiol*, 61, 1673–1685.

Guinane, C. M., Zakour, N. L. B., Tormo-Mas, M. A., Weinert, L. A., Lowder, B. V., Cartwright, R. A., ... & Fitzgerald, J. R. (2010). Evolutionary genomics of *Staphylococcus aureus* reveals insights into the origin and molecular basis of ruminant host adaptation. *Genome biology and evolution*, evq031.

- Guindon, S. y Gascue, O. (2003). A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Syst Bio* 52(5), 696–704.
- Hacker, J., Kaper, J.B. (2000). Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol*, 54, 641–679.
- Hanssen, A. M., & Sollid, J. U. E. (2006). *SCCmec* in staphylococci: genes on the move. *FEMS Immunology & Medical Microbiology*, 46(1), 8-20.
- Herron-Olson, L., Fitzgerald, J. R., Musser, J. M., & Kapur, V., (2007). Molecular Correlates of Host Specialization in *Staphylococcus aureus*. *PLoS ONE*. 2(10). e1120. <http://doi.org/10.1371/journal.pone.0001120>.
- Hiramatsu, K., Cui, L., Kuroda, M., & Ito, T. (2001). The emergence and evolution of methicillin-resistant *Staphylococcus aureus*. *Trends in microbiology*, 9(10), 486–493.
- Holden, M. T. G., Feil, E. J., Lindsay, J. A., Peacock, S. J., Day, N. P. J., Enright, M. C., ... Parkhill, J. (2004). Complete genomes of two clinical *Staphylococcus aureus* strains: Evidence for the rapid evolution of virulence and drug resistance. *Proceedings of the National Academy of Sciences of the United States of America*, 101(26), 9786–9791. <http://doi.org/10.1073/pnas.0402521101>.
- Holden, M. T. G., Hsu, L.-Y., Kurt, K., Weinert, L. A., Mather, A. E., Harris, S. R., ... Nübel, U. (2013). A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant *Staphylococcus aureus* pandemic. *Genome Research*, 23(4), 653–664. <http://doi.org/10.1101/gr.147710.112>
- Howe, R.A., Brown, N.M. y Spencer, R.C. (1996). The new threats of Gram positive pathogens: re-emergence of things past. *J Clin Pathol*, 49(6), 444–449.

- Huang, H., Flynn, N. M., King, J. H., Monchaud, C., Morita, M., & Cohen, S. H. (2006). Comparisons of Community-Associated Methicillin-Resistant *Staphylococcus aureus* (MRSA) and Hospital-Associated MRSA Infections in Sacramento, California. *Journal of Clinical Microbiology*, 44(7), 2423–2427. <http://doi.org/10.1128/JCM.00254-06>.
- Huang, T. W., Chen, F. J., Miu, W. C., Liao, T. L., Lin, A. C., Huang, I. W., ... & Lauderdale, T. L. Y. (2012). Complete genome sequence of *Staphylococcus aureus* M013, a pvl-positive, ST59-SCCmec type V strain isolated in Taiwan. *Journal of bacteriology*, 194(5), 1256-1257.
- Huijsdens, X. W., van Dijke, B. J., Spalburg, E., van Santen-Verheuevel, M. G., Heck, M. E., Pluister, G. N., ... de Neeling, A. J. (2006). Community-acquired MRSA and pig-farming. *Annals of Clinical Microbiology and Antimicrobials*, 5, 26. <http://doi.org/10.1186/1476-0711-5-26>.
- Huson H.D. & Bryant, D. (2006). Application of Phylogenetic Networks in Evolutionary Studies. *Mol Biol Evol*, 23(2), 254–267.
- Ito, T., Okuma, K., Ma, X.X., Yuzawa, H. y Hiramatsu, K. (2003). Insights on antibiotic resistance of *Staphylococcus aureus* from its whole genome: genomic island SCC. *Drug Resist. Updat*, 6(1), 41–52.
- Jain, R., Rivera, M. C., & Lake, J. A. (1999). Horizontal gene transfer among genomes: the complexity hypothesis. *Proceedings of the National Academy of Sciences*, 96(7), 3801–3806.
- Kanafani, Z.A & Fowler, V.G. (2006). *Staphylococcus aureus* infections: new challenges from an old pathogen. *Enferm Infec Microbiol Clin*, 24,182–193.

- Katayama, Y., Ito, T., & Hiramatsu, K. (2000). A New Class of Genetic Element, Staphylococcus Cassette Chromosome *mec*, Encodes Methicillin Resistance in *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, 44(6), 1549–1555.
- Keeling, P.J. & Palmer, J.D. (2008). Horizontal gene transfer in eukaryotic evolution. *Nat Rev Genet*, 9, 605–618.
- Kondo, Y., Ito, T., Ma, X. X., Watanabe, S., Kreiswirth, B. N., Etienne, J., & Hiramatsu, K. (2007). Combination of multiplex PCRs for staphylococcal cassette chromosome *mec* type assignment: rapid identification system for *mec*, *ccr*, and major differences in junkyard regions. *Antimicrobial Agents and Chemotherapy*, 51(1), 264–274.
- Larner-Svensson, H., Worning, P., Bartels, M. D., Hestbjerg Hansen, L., Boye, K., & Westh, H. (2013). Complete Genome Sequence of *Staphylococcus aureus* Strain M1, a Unique t024-ST8-IVa Danish Methicillin-Resistant *S. aureus* Clone. *Genome Announcements*, 1(3), e00336–13. <http://doi.org/10.1128/genomeA.00336-13>
- Lina, G., Durand, G., Berchich, C., Short, B., Meugnier, H., Vandenesch, F., ... & Enright, M. C. (2006). Staphylococcal chromosome cassette evolution in *Staphylococcus aureus* inferred from *ccr* gene complex sequence typing analysis. *Clinical microbiology and infection*, 12(12), 1175-1184.
- Lindsay, J. A. (2010). Genomic variation and evolution of *Staphylococcus aureus*. *International Journal of Medical Microbiology*, 300(2), 98-103.

- Lindsay, J. A., & Holden, M. T. (2004). *Staphylococcus aureus*: superbug, super genome?. *Trends in microbiology*, 12(8), 378–385.
- Lowy, D.F. (2003). Antimicrobial resistance: the example of *Staphylococcus aureus*. *J Clin Invest*, 111, 1265–1273.
- Malachowa, N., & DeLeo, F. R. (2010). Mobile genetic elements of *Staphylococcus aureus*. *Cellular and Molecular Life Sciences*, 67(18), 3057–3071.  
<http://doi.org/10.1007/s00018-010-0389-4>
- Martin, P.D., Lemey, P., Lott, M., Moulton, V., Posada, D., Lefevre, P. (2010). RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics*, 26, 2462–2463.
- McClintock, B. (1950). The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences*, 36(6), 344–355.
- McClintock, B. (1951, January). Chromosome organization and genic expression. In *Cold Spring Harbor Symposia on Quantitative Biology* (Vol. 16, pp. 13-47). Cold Spring Harbor Laboratory Press.
- Misiura, A., Pigli, Y. Z., Boyle-Vavra, S., Daum, R. S., Boocock, M. R., & Rice, P. A. (2013). Roles of two large serine recombinases in mobilizing the methicillin-resistance cassette SCCmec. *Molecular Microbiology*, 88(6), 1218–1229.  
<http://doi.org/10.1111/mmi.12253>
- Molina, A., Del Campo, R., Máiz, L., Morosini, María-Isabel, Lamas, A., Baquero, F. & Cantón, R. (2008). High prevalence in cystic fibrosis patients of multiresistant hospital-acquired methicillin-resistant *Staphylococcus aureus* ST228-SCC mec

- I capable of biofilm formation. *Journal of Antimicrobial Chemotherapy*, 62, 961–967.
- Moore, P. C., & Lindsay, J. A. (2001). Genetic variation among hospital isolates of methicillin-sensitive *Staphylococcus aureus*: evidence for horizontal transfer of virulence genes. *Journal of clinical microbiology*, 39(8), 2760–2767.
- Noto, M. J., Kreiswirth, B. N., Monk, A. B., & Archer, G. L. (2008). Gene Acquisition at the Insertion Site for *SCCmec*, the Genomic Island Conferring Methicillin Resistance in *Staphylococcus aureus*. *Journal of Bacteriology*, 190(4), 1276–1283. <http://doi.org/10.1128/JB.01128-07>.
- Rice, P., Longden, I. & Bleasby, A. (2000). EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genet*, 16(6), 276–277.
- Seong, W. J., Kwon, H. J., Kim, T. E., Lee, D. Y., Park, M. S., and Kim, J. H. (2012). Comparison of complete *rpoB* gene sequence typing and multi-locus sequence typing for phylogenetic analysis of *Staphylococcus aureus*. *J Gen Appl Microbiol*, 59, 335–343.
- Shore, C.A., Coleman, C.D. (2013). Staphylococcal cassette chromosome mec: Recent advances and new insights. *Int J Med Microbiol*, 303, 350–359.
- Spratt, B. G., Hanage, W. P., Li, B., Aanensen, D. M., & Feil, E. J. (2004). Displaying the relatedness among isolates of bacterial species—the eBURST approach. *FEMS Microbiology Letters*, 241(2), 129–134.
- Stegger, M., Driebe, E. M., Roe, C., Lemmer, D., Bowers, J. R., Engelthaler, D. M., Keim P.A., Andersen, P. S. (2013). Genome sequence of *Staphylococcus aureus*

strain CA-347, a USA600 methicillin-resistant isolate. *Genome announcements*, 1(4): e00517-13.

Stinear, T. P., Seemann, T., Pidot, S., Frigui, W., Reysset, G., Garnier, T., ... Cole, S. T. (2007). Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. *Genome Research*, 17(2), 192–200. <http://doi.org/10.1101/gr.5942807>

Tamura, K., Stecher, G., Peterson, D., Filipiński, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*, 30, 2725–2729.

Tsubakishita, S., Kuwahara-Arai, K., Baba, T., Hiramatsu, K. (2010a). Staphylococcal cassette chromosome mec-like element in *Micrococcus caseolyticus*. *Antimicrob Agents Chemother*, 54(4), 1469–1475.

Tsubakishita, S., Kuwahara-Arai, K., Sasaki, T., Hiramatsu, K. (2010b). Origin and molecular evolution of the determinant of methicillin resistance in staphylococci. *Antimicrob Agents Chemother*, 54(10), 4352–4359.

Velázquez-Meza, M.E. (2005). Surgimiento y diseminación de *Staphylococcus aureus* meticilinoresistente. *Salud Pública Méx.*, 47, 381–387.

Wang, L., & Archer, G. L. (2010). Roles of *CcrA* and *CcrB* in Excision and Integration of Staphylococcal Cassette Chromosome *mec*, a *Staphylococcus aureus* Genomic Island. *Journal of Bacteriology*, 192(12), 3204–3212. <http://doi.org/10.1128/JB.01520-09>.

- Wu, S., de Lencastre, H., Tomasz, A. (1998). Genetic organization of the *mecA* region in methicillin-susceptible and methicillin-resistant strains of *Staphylococcus sciuri*. *J Bacteriol*, 180(2), 236–242.
- Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci*, 13(5), 555–556.
- Zhang, K., McClure, J., Elsayed, S., Louie, T., Conly, J. (2005). Novel multiplex PCR assay for characterization and concomitant subtyping of staphylococcal cassette chromosome *mec* types I to V in methicillin-resistant *Staphylococcus aureus*. *J Clin Microbiol*, 43, 5026–5033.

# GLOSARIO

## Glosario de abreviaciones

- **attB**: sitio de unión del SCCmec en el cromosoma de *S. aureus*.
- **CA-MRSA**: Del inglés *Community-Acquired Methicillin-resistant Staphylococcus aureus* (*Staphylococcus aureus* resistente a meticilina asociado a comunidad).
- **CC**: Complejo clonal.
- **ccr**: Complejo de genes, formado por los genes *ccrA* y *ccrB*.
- **ccrA**: Gen que codifica para una recombinasa sitio específica.
- **ccrB**: Gen que codifica para una recombinasa sitio específica.
- **dN**: Mutaciones no sinónimas.
- **DR**: Del inglés *Direct Repeats* (secuencias de repetición directa).
- **dS**: Mutaciones sinónimas.
- **HA-MRSA**: Del inglés *Hospital-Associated Methicillin-resistant Staphylococcus aureus* (*Staphylococcus aureus* asociado a hospitales).
- **LA-MRSA**: Del inglés *Livestock-Associated Methicillin-resistant Staphylococcus aureus* (*Staphylococcus aureus* asociado a ganado).
- **LRT**: Del inglés *Likelihood Ratio Test* (Test de máxima verosimilitud).
- **mecA**: Gen que codifica para proteínas de bajo unión a penicilina (PBP2a), confiere resistencia.
- **MGE**: *Mobile Genetic Elements* (Elementos genéticos móviles).
- **MLST**: Del inglés *Multi Locus Sequence Typing* (base de datos).

- **MRSA:** Del inglés *Methicillin-resistant Staphylococcus aureus* (*Staphylococcus aureus* resistente a meticilina).
- **orfX:** Gen que codifica para una metiltransferasa ribosomal.
- **PBP:** Del inglés *Penicillin-binding proteins* (proteínas de unión a penicilina).
- **PBP2a o PBP2A:** Del inglés *Penicillin-binding proteins 2 a* (proteínas de baja unión a penicilina 2a).
- **Región J o junkyard:** Región hipervariable del *SCCmec*
- **SCC:** Del inglés *Staphylococcal Cassette Chromosome*
- **SCCmec:** Del inglés *Staphylococcal Cassette Chromosome*, que porta al complejo de genes *mec*, que confieren resistencia a múltiples antibióticos.
- **ST:** Del inglés *Sequence Type* (secuencia tipo).
- **TGH:** Transferencia genética horizontal.
- **ω:** Cociente de mutaciones no sinónimas (dN) y sinónimas (dS).

## Glosario de términos

**Alineamiento de secuencias:** En bioinformática, es una forma de representar y comparar dos o más secuencias de ADN, ARN, o estructuras primarias proteicas.

**Complejo Clonal:** Conjunto de bacterias que tienen al menos una variación en el perfil alélico.

**Homólogo:** Cuando dos secuencias de proteínas o de DNA son similares, decimos que son homólogas, ya que tienen el mismo ancestro, la misma función y la misma estructura.

**Intrínseco:** Que es propio o característico de la cosa que se expresa por sí misma y no depende de las circunstancias.

**Likelihood Ratio Test:** Es un Test que nos permite comparar modelos de sitios para tratar de inferir el tipo de selección que actúa sobre cada uno de los codones.

**Modelos de selección de sitio:** Nos permiten determinar la presión de selección por sitio, es decir por codones.

**PCR anidado:** La PCR anidada conocida como Nested PCR es una variante de la PCR convencional que comprende dos rondas de amplificación con distintos pares de cebadores en cada una con el fin de incrementar la sensibilidad y la especificidad de la detección. Primero, se realiza una reacción con los cebadores externos para amplificar una región de ADN más extensa, que contiene el segmento diana. Después, este producto de amplificación se utiliza como molde de una segunda PCR con los cebadores internos para amplificar la región específica.

**Perfiles alélicos:** Es un método altamente discriminatoria para aislados bacterianos, para esto se utilizan generalmente 7 genes de mantenimiento o genes *housekeeping*. Las variaciones en cada uno de los 7 alelos distintos define el perfil alélico o *sequence type* (secuencia tipo).

**Resistencia heterogénea:** Es cuando una pequeña porción de la población tolera los antibióticos; mientras que la mayor parte de la población, no.

**Secuencia de Inserción:** Son los transposones más simples, contienen un gen de una transposasa flanqueado por secuencias repetidas e invertidas (IR) que reconoce la transposasa.

**Secuencias de repetición directa (DR):** Los transposones siempre están flanqueados por secuencias repetidas directas (DR) que proceden del DNA al que se integraron (no confundir con las secuencias repetidas e invertidas de los extremos del transposón). Las transposasas producen cortes no enfrentados en los sitios de inserción de transposones, generándose extremos cohesivos. Tras la inserción del transposón, éste va a quedar flanqueado por huecos que al ser rellenados por la DNA polimerasa, van a originar las repeticiones directas.

**Selección positiva:** Es el mecanismo evolutivo mediante el cual nuevos mutantes poseen adecuaciones mayores que el promedio de la población, y las frecuencias de dichos mutantes se incrementan en la siguiente generación.

**Selección neutral:** Es la variación a nivel molecular que ocurre de manera aleatoria y no tiene importancia adaptativa.

**Selección purificadora:** Los nuevos mutantes poseen adecuaciones menores que el promedio de la población, y la frecuencia de estos mutantes disminuye en las siguientes generaciones.

**Transposón:** Los transposones son segmentos pequeños de DNA que pueden trasladarse (ser transpuestos) de una región de una molécula de DNA a otra.

## ANEXOS

### Anexo 1. Alineamientos con TRANALIGN

De la página <http://emboss.sourceforge.net/> se descargó el programa EMBOSS. Con la herramienta TRANALIGN, se alinearon las secuencias de DNA usando como referencia secuencias proteicas (Rice *et al.*, 2000), generando así alineamientos más robustos para su posterior análisis filogenético. Para construir los alineamientos, se crearon 2 archivos (ver figura 15); el primero contenía a las secuencias de nucleótidos, y el segundo las secuencias proteicas –es decir de amino ácidos–. En ambos archivos se debe de llevar el mismo orden consecutivo de las cepas. Finalmente, se cargaron ambos archivos en el programa, para posteriormente correrlo y obtener el alineamiento.

\*Nota: Para cada uno de los genes se creó los dos tipos de archivo.

### Anexo 2. Evaluación de la selección natural y la deriva génica

#### Modelos de sitio y el Test de máxima verosimilitud (LRT) para su comparación

El cociente dN/dS (radio  $\omega$ ) es un parámetro usado para medir la dirección y la intensidad de la selección natural que está actuando sobre la proteína. Este cociente se puede calcular para estimar la presión de selección que actúa sobre toda la secuencia o bien para calcular la presión de selección por sitio, es decir por codones (ver figura 16), esto mediante modelos de selección de sitio (ver tabla 4). Además con el radio test de máxima verosimilitud (*Likelihood Ratio Test*, por sus siglas en inglés *LRT*), podemos comparar los modelos de sitios para tratar de inferir el tipo de selección que actúa sobre cada uno de los codones. La Tabla 5 muestra cómo se obtuvo los *LRTs* para el gen *ccrA*, con una  $P=0.001$ , este procedimiento se usó para los genes restantes.

## Alineamientos con TRANALIGN

### a) Archivo con secuencias de nucleótidos

```
>TW20
ATGAAAAAGATAAAAATTGTTCCACTTATTTTAATAGTTGTAGTTGTCGGGTTTGGTATATATTT
TTATG...
>z172
ATGAAAAAGATAAAAATTGTTCCACTTATTTTAATAGTTGTAGTTGTCGGGTTTGGTATATATTT
TTATG...
>Bmb9393
ATGAAAAAGATAAAAATTGTTCCACTTATTTTAATAGTTGTAGTTGTCGGGTTTGGTATATATTT
TTATG...
>SA40
ATGAAAAAGATAAAAATTGTTCCACTTATTTTAATAGTTGTAGTTGTCGGGTTTGGTATATATTT
TTATG...
```

### b) Archivo de secuencias proteicas

```
>TW20
MKKIKIVPLILIVVVVGFGIYFYASKDKEINNTIDAIEDKNFKQVYKDSSYISKSDNGEVEMTE
RPIKIY...
>z172
MKKIKIVPLILIVVVVGFGIYFYASKDKEINNTIDAIEDKNFKQVYKDSSYISKSDNGEVEMTE
RPIKIY...
>Bmb9393
MKKIKIVPLILIVVVVGFGIYFYASKDKEINNTIDAIEDKNFKQVYKDSSYISKSDNGEVEMTE
RPIKIY...
>SA40
MKKIKIVPLILIVVVVGFGIYFYASKDKEINNTIDAIEDKNFKQVYKDSSYISKSDNGEVEMTE
RPIKIY...
```

Fig. 15|Alineamientos con TRANALIGN| a) Archivo con las secuencias de nucleótidos, de color rojo se indica el nombre de la cepa, seguido de su respectiva secuencia. b) Archivo con las secuencias proteicas, el color azul indica el nombre de la cepa, seguido de su secuencia y los tres puntos en ambas imágenes indican que la secuencia continua.

## Presión de selección en la secuencia

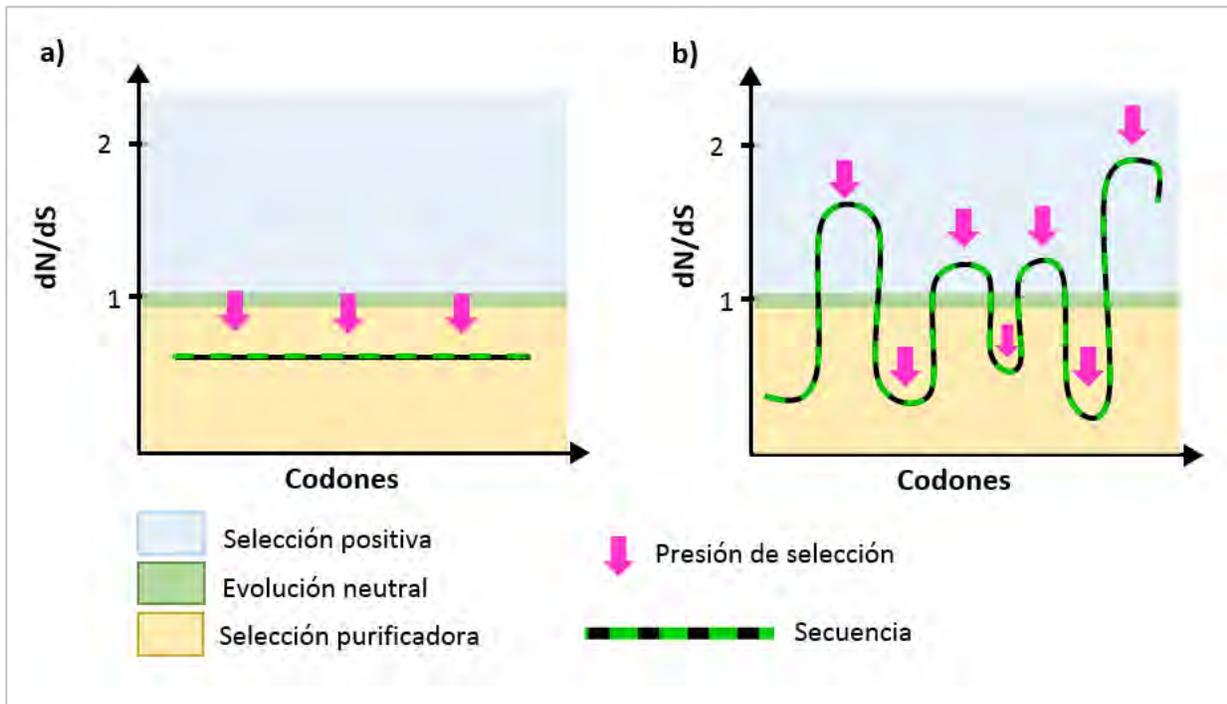


Fig. 16|Presión de selección en la secuencia| a) Se asume la misma presión de selección para todos los sitios, es decir una sola  $\omega$ , esto es dado por el modelo M0. Aunque el promedio de la selección de toda la secuencia sea negativo ( $dN/dS < 1$ ) como se ve en la figura "a", b) esto no quiere decir que no existan sitios particulares dentro de la secuencia que se encuentren seleccionados positivamente ( $dN/dS > 1$ ), dado por los modelos (M0 vs M3 y M1a vs M2a), como se muestra en la figura "b".

**Tabla 4. Información sobre los Modelos de sitio**

<b>M0</b>	Este modelo asume la misma presión de selección (una sola $\omega$ ) para todos los sitios
<b>M0 vs M3</b>	Este modelo permite determinar si hay variación a lo largo del gen; es decir, si el cociente $dN/dS$ es diferente para cada sitio (codón). En este modelo, M0 implica un sólo valor de $dN/dS$ para todo el gen y M3 implica que hay diferentes valores de $dN/dS$ para el gen.
<b>M1a vs M2a</b>	Permite determinar si hay sitios que se encuentran bajo selección positiva. Donde M1a permite variación en $dN/dS$ pero no con valores de $dN/dS$ mayores a uno: es decir, no hay selección positiva mientras que en M2a hay variación en los valores de $dN/dS$ y permite valores de $dN/dS$ mayores a uno (selección positiva).

**Tabla 5. Comparación entre los modelos a partir del *Likelihood Ratio Test (LRT)***

Modelo	Fórmula $2 \times (\ell_{\text{alta}} - \ell_0)$	<i>LRT</i>	Grado de Libertad	Valor de Ji-Cuadrada
M0 vs M3	$2((-8866.5050) - (-8995.226922))$	257.4438	4	18.4668
M1a vs M2a	$2((-8920.211798) - (-8920.211133))$	- 0.0013	2	13.8155

Tabla 5 | **Modelos de sitio** | Los datos presentados corresponden al gen *ccrA*. Los logaritmos naturales son dados por PALM, esto se toman y se sustituyen en la formula, los grados de libertad se calcula de la siguiente manera: al modelo con el mayor número de parámetros se le resta el modelo con el menor número de parámetros. Finalmente con los grados de libertad y con tú P (0.001) se busca el valor de Ji cuadrada en la tabla. Sí tu valor de *LRT* es mayor que su valor de Ji cuadrada entonces decimos que si es significativo y se corre el segundo test M1a vs M2a y se repite el procedimiento. Sí el segundo teste es significativo, este sugiere que hay sitios (codones) que están bajo selección positiva o diversificadora.

### **Anexo 3. Breve definición de los diagramas eBURST**

Los diagramas eBURST muestran las relaciones entre aislados estrechamente relacionados de una especie bacteriana o población. Utiliza un modelo simple, en el que un genotipo ancestral (fundador) incrementa su frecuencia en la población, y mientras lo hace, comienza a diversificar, produciendo grupos –que presentan genotipos relacionados al genotipo fundador–, cercanamente a él. Este conjunto de genotipos relacionados se refiere a menudo como un Complejo Clonal (CC).

El algoritmo BURST utiliza los *Sequence Types (STs)* y perfiles alélicos de la base de datos *Multi Locus Sequence Typing (MLST)* para formar los grupos de genotipos; para esto todos los miembros asignados a un grupo comparten alelos idénticos (> 6 de los 7 loci) con al menos otro miembro del grupo –generalmente se ocupan 7 genes, denominados, genes de mantenimiento o genes *housekeeping*–. Las variantes del fundador que difieren en sólo uno de los siete loci son conocidos como variantes de locus simple (del inglés *Single Locus Variant, SLV*). Para más información, pueden consultar la siguiente página <http://eburst.mlst.net/v3/instructions/>.

#### **Diagramas eBURST de *S. aureus***

Aquí se muestran el resto de los diagramas eBURST, cada complejo clonal genero su propio diagrama –presentándose un diagrama por complejo–. Los complejos fueron: el CC45 (ver figura 17), CC22 (ver figura 18), CC59 (ver figura 19), el CC80 (ver figura 20), el CC75 (ver figura 21), el CC133 (ver figura 22), el CC151 (ver figura 23), el CC398 (ver figura 24), el CC425 (ver figura 25), el nombre del complejo clonal corresponde al *ST* fundador.

La aparición de cada uno de estos complejos clonales en su respectivo diagrama nos sugiere que los genotipos fundadores –presentes en cada diagrama, indicados por los círculos azules–, han sido exitosos por lo que han podido diversificar, generando así propio complejo clonal.

Diagrama eBURST del CC45

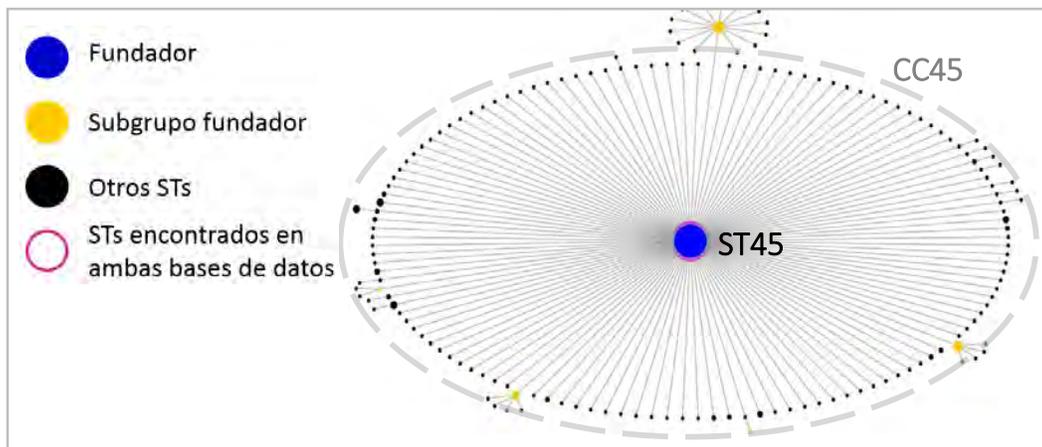


Fig. 17 | **Diagrama eBURST del CC45** | El óvalo punteado de color gris señala al complejo clonal 45, conformado por el ST45, siendo este ST fundador.

Diagrama eBURST del CC22

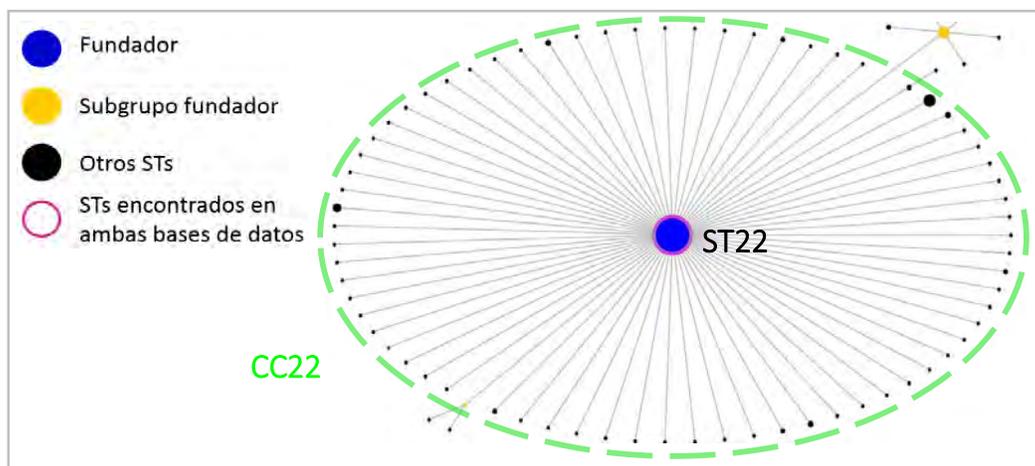


Fig. 18 | **Diagrama eBURST del CC22** | El óvalo punteado de color verde señala al complejo clonal 22, conformado por el ST22, siendo este ST fundador.

### Diagrama eBURST del CC59

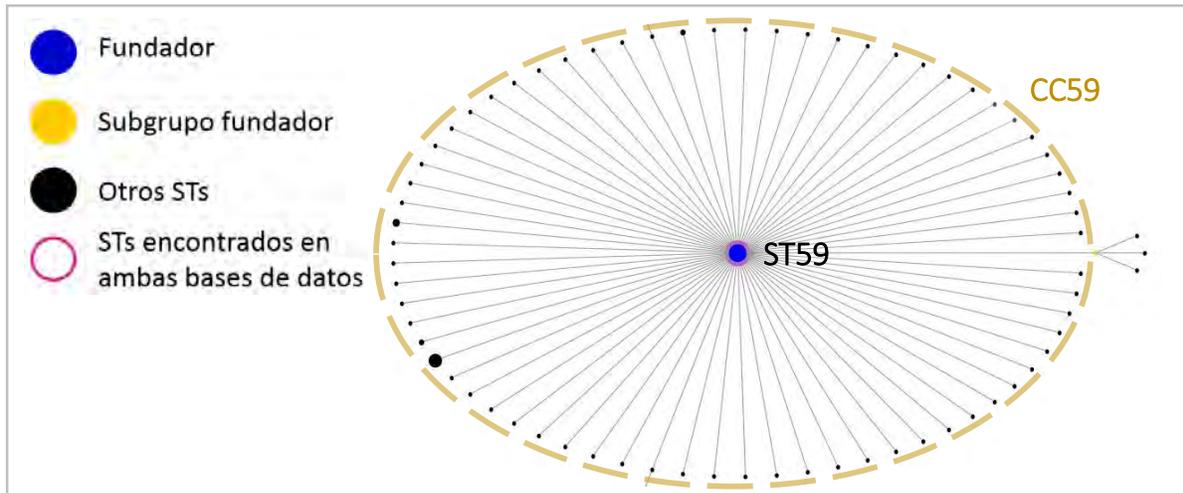


Fig. 19 | **Diagrama eBURST del CC22** | El óvalo punteado de color dorado señala al complejo clonal 59, conformado por el *ST59*, siendo este *ST* fundador.

### Diagrama eBURST del CC80

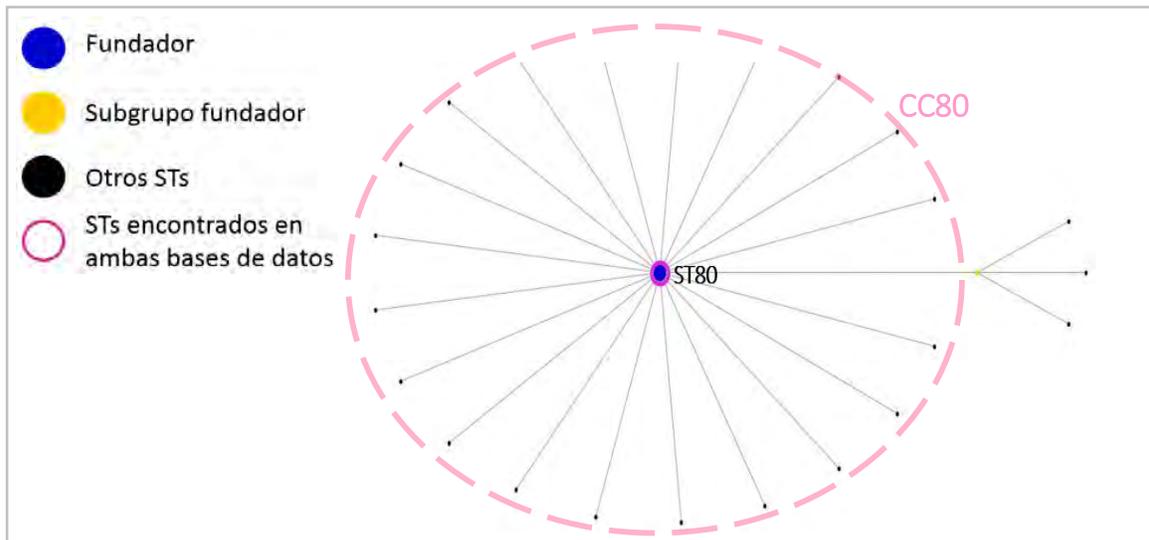


Fig. 20 | **Diagrama eBURST del CC80** | El óvalo punteado de color rosa señala al complejo clonal 80, conformado por el *ST80*, siendo este *ST* fundador.

### Diagrama eBURST del CC75

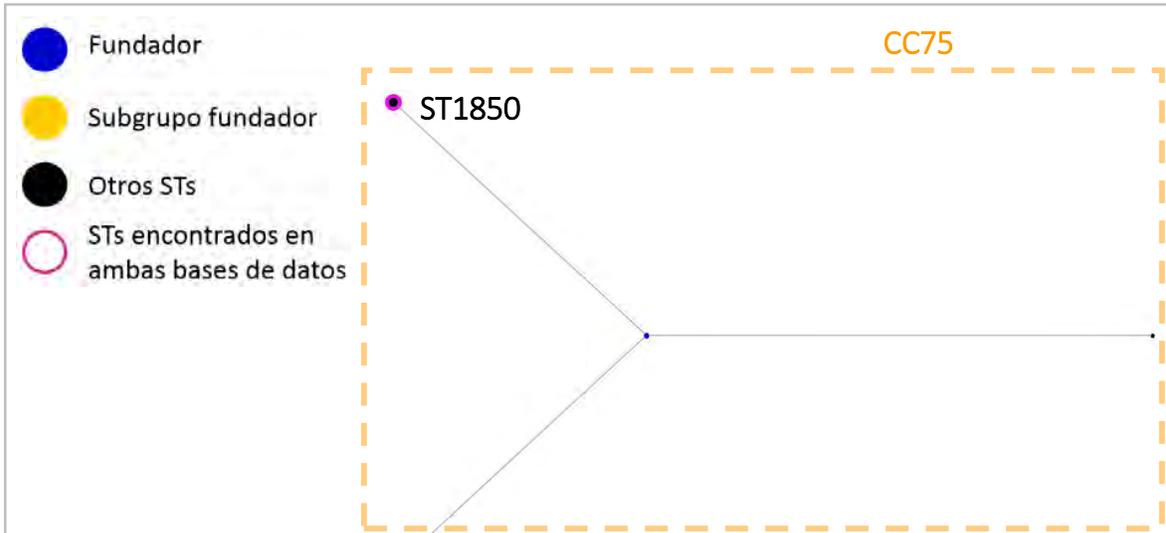


Fig. 21 | **Diagrama eBURST del CC75** | El rectángulo punteado de color naranja señala al complejo clonal 75, conformado por el *ST1850*, siendo este un *SLV*.

### Diagrama eBURST del CC133

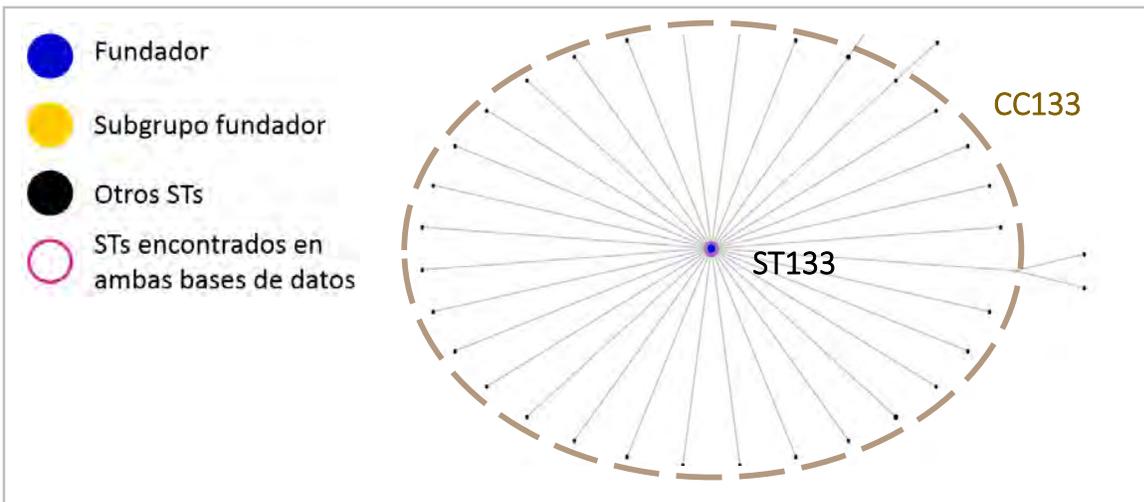


Fig. 22 | **Diagrama eBURST del CC133** | El óvalo punteado de color café señala al complejo clonal 133, conformado por el *ST133*, siendo este un *ST* fundador.

### Diagrama eBURST del CC151

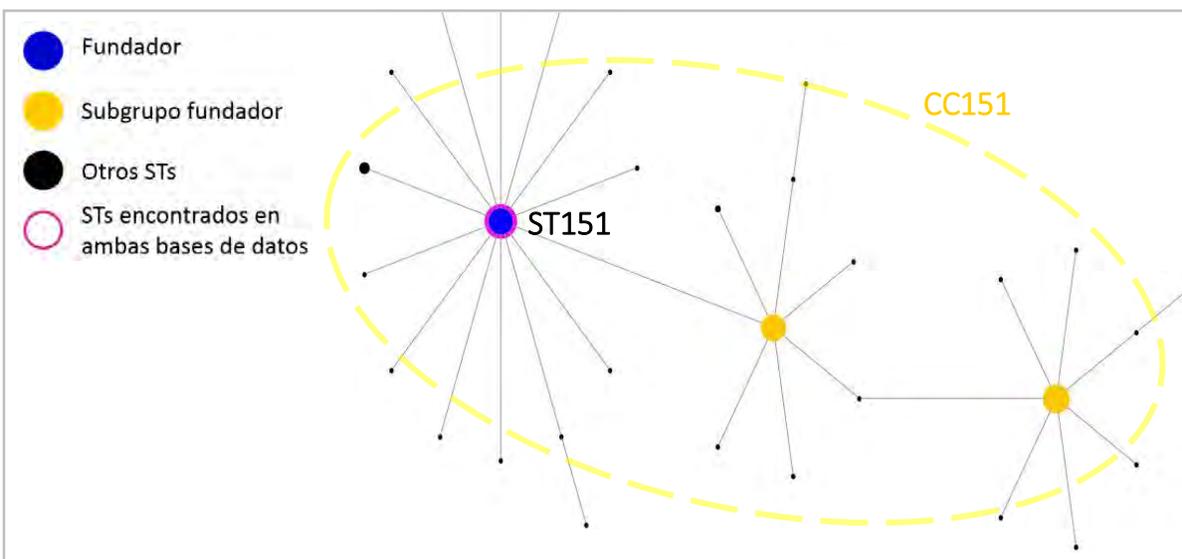


Fig. 23 | **Diagrama eBURST del CC151** | El óvalo punteado de color amarillo señala al complejo clonal 151, conformado por el *ST151*, siendo este un *ST* fundador.

### Diagrama eBURST del CC398

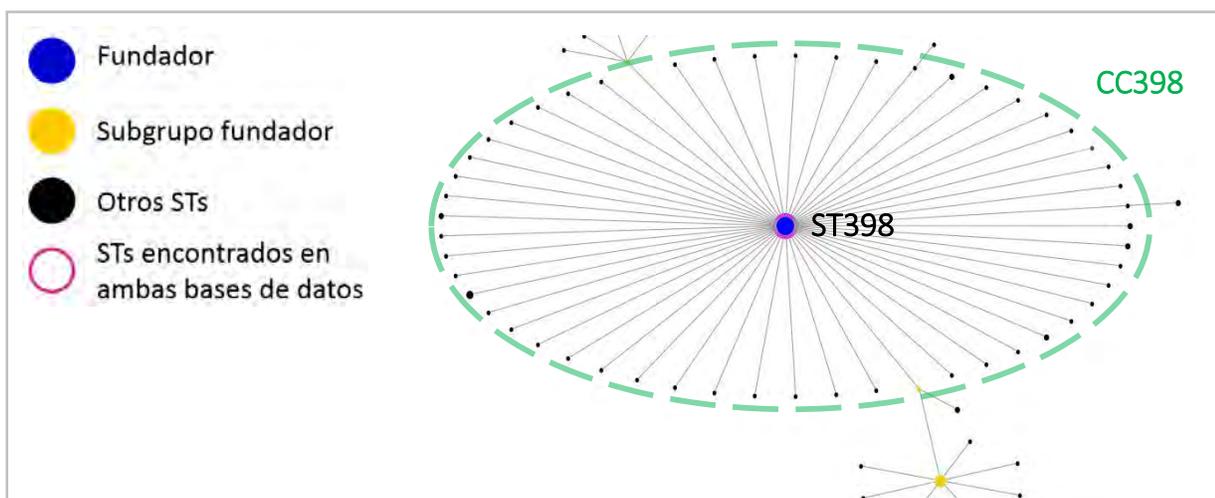


Fig. 24 | **Diagrama eBURST del CC398** | El óvalo punteado de color verde señala al complejo clonal 398, conformado por el *ST398*, siendo este un *ST* fundador.

## Diagrama eBURST del CC425

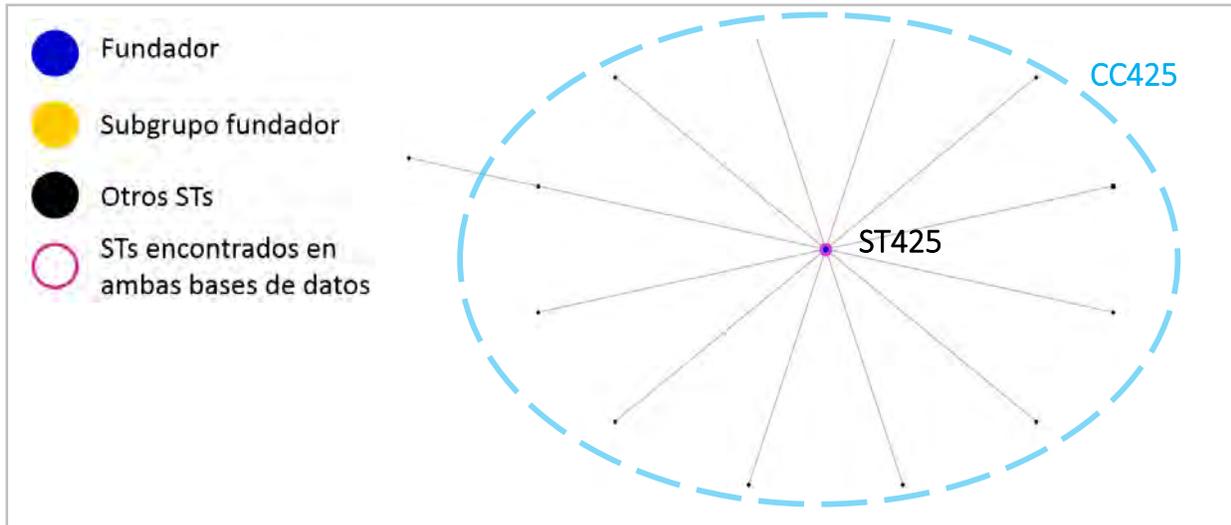


Fig. 25 |**Diagrama eBURST del CC425**| El óvalo punteado de color azul claro señala al complejo clonal 425, conformado por el *ST425*, siendo este un *ST* fundador.

## Anexo 4. Análisis de recombinación con RDP

RDP es una única herramienta que permite analizar grandes alineamientos de secuencias con una gama de diferentes métodos de detección de la recombinación. Para este estudio, se realizaron los test dados por default: RDP, GENECONV, Bootscan, MaxChi, Chimera y Siscan. Obteniendo un total de 3,988 posibles eventos de recombinación para el gen *ccrA* (ver figura 26).

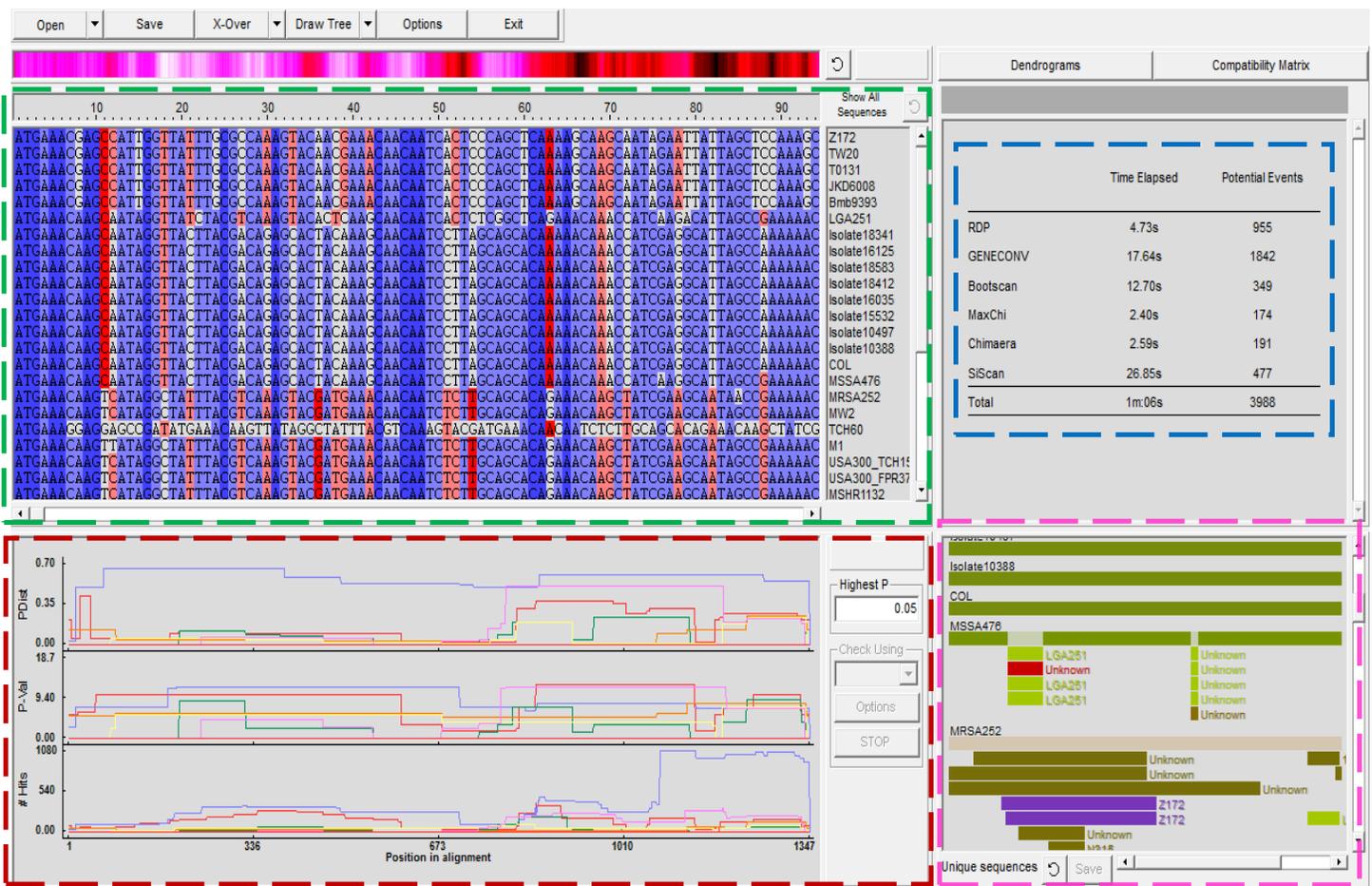


Fig. 26 | Análisis de recombinación con RDP para el gen *ccrA* | En la parte superior se encuentra la barra de herramientas, el rectángulo punteado verde indica las secuencias alineadas y sus respectivos nombres. El rectángulo azul señala el total de eventos de recombinación obtenidos por los distintos test, aplicados por RDP. El rectángulo rojo muestra los sitios de recombinación y finalmente el rectángulo rosa engloba las secuencias esquematizadas.

Para el gen *ccrB* hubo 3,104 posibles eventos de recombinación y finalmente el gen *mecA*, no presento ningún evento de recombinación (Ver figura 27).

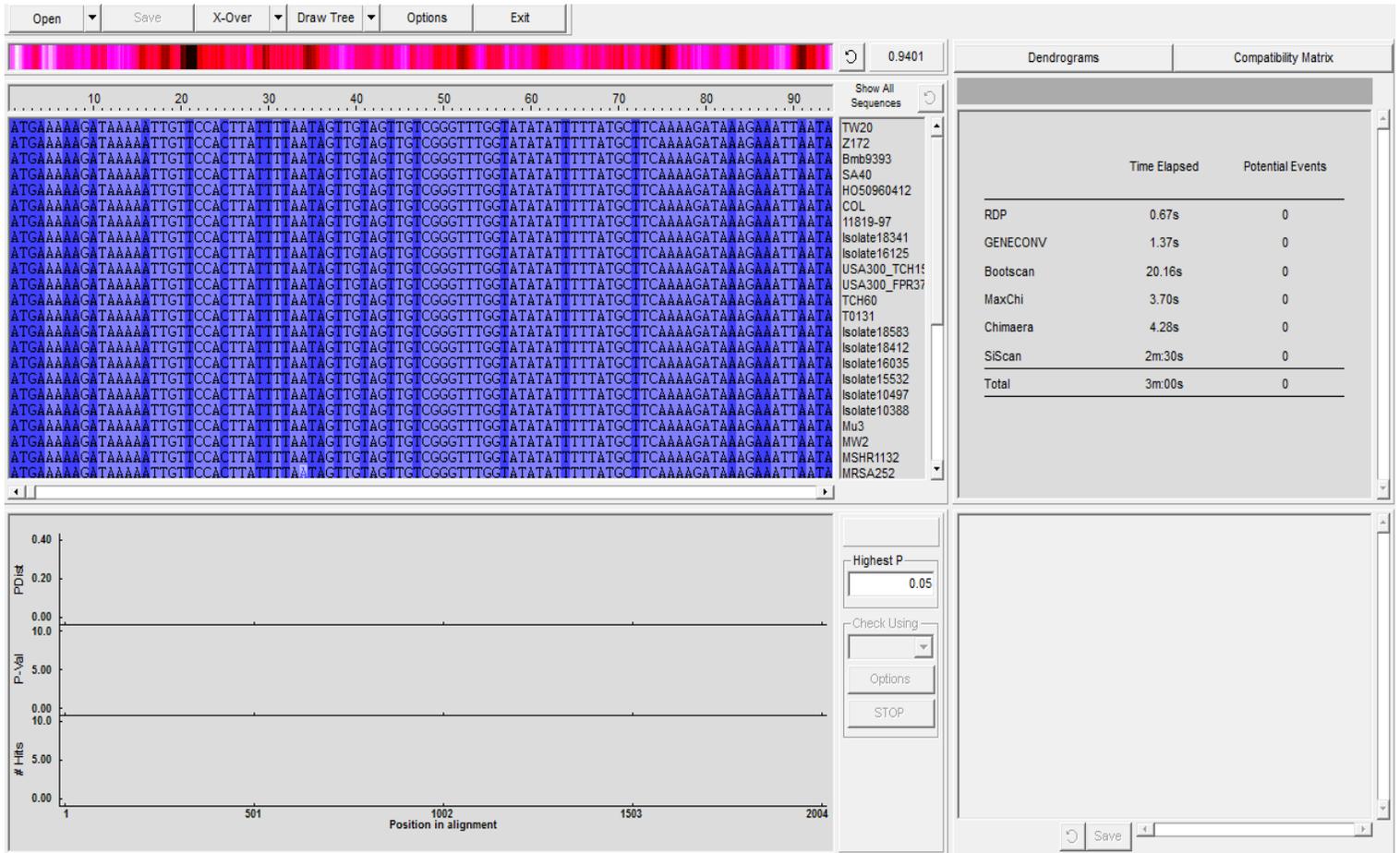


Fig. 27 |Análisis de recombinación con RDP para el gen *mecA*| Los seis test realizados por RDP (RDP, GENECONV, Bootscaan, MaxChi, Chimera y Siscan) dieron valores de cero por lo que indica que no hay eventos de recombinación para el gen *mecA*.