



UNIVERSIDAD DE BUENOS AIRES

FACULTAD DE FARMACIA Y BIOQUÍMICA

Tesis de maestría en Biología Molecular Médica

**“Estudio molecular y funcional del gen *cassette aadA1*”**

Autora: Lic. Angélica Galán Delgado

Directora: Dra. Daniela Centrón

Directora Adjunta: Dra. María Paula Quiroga

Año 2017

Laboratorio de investigaciones en mecanismos de resistencia a antibióticos

Instituto de Investigaciones en Microbiología y Parasitología Médica (IMPAM) UBA/CONICET

Universidad de Buenos Aires

A mi mamá y papá, a mis hermanas adoradas  
y a mis sobrinos Lucas y Mariana

## **AGRADECIMIENTOS**

Quiero agradecer a la Dra. Daniela Centrón por abrirme las puertas del laboratorio y poder desarrollar mis proyectos de investigación bajo su dirección, así mismo también agradezco a la doctora María Paula Quiroga por la asistencia en los diseños experimentales y revisión del documento.

Principalmente a mis padres y a mi familia por el apoyo constante para llevar a cabo mis proyectos personales y profesionales aunque eso signifique estar lejos de ellos, especialmente a mis hermanas Mónica y Nubia que son la otra parte de mi corazón y las extraño todos los días.

A mis amigos de Colombia que también a la distancia me apoyan y me dan ánimo para que siga adelante a pesar de las dificultades, en especial a mi mejor amiga Lorena Chaparro.

A gente del L8 en especial a Ceci Rodríguez, Eli, Gise, Ceci Quiroga, Sabri, Sofi, que se han llegado a convertir en amigas muy importantes desde que estoy trabajando en el laboratorio.

A toda la gente del IMPAM, en especial a la directora Cristina Cerquetti quien siempre tiene una sonrisa y una palabra de aliento.

A todos los chicos del becalario, pero especialmente a Cami, Nico y Gera, quienes son casi como mis hermanos y quienes no me dejan caer y me apoyan a pesar de todo.

Y en general a toda la gente con la que interactúe durante estos 6 años de larga travesía, a los que me dieron buenos y malos momentos, porque sin ellos no hubiera crecido y valorado la gran oportunidad de vivir en este hermoso país.

# ÍNDICE GENERAL

## DEDICATORIA

## AGRADECIMIENTOS

## ÍNDICE DE TABLAS

## ÍNDICE DE FIGURAS

## ABREVIATURAS

|   |    |
|---|----|
| Capítulo 1: Introducción .....  | 1  |
| 1.1- El rol de los integrones en la Transferencia Horizontal Genética.....  | 1  |
| 1.2- Clasificación de los integrones.....   | 3  |
| 1.3- Sitio de recombinación <i>attI</i> .....   | 5  |
| 1.4- Sitios <i>attC</i> .....   | 6  |
| 1.5- Características genéticas de los genes <i>cassettes</i> .....  | 13 |
| 1.6- Reacciones de recombinación en integrones.....   | 14 |
| 1.7- Identificación del gen <i>cassette aadA1</i> en el ambiente.....   | 17 |
| Capítulo 2: Hipótesis y Objetivos.....  | 21 |
| 2.1- Hipótesis .....  | 21 |
| 2.2- Objetivos.....   | 21 |
| 2.2.1- Objetivo general.....  | 21 |
| 2.2.2- Objetivos específicos.....   | 21 |
| Capítulo 3: Materiales y métodos .....  | 23 |
| 3.1- Análisis de bioinformática de los genes <i>cassette aadA-like</i> . .....  | 23 |
| 3.1.1- Estudio de frecuencia y arreglos de los alelos <i>aadA-like</i> . .....  | 23 |
| 3.1.2- Análisis de los arreglos de los genes <i>cassettes aadA-like</i> de integrones clínicos.....   | 24 |
| 3.1.3- Dispersión de las variantes de los sitios <i>attC</i> de los genes <i>cassette aadA-like</i> y del sitio <i>attC<sub>aadA1</sub></i> ..... | 24 |
| 3.2- Ensayos funcionales de recombinación para el gen <i>cassette aadA1</i> .....   | 26 |
| 3.2.1- Selección del modelo a estudiar .....  | 26 |

|  |    |
|--|----|
| 3.2.2- Plásmidos recombinantes y cepas utilizadas en este estudio .....  | 26 |
| 3.2.3- Reactivos utilizados de rutina .....  | 27 |
| 3.2.4- Condiciones de crecimiento de los cultivos bacterianos.....   | 28 |
| 3.2.5- Extracción de ADN genómico total y de ADN plasmídico.....   | 28 |
| 3.2.6- Amplificación por reacción en cadena de la polimerasa (PCR) .....   | 28 |
| 3.2.7- Secuenciación de ADN y análisis de secuencias. ....   | 30 |
| 3.2.8- Preparación de células competentes y transformación bacteriana .....  | 30 |
| 3.2.9- Ensayos de escisión de genes <i>cassettes in vivo</i> mediados por <i>Int1</i> .....  | 31 |
| 3.2.10- Ensayos de inserción de <i>cassettes</i> en el sitio <i>attI1 in vivo</i> mediados por <i>Int1</i> .....                               | 33 |
| Capítulo 4: Resultados y Discusión .....   | 36 |
| 4.1- Análisis de bioinformática de los alelos del gen <i>cassette aadA-like</i> . ....   | 36 |
| 4.1.1- Identificación de los alelos de los genes <i>cassettes aadA-like</i> . ....   | 36 |
| 4.1.2- Análisis de bioinformática aplicada a cada alelo <i>aadA-like</i> : abundancia y región variable de integración .....                   | 38 |
| 4.1.2.1- Abundancia de cada alelo <i>aadA-like</i> .....   | 38 |
| 4.1.2.2- Plataformas genéticas del gen <i>cassette aadA1</i> en integrones de clase 1 .....  | 39 |
| 4.1.2.3- Plataformas genéticas del gen <i>cassette aadA1</i> en integrones de clase 2. ....  | 43 |
| 4.1.3- Análisis de dispersión del gen <i>cassette aadA1</i> . ....   | 44 |
| 4.1.4- Plataformas genéticas del gen <i>cassette aadA2</i> . ....  | 52 |
| 4.1.5- Plataformas genéticas del gen <i>cassette aadA5</i> . ....  | 53 |
| 4.1.6- Plataformas genéticas de otros alelos de genes <i>cassettes aadA-like</i> menos frecuentes. ....  | 54 |
| 4.2- Estudio de los componentes genéticos de los genes <i>cassettes aadA-like</i> : marco abierto de lectura (ORF) y sitios <i>attCs</i> ..... | 56 |
| 4.2.1- Estructura genética de los ORFs de los alelos <i>aadA-like</i> . ....   | 56 |
| 4.2.2- Estructura genética de los sitios <i>attCs</i> de los alelos de los genes <i>cassettes aadA-like</i> . ....                             | 58 |
| 4.2.3- Dispersión de las variantes de los sitios <i>attCs</i> de los genes <i>cassettes aadA-like</i> . ...                                    | 63 |
| 4.3- Estructuras secundarias de los sitios <i>attC</i> usados en los estudios de recombinación...  | 65 |

|   |     |
|---|-----|
| 4.3.1- Descripción de la región LH en los sitios <i>attC</i> de los genes <i>cassettes</i> : <i>attC<sub>aadA1</sub></i> ,<br><i>attC<sub>aadA3</sub></i> , <i>attC<sub>aacA1</sub></i> y <i>attC<sub>dfrA1</sub></i> .....   | 66  |
| 4.3.2- Descripción de la región RH en los sitios <i>attC</i> de los genes <i>cassettes</i> : <i>attC<sub>aadA1</sub></i> ,<br><i>attC<sub>aadA3</sub></i> , <i>attC<sub>aacA1</sub></i> y <i>attC<sub>dfrA1</sub></i> .....   | 70  |
| 4.3.3- Secuencias espaciadoras de las regiones LH y RH .....  | 72  |
| 4.3.4- Bases extrahelicoidales de los sitios <i>attC<sub>aadA1</sub></i> , <i>attC<sub>aadA3</sub></i> , <i>attC<sub>aacA1</sub></i> y <i>attC<sub>dfrA1</sub></i> .....  | 73  |
| 4.3.5- Influencia sobre el proceso de recombinación de los diferentes sitios <i>attC</i><br>analizados: <i>attC<sub>aadA1</sub></i> , <i>attC<sub>aadA3</sub></i> , <i>attC<sub>aacA1</sub></i> y <i>attC<sub>dfrA1</sub></i> .....   | 75  |
| 4.4- Estudio funcional del gen <i>cassette aadA1</i> mediante ensayos de recombinación .....  | 79  |
| 4.4.1- Evaluación de la capacidad de Int1 para escindir a los genes <i>cassettes attI1-</i><br><i>aadA1-attC<sub>aadA1</sub></i> , <i>attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub></i> , <i>attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub></i> y <i>attC<sub>dhfr1</sub>-aadA1-</i><br><i>attC<sub>aadA1</sub></i> .....              | 79  |
| 4.4.1.1- Verificación del evento de escisión del 100% .....   | 90  |
| 4.4.2- Evaluación de la capacidad de Int1 para insertar el lado 5' de los genes <i>cassettes</i><br><i>attI1-aadA1-attC<sub>aadA1</sub></i> , <i>attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub></i> , <i>attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub></i> y <i>attC<sub>dhfr1</sub>-</i><br><i>aadA1-attC<sub>aadA1</sub></i> . ..... | 91  |
| 4.4.3- Evaluación de la capacidad de Int1 para insertar el lado 3' de los genes <i>cassettes</i><br><i>attI1-aadA1-attC<sub>aadA1</sub></i> , <i>attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub></i> , <i>attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub></i> y <i>attC<sub>dhfr1</sub>-</i><br><i>aadA1-attC<sub>aadA1</sub></i> .....   | 92  |
| Capítulo 5: Conclusiones .....  | 100 |
| Resumen .....   | 103 |
| Referencias bibliográficas.....   | 108 |
| Anexos .....  | 113 |

## ÍNDICE DE TABLAS

**Tabla 1.** Plásmidos usados en este estudio para realizar los ensayos de recombinación.

**Tabla 2.** Cebadores usados en este estudio para evaluar los ensayos de recombinación.

**Tabla 3.** Descripción de los alelos de los genes *cassette aadA-like* reportados en la base de datos INTEGRALL.

**Tabla 4.** Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 1 que incluyen al gen *cassette aadA1*.

**Tabla 5.** Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 2 que incluyen al gen *cassette aadA1*.

**Tabla 6.** Arreglos genéticos asociados *upstream* y *downstream* a la secuencia del sitio *attC<sub>aadA1</sub>* en 200 secuencias con un 100% de identidad y 100% de cobertura provenientes de GenBank.

**Tabla 7.** Arreglos genéticos asociados *upstream* y *downstream* a la secuencia del sitio *attC<sub>aadA1</sub>* en 12 secuencias sin nomenclatura con un 100% de identidad y 100% de cobertura provenientes de GenBank.

**Tabla 8.** Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 1 que incluyen al gen *cassette aadA2*.

**Tabla 9.** Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 1 que incluyen al gen *cassette aadA5*.

**Tabla 10.** Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 1 que incluyen a los genes *cassettes aadA6, aadA7, aadA1a, aadA1b, aadA11* y *aadA4*.

**Tabla 11.** Número de registros correspondientes a las secuencias determinadas como variantes de sitios *attCs* de los alelos *aadA-like*.

**Tabla 12.** Plásmidos con los genes *cassettes* usados en este estudio para realizar los ensayos de recombinación.

**Tabla 13.** Frecuencias de escisión estimadas para los genes *cassettes* estudiados en distintos ensayos de recombinación.

**Tabla 14.** Frecuencias de inserción en el sitio *attI1*, del lado 5' de los genes *cassettes* analizados, estimadas en los ensayos de recombinación

**Tabla 15.** Frecuencias de inserción en el sitio *att1*, del lado 3' de los *genes cassettes* analizados, estimadas en los ensayos de recombinación.



## ÍNDICE DE FIGURAS

**Figura 1.** Estructura general de los integrones de clase 1.

**Figura 2.** Estructura del sitio de recombinación *attI1*.

**Figura 3.** Secuencia consenso de un sitio *attC*.

**Figura 4.** Estructura de un sitio *attC* típico.

**Figura 5.** Estructura del gen *cassette*.

**Figura 6.** Estructuras secundarias de distintos sitios *attC* de algunos genes *cassettes* relevantes en la clínica.

**Figura 7.** Árbol filogenético de sitios *attCs*.

**Figura 8.** Diagrama ilustrativo del ensayo de escisión *in vivo*.

**Figura 9.** Diagrama ilustrativo del ensayo de inserción *in vivo* en el sitio *attI1*.

**Figura 10.** Esquema de un gen *cassette*.

**Figura 11.** Secuencia del sitio *attC<sub>aadA1</sub>*.

**Figura 12.** Árbol de máxima similitud entre las secuencias de los ORFs de los alelos *aadA-like*.

**Figura 13.** Alineamiento por ClustalW para los sitios *attC* de los distintos alelos *aadA-like*.

**Figura 14.** Árbol de máxima similitud de las secuencias de los sitios *attCs* de los genes *cassettes aadA-like*.

**Figura 15.** Estructura general de un sitio *attC* típico.

**Figura 16.** Estructura secundaria de los sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>*.

**Figura 16a.** Características del sitio *attC<sub>aadA1</sub>*

**Figura 16b.** Estructura secundaria de la hebra inferior del sitio *attC<sub>aadA1</sub>*

**Figura 16c.** Secuencia en doble cadena del sitio *attC<sub>aadA3</sub>*

**Figura 16d.** Estructura secundaria de la hebra inferior del sitio *attC<sub>aadA3</sub>*

**Figura 17.** Estructura secundaria del sitio *attC<sub>aacA1</sub>*.

**Figura 17 a.** Secuencia en doble cadena del sitio *attC<sub>aacA1</sub>* del gen *cassette aacA1*.

**Figura 17b.** Estructura secundaria de la hebra inferior del sitio *attC<sub>aacA1</sub>*.

**Figura 18.** Estructura secundaria del sitio *attC<sub>dfrA1</sub>*.

**Figura 18a.** Secuencia en doble cadena del sitio *attC<sub>dfrA1</sub>*.

**Figura 18b.** Estructura secundaria de la hebra inferior del sitio *attC* del gen *cassette dfrA1*.

**Figura 19.** Estructura de los genes *cassettes* usados para los ensayos de recombinación.

**Figura 20.** Esquema general del proceso de escisión del gen *cassette aadA1* para los diferentes arreglos construidos en los plásmidos usados en los ensayos de recombinación.

**Figura 21.** Geles de agarosa donde se indica con una flecha los amplicones de las recombinaciones luego de las escisiones

**Figura 22.** Gel de agarosa donde se observa la frecuencia de escisión (100%) del ensayo de recombinación con el clon pLQ445 *attC<sub>aacA1</sub>-aadA1- attC<sub>aadA1</sub>*.

**Figura 23.** Secuencia del fragmento de escisión secuenciado.

**Figura 24.** Gel de agarosa donde se muestra el resultado de la evaluación por PCR de las inserciones a 5' para los distintos ensayos de recombinación.

**Figura 25.** Gel de agarosa donde se muestra el resultado de la evaluación por PCR de las inserciones a 3' para los distintos ensayos de recombinación.

**Figura 26.** Porcentajes de recombinación de los genes *cassettes* estudiados en este trabajo.

## ABREVIATURAS

5'-CS: Región 5' conservada de los integrones

3'-CS: Región 3' conservada de los integrones

1L: Dominio de unión a IntI en un sitio *attC*, también llamado sitio *core*

2L: Dominio de unión a IntI en un sitio *attC*

1R: Dominio de unión a IntI en un sitio *attC*, también llamado sitio *inverse core*

2R: Dominio de unión a IntI en un sitio *attC*

A: Adenina

ADN: Ácido desoxiribonucleico

AN: Número de acceso al GenBank

ARN: Ácido ribonucleico

ARNm: ARN mensajero

*attI*: Sitio de recombinación de la integrasa provisto por el integrón

*attC*: Sitio de recombinación de la integrasa provisto por el *cassette*

*attC*::Tn-like: tipo de sitio de recombinación atípico y/o *cassette* inusual en el que el sitio *attC* está invadido por un elemento móvil.

$\Delta attI$ -like: tipo de sitio de recombinación atípico de la integrasa y/o *cassette* inusual en el que el sitio *attC* presenta identidad con una porción de un sitio *attI*; si aparecen con subíndices, éstos indican la longitud de la identidad incluyendo al triplete GTT.

$\Delta attC$ -like: tipo de sitio de recombinación atípico de la integrasa y/o *cassette* inusual en el que se observa un ORF en la región variable de un integrón pero presentan secuencias no compatibles con las de los sitios *attCs* o *attI* típicos.

BGNNFG: Bacilos Gram-negativos no fermentadores de la glucosa

C: Citocina

CS: Sitio *core* de los sitios *attC* y *attI1*, también corresponde de forma general al heptámero estándar de unión a IntI1, su secuencia consenso es GTTRRRY

ICS: Sitio *inverse core* de los sitios *attC* y *attI1*, su secuencia consenso es YRRRAAC

dNTPS: Desoxiribonucleótidos

## Capítulo 1: Introducción

G: Guanina

*intl*: Gen de una integrasa, cuando está seguido de un número indica que es el correspondiente a un tipo de integrasa en particular

IntI: Integrasa, cuando está seguido de un número indica que corresponde a un tipo de integrasa en particular

IPTG: Isopropyl- B-D- tiogalactopiranosido

IS: Secuencia de inserción

IR: Repeticiones invertidas

PCR: Reacción en cadena de la polimerasa

THG: Transferencia horizontal de genes

ORF: marco de lectura abierto

R: Purina

T: Timina

pb: Pares de bases

PM: Peso Molecular

kb: Kilopares de bases

ND: No determinado

SC: Secuencia consenso

Y: Pirimidina

ZV: Zona variable del integrón. Región donde se encuentran los genes *cassette*.

# Capítulo 1: Introducción

---

## 1.1- El rol de los integrones en la Transferencia Horizontal Genética

La transferencia horizontal genética (THG) es una de las estrategias evolutivas más eficientes de las bacterias (Rowe-Magnus y Mazel, 1999). La adquisición de material genético foráneo permite aumentar la variabilidad genética de un determinado genoma. La eficiencia con la cual las bacterias incorporan información genética refleja su capacidad para adaptarse a los cambios ambientales. Varios mecanismos están involucrados en el intercambio de material genético, pero el material genético transferido, puede o no ser asimilado al nuevo genoma. La recombinación homóloga permite el establecimiento de secuencias en taxones relacionados, por lo cual es poco probable la introducción de nuevos marcadores. Por el contrario, la recombinación sitio-específica, puede mediar el intercambio de marcadores diversos, aunque son dependientes de sitios blanco para el reconocimiento y la catálisis respectiva. Las integrasas desarrollaron un ingenioso mecanismo para permitir la asimilación de ADN foráneo, basado en el reconocimiento de dos pequeñas secuencias de ADN, identificadas como sitios *attI* y *attC*. El reconocimiento de estos sitios permite la escisión y la inserción, y por lo tanto la acumulación y/o reordenamiento de los genes de resistencia a antibióticos que se encuentran bajo la forma de genes *cassettes* en **integrones**.

Un **integrón** es un sistema de escisión y/o integración involucrado en la captura, movilización y expresión de genes (Recchia y Hall, 1995). Consiste en una estructura

## *Introducción*

genética que tiene la capacidad de incorporar genes bajo la forma de genes *cassettes*, a través de un mecanismo de recombinación sitio específica y transformarlos en genes funcionales al permitir su expresión génica (Mazel, 2006).

Según la definición actual, los integrones contienen en su estructura 3 elementos: un gen (*intI*) que codifica una **integrasa** perteneciente a la familia de las tirosina-recombinasas; un sitio primario de recombinación o ***attI*** y un promotor orientado en forma divergente a la integrasa (*outward*) llamado Pc, que dirige la transcripción de los genes *cassette* incorporados (Hall y col., 1991; Levesque y col., 1994; Recchia y Hall, 1997; Sundstrom y col., 1998; Rowe-Magnus y col., 1999).

El material genético exógeno que incorporan los integrones tiene una estructura particular y se denomina **gen *cassette***. Generalmente los genes que se encuentran bajo forma de genes *cassettes* no tienen promotor propio, así que usan el promotor del integrón para transcribirse. Su estructura se caracteriza por contener un ORF y una repetición imperfecta invertida en el extremo 3' del gen llamado **sitio *attC*** (o elemento de 59 pb). Los sitios *attC* tienen una gran variabilidad de secuencia y su función principal es servir como sitio diana de reconocimiento de la integrasa (Rowe-Magnus y Mazel, 2001).

En los aislamientos clínicos, la mayoría de los genes *cassettes* que se integran en este sistema codifican proteínas que en conjunto confieren resistencia a casi todas las familias de antibióticos ( $\beta$ -lactámicos, aminoglucósidos, cloranfenicol, trimetoprima,

## Introducción

sulfonamidas y otras familias de antibióticos) (Recchia y Hall, 1997). La mayoría de los microorganismos que contienen este tipo de integrones asociados a multirresistencia antibiótica son bacterias Gram negativas como las especies pertenecientes a la Familia *Enterobacteriaceae* (Stokes y Hall, 1989; White y col., 2001) y otras especies como *Vibrio cholerae* (Rowe-Magnus y col., 2001) y *Pseudomonas aeruginosa* (Rosser y Young, 1999). Posteriormente se han descrito también en bacterias Gram Positivas (Fluit y Schmitz, 1999; Nesvera y col., 1998) como *Corynebacterium glutamicum*.

### 1.2- Clasificación de los integrones

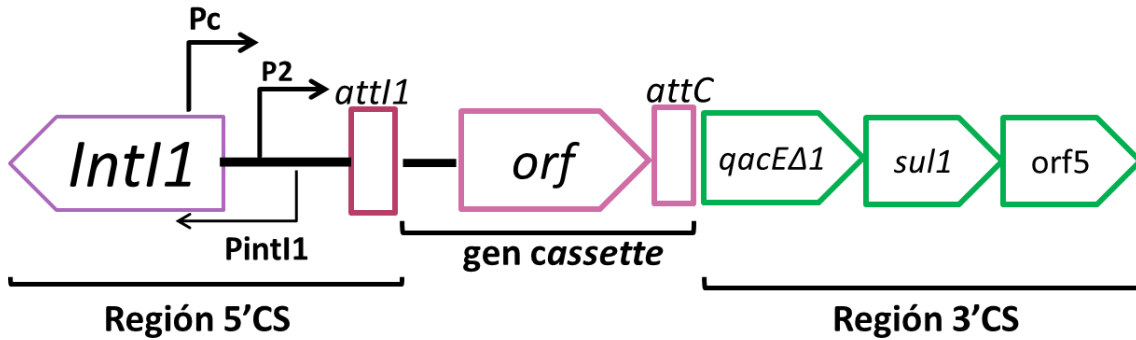
Hasta la fecha han sido identificadas 5 clases de integrones móviles, en base a la secuencia de sus integrasas (con un 40-58% de identidad). Todos ellos están asociados a multirresistencia antibiótica. Comparten la característica de localizarse en transposones, plásmidos y/o islas genómicas, lo cual les permite dispersarse en genomas bacterianos de diferentes especies y géneros, y así responder rápida y eficientemente a la presión antibiótica. Ellos son los integrones de clase 1, 2, 3, SXT y pRSV1.

Los integrones de clase 1 y 2 son los más exitosos dentro del ambiente hospitalario, mientras que los otros han sido detectados esporádicamente (Cambray y col., 2010). Los integrones de clase 1 son encontrados principalmente en aislamientos clínicos, están bien caracterizados y han sido los más estudiados. Su estructura básica consiste en una secuencia 5' conservada (5'CS) que contiene al gen *int1* que codifica para la integrasa de tipo 1 (Int1), encargada de catalizar la recombinación genética sitio específica; luego el

## Introducción

sitio *attI1* que es un sitio de recombinación donde se insertan los genes *cassette*; entre *intI1* y *attI1* están dos promotores: P<sub>c</sub> ubicado dentro del gen de la integrasa y del cual se identificaron diferentes variantes, y P<sub>2</sub>, un segundo promotor presente en aproximadamente 10% de estos integrones, cuya función generalmente compensa la actividad débil del promotor P<sub>c</sub>, aumentando la expresión de los genes *cassette* sin interferir con la expresión de la integrasa ; además se encuentra el promotor *PintI1* que promueve la expresión de *intI1* (Guérin y col., 2011 y Jové y col., 2010). La región variable es donde se localizan los genes *cassette* estructurales del integrón, en número variable que suele ser de 1 a 4, y que de acuerdo a estudios de secuenciación, son en su mayoría, pero no exclusivamente, genes que codifican resistencia a los antimicrobianos (Stokes y Hall, 1989; Collis y Hall, 1992). En el extremo 3' se encuentra la secuencia 3' conservada (3'CS), donde se localizan tres genes, *qacEΔ1*, *sul1* y *orf5*; el primero confiere resistencia a compuestos de amonio cuaternario que forman parte de algunos antisépticos y desinfectantes, el segundo confiere resistencia a sulfonamidas y el *orf5* es de función desconocida (Larouche y Roy, 2011) (Figura 1).



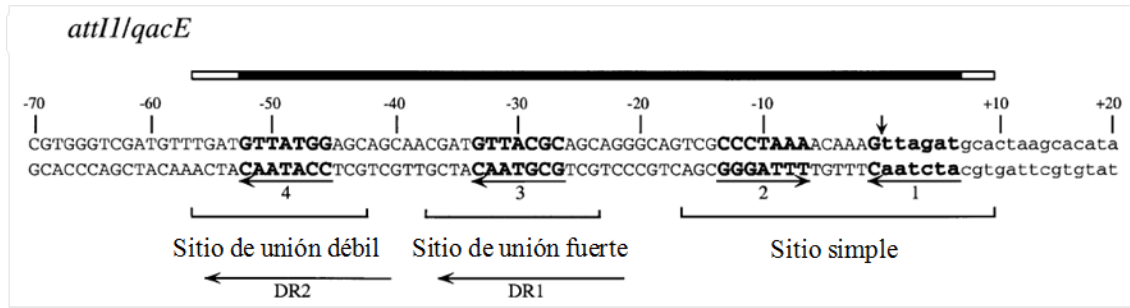


**Figura 1. Estructura general de los integrones de clase 1.** Los genes se muestran como flechas en la siguiente secuencia: *int1*, gen de la integrasa de tipo 1; *qacEΔ1*, gen de resistencia a antisépticos; *sul1*, gen de resistencia a sulfonamidas; *orf5*, gen de función desconocida. Los sitios de recombinación *attI1* y *attC* se muestran como rectángulos. Los genes *cassette* están compuestos por un gen y un sitio de recombinación *attC*. Se muestran los promotores *Pc*, *P2* y *Pint1*. Modificado de Larouche y Roy,2011.

### 1.3- Sitio de recombinación *attI*

El sitio *attI* se caracteriza por ser específico para cada una de las integrasas (Grindley y col., 2006). Es una región de 40 a 70 pb que en el caso del sitio *attI1* incluye: Un sitio simple que consiste en un par de dominios de unión a la integrasa inversamente orientados denominados CS “core site” e ICS “inverse core site” junto a otros dos sitios de unión a la integrasa que son repeticiones directas, denominados sitio de unión fuerte y débil, respectivamente. (Partridge y col., 2000). En el sitio *core* “CS” se produce la recombinación (Figura 2).

## Introducción



**Figura 2. Estructura del sitio de recombinación *attI1*.** En el sitio simple, se encuentra el sitio *core* “CS” (flecha 1) y el *inverse core site* “ICS” (flecha 2). La flecha vertical indica el punto de recombinación y las flechas horizontales indican las repeticiones directas (DR1 y DR2), las cuales se solapan con los sitios fuerte y débil de unión a la integrasa. Modificada de Partridge y col., 2000.

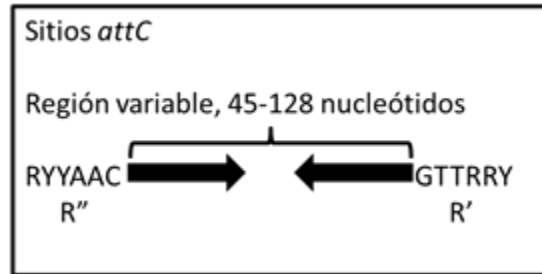
### 1.4- Sitios *attC*

El sitio *attC*, conocido también como elemento de 59 pares de bases, se encuentra *downstream* al gen *cassette*. Consiste en una secuencia palindrómica formada por 57 a 141 pb involucrada en el evento de integración y de escisión de los genes *cassette* (Recchia y Sherratt, 2002). La función de los sitios *attC* es proveer al gen *cassette* del sitio de recombinación para lograr la escisión o la integración del mismo en el genoma bacteriano (Cameron y col., 1986, Recchia y Hall 1995).

Los sitios *attC* tienen una baja similitud entre ellos y se diferencian tanto en tamaño como en secuencia. En una primera clasificación del año 1997 se agruparon por la longitud de su secuencia. La similitud de las secuencias de los sitios *attC* está localizada principalmente en 20 nucleótidos en sus extremos denominadas RH (*Right Hand*) y LH (*Left Hand*) por lado derecho e izquierdo, respectivamente, donde cada uno contiene a dos cajas de heptanucleótidos llamadas R' (GTRRRY, R: Purina y Y: Pirimidina) y R'' (RYYAAC) con un punto de recombinación ubicado entre la G y T en R' (Figura 3). Estas cajas son parte de dos sitios antiparalelos de recombinación “*core*” denominados: R”-L” y

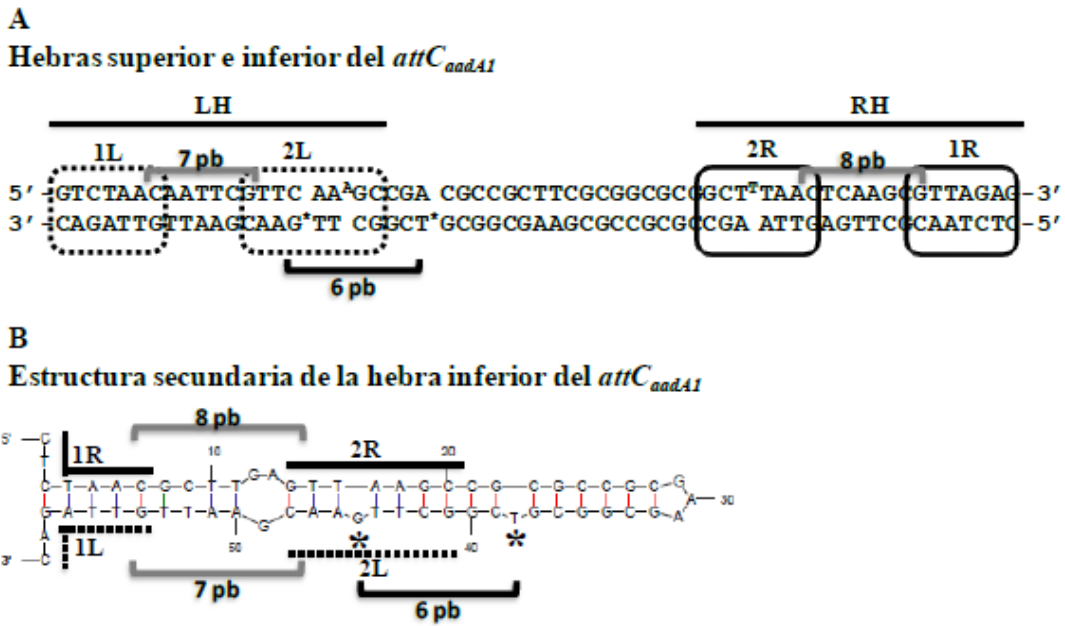
## Introducción

L'-R', ó 1L-2L y 2R-1R, respectivamente. La estructura general del sitio *attC* es palindrómica y permite que la hebra simple del *attC* adopte una forma de cadena doble por la unión de las regiones L'' a L' y R'' a R'. (Recchia y Hall, 1997; Quiroga y col., 2007; Hall y col., 1991; Collis y Hall 1992; Stokes y col., 1997).



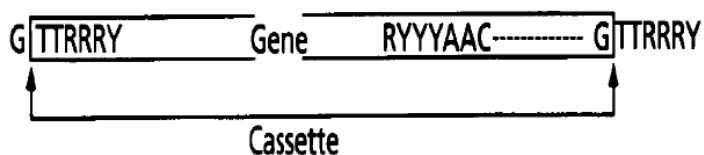
**Figura 3. Secuencia consenso de un sitio *attC*.** Se observan las dos regiones definidas como R'' RYYAAC, R: Purina e Y: Pirimidina) y la secuencia R' (GTTRRY), también llamadas 1L y 1R, respectivamente. Modificado de Mazel, 2006.

Cómo se mencionó anteriormente, cuando el sitio *attC* se encuentra bajo su estado de ADN simple cadena tiene la posibilidad de plegarse adquiriendo una estructura secundaria estable característica en forma de tallo y bucle (stem-loop). Dicha estructura secundaria es fundamental para que ocurra la recombinación sitio-específica mediada por la integrasa (Figura 4) (Johansson y col. 2004; Bouvier y col. 2005).



**Figura 4. Estructura de un sitio *attC* típico.** Se toma como referencia al sitio *attC<sub>aadA1</sub>*. Se indican los sitios simples LH y RH y los sitios 1L, 2L, 2R y 1R. Los superíndices A y T indican las bases A y T conservadas de los sitios 2L y 2R, respectivamente. Los asteriscos indican las bases extrahelicoidales. Las llaves grises marcan la longitud de las secuencias espaciadoras, y las negras marcan la distancia entre las bases extrahelicoidales típicas. Parte A: Hebras superior e inferior del sitio *attC<sub>aadA1</sub>*. Parte B: Estructura secundaria formada por el plegamiento de la hebra inferior del sitio *attC<sub>aadA1</sub>*.

Ha sido demostrado experimentalmente que la integrasa de tipo 1 o IntI1 reconoce las regiones 1L-1R, 2L-2R en la estructura secundaria del sitio *attC* y corta el ADN simple cadena entre la C y la A del sitio 1R de la cadena inferior (Stokes y col., 1997; Collis y col. 2001 Biskri y col. 2005; Bouvier y col. 2005; Bouvier y col., 2005). Lo cual implica que en la posterior inserción del gen *cassette* cuando éste se encuentra en su forma lineal abierto en ese punto, quedan *upstream* del ORF los 6 últimos nucleótidos de 1L (CS) **G'TTRRRY** del sitio *attC*, y río abajo del ORF se encuentra el resto del sitio *attC*, que finaliza en el primer nucleótido del CS **G'TTRRRY** (en negrita se identifican los nucleótidos que componen los extremos del gen *cassette*) (Figura 5) (Recchia y Hall, 1995).



**Figura 5. Estructura del gen *cassette*.** El largo del gen *cassette* está marcado por las flechas verticales. Tomado de Recchia y Hall, 1995.

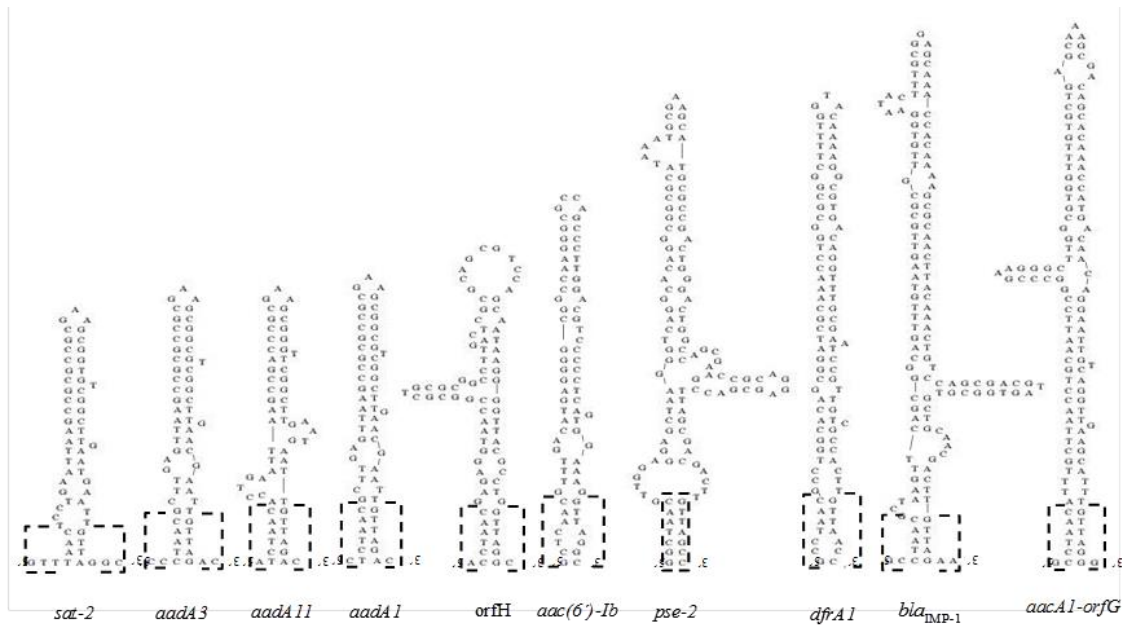
Adicionalmente, se ha demostrado que hay dos bases extrahelicoidales en la estructura secundaria conformada por la hebra inferior del sitio *attC*, mayoritariamente una G y una T, las cuales colaboran con la afinidad de la integrasa y su unión al sustrato (sitio *attC*) para indicar la direccionalidad en la cual se inserta posteriormente el gen *cassette* (Figura 4) (MacDonald y col., 2006; Bouvier y col., 2009). La primera de estas bases extrahelicoidales, que está dentro del sitio 2L, es más importante para determinar la elección de la cadena que recombinará mientras que la segunda base extrahelicoidal fortalece la elección (Bouvier y col., 2009). A su vez, se ha identificado una distancia espaciadora óptima para las dos bases extrahelicoidales, la cual es de 6 pb (Larouche y Roy, 2011). Se observó que ante la ausencia de bases extrahelicoidales, la presencia de una burbuja en la región espaciadora entre los sitios 1 y 2 es esencial para la recombinación (Bouvier y col., 2009).

Se identificó también que es imprescindible que se mantenga la complementariedad de bases que lleva a la formación de la estructura secundaria característica de los sitios *attC*, ya que cualquier modificación en la secuencia que afectase el correcto plegamiento de la estructura secundaria modificaría el proceso de

## Introducción

recombinación (Bouvier y col., 2005; MacDonald y col., 2006). Se demostró que la ruptura del apareamiento mediante mutaciones específicas en la región variable del *attC* correspondiente a la hebra inferior del ADN simple cadena, disminuye la frecuencia de recombinación (Bouvier y col., 2005). La recuperación del apareamiento con mutaciones complementarias restablece las frecuencias de recombinación mediadas por la integrasa.

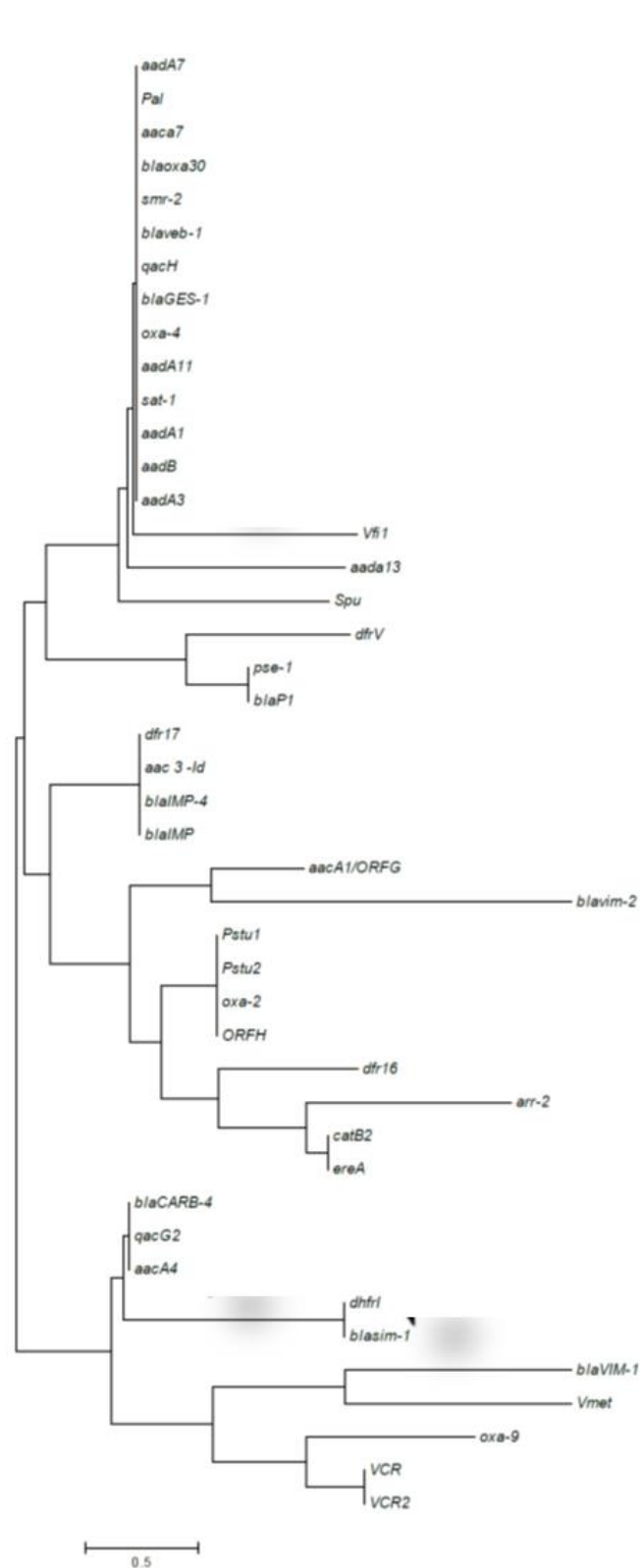
Los sitios *attC* de los genes *cassette* descritos en el ambiente hospitalario, difieren mucho entre sí. Por ejemplo, los sitios *attCs* de los genes *cassette bla<sub>OXA-5</sub>* y *bla<sub>OXA-9</sub>*, poseen un 79,3% de identidad, confiriendo ambos genes *cassette* resistencia a la misma familia de antibióticos. El sitio *attC* del gen *cassette sat2* tiene una longitud de 60 pb, mientras que el sitio *attC* de *bla<sub>IMP-1</sub>* posee 127 pb en su estructura (Figura 6) (Recchia y Hall, 1997). En base a esta diversidad, Recchia y Hall (1997) los separaron en grupos: en el grupo 1 incluyeron a aquellos sitios *attCs* que contienen entre 57 y 60 pb; en el grupo 2 a los formados por 70 a 72 pb, y en el grupo 3 incluyeron a los *attCs* de entre 127 y 141 pb, dejando fuera de estos grupo a sitios *attC* que son únicos en cuanto a su longitud.



**Figura 6. Estructuras secundarias de distintos sitios *attC* de algunos genes *cassettes* relevantes en la clínica.** Se muestran los sitios *attC* de los genes *cassettes* de importancia en la clínica, y las diferencias en sus longitudes y variantes de sus estructuras secundarias. Tomada de Quiroga, 2007.

### 1.4.1- Familias de sitios *attC*

El estudio filogenético de diversos sitios *attC* realizado por MP Quiroga (Tesis doctoral, 2012) permitió identificar 3 grandes grupos de sitios *attC* (Figura 7). El primer grupo contiene el mayor número de sitios *attC*, además de varios que son iguales entre si e incluye el sitio *attC* del gen *cassette aadA1*, el segundo grupo incluye a los sitios *attC* de los genes *cassette bla<sub>VIM-2</sub>* y *aac(6')-IId* de alta importancia clínica, mientras que el sitio *attC* del gen *cassette dfrA1* pertenece al grupo 3 donde también están incluidos los sitios VCR (*Vibrio cholerae repeats*).



**Figura 7. Árbol filogenético de sitios *attCs*.** Las ramas del árbol muestran a las familias de sitio *attCs*. Tomado de: MP Quiroga (Tesis doctoral, 2012).



## Introducción

### 1.5- Características genéticas de los genes *cassettes*

Los genes *cassettes* fueron descritos a través de análisis de mapas de restricción, heteroduplex y análisis de los plásmidos R388 y pSA (Ward y Ginsted, 1982), y a través del estudio de transposones como el Tn21, Tn2603 y Tn2424 en el estudio de la resistencia antibiótica. Actualmente hay descritos más de 130 genes *cassette* de resistencia a antibióticos, pero también hay otros con una función aún no determinada (Partridge y col., 2009). Como se mencionó anteriormente, un gen *cassette* se caracteriza por tener un ORF y el sitio *attC* responsable de su movilidad. A diferencia de otros elementos genéticos como los transposones, no están flanqueados por repeticiones invertidas y no codifican para ninguna enzima responsable para su transposición. Su longitud varía entre 262 a 1549 pb, el ORF puede estar flanqueado por secuencias no codificantes, en general no tienen promotor propio por lo cual éstos genes *cassettes* solo se expresan cuando están integrados en la zona variable de los integrones. Estos elementos son altamente eficientes en su movilización intra e inter genómica. La integración y escisión de los genes *cassettes* es *dependiente* por completo de integrasas de integrones y de los sitios de recombinación ubicados en el mismo gen *cassette* y en el sitio de integración. Esto pone en evidencia la importancia de la arquitectura en la que se encuentran los genes *cassettes* cuando se localizan en la zona variable de los integrones, donde *upstream* del ORF puede quedar un sitio *attI* o el sitio *attC* del gen *cassette* que lo precede, y *downstream* se encuentra su propio sitio *attC*. Los genes *cassettes* también pueden encontrarse libres en el citoplasma en su forma circularizada, pero son incapaces de replicarse por sí mismos (Collis y Hall, 1992).

## Introducción

### 1.6- Reacciones de recombinación en integrones

Para la integración o escisión de genes *cassettes*, la integrasa de tipo 1 (IntI1) cataliza varias reacciones de recombinación específicas, entre un sitio *attI1* y un sitio *attC*, y recombinaciones entre dos sitios *attC*.

La reacción de recombinación más estudiada ha sido la mediada por la integrasa de clase 1, estableciendo 5 tipos diferentes de reacciones. Tres corresponden a los diferentes eventos de recombinación posibles entre los sitios normales de recombinación (*attIxattC*, *attCxattC* y *attIxattI*) y las otras dos corresponden entre un sitio *attC* o *attI* y secuencias no específicas que contienen GTT (Hall y col., 1999). En un experimento previo, usando un plásmido que contiene un fragmento de integrón con los genes *cassettes aacC1*, *orfE* y *aadA2* se demostró que este último tiene una diferencia entre la frecuencia de escisión con respecto a la del gen *cassette aacC1* y que la recombinación entre diferentes pares de sitios (*att*) se produce en muy diferentes frecuencias. A través de la evaluación de eventos de recombinación se demostró que la eficiencia de la recombinación entre dos sitios *attI* es menor en comparación a las reacciones que involucran dos sitios *attC* o un sitio *attI* y un *attC*, la cual se ha definido como la reacción de recombinación más eficiente (Collis y Hall, 1992).

Otros estudios sobre esta recombinación, específicamente de las frecuencias de recombinación han estudiado la eficiencia de la integrasa de tipo 1 y otras integrasas para integrar o escindir ciertos genes *cassettes* de resistencia. En un estudio sobre la especie

## Introducción

*Nitrosomonas europaea* se evaluó la capacidad de su integrasa (**IntI<sub>Neu</sub>**) de integrar o escindir ciertos genes *cassettes* de resistencia con distintos sitios *attI* y *attC*, demostrando que la eficiencia de la escisión no sólo depende del sitio *attC* del gen *cassette* sino del sitio *att* *upstream* del mismo, teniendo algunos sitios *attC* efectos negativos sobre cualquier posición (Leon y Roy, 2003).

Más adelante, al comparar la eficiencia de escisión de genes *cassettes* mediada por varios tipos de integrasas, se determinó un efecto pronunciado de la presencia de las bases extrahelicoidales presentes en los sitios *attCs* *upstream* y *downstream* de los genes *cassettes*, que no afectan el sitio de unión de la integrasa pero si la reacción de recombinación. También se describió que la integrasa de tipo 1 es más eficiente en la escisión de *cassettes* con estructuras de *attCs* de tipo T-N<sub>6</sub>-G o T-N<sub>6</sub>-C (Larouche y Roy, 2011).

A partir de análisis de bioinformática y secuencias de genes *cassettes* de nuestro laboratorio evidenciamos la presencia de tres grupos de genes *cassettes* inusuales a los que denominamos  $\Delta attI$ -like,  $\Delta attC$ -like y *attC*:Tn-like, y se estudiaron a nivel molecular y funcional (MP Quiroga, Tesis doctoral 2012). Se comprobó la completa funcionalidad del gen *cassette attI1-bla<sub>GES-1</sub>- $\Delta attI1$ <sub>3pb</sub>* y se consideró que puede extrapolarse a otros genes *cassettes* inusuales del tipo  $\Delta attI$ -like, siendo sus frecuencias de recombinación equivalentes a las de genes *cassettes* típicos.

## Introducción

Se investigó la funcionalidad de genes *cassettes* inusuales del segundo grupo, y se tomó como modelo al paradigmático módulo *orfX-ybfA-ybfB-ybgA* localizado en la zona variable de Tn7::In2-8. Se llevaron a cabo ensayos de recombinación en los cuales se observó que *IntI1* escinde al módulo completo de dos formas (70% y 3%) y también a *orfX* (28%), *ybfA-ybfB-ybgA* (3%) y a *ybgA* (10%) por separado. Se observó que para ello reconoce 7 sitios de recombinación, cuatro de los cuales describimos por primera vez, y se reportó el primer evento de recombinación mediado por *IntI1* que no involucra sitios típicos. También se mostró que el gen *cassette* inusual *attI1-aadB-attC::S.ma.I2* que posee al intrón *S.ma.I2* inserto en el sitio *attC*, es escindido (12%) en mayor frecuencia que algunos genes *cassettes* típicos y es insertado (3%) nuevamente en un sitio *attI1* con una frecuencia similar a la de genes *cassettes* típicos. Por otro lado, no detectamos recombinaciones en el sitio *attC* invadido por *ISPa119* del gen *cassette* inusual *attC-aacA7-attC::ISPa119*, lo cual pone de manifiesto como por otro lado, los integrones actúan de reservorio genómico no solo de nuevas estructuras, sino también de elementos móviles que se garantizan su persistencia gracias al éxito adaptativo de los integrones. Al analizar genes *cassettes* típicos de relevancia clínica tales como *bla<sub>VIM-2</sub>*, *aac(6')-IIa*, *dfrA1* y *aadB*, se observó que el más exitoso en los eventos de inserción es el gen *cassette attI1-bla<sub>VIM-2</sub>-attC*, cuya frecuencia es igualada sólo por el gen *cassette* inusual *attI1-aadB-attI1*. El gen *cassette bla<sub>VIM-2</sub>*, que confiere resistencia a los carbapenemes, presenta un sitio *attC* característico que contiene una C en vez de la G extrahelicoidal típica, y ese fue el único trabajo en el que se evaluó su movilidad hasta el momento (MP Quiroga, Tesis doctoral 2012).

## Introducción

Además en ese estudio, se evaluó la movilidad del gen *cassette aadB* en diferentes arquitecturas y se calcularon las frecuencias de recombinación de diferentes arquitecturas de genes *cassettes*, tanto *attI1-ORF-attC* como *attC-ORF-attC*, para el gen *cassette aadB* que otorga resistencia a Gentamicina. Estas arquitecturas son características de genes *cassettes* localizados en la primera posición de la zona variable de un integrón y en la siguiente posición del mismo, respectivamente. Allí también se estudiaron otros genes *cassettes* de relevancia en la clínica en la arquitectura *attI1-ORF-attC*. En ese estudio se observó una variación en las frecuencias de recombinación, con valores de recombinación cercanos al 80% para la estructura *attI1-aadB-attC* y de 50% para la estructura *attC-aadB-attC*, lo cual evidenció cómo se ven afectadas las frecuencias de recombinación cuando varía el sitio *att* localizado *upstream* de un gen *cassette*. Notamos que la identidad del sitio de recombinación que está *downstream* de los genes *cassettes* tiene una gran influencia en las frecuencias de escisión, mientras las frecuencias de inserción netas parecen estar más determinadas por el sitio de inserción (Quiroga MP, Tesis doctoral 2012).

### 1.7- Identificación del gen *cassette aadA1* en el ambiente

La mayoría de los genes de resistencia antibiótica que se encuentran en la clínica actualmente se originaron a partir de los productores de antibióticos, principalmente actinomicetos (Davies, 1994) que son bacterias Gram positivas, ambientales. Diversos elementos genéticos de las bacterias, como plásmidos, transposones e integrones son considerados de gran importancia por su rol en la diseminación de los genes de resistencia

## *Introducción*

antibiótica a través de la THG de la clínica al ambiente y viceversa (Bennett, 2008; Harbottle y col., 2006; Summers, 2006; Walsh, 2006). Es probable que la coexistencia prolongada en el ambiente entre bacterias sensibles y resistentes haya permitido que estos genes se diseminen por THG entre diversas bacterias Gram negativas y Gram positivas que ocuparon el mismo nicho que los actinomicetos. Para comprobar esta hipótesis se estudiaron cepas ambientales de muestras de suelo y agua dónde se encontraron aislamientos de bacterias resistentes a antibióticos. Las bacterias resistentes a varios antibióticos han sido caracterizadas entre distintas comunidades bacterianas del suelo siendo algunas de ellas resistentes a más de 2 antibióticos (Esiobu y col., 2002; Riesenfeld y col., 2004; D'Costa y col., 2006). Esto indica que probablemente las bacterias ambientales son reservorios naturales de genes de resistencia para aquellas bacterias que se aíslan a partir de muestras clínicas.

Para comprobar esta última idea sobre el origen de la resistencia antibiótica en la clínica, han sido realizados diferentes estudios abocados a buscar e identificar comunidades de bacterias de sedimentos de permafrost, los cuales por sus condiciones físicoambientales no han sido expuestos a factores antropogénicos.

Entre estos estudios, se encuentran los realizados por Mindlin y col., 2009 dónde se tomaron muestras de diferentes orígenes, tiempo geológico y profundidades provenientes de *permafrost* (Mindlin y col., 2009). A partir de estas muestras, describieron aislamientos con resistencia a componentes de mercurio altamente homólogos a los de

## Introducción

bacterias actuales, así como también se encontraron bacterias con resistencia a diferentes antibióticos. Específicamente, se estudió más en profundidad la resistencia a estreptomicina para determinar su estructura molecular. En este mismo estudio a través de la amplificación y posterior secuenciación del gen de 16s rRNA presente en las muestras de *permafrost*, con una una longevidad de más de 1 millón de años. Se identificaron diferentes géneros bacterianos tales como *Acinetobacter*, *Pseudomonas*, *Xanthomonas*, *Bacillus*, *Arthrobacter*, *Micrococcus*, *Flavobacterium*, *Stenotrophomonas*, *Sphingobacterium*, *Psychrobacter* y *Paenibacillus*. Asimismo, se encontró que la resistencia a estreptomicina estaba mediada por los genes *strA* y *strB* y por los genes *cassette* de la familia *aadA*, cuya resistencia se encuentra altamente distribuida en las bacterias actuales.

Otro estudio, es el realizado por Petrova y col., 2011, en el cual describieron al transposón Tn5045 en una cepa de la especie *Pseudomonas* sp. Tik3 proveniente de muestras de *permafrost* de la región de costa del mar de Laptev (Petrova y col., 2011). Dicho transposón contiene elementos de resistencia a estreptomicina/espectinomicina y sulfonamidas, y corresponde al primer transposón portador de un integrón asociado a la resistencia antibiótica. El integrón de clase 1 presente en Tn5045, posee el gen *cassette aadA2* en su zona variable, con características genéticas muy parecidas a los que se describen en la actualidad. Este transposón se encontró ubicado en el cromosoma de la bacteria *Pseudomonas* sp. Tiki, y se cree que el integrón que porta es el primer integrón de resistencia antibiótica originado en la era preantibiótica.

## *Introducción*

Se han descrito dos ramas de integrones portadores de la familia de genes *cassettes aadA*, estas son *aadA1* y *aadA2* (comparten una identidad de secuencia de 89%) (Brown y col., 1996; Liebertet y col., 1999; Riccio y col., 2005). El integrón denominado In2 que contiene un único *gen cassette aadA1* ha sido propuesto como el integrón ancestro de la familia de *aadA1* (Brown y col., 1996). Por otro lado, el integrón denominado InC, correspondiente a un hipotético In2-*like* ancestro de la familia *aadA2* y portador de un único *gen cassette aadA2* (Bissonnette and Roy, 1992), este integrón InC aún no ha sido identificado.

Es evidente al observar los datos que nos arrojan estos estudios y la alta diseminación de la resistencia a estreptomicina/espectinomicina en los aislamientos clínicos actuales y sus determinantes genéticos, que estas plataformas genéticas son claves para entender la THG entre las bacterias del ambiente y de la clínica, dado que conservan una gran similitud en sus estructuras genéticas. Además, estos estudios evidencian que los genes de resistencia a antibióticos aparecieron mucho antes del uso de la terapia antibiótica.



# Capítulo 2: Hipótesis y Objetivos

---

## 2.1- Hipótesis

Estudios de revisión bibliográfica evidencian la diseminación global del gen *cassette aadA1*. Sin embargo, aún no se ha descrito en su totalidad su rol a nivel de diversidad, evolución y funcionalidad dentro del ámbito de los integrones clínicos y a nivel de reservorio dentro de los integrones ambientales. Basándonos en estos conceptos, proponemos que el gen *cassette aadA1* puede ser tomado como modelo de estudio, ya que se ha descrito como un determinante de resistencia que habría estado presente previo a la era antibiótica y se habría transmitido a través de eventos de THG a las bacterias clínicas actuales. Dicho gen *cassette* se encuentra ampliamente distribuido en los aislamientos clínicos aportando a la problemática de la multiresistencia antibiótica.

## 2.2- Objetivos

### 2.2.1- Objetivo general

Identificar y determinar las características moleculares y funcionales del gen *cassette aadA1*.

### 2.2.2- Objetivos específicos

**2.2.2.1-** Analizar la evolución y variabilidad de la familia del gen *cassette aadA* mediante estudios de bioinformática.

**2.2.2.2-** Determinar la funcionalidad del gen *cassette aadA1* a través de ensayos experimentales de recombinación mediados por la integrasa de tipo 1

### *Hipótesis y objetivos*

-Definir las características de las estructuras secundarias de los diferentes sitios *att* involucrados en los ensayos experimentales de recombinación y su influencia sobre los mismos

- Estimar la frecuencia de escisión mediadas por la integrasa tipo 1 para el *cassette aadA1* con diferentes sitios *att* (*attI1* y *attCs*) *upstream* del mismo.

- Estimar las frecuencias de inserción mediadas por la integrasa de tipo 1 para el *cassette aadA1* con diferentes sitios *att* (*attI1* y *attCs*) *upstream* del mismo.

# Capítulo 3: Materiales y métodos

---

## 3.1- Análisis de bioinformática de los genes *cassette aadA-like*.

### 3.1.1- Estudio de frecuencia y arreglos de los alelos *aadA-like*.

A través de la búsqueda de los distintos alelos de los *cassettes aadA-like* reportados en la base de datos INTEGRALL (<http://integrall.bio.ua.pt>), se realizó una recopilación de sus secuencias incluyendo su marco de lectura abierto (ORF) y su sitio *attC*, tanto si estuviese anotado en dicha base de datos o anotado en la base de datos GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>). A partir de los datos contenidos en INTEGRALL se estudiaron cuales genes *cassettes* son los más frecuentes, cuales tienen un mayor número de arreglos en la zona variable (ZV) del integrón que los contiene y a cuales genes *cassettes* pueden estar asociados.

Con el fin de evaluar si hay conservación de las secuencias que codifican para el gen de resistencia o específicamente para aquellas que están involucradas en la recombinación, se analizó la secuencia de los distintos *cassettes* completos (ORF+*attC*), el ORF y el sitio *attC* por separado a través de un alineamiento múltiple de secuencias usando ClustalW y adicionalmente un análisis de filogenia por máxima similitud, con los programa MEGA6 y CLC *Sequence Viewer* 7.6.

### **3.1.2- Análisis de los arreglos de los genes *cassettes aadA-like* de integrones clínicos**

A través de la búsqueda de los distintos alelos del gen *cassette aadA-like* reportados en la base de datos de INTEGRALL, se realizó un análisis de los arreglos genéticos en los cuales estuvieran contenidos cada uno de estos alelos, con el fin de evidenciar cuál de estos alelos es el más frecuente, a qué tipo de integrasa están asociados, su posición en el en la zona variable del integrón y los genes *cassettes* a los cuales está asociado.

Para evaluar si el gen *cassette aadA1* y los genes *cassettes aadA-like* más frecuentes (*aadA2* y *aadA5*) se encuentran preferencialmente ubicados en una posición en particular de la zona variable del integrón, identificamos en los registros obtenidos en INTEGRALL cuántos de ellos tenían dichos genes *cassettes* y sus posiciones.

Para ello validamos los registros obtenidos con el test exacto de Fisher realizando una comparación entre las distintas posiciones encontradas para el gen *cassette aadA1*, *aadA2* y *aadA5* en integrones de clase 1 e integrones de clase 2 comparándolas todas contra todas para evidenciar si había una diferencia significativa entre la preferencia de aparición de las posiciones en las que encontramos a estos genes *cassettes*.

### **3.1.3- Dispersión de las variantes de los sitios *attC* de los genes *cassette aadA-like* y del sitio *attC<sub>aadA1</sub>***

A partir de los resultados obtenidos del análisis de los sitios *attC* identificados para los distintos alelos *aadA-like*, se evaluó la dispersión de cada una de las variantes identificadas

## *Materiales y métodos*

en la base de datos de GenBank, con el fin de identificar cuál de estas variantes es la más frecuente. Adicionalmente, usamos una secuencia de referencia del sitio *attC* del gen *cassette aadA1* GenBank Accession Number: KR827036.1. Usada por Quiroga, MP 2012 Tesis Doctoral y Roy, 2011 para evaluar su dispersión. Para lograr este objetivo se realizó un MEGABLAST de esta secuencia en particular en la base de datos GenBank y estudiamos 200 secuencias que tuvieran un 100% de identidad, con el fin de ver si dicha secuencia siempre se encuentra asociada al gen *cassette aadA1*, a genes de resistencia similar o a otros genes, adicionalmente ver si están contenidos en integrones, si éstos son de clase 1 o de clase 2 y determinar a que otros *cassettes* de resistencia antibiótica están asociados.

Adicionalmente, se recopiló la información contenida en el output del MEGABLAST de la secuencia del sitio *attC* del gen *cassette aadA1* con el fin de aplicarle un filtro para el total de las secuencias con 100% de identidad y 100% de cobertura del sitio *attC* del gen *cassette aadA1*, y determinar si están relacionadas otro gen *cassette* distinto y/o identificar si estaba asociada a otras estructuras genéticas alejadas de los integrones y la resistencia antibiótica. Para ello empleamos un “Script” diseñado por la Lic. Verónica Álvarez (Estudiante Doctoral de nuestro laboratorio, trabaja en Bioinformática) que permitió obtener los números de acceso de aquellas secuencias que no estuvieran anotadas de tal manera que nos dieran como resultado algo relacionado a adenytransferasas, resistencia a estreptomicina/espectinomicina, integrones y específicamente el gen *cassette aadA1*.

### **3.2- Ensayos funcionales de recombinación para el gen *cassette aadA1***

#### **3.2.1- Selección del modelo a estudiar**

Se realizó la selección del modelo a estudiar en una estructura genética con diferentes sitios de recombinación (*att*) *upstream* del gen *cassette aadA1*. En el estudio de tesis doctoral de MP Quiroga, 2012 se observó una variación en las frecuencias de recombinación para las distintas estructuras analizadas *attI1-ORF-attC* como *attC-ORF-attC*, por lo cual el presente estudio profundizó sobre cómo se ven afectadas las frecuencias de recombinación cuando varía el sitio *att* localizado *upstream* de un *cassette*. Para ello se tomó como modelo al gen *cassette aadA1*, que confiere resistencia a estreptomicina/espectinomicina.

#### **3.2.2- Plásmidos recombinantes y cepas utilizadas en este estudio**

En esta sección se describen los clones utilizados durante la tesis para llevar a cabo los ensayos de recombinación. Utilizamos las cepas de *E. coli* DH10B o TOP10 como huésped para los clones generados. Realizamos las extracciones plasmídicas y las transformaciones como se detallará más adelante.

Los plásmidos recombinantes denominados como **pLQ423**, **pLQ443**, **pLQ445** y **pLQ426** fueron cedidos por el Dr. Paul H. Roy y la Dra. Nancy Messier (Gravel y col., 1998a). El plásmido **PMI1-1 (AS)2** portador de la integrasa de tipo 1 y un sitio *attI1* y fue construido en el laboratorio (Quiroga. MP, 2012). La selección de los clones contenedores de los diferentes sitios *attC upstream* del gen *cassette aadA1* se basó en la elección de un representante por

## Materiales y métodos

grupo de los distintos grupos que se observaron en el análisis filogenético de distintos sitios *attC* realizado por MP Quiroga (Tesis doctoral, 2012). (Figura 7, capítulo Introducción)

Para la realización de este estudio se usaron los clones y cepas detalladas en la tabla 1: **pLQ423: *attI1-aadA1-attC<sub>aadA1</sub>***, **pLQ443: *attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>***, **pLQ445: *attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>*** y **pLQ426: *attC<sub>dhfrA1</sub>-aadA1-attC<sub>aadA1</sub>***. Adicionalmente se emplearon los generados por la doctora María Paula Quiroga en su tesis doctoral. Para realizar los ensayos de recombinación se usó la cepa *E.coli* TOP 10.

**Tabla 1.** Plásmidos usados en este estudio para realizar los ensayos de recombinación.

| Plásmidos   | Descripción  | Fuente              |
|-------------|--|---------------------|
| pLQ423      | <i>attI1-aadA1-attC<sub>aadA1</sub></i> de Tn21 clonado en el vector pACYC184, Cl <sup>R</sup>                       | P. Roy y N. Messier |
| pLQ443      | <i>attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub></i> clonado en el vector pACYC184, Cl <sup>R</sup>                | P. Roy y N. Messier |
| pLQ445      | <i>attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub></i> clonado en el vector pACYC184, Cl <sup>R</sup>                | P. Roy y N. Messier |
| pLQ426      | <i>attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub></i> clonado en el vector pACYC184, Cl <sup>R</sup>                | P. Roy y N. Messier |
| PMI1-1(AS)2 | <i>intI1-attI1</i> (pLQ3poly:: <i>attI1</i> – <i>attI1</i> S al M13F) clonado en el vector pMAL-c2, Amp <sup>R</sup> | Quiroga MP.         |

Amp<sup>R</sup> y Cl<sup>R</sup> indican resistencia a ampicilina y cloranfenicol, respectivamente.

### 3.2.3- Reactivos utilizados de rutina

Se emplearon los siguientes medios de cultivo en los ensayos realizados: caldo Luria Bertani (LB, Difco), LB agar. En los casos en los que fue necesario, se adicionaron a los medios de cultivo los siguientes antibióticos en las concentraciones indicadas: ampicilina 100 µg/ml y/o cloranfenicol 50 µg/ml, ambos de Sigma.

Los reactivos utilizados en los ensayos de biología molecular eran de las marcas Invitrogen, Promega, Fermentas, InBio, New England BioLabs y Roche. Utilizamos los siguientes Kits comerciales: Kits de extracción plasmídica “QIAprep Spin Miniprep” y kit de

## *Materiales y métodos*

purificación de ADN “PCR NucleoSpin Extract II” (Macherey-Nagel). Los reactivos químicos utilizados de rutina eran de las marcas Merck, Invitrogen y Sigma.

### **3.2.4- Condiciones de crecimiento de los cultivos bacterianos**

Se repicaron los aislamientos bacterianos en los medios sólidos LB agar, y se incubaron en placas a 37°C durante 24hs. Posteriormente, para la extracción del ADN de los aislamientos se tomó una colonia y se realizó un cultivo en medio LB líquido suplementado con el antibiótico correspondiente, el cual se incubó durante 18hs a 37°C con agitación a 180-200 revoluciones por minuto (rpm).

### **3.2.5- Extracción de ADN genómico total y de ADN plasmídico**

Para la extracción del ADN por el método de hervido, se coloca una colonia en 20 µl de H<sub>2</sub>O destilada, la incubamos a 100°C durante 5 min, luego a 4°C durante 1 min y finalmente se centrifugó la muestra a 13000 rpm durante 30 seg. La extracción de ADN plasmídico se realizó con el kit comercial “QIAprep Spin Miniprep” o a través de la técnica de extracción manual “Miniprep” (Sambrook y Russell, 2011) y “Megaplasmid” (Herrick y col., 2007)

### **3.2.6- Amplificación por reacción en cadena de la polimerasa (PCR)**

En las reacciones de amplificación por PCR se utilizaron los cebadores específicos para cada caso (Tabla 2). Las mezclas de reacción se llevaron a cabo en un volumen final de 25 µl que contenía 1 µM de cada cebador, 0,2 µl de cada dNTP, 5µl de buffer que genera una concentración final de 1,5 mM de MgCl<sub>2</sub> y 1.25 U de Taq ADN polimerasa (GoTaq de



## *Materiales y métodos*

Promega y Taq polimerasa de InBio). Se agregó entre 5-10 ng de ADN a partir de una dilución de 10ng/ $\mu$ l, o 5 $\mu$ l de la preparación de hervido de colonia, según el caso, preparado se describió en el punto anterior.

**Tabla 2.** Cebadores usados en este estudio para evaluar los ensayos de recombinación.

| <b>Cebador</b> | <b>Secuencia 5' a 3'</b> | <b>Fuente</b>                   |
|----------------|--------------------------|---------------------------------|
| PACYC184 5'    | TGTAGCACCTGAAGTCAGCC     | Gravel y col., 1998b            |
| PACYC184 3'N   | GTGATGTCGGCGATATAGGC     | Ramírez MS tesis doctoral, 2008 |
| M13Fndel       | CATATGGTAAAACGACGGCCAG   | Quiroga MP tesis doctoral, 2012 |
| M13Rndel       | CATATGCAGGAAACAGCTATGAC  | Quiroga MP tesis doctoral, 2012 |
| aadA1          | TCGATGACGCCAACTAC        | Ramírez MS tesis doctoral, 2008 |
| aadAR          | CGCAGATCAGTTGGAAG        | Ramírez MS tesis doctoral, 2008 |

Las amplificaciones por PCR se llevaron a cabo en un ciclador térmico (Eppendorf Mastercycler, MJ Research PTC-200 o Labnet MultiGene™ Gradient), utilizando el siguiente ciclo: I) 95°C – durante 5 min., II) 95°C – durante 1min., III) temperatura de apareamiento de los cebadores (Ta) según corresponda durante 30 seg., IV) 72°C – durante 1 min./Kb, V) Con repetición de los pasos II, III y IV 35 veces, y VI) por último realizamos una extensión final a 72°C – durante 5 o 10 min.

El análisis de los productos de PCR obtenidos se realizó por electroforesis en geles de agarosa 0.8% a 1.5% (preparados con buffer TBE 1X y Bromuro de Etidio para posterior visualización de las bandas al UV). Se utilizó el buffer de corrida TBE 1X y realizamos las corridas a 100V-110V durante 40-60 min. y se visualizaron las bandas con lámpara de luz Ultra Violeta (UV).

### **3.2.7- Secuenciación de ADN y análisis de secuencias.**

Para confirmar algunos de los resultados obtenidos, se secuenció en ambas cadenas con el equipo ABIPrism 3100 BioAnalyzer los productos de amplificación por PCR, así como los productos de recombinación obtenidos en los ensayos.

El análisis de las secuencias se realizó con los siguientes programas:

- Bioedit y Sequencher (Gene Codes Corporation) para corregir las secuencias obtenidas y armar los contigs.
- Algoritmo Blastn del programa Basic Local Alignment Search Tool (BLAST) del sitio web PubMed (<http://www.ncbi.nlm.nih.gov/pubmed/>), dejando los parámetros estándar para comparar las secuencias obtenidas con las de referencia.

### **3.2.8- Preparación de células competentes y transformación bacteriana**

La preparación de células competentes se realizó usando una modificación del protocolo desarrollado por Sambrook y Russell (2001):

Se toma una colonia bacteriana de la célula a ser transformada y se incuba en 5 ml de LB líquido a 37°C durante 18hs con agitación a 180 rpm. A partir de ahí, se toma una alícuota 1/100 del crecimiento bacteriano y se incuba a 37°C con agitación a 180 rpm durante aproximadamente 3 hs (hasta alcanzar una D.O. de 0,5). Luego se centrifuga todo el cultivo a 14.000 rpm durante 1 min. Posteriormente se hace un lavado con CaCl<sub>2</sub> 100 mM estéril tomando una alícuota 1/12,5 del volumen inicial y se mantiene a 4°C durante 30 min. Se centrifuga a 14000 rpm durante 15 min. Se procede a resuspender las células en una alícuota de ¼ del volumen inicial de CaCl<sub>2</sub> 100 mM estéril y se realiza el protocolo de transformación o

### *Materiales y métodos*

se adiciona un 10% de glicerol para guardarlas a  $-80^{\circ}\text{C}$  por un máximo de 6 meses. Es fundamental mantener la esterilidad durante todo el protocolo para evitar la contaminación de las células.

Las transformaciones de ADN plasmídico dentro de una célula bacteriana se hacen siguiendo una modificación del protocolo de transformación desarrollado por Sambrook y Russell (2001):

Se toman 50  $\mu\text{l}$  de células y se les adiciona 1-2  $\mu\text{l}$  (50ng/ $\mu\text{l}$ ) de ADN plasmídico. Se mantiene esta mezcla en hielo durante 30 min. y luego se realiza el shock térmico a  $42^{\circ}\text{C}$  durante 90 segundos. Posteriormente, se adicionan 200  $\mu\text{l}$  de medio LB y se incuba durante 1 hora a  $37^{\circ}\text{C}$ . Finalmente se plaquean 50  $\mu\text{l}$  de las células tratadas en LB agar suplementado con el antibiótico apropiado para la selección de las células transformadas.

#### **3.2.9- Ensayos de escisión de genes *cassettes in vivo* mediados por *Int1***

Realizamos los ensayos *in vivo* de escisión de *cassettes* (Figura 8) siguiendo la metodología descrita por Gravel y col. (1998b) que describimos a continuación:

**Día 1:** Cultivo de células conteniendo el plásmido pMI-1(AS)2 en 5 ml de LB líquido con ampicilina (100 $\mu\text{g}/\mu\text{l}$ ) e incubarlas durante 18 hs. a  $37^{\circ}\text{C}$  con agitación.

**Día 2:** Tomar una alícuota del cultivo para preparar células competentes y luego transformar 50 $\mu\text{l}$  de esas células con 2  $\mu\text{l}$  de extracción plasmídica del plásmido recombinante que contiene el gen *cassette* cuya escisión se quiere evaluar (pLQ423, pLQ443, pLQ445, pLQ426). Plaquear en LB agar suplementado con ampicilina(100 $\mu\text{g}/\mu\text{l}$ ) y cloranfenicol (50 $\mu\text{g}/\mu\text{l}$ ). Incubar 18 hs. a  $37^{\circ}\text{C}$ .

## *Materiales y métodos*

**Día 3:** Hacer un cultivo de una colonia doble transformante en 5 ml de LB líquido suplementado con ampicilina y cloranfenicol. Incubar 18 hs. a 37°C con agitación.

**Día 4:** Tomar una alícuota del cultivo (10%) para llevar a cabo el ensayo de escisión propiamente dicho en 5ml finales de LB líquido suplementado con ampicilina y cloranfenicol. Incubar 3 hs. a 37°C con agitación (hasta alcanzar una D.O. de 0,5). Inducir la expresión del gen *int11* adicionando 0,5mM de IPTG (isopropil- $\beta$ , D-tiogalactopiranosido) y seguir incubando durante 18hs a 37°C con agitación.

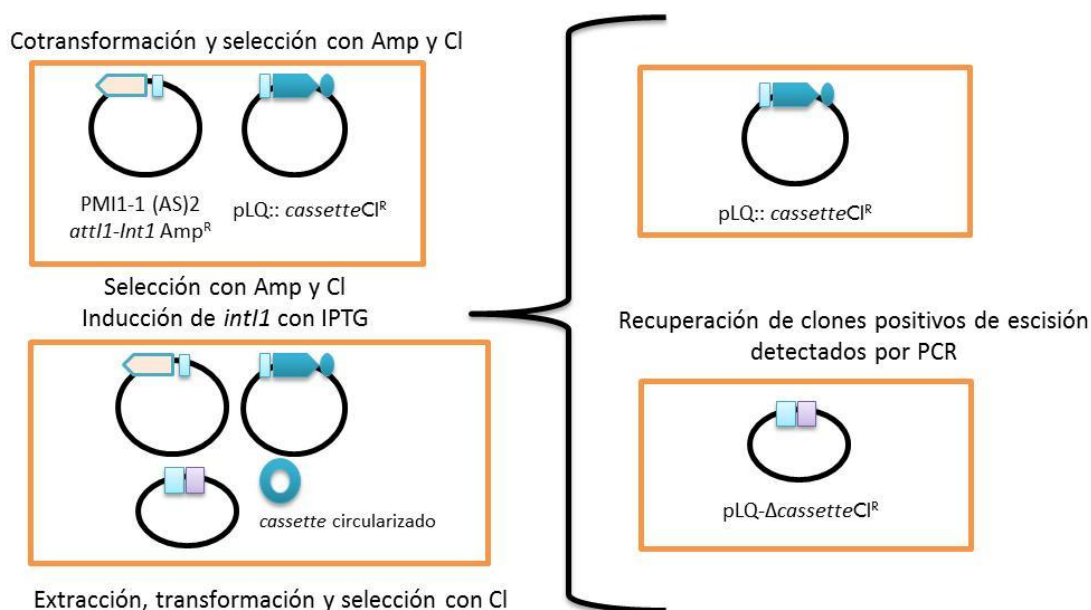
**Día 5:** Realizar la extracción plasmídica y luego realizar amplificaciones por PCR con los cebadores correspondientes para determinar si ha ocurrido la escisión del gen *cassette* que queremos evaluar.

**Día 6:** Corrida en gel de agarosa de las amplificaciones por PCR. Si ha ocurrido la escisión, transformar 50  $\mu$ l células competentes de *E.coli* TOP10 con 2 $\mu$ l de la extracción plasmídica (del día 5 del ensayo) para luego analizar la frecuencia de escisión. Seleccionar en LB agar suplementado con cloranfenicol. Incubar 18 hs. a 37°C.

**Día 7:** Realizar amplificaciones por PCR (utilizando hervidos individuales de 30 colonias de la transformación del día 6 del ensayo) con los mismos cebadores usados en la PCR anterior.

**Día 8:** Corrida en gel de agarosa de las amplificaciones por PCR y en base a los resultados calcular la frecuencia de escisión del gen *cassette* analizado. Para determinar la frecuencia de recombinación del gen *cassette*, se saca un promedio de las frecuencias obtenidas en al menos 3 ensayos independientes. Los productos de amplificación positivos se confirman por secuenciación al igual que los plásmidos recuperados.

## Materiales y métodos



**Figura 8.** Diagrama ilustrativo del ensayo de escisión *in vivo*. Amp<sup>R</sup> y Cl<sup>R</sup> indican resistencia a ampicilina y cloranfenicol, respectivamente.

### 3.2.10- Ensayos de inserción de *cassettes* en el sitio *attI1* *in vivo* mediados por *IntI1*

Los ensayos *in vivo* de inserción de genes *cassettes* (Figura 9) tienen pequeñas variaciones con respecto al de escisión que describimos en el apartado anterior:

**Día 1:** Cultivo de células conteniendo el plásmido pMI-1(AS)2 en 5 ml de LB líquido con ampicilina e incubarlas durante 18 hs. a 37°C con agitación.

**Día 2:** Tomar una alícuota del cultivo para preparar células competentes y luego transformar 50µl de esas células con 2 µl de extracción plasmídica del plásmido recombinante que contiene el *cassette* (pLQ423, pLQ443, pLQ445, pLQ426) cuya inserción en el sitio *attI1* se quiere evaluar. Plaques en LB agar suplementado con ampicilina y cloranfenicol. Incubar 18 hs. a 37°C.

## *Materiales y métodos*

**Día 3:** Preparar un cultivo de una colonia doble transformante en 5 ml de LB líquido suplementado con ampicilina y cloranfenicol. Incubar 18 hs. a 37°C con agitación.

**Día 4:** Tomar una alícuota del cultivo (10%) para llevar a cabo el ensayo de inserción propiamente dicho en 5ml finales de LB líquido suplementado con ampicilina y cloranfenicol. Incubar 3 hs. a 37°C con agitación (hasta alcanzar una D.O. de 0,5). Inducir la expresión del gen *int11* adicionando 0,5mM de IPTG (isopropil- $\beta$ , D-tiogalactopiranosido) y seguir incubando durante 18hs a 37°C con agitación.

**Día 5:** Realizar la extracción plasmídica y luego realizar amplificaciones por PCR con las combinaciones adecuadas de cebadores para determinar si ha ocurrido la inserción del *cassette* en el sitio *att11* en los lados 5' y 3'.

**Día 6:** Corrida en gel de agarosa de las amplificaciones por PCR. Si ha ocurrido la inserción, transformar 50  $\mu$ l células competentes de *E.coli* TOP10 con 2 $\mu$ l de la extracción plasmídica (del día 5 del ensayo) para luego analizar la frecuencia de inserción. Seleccionar en LB agar suplementado con ampicilina. Incubar 18 hs. a 37°C.

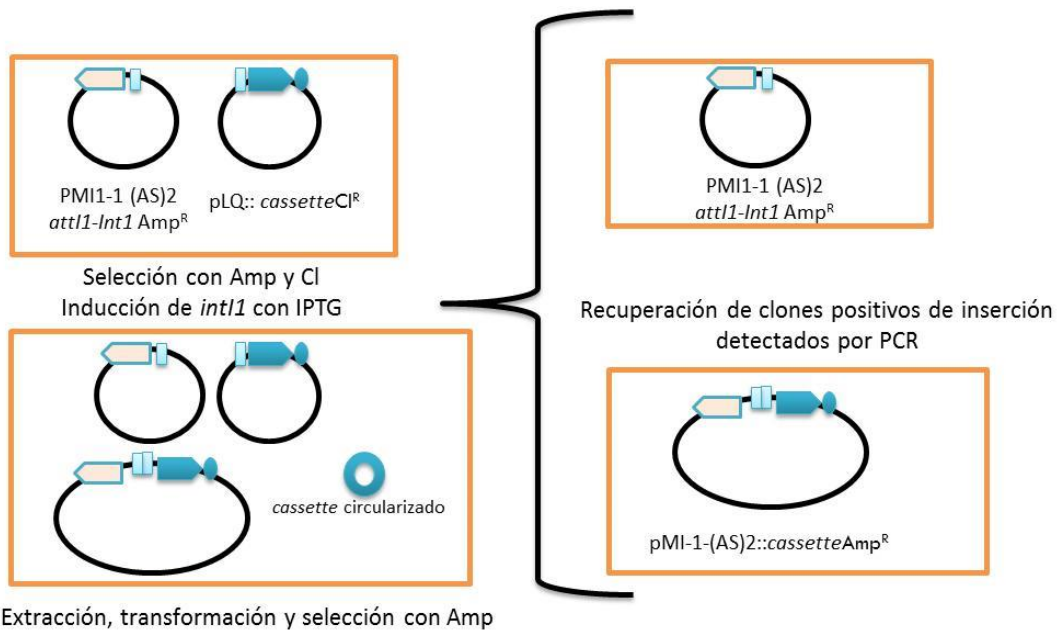
**Día 7:** Realizar amplificaciones por PCR (utilizando hervidos individuales de 30 colonias de la transformación del día 6 del ensayo) con los mismos cebadores usados en las PCRs anteriores.

**Día 8:** Corrida en gel de agarosa de las amplificaciones por PCR y en base a los resultados calcular las frecuencias de inserción del *cassette* analizado. Para determinar la frecuencia de recombinación del *cassette*, se saca un promedio de las frecuencias obtenidas en al menos 3 ensayos independientes. Los productos de amplificación positivos se confirman por

## Materiales y métodos

secuenciación al igual que los plásmidos recuperados para terminar de confirmar las inserciones netas o de los lados 5' o 3'.

Tanto en los ensayos de escisión como de inserción, como control positivo realizamos en paralelo ensayos para el gen *cassette attI-aadB-attC* en el clon PAO1BC1 en las mismas condiciones, en presencia de los clones que contienen al gen *intI1*. Como control negativo, realizamos los ensayos sin la presencia del gen *intI1*.



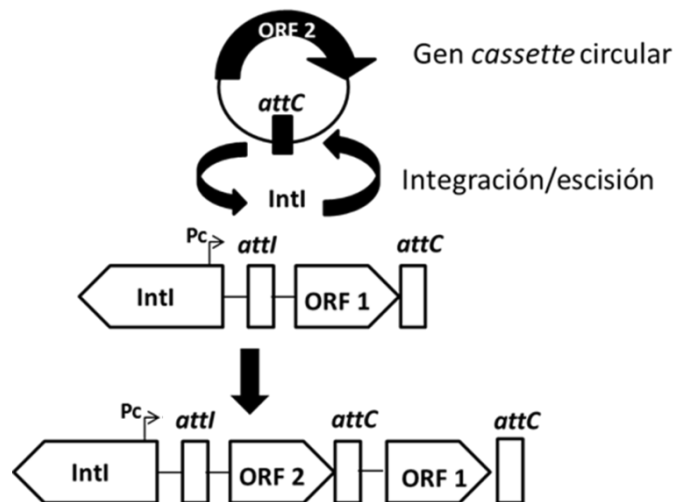
**Figura 9.** Diagrama ilustrativo del ensayo de inserción *in vivo* en el sitio *attI1*. Amp<sup>R</sup> y Cl<sup>R</sup> indican resistencia a ampicilina y cloranfenicol, respectivamente

# Capítulo 4: Resultados y Discusión

## 4.1- Análisis de bioinformática de los alelos del gen *cassette aadA-like*.

### 4.1.1- Identificación de los alelos de los genes *cassettes aadA-like*.

A través de la búsqueda en la base de datos INTEGRALL encontramos 1.079 registros de secuencias totales o parciales de genes *cassettes aadA-like* (Junio de 2015). Para los objetivos de nuestro análisis requeríamos de la secuencia completa del gen *cassette*, conformado por un ORF y el sitio de recombinación *attC*, el cual se observa en el esquema general de la Figura 10. Para este estudio usamos como referencia el gen *cassette aadA1* (AN: KR827036.1, ORF: 852pb y su correspondiente sitio *attC*: 53pb).



**Figura 10.** Esquema de un gen *cassette*. El gen *cassette* consiste en un ORF simple, generalmente sin promotor y su respectivo sitio de recombinación denominado *attC*. En su forma libre el gen *cassette* se encuentra circularizado y es recombinado en un integrón por medio de una integrasa. Modificado de: Stokes y col., 2009.

Por esa razón, descartamos 39 registros de secuencias parciales que correspondían a los genes *cassettes aadA10*, *aadA1Δ6*, *aadA1e*, *aadA5D6* y *aadA6/aadA10*. En total



## Resultados y discusión

identificamos 37 alelos de genes *cassettes aadA-like* contenidos en 1.040 registros (Tabla 3). En la Tabla 3 se muestran los nombres de los alelos encontrados, la cantidad de registros de cada uno y la integrasa o integrasas asociadas a ellos. Es importante destacar, que únicamente el alelo del gen *cassette aadA1* está asociado a integrasas de tipo 1 y tipo 2. El resto de los alelos está asociado únicamente a integrasas de tipo 1, incluyendo en una menor cantidad a integrasas de tipo 1 truncadas.

**Tabla 3. Descripción de los alelos de los genes *cassette aadA-like* reportados en la base de datos INTEGRALL.**

| <i>Genes cassettes aadA-like</i> | Número de registros | Integrasa asociada                                       | Número de arreglos en la zona variable del integrón      |
|----------------------------------|---------------------|--|--|
| <i>aadA</i>                      | 6                   | <i>intI1</i>   | 1  |
| <i>aadA1</i>                     | 468                 | <i>intI1</i> n=399/ <i>ΔintI1</i> n=5/ <i>intI2</i> n=69 | <i>intI1</i> n=174/ <i>ΔintI1</i> n=5/ <i>intI2</i> n=19 |
| <i>aadA11</i>                    | 12                  | <i>intI1</i>   | 9  |
| <i>aadA11c</i>                   | 1                   | <i>intI1</i>   | 1  |
| <i>aadA12</i>                    | 3                   | <i>intI1</i>   | 1  |
| <i>aadA13</i>                    | 14                  | <i>intI1</i>   | 7  |
| <i>aadA15</i>                    | 2                   | <i>intI1</i>   | 2  |
| <i>aadA16</i>                    | 4                   | <i>intI1</i>   | 4  |
| <i>aadA17</i>                    | 1                   | <i>intI1</i>   | 1  |
| <i>aadA1a</i>                    | 15                  | <i>intI1</i>   | 11   |
| <i>aadA1b</i>                    | 7                   | <i>intI1/ΔintI1</i>                                      | 5  |
| <i>aadA1bq</i>                   | 1                   | <i>ΔintI1</i>  | 1  |
| <i>aadA1bs</i>                   | 1                   | <i>intI1</i>   | 1  |
| <i>aadA1c</i>                    | 6                   | <i>intI1</i>   | 1  |
| <i>aadA1k</i>                    | 1                   | <i>intI1</i>   | 1  |
| <i>aadA2</i>                     | 273                 | <i>intI1/ΔintI1</i>                                      | 120  |
| <i>aadA21</i>                    | 2                   | <i>intI1</i>   | 2  |
| <i>aadA22</i>                    | 5                   | <i>intI1</i>   | 2  |
| <i>aadA23</i>                    | 2                   | <i>intI1</i>   | 1  |
| <i>aadA23b</i>                   | 2                   | <i>intI1</i>   | 1  |
| <i>aadA24</i>                    | 5                   | <i>intI1</i>   | 3  |
| <i>aadA2a</i>                    | 3                   | <i>intI1</i>   | 3  |
| <i>aadA2c</i>                    | 1                   | <i>intI1</i>   | 1  |
| <i>aadA3</i>                     | 1                   | <i>intI1</i>   | 1  |
| <i>aadA3c</i>                    | 2                   | <i>intI1</i>   | 1  |
| <i>aadA4</i>                     | 11                  | <i>intI1</i>   | 7  |
| <i>aadA4c</i>                    | 1                   | <i>intI1</i>   | 1  |
| <i>aadA5</i>                     | 137                 | <i>intI1</i>   | 28   |
| <i>aadA5q</i>                    | 1                   | <i>intI1</i>   | 1  |
| <i>aadA5r</i>                    | 1                   | <i>intI1</i>   | 1  |
| <i>aadA6</i>                     | 20                  | <i>intI1</i>   | 14   |
| <i>aadA6j</i>                    | 1                   | <i>intI1</i>   | 1  |
| <i>aadA7</i>                     | 16                  | <i>intI1</i>   | 7  |
| <i>aadA7-like</i>                | 1                   | <i>intI1</i>   | 1  |
| <i>aadA8</i>                     | 4                   | <i>intI1</i>   | 2  |
| <i>aadA8b</i>                    | 2                   | <i>intI1</i>   | 2  |
| <i>aadA9</i>                     | 8                   | <i>intI1</i>   | 4  |

#### **4.1.2- Análisis de bioinformática aplicada a cada alelo *aadA-like*: abundancia y región variable de integrón**

##### **4.1.2.1- Abundancia de cada alelo *aadA-like***

Al analizar la cantidad de reportes de cada alelo de los genes *cassettes aadA-like*, se identificó que la mayor cantidad de reportes correspondía al alelo del gen *cassette aadA1* (AN: KR827036.1) con un total de 468 registros. De éstos, 399 corresponden a integrones de clase 1 con 174 arreglos distintos, 5 arreglos en integrones con integrasas deletadas y un total de 69 registros en integrones de clase 2 con 19 arreglos distintos (Tabla 3). En segundo lugar encontramos al alelo del gen *cassette aadA2* (AN: KF914298.1) con 273 registros, todos asociados a integrones de clase 1 con un total de 120 arreglos distintos. En tercer lugar, identificamos al alelo del gen *cassette aadA5* (AN: KP642031.1) con 137 registros y un total de 28 arreglos distintos.

Llamativamente, los valores de los demás alelos bajan en un orden de magnitud, encontrando a los genes *cassettes aadA6* (AN AF140629.1), *aadA7* (AN: AF224733.1), *aadA1a* (AN: DQ221760.1), *aadA13* (AN: HM175875.1), *aadA11* (AN: GQ900749.1) y *aadA4* (AN: AF364344.1) con valores entre los 10 y 20 registros. El resto de alelos presentaron una cantidad menor de registros y de arreglos.

Con respecto a estos resultados, se puede observar que el alelo del gen *cassette aadA1* se destaca no solo por su prevalencia, sino por su diseminación tanto en integrones de

## Resultados y discusión

clase 1 como de clase 2, y por la gran cantidad de arreglos encontrados para ambas asociaciones.

### 4.1.2.2- Plataformas genéticas del gen *cassette aadA1* en integrones de clase 1

Como se mencionó anteriormente el gen *cassette aadA1* (AN: KR827036.1) fue el más frecuente en los reportes de la base de datos INTEGRALL, con un total de 468 registros (Tabla 3). Se usó esta misma base de datos para obtener la información de cada uno de los registros del gen *cassette*, analizando en detalle la plataforma genética que lo contiene (integrón de clase 1, 2, u otro) (Tabla 4).

**Tabla 4. Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 1 que incluyen al gen *cassette aadA1*.** Se muestran los 5 arreglos de genes *cassettes* más frecuentes y los restantes se muestran en la Tabla completa (Anexo 1).

| Integrasa    | Estructura de la zona variable   | Cantidad | Nombre del integrón |
|--------------|----------------------------------|----------|---------------------|
| <i>intI1</i> | <i>dfrA1-aadA1</i>               | 79       | In361               |
| <i>intI1</i> | <i>aadA1</i>                     | 74       | In2                 |
| <i>intI1</i> | <i>aacA4-catB8-aadA1</i>         | 14       | In1055              |
| <i>intI1</i> | <i>bla<sub>OXA-1</sub>-aadA1</i> | 12       | In757               |
| <i>intI1</i> | <i>aadB-aadA1</i>                | 5        | In822               |

Como se puede observar en la Tabla 4, el gen *cassette aadA1* está en su mayoría asociado al gen *cassette dfrA1* con un total de 79 registros con el gen *cassette aadA1 downstream* del gen *cassette dfrA1*, en el denominado In361. En segundo lugar de frecuencia de arreglos en la zona variable, encontramos a In2, formado por *aadA1* como único gen *cassette* en la zona variable con un total de 74 registros, y en menor medida asociado a otros genes *cassettes* como *aac(6′)-Ib*, *catB8*, *bla<sub>OXA-1</sub>* y *aadB* en los integrones In1055, In757 e In822, respectivamente. Con la gran cantidad de arreglos en los que se encuentra este gen *cassette*, de los 468 registros encontrados, 156 de ellos contienen al gen *cassette aadA1* en la

## *Resultados y discusión*

segunda posición de la zona variable del integrón, lo que representa un 33,3% de frecuencia de aparición en la segunda posición de la zona variable, siendo el arreglo más frecuente de los integrones de clase 1 para este análisis en particular.

Para comprobar si hay una tendencia a que el gen *cassette aadA1* se encuentre en una posición en particular de la zona variable del integrón, realizamos el análisis estadístico correspondiente encontrando una diferencia altamente significativa al comparar las posiciones 1, 2 y 3 con las restantes ( $P < 0,0001$ ), lo cual muestra que existe una selección para que este gen *cassette* se ubique en estas posiciones. Al compararlas con otras posiciones donde no hay tantos registros como las posiciones 6, 7 y 8 es evidente que se encuentra más frecuentemente en los primeros lugares de la zona variable del integrón. En aquellas posiciones donde los registros tenían frecuencias similares (posiciones 6 y 8 por ejemplo) no hubo diferencia estadística significativa ( $P=1$ ), lo que indica que para aquellas posiciones más alejadas de la región 5' conservada del integrón no hay una tendencia o marcada preferencia de que el gen *cassette* se ubique en alguno de estos lugares.

El nivel de resistencia a un determinado antibiótico, codificado por un gen *cassette* de resistencia, depende de su posición en el integrón, más cercana o lejana del promotor común. Collis y Hall, en el año 1995 demostraron que las bacterias manifiestan niveles de resistencia más elevados cuando el gen *cassette* de resistencia se ubica en la primera posición de la zona variable, esto es, muy cercano al promotor y estos niveles se reducen a medida que los genes *cassettes* se localizan en posiciones más distantes con respecto al

## *Resultados y discusión*

promotor. Existe una gran variabilidad en el nivel de expresión de los genes *cassettes* dentro de un integrón, la cual está relacionada con estructuras genéticas intrínsecas tanto de los integrones como de los genes *cassettes*. En su mayoría, éstos no tienen promotor propio y esta expresión se ha estudiado principalmente a nivel transcripcional. La transcripción de estos genes se lleva a cabo por el promotor Pc (anteriormente llamado P1 o P<sub>ant</sub>) y ocasionalmente por un segundo promotor P2, ambos localizados en la zona 5'CS de los integrones de clase 1. La expresión de los genes *cassettes* varía de acuerdo a las secuencias particulares de las variantes de dichos promotores. Pc está localizado dentro de *intl1* aproximadamente 250pb *upstream* del primer gene *cassette* de la zona variable y P2 se encuentra 90pb *downstream* de Pc. Adicionalmente el promotor denominado Pint1, que media la expresión del gen *intl1*, está localizado *downstream* del promotor Pc, lo que resulta en la convergencia de ambos y en que Pint1 y P2 se superponen; *upstream* del sitio *att1*. (Guerín y col., 2011, Lévesque y col., 1994; Bunny y col., 1995; Collis y Hall, 1995; Recchia y Hall, 1995). Cuando hay múltiples genes *cassettes*, el nivel de expresión se puede medir por el nivel de resistencia antibiótica y puede ser afectado por la distancia entre Pc y el gen de resistencia, debido a un efecto posicional que se cree que se debe a una terminación prematura de la transcripción en los extremos 3' de los genes *cassettes* donde las repeticiones invertidas de los sitios *attC* forman estructuras de tallo y bucle (Hanau-Berçot y col., 2002). Este tipo de estructuras pueden silenciar completamente un gen adyacente *downstream*. Al analizar nuestros resultados teniendo en cuenta estos antecedentes, podemos inferir que el gen *cassette aadA1* se encontrará generalmente entre los que presenten los niveles de expresión más elevados dentro de cada integrón que lo contenga.

## Resultados y discusión

La traducción génica es otro paso limitante para la expresión ya que requiere la presencia de una región de iniciación de la traducción (TIR) que consiste en el codón de iniciación (ATG, GTG o TTG), una secuencia de Shine-Dalgarno (SD) (usualmente 5-AAGGAGG-3) y una distancia entre ambos de 5 a 15 nt. En el trabajo de Hanau-Berçot y col., 2002 se reportó que tanto *aadA1* como *aadA2* tienen 1 secuencia SD con una separación óptima de 6 nucleótidos con respecto al codón de inicio, lo cual podría favorecer su traducción. Llamativamente el 25% de los genes *cassettes* carecen de una TIR definida (Hanau-Berçot y col., 2002). Para compensar esta ausencia, el sitio *attI1* colabora con la presencia constante de un pequeño ORF dentro del mismo, el cual tiene una TIR propia que posee su secuencia SD (GGAG) 8 nt antes del codón de iniciación del ORF y se extiende hasta un codón de terminación frecuentemente localizado dentro del CS 5-GTTRRRY-3. De esta manera, la inserción del gen *cassette* en el sitio *attI1* resulta en la formación de un codón de terminación en fase con el pequeño ORF, el cual codifica un péptido de 11 aminoácidos y es denominado ORF11. O sea que la expresión de varios genes *cassettes* de resistencia en los integrones de clase 1 dependen del ORF11 que se transcribe en conjunto a partir de Pc y/o P2, y promueve su traducción por acoplamiento (Depardieu y col., 2007). En algunos casos se genera un péptido de mayor tamaño como en *bla<sub>OXA-1</sub>*, o incluso fusionado con el producto del gen *cassette* inserto en el sitio *attI1* como ocurre con *aacC1* en Tn1696. En el caso del *cassette aadA1*, su CS está compuesto por la secuencia 5-GTTAAAC-3 el cual forma el codón de terminación TAA para ORF11, que luego de la traducción llevará a la generación del

## *Resultados y discusión*

péptido de 11 aminoácidos, sin que éste quede fusionado a *aadA1* lo cual hubiera podido afectar su capacidad para actuar contra los antibióticos.

En este trabajo en particular se evaluó la dispersión de los genes *cassettes aadA-like* en integrones clínicos y a partir del análisis posicional que realizamos de dichos genes *cassettes*, confirmamos que tienden a estar en las primeras posiciones de la zona variable de los integrones donde están contenidos. Teniendo en cuenta estos resultados y con la información que se mencionó en los párrafos anteriores, podemos inferir que van a tener una mayor expresión al estar más cerca de los promotores, lo cual habría que confirmar realizando estudios a nivel transcripcional y/o traduccional.

### **4.1.2.3- Plataformas genéticas del gen *cassette aadA1* en integrones de clase 2.**

Con respecto al gen *cassette aadA1* en el contexto de los integrones de clase 2, a diferencia de lo descrito anteriormente para los integrones de clase 1, no se encontró como único gen *cassette* en la zona variable de los mismos. El arreglo más frecuente corresponde al integrón In2-4 en el cual *aadA1* se ubica *downstream* de los dos genes *cassette*, *dfrA1* y *sat2*, con un total de 30 registros con esta estructura; y en segundo lugar está asociado únicamente al gen *cassette sat2* en el integrón In2-6 con 7 registros. Con los datos obtenidos en este análisis para los integrones de clase 2, se detectó que el gen *cassette sat2* es el más frecuentemente asociado al gen *cassette aadA1*, con un total de 67 de los 71 registros totales (Tabla 3) y como se describe en detalle en la Tabla 5 aparece en 4 de los 5 arreglos de genes *cassettes* más frecuentes de la zona variable de dichos integrones. En cuanto a la posición

## Resultados y discusión

dentro de la zona variable en integrones de clase 2, al comparar las distintas posiciones del gen *cassette aadA1* encontramos una diferencia altamente significativa ( $P < 0,0001$ ) entre las posiciones 2 y 3 con las restantes de la zona variable del integrón, lo que indica una preferencia a estar ubicado en dichas posiciones. Así mismo, encontramos que al igual que para el análisis de integrones de clase 1, no hay diferencias significativas ( $P=1$ ) entre los genes *cassettes* que se encuentran en posiciones más distantes de la zona variable del integrón, específicamente desde la cuarta hasta la séptima posición.

**Tabla 5. Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 2 que incluyen al gen *cassette aadA1*.** Se muestran los 5 arreglos más frecuentes y los restantes se muestran en la Tabla completa (Anexo 2).

| Integrasa    | Estructura de la zona variable | Cantidad | Nombre del integrón |
|--------------|--------------------------------|----------|---------------------|
| <i>intI2</i> | <i>dfrA1-sat2-aadA1</i>        | 30       | In2-4               |
| <i>intI2</i> | <i>sat 2-aadA1</i>             | 7        | In2-6               |
| <i>intI2</i> | <i>dfrA1-sat2-aadA1-orfX</i>   | 6        | *In2-48             |
| <i>intI2</i> | <i>estX-sat2-aadA1</i>         | 6        | In2-24              |
| <i>intI2</i> | <i>dfrA1-aadA1</i>             | 3        | In2-23              |

\*No especificado en la base de datos, se le asigna el número consecutivo al último integrón anotado en INTEGRALL.

### 4.1.3- Análisis de dispersión del gen *cassette aadA1*.

Para determinar en profundidad la dispersión del gen *cassette aadA1*, se realizó un análisis por MEGABLAST contra las secuencias depositadas en la base de datos GenBank. En este análisis se tomó la secuencia de referencia del sitio *attC<sub>aadA1</sub>*, específicamente la secuencia de la hebra inferior subrayada en la Figura 11 (AN: KR827036.1).

Sitio *attC<sub>aadA1</sub>* doble cadena

```
5' GTCTAACAATTCGTTCAAGCCGACGCCGCTTCGCGGCGGGCTTAACTCAAGCGTTAAAC 3'  
3' CAGATTGTTAAGCAAGTTCGGCTGCGGCGAAGCGCCGCGCCGAATTGAGTTCGCAATTTG 5'
```

Hebra inferior sitio *attC<sub>aadA1</sub>*

```
5' -GTTTAACGCTTGAGTTAAGCCGCGCCGCGAAGCGGCGTCGGCTTGAACGAATTGTTAGAC-3'
```



## Resultados y discusión

**Figura 11.** Secuencia del sitio *attC<sub>aadA1</sub>*. Se muestra la secuencia en forma de doble cadena y posteriormente la hebra inferior en dirección 5' a 3'. Se subraya la secuencia usada para realizar el análisis de dispersión descrito anteriormente (Roy, 2011. Quiroga, 2012).

A partir de los 918 resultados obtenidos desde GenBank, seleccionamos 200 secuencias con un 100% de identidad y 100% de cobertura, dentro de las cuales se estudió si estaban relacionadas específicamente al gen *cassette aadA1*, a genes *cassettes* de resistencias similares y/o a integrones.

Como resultado del análisis de las 200 secuencias encontramos un total de 86 arreglos distintos conteniendo a nuestra secuencia de referencia (sitio *attC<sub>aadA1</sub>*), muchos de los cuales están anotados de manera diferente en la base de datos GenBank. Encontramos que los arreglos más frecuentes corresponden a los integrones de clase 1 In361 e In2, que contienen al gen *cassette dfrA1 upstream* del gen *cassette aadA1* (33 registros) y al gen *cassette aadA1* como único gen *cassette* en la zona variable del integrón (32 registros), respectivamente (Tabla 6). Este resultado confirma lo descrito en el análisis de los datos obtenidos a partir de INTEGRALL. Del total de los registros, 71 corresponden a integrones de clase 1, 11 a integrones de clase 2, 2 que definimos como *Δint1* tras confirmar que presentan integrasas parcialmente deletadas y 2 en donde la secuencia de la integrasa no está incluida en la secuencia del registro y dicha secuencia no presentaba rastros de las regiones conservadas de los integrones que permitieran asociarlos a alguna clase de integrón en particular.

En cuanto a la asociación de la secuencia del sitio *attC<sub>aadA1</sub>* a otros genes *cassettes* de resistencia antibiótica o relacionados encontramos que, en secuencias únicas, parte del sitio

## Resultados y discusión

*attC* del gen *cassette aadA1* (49 nucleótidos) está asociado a varios genes *cassettes* como *aacA3*, *aadA23*, *aadA8*, *aadA3*, *aadA24-like* y *qacH*. Esto se condice con los resultados que reportamos en el punto 4.2.2. (correspondiente al análisis denominado: Estructura genética de los sitios *attCs* de los alelos de los genes *cassettes aadA-like*, el cual está desarrollado en un apartado posterior de este capítulo) donde vimos a través del análisis filogenético de la secuencia del sitio *attC<sub>aadA1</sub>* que dicho sitio, el cual denominamos como variante 1 (Figura 13) constituye el sitio de recombinación de 15 alelos de *aadA1* reportados en INTEGRALL, sin embargo hay un único reporte en nuestro análisis para los genes *cassettes aacA3* y *qacH*, para los cuales habría que estudiar en detalle las características de los sitios *attCs* de todos los reportes de ambos genes *cassettes*, ya que podrían estar asociados con otros sitios *attCs*.

**Tabla 6. Arreglos genéticos asociados *upstream* y *downstream* a la secuencia del sitio *attC<sub>aadA1</sub>* en 200 secuencias con un 100% de identidad y 100% de cobertura provenientes de GenBank. Se muestra el gen de la integrasa a la cual están asociadas y la cantidad de registros totales de cada uno de los arreglos.**

| Estructura/arreglos <i>upstream</i> y <i>downstream</i> del sitio <i>attC<sub>aadA1</sub></i>                                 | Integrasa    | Cantidad de registros |
|---|--------------|-----------------------|
| <i>dfrA1-aadA1</i>  | <i>intI1</i> | 33                    |
| <i>aadA1</i>  | <i>intI1</i> | 32                    |
| <i>ι-psp-aadA2-cmlA1-aadA1-qacH-IS440-sul3</i>  | <i>intI1</i> | 13                    |
| <i>catB8-aadA1</i>  | <i>intI1</i> | 6                     |
| <i>dfrA1-aadA1</i>  | <i>intI2</i> | 7                     |
| <i>dfrA12-sat2-aadA1</i>  | <i>intI2</i> | 18                    |
| <i>aacC1-ORFX-orfY-aadA1</i>  | <i>intI1</i> | 3                     |
| <i>ι-aacC1/aacA1-orfP-orfQ-aadA1-qacEΔ1-sul1-orf5-orf-resx-Trb1-IRTn6018-?</i>  | <i>intI1</i> | 3                     |
| <i>ι-dfrA12-orfF-aadA2-cmlA1-aadA1-qacH-IS440?</i><br>(hibrida en <i>aadA1</i> y en parte de <i>qacH</i> no en <i>aadA2</i> ) | <i>intI1</i> | 3                     |
| <i>ι-IRTn501-orf5Δ-sul1-qacEΔ1-aadA1-int1-IR1-tnpR-tnpA?</i>  | <i>intI1</i> | 2                     |
| <i>aacC1/aacA1-aacC1/aacCA1-orfP-orfQ-aadA1-qacEΔ-sul1-orfF</i>   | <i>intI1</i> | 2                     |
| <i>bla<sub>IMP10</sub>-aacA31-aadA1</i>   | <i>intI1</i> | 2                     |

## Resultados y discusión

|   |               |   |
|---|---------------|---|
| <i>bla<sub>VIM26</sub>-aacA7 (aac(6'))-lb-dfrA1</i> -similar a <i>aadA1</i> no funcional debido a una fusión en el extremo 5'- <i>qacEΔ1-sul1</i> | <i>int11</i>  | 2 |
| <i>sat2-aadA1</i>   | <i>Int12</i>  | 2 |
| <i>dfrB4-catB3-aadA1</i>  | <i>int11</i>  | 2 |
| <i>aac(6')-lb-ereA-bla<sub>OXA21</sub>-aadA1</i>  | <i>int11</i>  | 1 |
| ¿ <i>aac(6')-dfrA1-aadA1-ermE?</i>  | <i>int11</i>  | 1 |
| <i>aac(6')-lb-aac(3')-l-orfx-aad(3')-qacEΔ1-sul1-orf5</i>   | <i>int11</i>  | 1 |
| ¿ <i>aac(6')-lb-catB8-aadA1-qacEΔ1-sul1-ISCR1?</i>  | <i>int11</i>  | 1 |
| ¿ <i>aac(6')-lb-catB8-aadA1-qacEΔ1-sul1-ISCR1-tnpU-arma-tnpD-mel-mph2</i> -Transposasa-proteína hipotética?                                       | <i>int11</i>  | 1 |
| <i>aacA3-arr3-aadA1-TniR-TniQ-TniB-TniA</i> (Pega en una parte del <i>attC</i> del <i>aacA3</i> )   | <i>int11</i>  | 1 |
| <i>aacC1-orfX-orfY-aadA1-qacEΔ1</i>   | <i>int11</i>  | 1 |
| ¿ <i>aadA1-insA-insB</i> -proteína hipotética- <i>catA1-insB-insA-cusS?</i>   | <i>int11</i>  | 1 |
| ¿ <i>aadA23?</i>  | <i>int11</i>  | 1 |
| ¿ <i>aadA8?</i>   | <i>int11</i>  | 1 |
| ¿ <i>aadB-cat-oxa10-aadA1?</i>  | <i>int11</i>  | 1 |
| ¿Cloranfenicol transposasa acetiltransferasa – acetiltransferasa?   | <i>int11</i>  | 1 |
| ¿Proteína de resistencia a aminoglicósidos- <i>istb1?</i>   | <i>Δint11</i> | 1 |
| <i>aacC1/aacCA1-aacC1/aacA1-orfP-orfQ-aadA1-qacEΔ1-sul1</i>   | <i>int11</i>  | 1 |
| <i>aacC1/aacCA1-aacC1/aacA1-orfP-orfP-orfQ-aaadA1-qacE-sul1-orf5</i>  | <i>int11</i>  | 1 |
| <i>aacC1-orfp-orfQ-aadA1-qacEΔ1-sul1-orf5</i>   | <i>int11</i>  | 1 |
| ¿ <i>aacA7-dfrA1-aadA1-smr-ISPα21?</i>  | <i>int11</i>  | 1 |
| ¿ <i>aacC1-orFP-orfQ-aadA?</i>  | <i>int11</i>  | 1 |
| ¿ <i>bla<sub>VIM1</sub>-aac(6')-lb-dfrII-aadA1?</i>   | <i>int11</i>  | 1 |
| ¿ <i>dfrA15-aadA1?</i>  | <i>int11</i>  | 1 |
| <i>bla<sub>IMP33</sub>-aac(6')-lb-bla<sub>OXA2</sub>-aadA1- qacEΔ1-sul1</i>   | <i>int11</i>  | 1 |
| <i>dfrA1-catB2-sat2-aadA1-ybeA-ybfA-ybfB-ybgA</i>   | <i>Int12</i>  | 1 |
| <i>sat2-aadA1-YebA</i>  | <i>int11</i>  | 1 |
| ¿ <i>arr3-aadA1-TniR-TniB-TniA?</i>   | <i>int11</i>  | 1 |
| ¿ <i>dhfr-arr2-cm1A-oxa10-aadA1-qacEΔ-sul1-ISCR1?</i>   | <i>int11</i>  | 1 |
| <i>bla<sub>IMP1</sub>-aac(6')-lb_3-aacA1/gcuG3-gcuG3-aadA1a</i>   | <i>int11</i>  | 1 |
| <i>bla<sub>IMP1</sub>-aacA28-aadA1a</i>   | <i>int11</i>  | 1 |
| <i>bla<sub>GES11</sub>-aac(6')-lb-dfrA7-tnpA-aadA1-qacEΔ</i>  | <i>int11</i>  | 1 |
| <i>bla<sub>IMP6</sub>-aac(6')-lae-aadA1-qacEΔ1</i>  | <i>int11</i>  | 1 |
| <i>bla<sub>VIM-2</sub>-aadA1-qacEΔ1</i>   | <i>int11</i>  | 1 |

## Resultados y discusión

|   |                 |   |
|---|-----------------|---|
| <i>bla<sub>VIM1</sub>-aacA7-dfrI</i> -similar a <i>aadA1</i> no funcional debido a una fusión en el extremo 5'- <i>smr-ISPα21-tnpA</i>                                  | <i>int11</i>    | 1 |
| <i>bla<sub>IMP6</sub>-aacA7-aadA1-qacEΔ1</i>  | <i>int11</i>    | 1 |
| <i>bla<sub>VIM1</sub>-aacA7-dfrA1-aadA1-qacEΔ1</i>  | <i>int11</i>    | 1 |
| <i>bla<sub>VIM1</sub>-aac(6')-lb-aadA1-qacEΔ1-sul1</i>  | <i>int11</i>    | 1 |
| <i>catB8-aadA1</i>  | <i>int11</i>    | 1 |
| ¿ <i>dfrA12-orfF-aadA2-aadA2-cm1A-aadA1-qacH-tnpA-sul3</i> (no pega en el <i>aadA2</i> , sino en el <i>attC</i> de <i>aadA1</i> , y en el <i>attC</i> de <i>qacH</i> )? | <i>int11</i>    | 1 |
| ¿ <i>dfrA14-arr2-bla<sub>OXA10</sub>-aadA1</i> ?  | <i>int11</i>    | 1 |
| ¿ <i>dfrA15-aadA1-qacEΔ1</i> ?  | <i>int11</i>    | 1 |
| <i>dfrA16-bla1pb-aadA2-cmLA1-aadA1</i> (sólo pega en <i>aadA1</i> , no en <i>aadA2</i> )  | <i>int11</i>    | 1 |
| ¿ <i>dfrA16-rifampicina ADP-cm1A7-bla<sub>OXA10</sub>-aadA1-qacEΔ1-sul1-ISCR1</i> ?   | <i>int11</i>    | 1 |
| ¿ <i>dfrA1-aadA1-emrE-sul1-tnIB</i> ?   | <i>int11</i>    | 1 |
| ¿ <i>dfrA1-aadA1-IS440-sul3</i> ?   | <i>int11</i>    | 1 |
| <i>dfrA1-aadA1-emrE-sul1-tnIB</i>   | <i>int11</i>    | 1 |
| ¿ <i>dfrA1-aadA1</i> -proteína hipotética- <i>sul1</i> ?  | <i>int11</i>    | 1 |
| <i>dfrA1-sat1-aadA1</i>   | <i>Int12</i>    | 1 |
| <i>dfrA1-sat2-aadA1</i>   | No está anotada | 1 |
| <i>dfrA1-sat2-aadA1-ybeA</i>  | <i>Int12</i>    | 1 |
| <i>dfrA1-sat2-aadA1-ybfA-ybfB-ybgA</i>  | <i>Int12</i>    | 1 |
| <i>dfrB4-catB3-bla<sub>OXA10</sub>-aadA1</i>  | <i>int11</i>    | 1 |
| ¿Dihidrofolato reductasa- Estreptomina 3'adenilil transferasa-dihidropteroato sintasa- <i>ISCR1</i> ?   | <i>int11</i>    | 1 |
| ¿ <i>estX-aadA1</i> -región codificante no determinada?   | <i>int11</i>    | 1 |
| ¿dihidrofolato reductasa- <i>gap-aadA1</i> ?  | <i>int12</i>    | 1 |
| ¿ <i>estX-aadA1-qacEΔ1</i> ?  | <i>int11</i>    | 1 |
| ¿ <i>estX-sat2-aadA1</i> ?  | <i>Int12</i>    | 1 |
| ¿ <i>InsAB</i> -proteína de resistencia a aminoglicosidos-integrasa- <i>tnpR</i> ?  | <i>int11</i>    | 1 |
| ¿proteína hipotética-dihidropterato sintasa-proteína hipotética-transposasa putativa- <i>aadA1-cm1A- aadA2</i> - hidrolasa- <i>int11</i> - resolvasa <i>tnpR</i> ?      | <i>int11</i>    | 1 |
| ¿proteína hipotética-proteína de resistencia a bromuro de etidio-adenilil transferasa-transportador MFS?  | <i>int11</i>    | 1 |
| ¿ <i>IS26-tnp4-bla<sub>VEB1</sub>-aadB-arr2-cm1A5-bla<sub>OXA10</sub>-aadA1-qacEΔ1-sulΔ</i>   | No está anotada | 1 |
| ¿ <i>IS-int1-XerD-aadA1-InsA-insB-IS-CatA</i> ?   | <i>int11</i>    | 1 |

## Resultados y discusión

|   |               |   |
|---|---------------|---|
| <i>¿istB-sul1-EmrE-aadA1b-dfrA1-int1-strA?</i>  | <i>int11</i>  | 1 |
| ¿proteína de membrana interrumpida?   | <i>int11</i>  | 1 |
| ¿proteína hipotética- proteína de resistencia a aminoglicosidos?  | <i>int11</i>  | 1 |
| Fosforibosiltranilato isomerasa-proteína de resistencia a bleomicina- <i>bla<sub>NDM-1</sub></i> - ISAb125- <i>aadA1</i> - dihidropteroato sintasa - IS6100 | <i>int11</i>  | 1 |
| ¿SFM transportador-adenilil tranferasa?   | <i>int11</i>  | 1 |
| ¿tnpMTn21- <i>int11</i> -Estreptomicina adenililtransferasa- <i>qacEΔ-sul1</i> ?  | <i>int11</i>  | 1 |
| ¿Tn402- <i>sul1-qacEΔ1-aadA1-orfX-orfX(A)-aacC1-aac(6')-Ib-Int11-IS26-tnpA-14base terminal repeat of IS26-Int11-dfrA1-aadA1-IS1R-inSB-inSA?</i>             | <i>int11</i>  | 1 |
| ¿ <i>tnpM-IntR-dfrA15-aadA-qacEΔ1-sull2-tnpA21?</i>   | <i>int11</i>  | 1 |
| ¿Estreptomicina 3''O-adenilil transferasa -3'N-aminoglicosido acetiltransferasa -GroEl (chaperona molecular)  | <i>int11</i>  | 1 |
| ¿ <i>sul3-sul1-qacE-aadA1-int12-strB?</i>   | <i>Int12</i>  | 1 |
| ¿ <i>tnp26-Integrasa-aac(6')-Ib-catB8-aadA1-qacEΔ1-sul1-ISCR1?</i>  | <i>Δint11</i> | 1 |

\*\*Las integrasas que figuran como no anotadas, no fueron posibles de predecir porque la anotación de la secuencia en particular solo contenía la información contenida en esta Tabla.

¿? Las anotaciones marcadas con signos de interrogación corresponden a registros donde la anotación no contenía secuencias correspondientes a las regiones conservadas de los integrones.

Con respecto a la asociación del gen *cassette aadA1* a diferentes integrasas, encontramos en GenBank que en su mayoría está asociado a integrones de clase 1 con un porcentaje de 84% (n=71). Este resultado concuerda con el obtenido a partir de la base de datos INTEGRALL, donde se ve que también hay una mayor asociación a integrones de clase 1. Con respecto a integrones de clase 2, el porcentaje de arreglos de integrones de clase 2 corresponde a un 13% (n=11) de las secuencias y de un 1,5% (n=2) a las *Δint11*.

La mayoría de los genes *cassettes* de *aadA1* están asociados *downstream* a *dfrA1*, como único gen *cassette* de la ZV, o también aparece asociado al *cassette catB8* en integrones de clase 1. Sin embargo, un porcentaje importante de los arreglos (42,5%)

## Resultados y discusión

correspondientes a 85 secuencias de las 200 totales son arreglos únicos, lo que no permite determinar una posición en particular del gen *cassette aadA1* más que las ya mencionadas. Para el caso de integrones de clase 2 también se lo encuentra *downstream* del gen *cassette dfrA1* en el 3,5% de los casos (n=7), pero también está asociado al gen *cassette sat2* con la estructura *dfrA1-sat2-aadA1* como la más predominante en las secuencias analizadas con un 9% correspondiente a 18 de las 200 secuencias analizadas.

**Al igual que lo reportado en INTEGRALL, el gen *cassette aadA1* está estrechamente relacionado al gen *cassette dfrA1* al encontrarse asociados en la mayoría de los registros a excepción de cuando el gen *cassette aadA1* se encuentra como único gen *cassette* en la zona variable del integrón. Su asociación se mantiene tanto en integrones de clase 1 como de clase 2, lo que nos da una idea de que estos dos genes *cassettes* podrían estar altamente asociados en los integrones clínicos. Las bajas frecuencias de recombinación, tanto de escisión como de inserción del gen *cassette dfrA1* en contraposición con las del gen *cassette aadA1* obtenidas en este y otro trabajos, se correlacionan con esta hipótesis.**

Al observar que los resultados del análisis de 200 secuencias indican que el sitio *attC<sub>aadA1</sub>* está únicamente asociado a este alelo en particular, estudiamos el total de las secuencias que contienen un 100% de identidad y 100% de cobertura de la secuencia del sitio *attC<sub>aadA1</sub>* en la base de datos de GenBank. Obtuvimos un total de 918 secuencias para analizar si alguna de ellas correspondía a otro gen *cassette* distinto a lo que ya habíamos observado o adicionalmente identificar si estaba asociada a otras estructuras genéticas diferentes de los integrones o a otros genes *cassettes* de resistencia antibiótica. Para ello empleamos un *script* diseñado por la Lic. en Análisis de Sistemas Verónica Álvarez (Estudiante Doctoral de nuestro laboratorio) que permitió obtener los números de acceso de aquellas

## Resultados y discusión

secuencias que no estuvieran anotadas como relacionadas a adeniltransferasas, resistencia a estreptomycin/espectinomycin, integrones y específicamente el gen *cassette aadA1*. A partir de estas condiciones, obtuvimos un total de 12 secuencias las cuales analizamos en detalle en la Tabla 7.

**Tabla 7. Arreglos genéticos asociados *upstream* y *downstream* a la secuencia del sitio *attC<sub>aadA1</sub>* en 12 secuencias sin nomenclatura con un 100% de identidad y 100% de cobertura provenientes de GenBank. Se muestra la integrasa a la cual están asociadas y la cantidad de registros totales de cada uno de los arreglos.**

| Estructura/arreglos <i>upstream</i> y <i>downstream</i> del sitio <i>attC<sub>aadA1</sub></i>                                  | Integrasa    | Cantidad de registros |
|--|--------------|-----------------------|
| ¿ <i>dhfrA1-aadA1-qacE</i> -dihidropteroato sintasa- <i>sul1</i> -acetiltransferasa putativa?                                  | <i>intI2</i> | 1                     |
| <i>aadB-aadA1-qacEΔ1sul1-ISCR1-bla<sub>CTXM-2</sub></i> (integrón complejo)  | <i>intI1</i> | 1                     |
| ¿Proteína de resistencia a aminoglicósidos-transposasa IS1?  | <i>intI1</i> | 1                     |
| ¿ <i>catB8-aadA1</i> ?   | <i>intI1</i> | 2                     |
| ¿Secuencia parcial <i>aadA24-like</i> ?  | <i>intI1</i> | 1                     |
| <i>bla<sub>IMP10</sub>-aacA31-aadA1a-qacEΔ1-sul1-orf5</i>  | <i>intI1</i> | 1                     |
| InsA-InsB- <i>aadA</i> -IntR-tnpR  | <i>intI1</i> | 1                     |
| ¿ <i>aadB-aadA1-qacEΔ1-sul1-ISCR1-bla<sub>CTXM-2</sub></i> ? (integrón complejo)   | <i>intI1</i> | 1                     |
| <i>aadA-Orf3/QacEΔ1</i> proteína fusionada   | <i>intI1</i> | 1                     |
| Sitios específicos de integración de genes en <i>hot spots</i> de recombinación, flanqueando <i>aadA</i> en transposones Tn21* | No anotada** | 2                     |

\*Se puso la información tal y como estaba anotada en la base de datos de GenBank

\*\*La integrasa que figura como no anotada, no fue posible de predecir porque la anotación de la secuencia en particular solo contenía la información contenida en esta Tabla.

¿? Las anotaciones marcadas con signos de interrogación corresponden a registros donde la anotación no contenía secuencias correspondientes a las regiones conservadas de los integrones.

Pudimos observar que existen discrepancias en la anotación de las secuencias, ya sea por la coexistencia de diferentes nomenclaturas o por ejemplo cuando se refiere a secuencias de genomas completos pueden estar anotadas de manera errónea o directamente no anotadas, requiriéndose una curación previa de las mismas lo cual dificulta el análisis de los resultados. De todas formas, al hacerle un alineamiento a cada una de estas secuencias resultantes de este filtro o *script* con respecto a GenBank observamos que todas

## Resultados y discusión

siguen estando asociadas al gen *cassette aadA1*, que como ya vimos en el análisis de alelos de *aadA1-like* puede estar anotado de distintas maneras, aunque hagan referencia a la misma secuencia.

Según nuestros resultados y el análisis realizado a través de las secuencias del GenBank del sitio *attC<sub>aadA1</sub>* con un 100% de identidad y 100% de cobertura, demostramos que la dispersión del gen *cassette aadA1* es muy alta y está altamente asociada a los integrones clínicos.

### 4.1.4- Plataformas genéticas del gen *cassette aadA2*.

A partir de los datos descritos en la Tabla del conteo total de arreglos de los datos obtenidos a partir de INTEGRALL (Tabla 3) encontramos que el *gen cassette aadA2* fue el segundo *gen cassette* en frecuencia de aparición, con un total de 273 registros y 120 arreglos distintos de *genes cassettes* en la zona variable del integrón donde están contenidos. A diferencia del gen *cassette aadA1*, el gen *cassette aadA2* solo está asociado con integrones de clase 1. A continuación, presentamos los arreglos más frecuentes (Tabla 8).

**Tabla 8. Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 1 que incluyen al gen *cassette aadA2*.** Se muestran los 5 arreglos más frecuentes y los restantes se muestran en la Tabla completa (Anexo 3).

| Integrasa    | Estructura de la zona variable | Cantidad | Nombre del integrón |
|--------------|--------------------------------|----------|---------------------|
| <i>intI1</i> | <i>aadA2</i>                   | 70       | In127               |
| <i>intI1</i> | <i>dfrA12-orfF-aadA2</i>       | 53       | In27                |
| <i>intI1</i> | <i>aadB-aadA2</i>              | 6        | In293               |
| <i>intI1</i> | <i>sat-<i>psp</i>-aadA2</i>    | 4        | *In1089             |
| <i>intI1</i> | <i>aacA4-aadA2</i>             | 3        | In821               |

\*No especificado en la base de datos, se le asigna el número consecutivo al último integrón anotado en INTEGRALL.



## Resultados y discusión

Como se puede ver en la Tabla 8, la mayoría de registros corresponden al gen *cassette aadA2* como único gen *cassette* en la zona variable del integrón (n=70), seguido del arreglo de los genes *cassettes dfrA12* y el gen *orfF upstream* al gen *cassette aadA2* (n=53). En menor frecuencia está asociado a los genes *cassettes aadB, sat, psp* y *aac(6')-Ib* con bajo número de registros (n=13). Con respecto a su posición dentro de la zona variable del integrón encontramos que hay una diferencia altamente significativa ( $P < 0,0001$ ) al comparar la posición 3 con las demás posiciones, lo que nos indica claramente que hay una tendencia a que este gen *cassette* se ubique en la tercera posición de la zona variable de los integrones de clase 1, sin embargo no hay diferencia significativa al comparar esta posición con la primera posición lo que indica que puede ubicarse sin ninguna preferencia en alguna de estas dos posiciones. La comparación entre las posiciones más lejanas, cuarta y quinta, tampoco nos da una diferencia significativa ( $p=0.4177$ ) lo que nos indica que no hay una tendencia a ubicarse en ninguna de estas dos posiciones en particular.

### 4.1.5- Plataformas genéticas del gen *cassette aadA5*.

El gen *cassette aadA5* fue el tercer alelo más frecuente de los genes *cassettes aadA-like* con un total de 137 registros y 28 arreglos de la zona variable (Tabla 3).

Como se puede observar en la Tabla 9, la mayor cantidad de arreglos corresponden a In785 con el gen *cassette dfrA17 upstream* del gen *cassette aadA5*. En menor cantidad se encuentra como único gen *cassette* en la región variable, mientras que en los consiguientes arreglos se encuentra asociado a los genes *cassettes dfrA17* y *dfrA1*. Para el gen *cassette aadA5* encontramos que hay una diferencia altamente significativa ( $P < 0,0001$ ) cuando

## Resultados y discusión

comparamos la posición 2 con las restantes, lo que nos indica que hay una clara tendencia a que el mismo se encuentre en esta posición en integrones de clase 1. Para las posiciones 1 y 3 no hubo diferencia estadística significativa ( $p=1$ ) ya que se reportó el mismo número de registros para ambas posiciones.

**Tabla 9. Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 1 que incluyen al gen *cassette aadA5*. Se muestran los 5 primeros que determinamos como los más frecuentes y los restantes se muestran en la Tabla completa (Anexo 4).**

| Integrasa    | Estructura de la zona variable  | Cantidad | Nombre del integrón |
|--------------|---------------------------------|----------|---------------------|
| <i>intI1</i> | <i>dfrA17-aadA5</i>             | 100      | In785               |
| <i>intI1</i> | <i>aadA5</i>                    | 9        | In681               |
| <i>intI1</i> | <i>dfrA1-aadA5</i>              | 3        | *In1090             |
| <i>intI1</i> | <i>bla<sub>IMP1</sub>-aadA5</i> | 2        | In892               |
| <i>intI1</i> | <i>bla<sub>OXA1</sub>-aadA5</i> | 1        | *In1091             |

\*No especificado en la base de datos, se le asigna el número consecutivo al último integrón anotado en INTEGRALL.

### 4.1.6- Plataformas genéticas de otros alelos de genes *cassettes aadA-like* menos frecuentes.

En menor proporción se reportaron a los genes *cassettes aadA6* (0,86%), *aadA7* (1,15%), *aadA1a* (0,57%), *aadA1b* (0,19%), *aadA11* (0,38%) y *aadA4* (0,48%) asociados a integrones de clase 1. Para el gen *cassette aadA6* el arreglo más frecuente corresponde a la plataforma *aadA6-orfD*, para el gen *cassette aadA7* la plataforma genética más frecuente es *aaC5-aadA7* y para los genes *cassettes* restantes cada uno de ellos se encuentra como único gen *cassette* en la ZV del integrón. En la Tabla 10 se detallan a modo de resumen los arreglos más frecuentes para cada uno de estos alelos.

**Tabla 10. Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 1 que incluyen a los genes *cassettes aadA6, aadA7, aadA1a, aadA1b, aadA11 y aadA4*. Se muestra un resumen de los registros más frecuentes y los restantes se muestran en la Tabla completa (Anexo 5).**

| Integrasa    | Estructura de la zona variable | Cantidad | Nombre del integrón |
|--------------|--------------------------------|----------|---------------------|
| <i>intl1</i> | <i>aadA6-orfD</i>              | 7        | In51                |
| <i>intl1</i> | <i>aacA7-aadA6-orfD</i>        | 2        | In446               |
| <i>intl1</i> | <i>aacC5-aadA7</i>             | 6        | In104               |
| <i>intl1</i> | <i>aadA7</i>                   | 6        | In142               |
| <i>intl1</i> | <i>aadA1a</i>                  | 4        | In2                 |
| <i>intl1</i> | <i>estX-aadA1a</i>             | 2        | In117               |
| <i>intl1</i> | <i>aadA1b</i>                  | 2        | In93                |
| <i>intl1</i> | <i>aadA11</i>                  | 4        | In925               |
| <i>intl1</i> | <i>aadA4</i>                   | 5        | In138               |

A diferencia de los otros genes *cassettes* descritos previamente, la frecuencia de este grupo de genes *cassettes* varía entre los 10 y 20 registros, con una menor cantidad de arreglos (n=56 para 6 genes *cassettes* juntos), resumidos en la Tabla 7. El gen *cassette aadA6* está asociado *upstream* al gen *cassette orfD* y aparece en la primera o segunda posición; el gen *cassette aadA7* está *downstream* al gen *cassette aacC5* o se encuentra como único gen *cassette* en la zona variable; mientras que el gen *cassette aadA1a* está asociado al gen *cassette estX* o se encuentra como único gen *cassette* en la zona variable. Por su lado, cada uno de los genes *cassettes aadA1b, aadA11 y aadA4* se encuentran como únicos genes *cassettes* dentro de su respectiva zona variable.

**Según nuestros resultados, el gen *cassette aadA1* es el más disperso y frecuente de los genes *cassettes aadA-like*, presentando una mayor variabilidad en cuanto a la cantidad de arreglos de la zona variable del integrón y es el único alelo *aadA-like* presente en dos clases de integrones, clase 1 y de clase 2. Todas estas características evidencian su gran contribución a la multirresistencia antibiótica en los integrones clínicos.**

#### **4.2- Estudio de los componentes genéticos de los genes *cassettes aadA-like*: marco abierto de lectura (ORF) y sitios *attCs*.**

Con el fin de estudiar la variabilidad y la evolución de los genes *cassettes aadA-like* se seleccionaron los ORFs de los 37 alelos identificados (ver Tabla 3) usando sus secuencias provenientes de la base de datos INTEGRALL (Tabla 3) y sus respectivos sitios *attCs*. Se seleccionaron estas secuencias porque contenían el ORF completo con un codón de inicio y un codón de terminación definido. Tomamos como gen *cassette* de referencia *aadA1* al AN KR827036.1 de 852pb. Estos 37 alelos *aadA-like* no tienen el mismo tamaño, pero los seleccionamos, ya que contenían los registros completos en las bases de datos INTEGRALL y GenBank. A partir de estos pares de secuencias, se realizó el análisis de alineamiento de secuencias con el algoritmo CLUSTALW para todos los marcos abiertos de lectura u ORFs de cada uno de los alelos y para sus correspondientes sitios *attCs* por separado y así se identificó si su diversificación corresponde a un proceso que se da en paralelo o por separado.

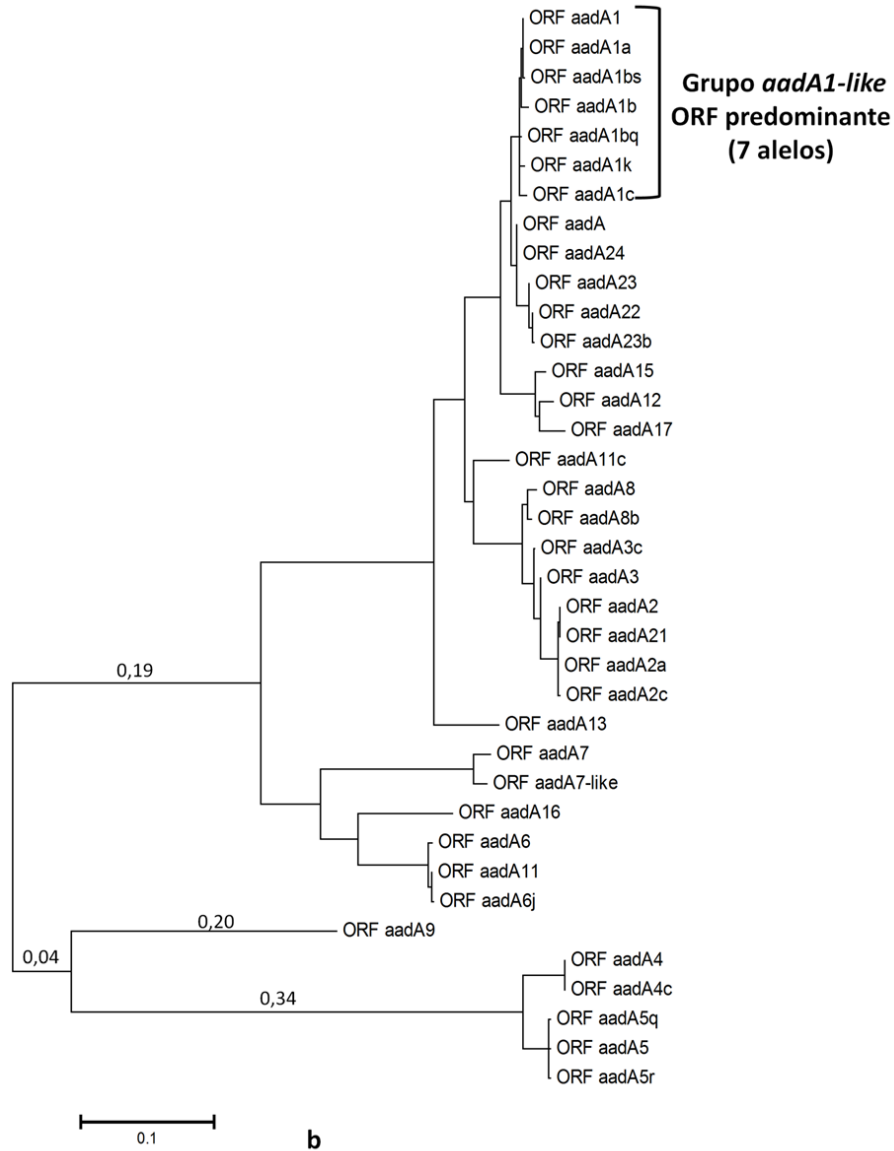
##### **4.2.1- Estructura genética de los ORFs de los alelos *aadA-like***

Al realizar el alineamiento con CLUSTALW y el árbol de máxima similitud para los ORFs de los alelos *aadA-like*, se observa la formación de algunos grupos, destacando el grupo que incluye a los que denominamos como alelos *aadA1-like* (Anexo 6 y Figura 12). Dentro del alineamiento observamos que el ORF correspondiente al gen *cassette aadA* es el más largo, con 1008 nucleótidos, que nuestro gen *cassette* de referencia el *aadA1* que tiene 792 nucleótidos y el ORF del gen *cassette aadA7-like* es el más corto, con 773 nucleótidos, lo cual

### *Resultados y discusión*

deja en evidencia que los ORFs de los genes *cassettes aadA-like* han sufrido múltiples inserciones y/o deleciones.

Con respecto a la conservación de las secuencias, al realizar el alineamiento encontramos 128 nucleótidos conservados, que a partir del porcentaje de los mismos con respecto a la longitud de cada uno de los ORFs determinamos un promedio de 16% de conservación y el menor porcentaje de conservación correspondió al ORF del gen *cassette aadA* (12%).



**Figura 12.** Árbol de máxima similitud entre las secuencias de los ORFs de los alelos *aadA-like*. El grupo que contiene al ORF predominante incluye a 7 alelos no idénticos denominados *aadA1-like*. Se observa una mayor diversidad de formación de grupos para el árbol correspondiente a los ORFs que para el correspondiente a los sitios *attC* de los alelos *aadA-like*. (Rango: 792 a 1008 nucleótidos).

#### 4.2.2- Estructura genética de los sitios *attC*s de los alelos de los genes *cassettes aadA-like*.

Al analizar el alineamiento de los sitios *attC*s correspondientes a los 37 alelos de *aadA-like* identificados en INTEGRALL, encontramos un total de 28 bases conservadas que al calcular un porcentaje entre la longitud del sitio *attC* más largo correspondiente al sitio *attC*

## Resultados y discusión

del gen *cassette aadA9* (61 pb) y el más corto, correspondientes a los sitios *attCs* de los genes *cassettes aadA4* y *aadA4C* (49pb) representa un 51% de conservación de la secuencia.

Se identificaron 15 variantes del sitio *attC* asociadas a los 37 alelos *aadA-like* (Figura 13). Como era de esperar, se observó mayor identidad de secuencias entre las bases nucleotídicas involucradas en la recombinación específicamente en las regiones 1L y 1R (RYYAAC y GTTRRRY, respectivamente), las cuales son clave en los procesos de recombinación mediados por las integrasas de integrón. Con respecto a la región 1L, 11 variantes del sitio *attC* (variantes 1 a 6, 8 a 11 y 13) tienen la secuencia consenso de dicha región, identidad de 100% con la región 1L del sitio *attC<sub>aadA1</sub>* y el total de dichas 11 variantes mencionadas están asociados a 30 alelos *aadA-like*. En el caso de la región 1R, encontramos que 2 secuencias correspondientes a dos alelos que se encuentran dentro de la variante 1 (*aada23b* y *aadA3*) y la variante 9 correspondiente al alelo *aadA6* tienen la secuencia consenso de la región 1R. A su vez, 13 variantes (1 a 6, 6 a 8 y 10 a 15) presentan modificaciones en la secuencia consenso de igual forma (cambiando la última base por una purina) y 2 variantes (5 y 9) sufrieron deleciones que modifican la SC.

Con respecto a las regiones 2L y 2R (GTTCRARY y RYTYAAC, respectivamente), asociadas con la formación de la base extrahelicoidal más relevante, 8 variantes (1, 2, 4, 5, 11 a 13 y 15) tienen la SC de la región 2L, todas con un 100% de identidad con la secuencia de la región 2L del sitio *attC<sub>aadA1</sub>*. En cambio, 5 variantes (3, 6, y 8 a 10) sufrieron modificaciones en su SC pero cumplen con la regla de complementariedad de las bases que normalmente se

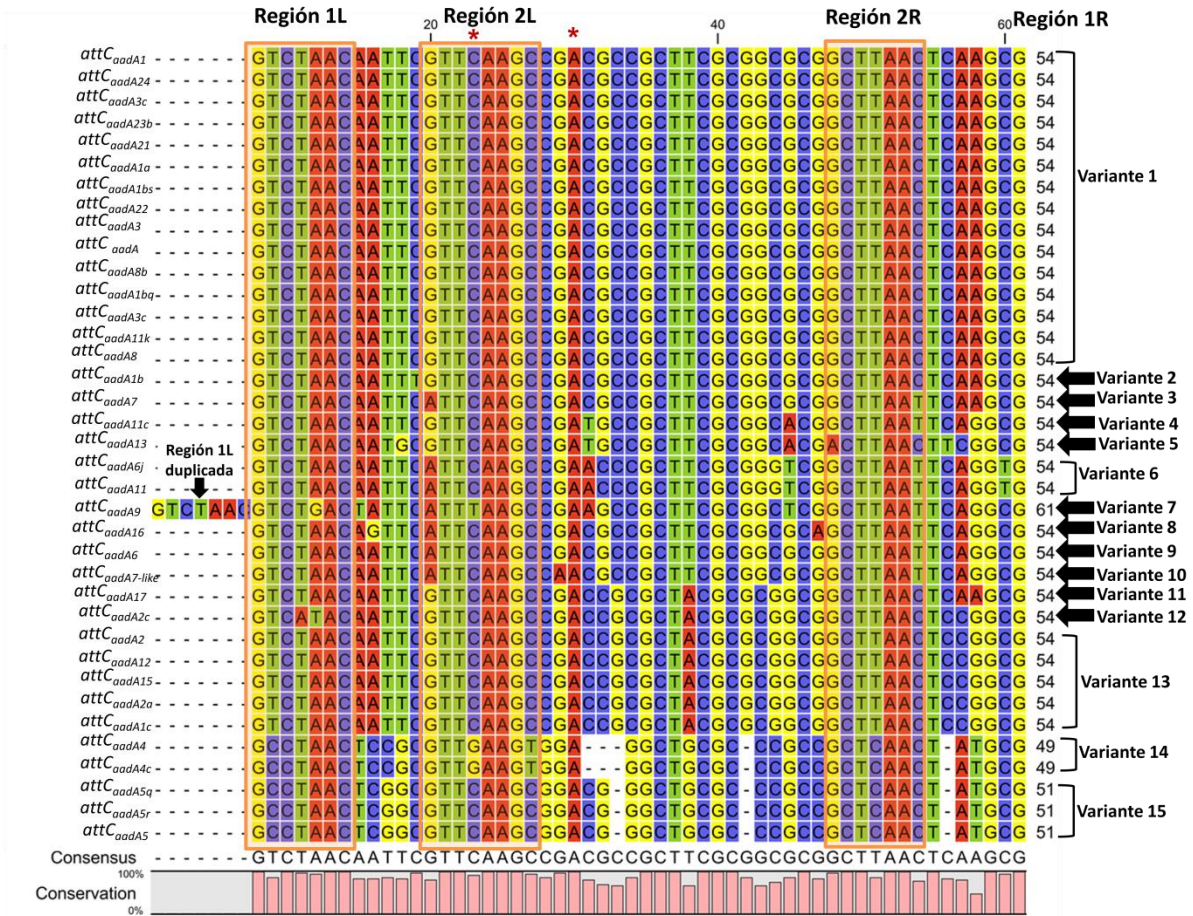
### *Resultados y discusión*

encuentran apareadas en la estructura secundaria que forma el sitio *attC*. Finalmente 2 variantes (7 y 14 asociadas a los alelos *aadA9*, *aadA4* y *aadA4c* respectivamente) perdieron la base extrahelicoidal característica de 2L, lo cual es de esperar que afecte significativamente las eficiencias de recombinación de los genes *cassettes* correspondientes.

Adicionalmente en el sitio *attC* de *aadA9* (variante 7) observamos una duplicación en tándem del sitio 1L que podría afectar el correcto apareamiento de las secuencias en la hebra inferior para permitir la formación de la estructura recombinogénica en forma de tallo y bucle. Se observaron además, 2 variantes (14 y 15) con pequeñas deleciones dentro de las secuencias espaciadoras que afectarán la formación de la burbuja entre los sitios 2L y 2R.



## Resultados y discusión

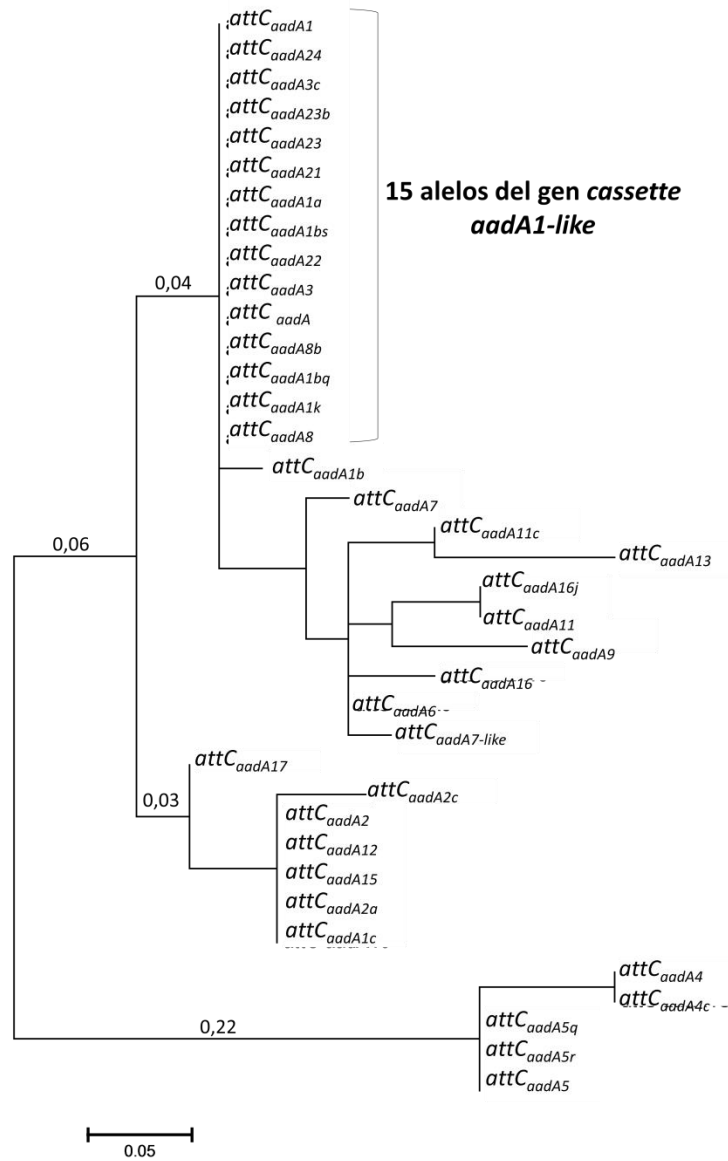


**Figura 13. Alineamiento por ClustalW para los sitios *attC* de los distintos alelos *aadA-like*.** Se consideró cada sitio *attC* desde la región 1L hasta la G de la región 1R. Las variantes de sitios *attC* se encuentran numeradas del 1 al 15. Las llaves y flechas indican las distintas variantes de los sitios *attC* que se detectaron. Los rectángulos anaranjados indican las regiones consenso asociadas a recombinación (Región 1L: RYYYAAC y 1R: GTTRRRY); se muestra hasta la G de esta última; región 2L: GTTCRARY y 2R: RYTYAAC. Los asteriscos rojos indican la localización de las bases extrahelicoidales. Adicionalmente se muestra una región 1L repetida en una de las variantes.

A partir del análisis del árbol de máxima similitud de los sitios *attC* observamos que el sitio *attC* predominante (llamado variante 1 en la Figura 13) está más cercanamente relacionado a 15 alelos de *aadA-like* (Figura 14). Con respecto al árbol de máxima similitud de los ORFs (Figura 12), la formación de grupos es similar con respecto a la observada para los sitios *attC*, donde sigue habiendo un ORF predominante denominado grupo *aadA1-like* más cercanamente relacionado a 7 alelos, pero se observa una mayor diversidad en la formación

### *Resultados y discusión*

de grupos. Como se pudo observar, en el árbol de máxima similitud de los ORF's de los alelos *aadA-like* se observa una mayor diversidad de formación de grupos que para el correspondiente de los sitios *attC* (Figura 14), lo cual muestra una estabilidad mayor en el mantenimiento de las secuencias correspondientes a los sitios de recombinación y más flexibilidad en la región codificante. En base a estos resultados inferimos que si bien algunos sitios *attCs* y ORFs de los alelos *aadA-like* se encuentran más relacionados entre sí, hay evidencias de la ocurrencia de intercambios entre ellos para formar genes *cassettes*.



**Figura 14.** Árbol de máxima similitud de las secuencias de los sitios *attC* de los genes *cassettes aadA-like*. El sitio *attC* predominante incluye a 15 alelos *aadA-like*.

#### 4.2.3- Dispersión de las variantes de los sitios *attC* de los genes *cassettes aadA-like*

Realizando un análisis de la dispersión de las 15 variantes que encontramos para los diferentes sitios *attC* de los 37 alelos *aadA-like*, usamos la herramienta BLAST de NCBI (<http://blast.ncbi.nlm.nih.gov/>) para evaluar cual o cuales de estas variantes son las más frecuentes en las secuencias contenidas en GenBank (Tabla 11). Para ello, con los resultados

## Resultados y discusión

arrojados por el programa, se realizó un filtro con el fin de saber cuáles de ellos tienen un 100% de identidad con nuestras secuencias, definidas como variantes de los sitios *attC*s de los 37 alelos *aadA-like* (Figura 10).

**Tabla 11. Número de registros correspondientes a las secuencias determinadas como variantes de sitios *attC*s de los alelos *aadA-like*.**

| <b>Sitio <i>attC</i></b> | <b>Registros</b> | <b>Genes <i>cassettes</i> asociados</b>  |
|--------------------------|------------------|--|
| Variante 1               | 553              | <i>aadA1, aadA24, aadA3c, aadA23b, aadA21, aadA1a, aadA1bs, aadA22, aadA3, aadA, aadA8b, aadA1bq, aadA3c, aadA11k, aadA8</i> |
| Variante 2               | 1                | <i>aadA1b</i>  |
| Variante 3               | 21               | <i>aadA7</i>   |
| Variante 4               | 1                | <i>aadA11c</i>   |
| Variante 5               | 4                | <i>aadA13</i>  |
| Variante 6               | 7                | <i>aadA6j, aadA11</i>  |
| Variante 7               | 3                | <i>aadA9</i>   |
| Variante 8               | 17               | <i>aadA16</i>  |
| Variante 9               | 34               | <i>aadA6</i>   |
| Variante 10              | 2                | <i>aadA7-like</i>  |
| Variante 11              | 1                | <i>aadA17</i>  |
| Variante 12              | 1                | <i>aadA2c</i>  |
| Variante 13              | 349              | <i>aadA2, aadA12, aadA15, aadA2a</i>   |
| Variante 14              | 168              | <i>aadA4, aadA4c</i>   |
| Variante 15              | 148              | <i>aadA5q, aadA5r, aadA5</i>   |

Con estos resultados podemos ver que el alelo del gen *cassette aadA1* es el más disperso en los registros contenidos en el GenBank, esto corresponde con nuestros resultados obtenido de INTEGRALL ya que el grupo que definimos como alelos *aadA1-like* contiene un total de 15 alelos (Figura 10) que comparten el mismo sitio *attC*, lo cual explicaría que puedan intercambiar de sitio *attC* entre ellos y poder generar nuevas variantes y por ende más dispersión. Las siguientes variantes con mayor número de registros

## Resultados y discusión

corresponden a otras que comparten distintos alelos como el caso del grupo de los alelos *aadA2-like* pero que no están tan dispersos como el grupo descrito anteriormente.

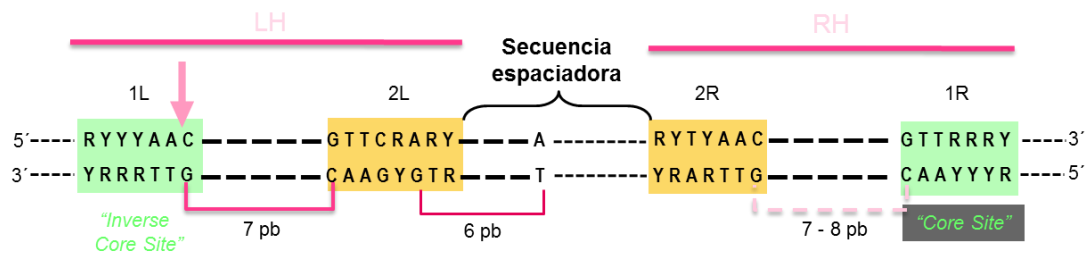
### 4.3- Estructuras secundarias de los sitios *attC* usados en los estudios de recombinación

Con el fin de analizar la relación entre la formación de estructuras secundarias de los sitios *attC* de los genes *cassettes* estudiados en este trabajo, *attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>*, *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>*, y su rol sobre las frecuencias de recombinación, analizamos por Bioinformática cada uno de dichos sitios *attC*. Para ello, a través del programa Mfold (<http://unafold.rna.albany.edu/?q=mfold>) realizamos el ensamblado de las estructuras secundarias de la hebra inferior del sitio *attC*, que constituye el sitio blanco de la integrasa IntI1 para la recombinación sitio específica.

La estructura general de un sitio *attC* típico (Figura 15) se caracteriza por tener una secuencia imperfecta repetida invertida que se encuentra *downstream* de un marco de lectura abierto (Figura 15). Por lo general dichos marcos de lectura abiertos corresponden a genes de resistencia antibiótica en los integrones de clase 1 aislados de la clínica. Los sitios *attC* constituyen una diversa familia de secuencias que funcionan como sitio de reconocimiento sitio-específico de la integrasa y varían en una longitud de 57 a 141 pb. La secuencia que tienen en común los sitios *attC* corresponde principalmente a 20 nucleótidos en sus extremos denominadas *Right Hand* (RH) y *Left Hand* (LH) por lado derecho e izquierdo, respectivamente (Figura 15). Cada una de estas regiones contiene dos cajas de nucleótidos llamadas: i) 1R o *core site* cuya secuencia consenso es 5'-GTTRRRY-3', y 2R cuya secuencia

## Resultados y discusión

consenso es 5'-RYTYAAC-3' en la región RH y 1L o *inverse core site* cuya secuencia consenso es: 5'-RYYYAAC-3' y ii) 2L cuya secuencia consenso es 5'GTTCRARY-3' en la región LH. En la caja 1R se encuentra el sitio de recombinación mediado por IntI1, ubicado entre la G y T (Figura 15). Estas cajas son parte de dos sitios antiparalelos de recombinación "core" denominados: 1L-2L y 2R-1R, respectivamente. La estructura general del sitio *attC* es palindrómica y permite que la hebra simple del *attC* adopte una forma de cadena doble por la unión de las regiones 2L a 2R y 1L a 1R.



**Figura 15. Estructura general de un sitio *attC* típico.** Se muestran las regiones 1L o *inverse core site* y 1R o *core site*, y las regiones 2L y 2R. Entre estas dos últimas regiones se encuentra la secuencia espaciadora. Tomado de: Gambino, Anahí. Tesis de licenciatura FCEyN 2016.

### 4.3.1- Descripción de la región LH en los sitios *attC* de los genes *cassettes*: *attC<sub>aadA1</sub>*,

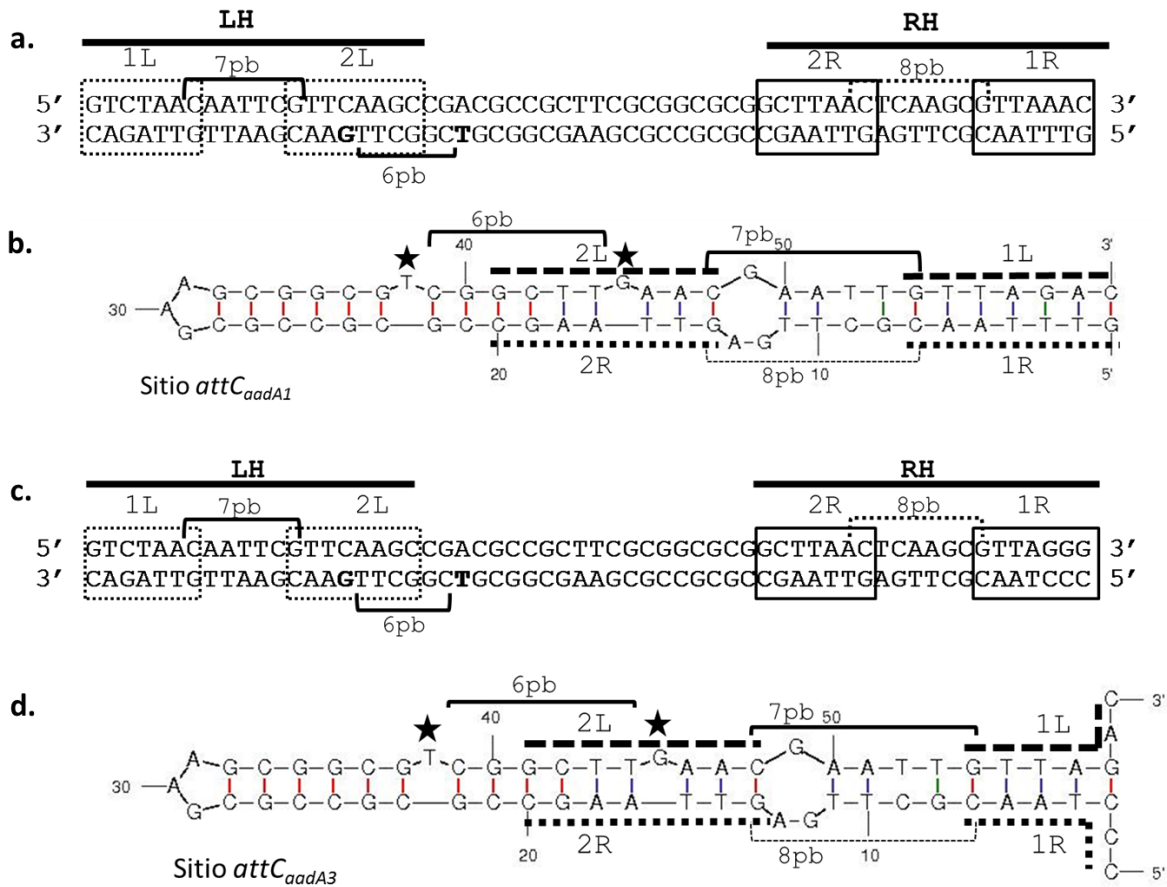
#### *attC<sub>aadA3</sub>*, *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>*

Como se mencionó anteriormente, la región LH en un sitio *attC* típico (Figura 15) comprende a su vez a dos cajas de nucleótidos, una de 7 nucleótidos denominada 1L o *inverse core site* (ICS) y una de 8 nucleótidos denominada 2L. Con respecto a la primera de ellas, la región 1L tiene una secuencia consenso 5'-RYYYAAC-3', siendo R=Purina: A y G, e Y= Pirimidina: C y T. Al hacer un análisis comparativo entre los sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>* (Figura 16) se observó que tienen un 95% de identidad en la totalidad de su secuencia nucleotídica,

## Resultados y discusión

con una diferencia en los 4 primeros nucleótidos de la misma. Con respecto a la secuencia de la región 1L de los dos sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>* la cual presenta la secuencia 5'-GTCTAAC-3' se ajusta completamente a la secuencia consenso 5'-RYYAAC-3'. En cambio, para la región 1L del sitio *attC<sub>aacA1</sub>* (Figura 17) identificamos la secuencia 5'-CCCTAAC-3' (se subraya el nucleótido que se modificó con respecto a la secuencia consenso). Dicha secuencia posee un cambio del primer nucleótido de la secuencia consenso donde se observa la transversión de la Purina: G correspondiente al primer nucleótido de la secuencia consenso por la Pirimidina: C. Para el caso de la región 1L del sitio *attC<sub>dfrA1</sub>* (Figura 18), observamos que en la secuencia de la región 1L respectiva, correspondiente a 5'-GGTTAAC-3' (se subraya el nucleótido que se modificó con respecto a la secuencia consenso), hay un cambio del segundo nucleótido de la secuencia, donde se produjo la transversión de la Pirimidina: A por la Purina: G. Este análisis indica que de las 4 regiones 1L de los 4 sitios *attC* estudiados *attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>*, *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>* los sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>* coinciden exactamente con la secuencia consenso de la región 1L, mientras que los sitios *attC<sub>aacA1</sub>* (Figura 17) y *attC<sub>dfrA1</sub>* (Figura 18) presentan la transversión de un nucleótido cada uno.

## Resultados y discusión



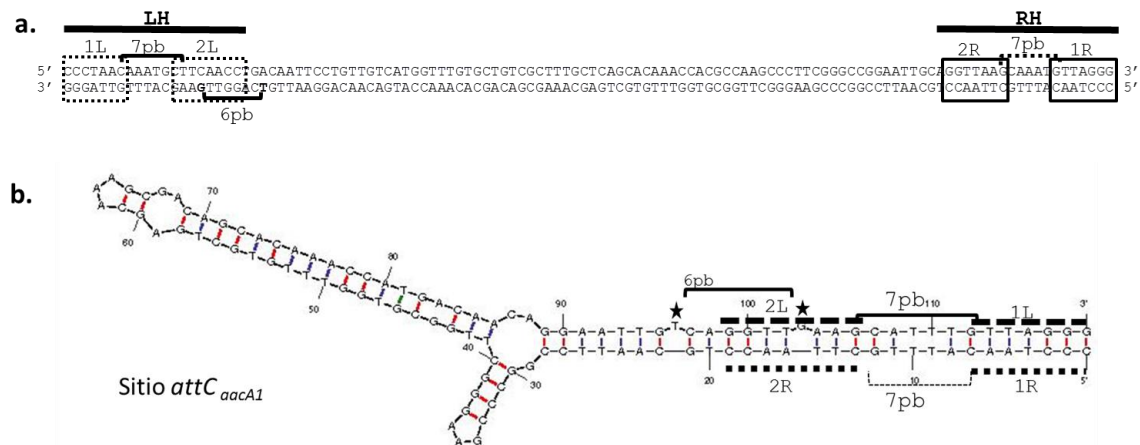
**Figura 16. Estructura secundaria de los sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>*.** a. Características del sitio *attC<sub>aadA1</sub>*. b. Estructura secundaria de la hebra inferior del sitio *attC<sub>aadA1</sub>*. c. Secuencia en doble cadena del sitio *attC<sub>aadA3</sub>*. d. Estructura secundaria de la hebra inferior del sitio *attC<sub>aadA3</sub>*. En a, que corresponde al sitio *attC<sub>aadA1</sub>*, y c, que corresponde al sitio *attC<sub>aadA3</sub>*, se muestran con líneas negras las regiones LH y RH y los rectángulos corresponden a las regiones 1L, 2L, 2R y 1R. En negrita se indican los nucleótidos que darán lugar a las bases extrahelicoidales. En b, que corresponde al sitio *attC<sub>aadA1</sub>*, y d, que corresponde al sitio *attC<sub>aadA3</sub>* se muestran en líneas punteadas las subregiones 2L y 1L, en puntos las regiones 2R y 1R, y con estrellas las bases correspondientes a bases extrahelicoidales.

Continuando con la descripción de la región LH, como se mencionó anteriormente dicha región consiste en dos cajas de nucleótidos denominadas: 1L con 7 nucleótidos y 2L con 8 nucleótidos. Con respecto a esta última, la región 2L presenta una secuencia consenso 5'-GTTCRARY-3'. La secuencia de la región 2L de los sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>* es 5'-GTTC AAGC-3' (Figura 16) y se ajusta a la secuencia consenso de la región 2L del sitio *attC* típico. La



## Resultados y discusión

región 2L del sitio *attC<sub>aacA1</sub>* (Figura 17) tiene la secuencia 5'-CTTCAACC-3' (se subraya el nucleótido que se modificó con respecto a la secuencia consenso) en la cual el primer nucleótido sufre una transversión de una Purina: G por una Pirimidina: C y los subsiguientes nucleótidos corresponden a los de la secuencia consenso de la región 2L.



**Figura 17. Estructura secundaria del sitio *attC<sub>aacA1</sub>*.** a. Secuencia en doble cadena del sitio *attC<sub>aacA1</sub>* del gen *cassette aacA1*. Se muestran con líneas negras las regiones LH y RH y los cuadrados corresponden a las regiones 1L, 2L, 2R y 1R. b. Estructura secundaria de la hebra inferior del sitio *attC<sub>aacA1</sub>*. Se muestran en líneas punteadas las regiones 2L y 1L y en puntos las regiones 2R y 1R, y con estrellas las bases correspondientes a bases extrahelicoidales.

La región 2L del sitio *attC<sub>dfrA1</sub>* (Figura 21) tiene la secuencia 5'-GCAGCAAC-3' (se subrayan los nucleótidos que se modificaron con respecto a la secuencia consenso) en la cual cambian del segundo al quinto nucleótido con respecto a la secuencia consenso de la región 2L. Estas modificaciones consisten en una sustitución de una Timina, T, por una Citosina, C, en el segundo nucleótido, una transversión de una Timina, T por una Purina, A, en el tercer nucleótido, una transversión de una Citosina, C a una Purina G, en el cuarto nucleótido y en el quinto nucleótido dónde debería haber alguna de las dos Purinas, hay una Pirimidina, C.

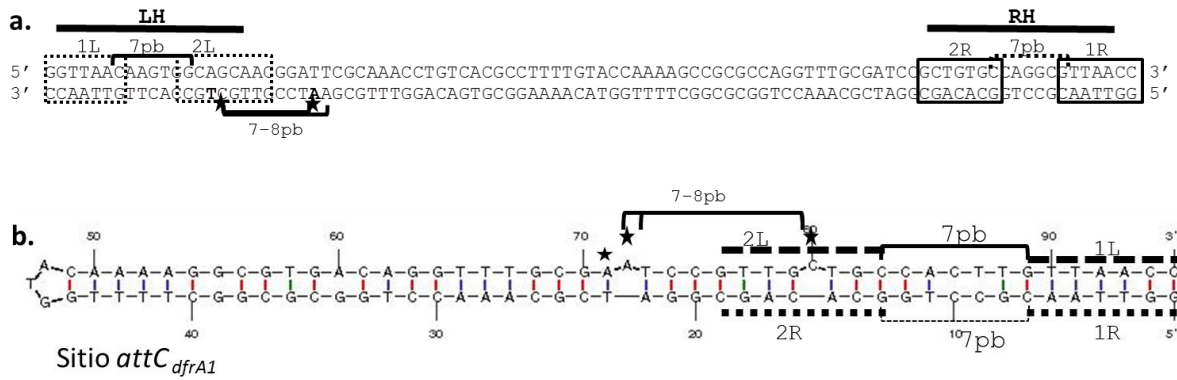
## Resultados y discusión

### 4.3.2- Descripción de la región RH en los sitios *attC* de los genes *cassettes*: *attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>*, *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>*

La región RH al igual que la región LH está compuesta por dos subregiones, éstas consisten en 2 cajas de heptanucléotidos denominadas 1R y 2R. La primera de ellas se denomina región 1R o *core site* (CS) cuya secuencia consenso es 5'-GTTRRRY-3'. La secuencia del sitio *attC<sub>aadA1</sub>* cuya secuencia de la región 1R o es 5'-GTAAAC-3'-se ajusta a la secuencia consenso. La región 1R o *core site* del sitio *attC<sub>aadA3</sub>* con la secuencia 5'-GTTAGGG-3' tiene una transversión de la Purina, G con respecto a la Pirimidina: C de la secuencia consenso en el último nucleótido. La secuencia de la región 1R o *core site* del sitio *attC<sub>aacA1</sub>* es igual a la secuencia del *core site* del sitio *attC<sub>aadA3</sub>* donde hay una transversión de la Purina, G a la Pirimidina, C de la secuencia consenso en el último nucleótido.

La región 1R del sitio *attC<sub>dfrA1</sub>* tiene la secuencia 5'-GTTAACC-3' (se subraya el nucleótido que se modificó con respecto a la secuencia consenso) en la cual, el penúltimo nucleótido donde la secuencia consenso dice que debería haber alguna de las dos Purinas: A o G hay una transversión a la Pirimidina: C.

## Resultados y discusión



**Figura 18. Estructura secundaria del sitio *attC*<sub>dfrA1</sub>.** **a. Secuencia en doble cadena del sitio *attC*<sub>dfrA1</sub>.** Se muestran con líneas negras las regiones LH y RH y los cuadrados corresponden a las regiones 1L, 2L, 2R y 1R. **b. Estructura secundaria de la hebra inferior del sitio *attC* del gen *cassette dfrA1*.** Se muestran en líneas punteadas las subregiones 2L y 1L, en puntos las regiones 2R y 1R, y con estrellas las bases correspondientes a bases extrahelicoidales.

Como se mencionó anteriormente, la región RH del sitio *attC* típico consiste en 2 cajas de heptanucleótidos denominadas 1R y 2R. Con respecto a esta última, la región 2R tiene la secuencia consenso 5'-RYTYAAC-3'. Al analizar la secuencia de la región 2R de los sitios *attC*<sub>aadA1</sub>, *attC*<sub>aadA3</sub> y *attC*<sub>aacA1</sub>, vemos que tienen la secuencia 5'-GCTTAAC-3' lo que les da una identidad de secuencia entre ellos del 100% y siguen la secuencia consenso de la región 2R. Sin embargo, la secuencia de la región 2R del sitio *attC*<sub>dfrA1</sub> tiene la secuencia 5'-GCTGTGC-3' (se subrayan los nucleótidos que se modificaron con respecto a la secuencia consenso) y presenta modificaciones con respecto a la secuencia consenso a partir del cuarto nucleótido, con una transversión de la Pirimidina correspondiente por la Purina: G, así como en el quinto nucleótido de la Purina: G por una Pirimidina: T y una sustitución de la Purina: A por la Purina: G, respectivamente. El último nucleótido mantiene la identidad del nucleótido final de secuencia consenso de la región 2R.

## Resultados y discusión

Al observar las características de las secuencias de las regiones 2L y 2R se confirma que es más importante la similitud de secuencia entre ellas que con respecto a las secuencias consenso (Mazel y col., 2009). A pesar de haber casos puntuales con mutaciones, al observar la estructura secundaria de los sitios *attC* analizados, se mantiene la complementariedad a lo largo de toda la región 2R y 2L (Figuras 19 a 23), con excepción de una posición desapareada pero manteniendo la base extrahelicoidal característica de la región 2L.

### 4.3.3- Secuencias espaciadoras de las regiones LH y RH

Al momento de formarse la estructura secundaria de la hebra inferior de un sitio *attC* típico, se produce el alineamiento de las regiones 1L-1R y 2R-2L, lo cual da como resultado la formación de dos regiones espaciadoras no complementarias, en conjunto denominadas como secuencias espaciadoras, espaciador central no apareado o UCS por sus siglas en inglés (*Unpaired Central Spacer*) (Figura 15). En los sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>* las secuencias espaciadoras entre las regiones 1L y 2L corresponden a 7 nucleótidos en ambos sitios, mientras que entre las regiones 1R y 2R es de 8 nucleótidos. Al analizar la estructura secundaria de estos mismos sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>* (Figura 16) se observa la formación de una pequeña burbuja contigua a las regiones 2L y 2R que involucra a los nucleótidos 12, 13 y 49 (la numeración es de la hebra inferior en dirección 5' a 3', Figura 16). Además, se observa que una Citosina y una Timina (nucleótidos 9 y 52) de estas secuencias permanecen desapareadas. La distancia de estos nucleótidos no apareados hacia la unión de las regiones 1L-1R es de 2 nucleótidos y hacia la unión de las regiones 2L-2R es de 4/5 nucleótidos. Dicho conteo se realiza teniendo en cuenta a la base nucleotídica que se incluye en el conteo de la

## Resultados y discusión

región espaciadora y que pertenece también a los sitios 1L, 2L, 1R o 2R, según el caso (Figura 16).

En el sitio *attC<sub>aacA1</sub>* (Figura 17) la secuencia espaciadora entre las regiones 1L-1R y 2R-2L contiene 7 nucleótidos y no se observa formación de una burbuja en el plegamiento de la estructura secundaria. Con respecto a la secuencia de la región espaciadora no se aparean los nucleótidos 9 y 10 correspondientes a dos Timinas que no se aparean con los nucleótidos 109 y 110 que también corresponden a dos Timinas (la numeración es de la hebra inferior en dirección 5' a 3', Figura 17). Estos nucleótidos desapareados están a una distancia de 2 nucleótidos de la región 1L-1R y a una distancia de 3 nucleótidos de la región 2L y 2R.

La secuencia de la región espaciadora entre las regiones 1L-1R y 2R-2L del sitio *attC<sub>dfrA1</sub>* (Figura 18) tiene 7 nucleótidos y no forma burbuja en esta región del sitio *attC*. Sin embargo, en las secuencias de la región espaciadora no se aparean los nucleótidos 9 y 10, correspondientes a Pirimidinas específicamente Citosinas, con los nucleótidos 86 y 87 Citosina y Timina, respectivamente (la numeración es de la hebra inferior en dirección 5' a 3', Figura 17). Estas bases no apareadas están a una distancia de la región de unión de 1L-1R de 2 nucleótidos y la región de unión 2L-2R de 3 nucleótidos.

### 4.3.4- Bases extrahelicoidales de los sitios *attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>*, *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>*

El reconocimiento por parte de la integrasa para realizar el proceso de recombinación se caracteriza por la unión sitio-específica a la hebra inferior de un sitio *attC*, y la base extrahelicoidal resultante a partir del plegamiento de dicha hebra lo cual es importante para

## *Resultados y discusión*

el reconocimiento de distintas integrasas (Johansson y col., 2004; Mc Donald y col., 2006). Algunas integrasas parecen tener una preferencia por ciertos sitios *attC* con respecto a otros, lo que puede estar relacionado con el reconocimiento de las bases extrahelicoidales (Larouche y Roy, 2011).

Como se venía describiendo, dentro de las características estructurales de importancia para la recombinación está la aparición de bases extrahelicoidales que consisten en bases simples que se ubican en la zona de asociación de 2R y 2L. De estas bases extrahelicoidales, la G (Purina, Guanina) presente en la región 2L y la T (Pirimidina, Timina) presente 6 a 8 nucleótidos hacia la parte distal del tallo, son particularmente conservadas (Bouvier y col., 2009).

La base extrahelicoidal proximal, presente en la región 2L, es responsable por la elección de la hebra, mientras que la otra base extrahelicoidal no es esencial para la elección de la hebra, pero sí para la eficiencia de recombinación.

Para los sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>* (Figuras 19) las bases extrahelicoidales corresponden a una G (Purina, Guanina) en la posición 45 y una T (Pirimidina, Timina) en la posición 38 con una distancia de 6 pares de bases entre ellas (la numeración es de la hebra inferior en dirección 5' a 3').

En el sitio *attC<sub>aacA1</sub>* (Figura 17) las bases extrahelicoidales son una G (Purina, Guanina) en la posición 103 y una T (Pirimidina, Timina) en la posición 96 con una distancia de 6 pares

## Resultados y discusión

de bases entre ellas. La secuencia restante de este sitio *attC* está dividida en dos ramas, una de 11 nucleótidos de longitud y otra de 46 nucleótidos de longitud (Figura 17).

Finalmente, en el sitio *attC<sub>dfrA1</sub>* (Figura 18) las bases extrahelicoidales son una C (Pirimidina, Citosina) en la posición 80 y una A (Purina, Adenina) en la posición 71 o 72, con una distancia de 7 u 8 pares de bases entre ellas.

### 4.3.5- Influencia sobre el proceso de recombinación de los diferentes sitios *attC* analizados:

#### ***attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>*, *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>***

En los sitios *attC* analizados *attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>*, *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>*, con respecto a las secuencias consenso de la región 1R o *core site* encontramos que el sitio *attC<sub>aadA1</sub>* es el único que se ajusta completamente a la secuencia consenso 5'-GTTRRY-3'. Los tres sitios *attC* restantes (*attC<sub>aadA3</sub>*, *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>*) se caracterizan por tener una transversión de una Purina a Pirimidina en uno de sus nucleótidos, pero no en aquellos nucleótidos donde la integrasa realiza el corte en el proceso de recombinación 5'-G'TT-3' (El punto de corte es entre la G y la T).

Con respecto a las secuencias consenso de la región 1L, los sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>* tienen un 100% de identidad de secuencia entre sus regiones 1L 5'-GTTAAC-3' y coinciden exactamente con la secuencia consenso de la región 1L, mientras que los correspondientes a los sitios *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>* presentan la transversión de un nucleótido cada uno.

## Resultados y discusión

Al revisar las bases extrahelicoidales de todos los sitios *attC* estudiados: *attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>*, *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>* encontramos que la mayoría de ellas corresponden a Guanina y Timina a excepción de lo observado en el sitio *attC<sub>dfrA1</sub>* (Figura 18), que tiene como bases extrahelicoidales una Citosina y una Adenina.

La diferencia más evidente y significativa con respecto a las bases extrahelicoidales dentro de los sitios *attC* analizados es el espaciado entre las mismas, y en aquellos sitios *attC* que comparten las mismas bases extrahelicoidales y adicionalmente presentan la misma distancia de nucleótidos entre ellas: *attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>*, *attC<sub>aacA1</sub>*, las diferencias más significativas corresponden a la longitud del sitio *attC*, lo que al parecer no tendría influencia sobre las frecuencias de recombinación, al menos para los sitios *attC* que analizamos en este trabajo.

Como ha sido descrito por otros autores (Bouvier y col., 2009; Larouche y Roy, 2011; Loot y col., 2014), los sitios *attC* independientemente de su secuencia, presentan determinantes de reconocimiento estructurales que deben existir dentro los mismos con el fin de facilitar la recombinación. El primero de estos determinantes de reconocimiento estructural y cómo describimos anteriormente, se refiere al alineamiento entre las regiones 1L-1R y 2R-2L lo que da como resultado la formación de la región espaciadora no complementaria. El segundo y fundamental determinante de reconocimiento estructural de importancia para la recombinación es la aparición de bases extrahelicoidales, las cuales son bases simples que se ubican en la zona de asociación de las regiones 2R-2L. De estas bases



## Resultados y discusión

extrahelicoidales, la G (Purina: Guanina) presente en la región 2L y la T (Pirimidina: Timina) presente 6 a 8 nucleótidos hacia la parte distal del tallo son particularmente conservadas (Bouvier y col., 2009).

Al observar los determinantes de reconocimiento estructural de importancia para la recombinación, podemos decir que los sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>* (Figura 16) mantienen una identidad en cuanto a sus bases extrahelicoidales G y T, así como la distancia entre ellas (6 pares de bases), lo cual podría estar contribuyendo a aumentar las frecuencias de escisión específicamente. Con respecto al sitio *attC<sub>aacA1</sub>* (Figura 17) podemos ver que las bases extrahelicoidales conservan la misma identidad G y T y el mismo espaciado entre ellas (6 pares de bases) y presentan valores muy similares de escisión cuando este sitio *attC* está en la estructura *attC<sub>aacA1</sub>-aadA1attC<sub>aadA1</sub>*, resultado de escisión que se chequeó por secuenciación al obtener un número tan alto y cercano a los valores del ensayo de recombinación realizado a partir del clon con la estructura *attC<sub>aadA3</sub>-aadA1attC<sub>aadA1</sub>* con un porcentaje de escisión del 98% (Ver el ítem 4.3.1: Evaluación de la capacidad de Int1 para escindir a los genes *cassettes attI1-aadA1-attC<sub>aadA1</sub> attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>, attC<sub>aacA1</sub>-aadA1- attC<sub>aadA1</sub> y attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>* del capítulo Resultados y Discusión).

Todos los sitios *attC* que analizamos presentan bases extrahelicoidales, por lo que nos resulta más relevante este análisis que el de la función de la región espaciadora. La influencia de la región espaciadora de los sitios *attC* es una de las características marcadas que influyen sobre el proceso de recombinación principalmente por su orientación y la burbuja que se

## *Resultados y discusión*

forma por los nucleótidos desapareados de la misma (Bouvier y col., 2009). El no apareamiento y la asimetría que se genera a partir de la región espaciadora son esenciales para que se ensamblen correctamente y para asegurar el éxito de los siguientes pasos de recombinación de simple cadena (Bouvier y col., 2009). Cuando las bases extrahelicoidales están ausentes, los sitios *attC* se diferencian en la formación o no de burbujas durante la formación de la estructura secundaria. Para nuestros clones, ensayos de recombinación y teniendo en cuenta estos resultados, al parecer la identidad y características de la región espaciadora no estarían involucradas en las frecuencias de recombinación observadas.

Con respecto a la región espaciadora, se ha descrito que esta burbuja que se forma a partir del plegamiento de la hebra en su forma de estructura secundaria, es esencial para la recombinación en ausencia de bases extrahelicoidales. En un ensayo experimental particular del trabajo de Bouvier y col., 2009, decidieron remover las bases extrahelicoidales y mutaron la región espaciadora de diferentes sitios *attC* con la finalidad de lograr que esta estuviera totalmente apareada, determinando que el proceso de recombinación fue drásticamente afectado, disminuyendo su frecuencia.

Algo que es importante destacar es que dos sitios *attC*: *attC<sub>aadA3</sub>*: 60 pb y *attC<sub>aacA1</sub>*: 118 pb muy diferentes en longitud y en su región espaciadora (Figura 16 y 20) tengan frecuencias de recombinación muy parecidas, lo que nos indica que la importancia de la identidad de las bases extrahelicoidales y el espaciado entre las mismas, es la característica fundamental que influye sobre la frecuencia de recombinación y específicamente la escisión

## Resultados y discusión

de un *gen cassette* en particular para nuestro estudio. Esto coincide con lo reportado por Laurouche y Roy, 2011 quienes en variados ensayos de recombinación realizados para el sitio *attC<sub>dfrA1</sub>* (Figura 18) cuyas bases extrahelicoidales son C y A con una distancia entre ellas de 7 a 8 pb comprobaron que este sitio con estas bases extrahelicoidales y un espaciado mayor a 6pb entre ellas tenían un marcado descenso en las frecuencias de escisión, tal como lo pudimos ver en nuestros resultados experimentales.

### 4.4- Estudio funcional del *gen cassette aadA1* mediante ensayos de recombinación

#### 4.4.1- Evaluación de la capacidad de *IntI1* para escindir a los genes *cassettes attI1-aadA1-*

*attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>*, *attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>* y *attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>*

Para este apartado, realizamos los ensayos *in vivo* de escisión de genes *cassettes*, mediados por *IntI1*, utilizando en cada caso al clon PMI1-1 (AS)2 (*intI1-attI1*) que contiene al gen de la integrasa de tipo 1 cotransformando en *E. coli* TOP10 con cada uno de los siguientes clones: pLQ423 (*attI1-aadA1-attC<sub>aadA1</sub>*), pLQ443 (*attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>*), pLQ445 (*attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>*) o pLQ426 (*attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>*), respectivamente (Tabla 12), siguiendo la metodología descrita en el capítulo de materiales y métodos (Apartado 3.2.9: Ensayos de escisión de genes *cassettes in vivo* mediados por *IntI1*), en experimentos independientes. La estructura de los genes *cassettes* clonados en cada uno de estos plásmidos corresponde a estructuras típicas observadas en diferentes arreglos de integrones de la clínica. En cuanto a la importancia del *gen cassette aadA1* como ya se ha descrito previamente, es uno de los genes *cassettes* de resistencia antibiótica que se ha detectado en los aislamientos con integrons de clase 1 más antiguos disponibles hasta el

## Resultados y discusión

momento y adicionalmente está ampliamente diseminado en integrones clínicos, otorgando resistencia a estreptomycin y espectinomycin.

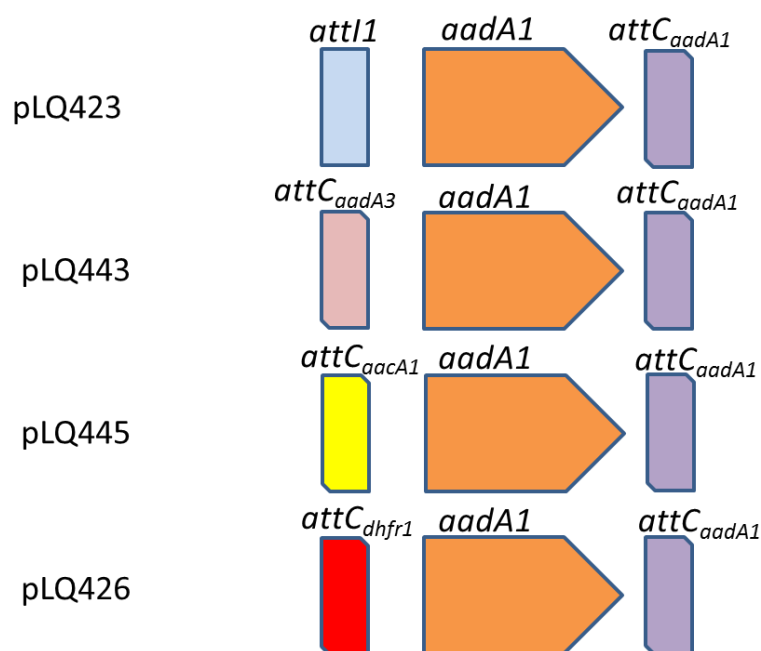
**Tabla 12.** Plásmidos con los genes *cassettes* usados en este estudio para realizar los ensayos de recombinación.

| Plásmidos   | Descripción  | Fuente              |
|-------------|--|---------------------|
| pLQ423      | <i>attI1-aadA1-attC<sub>aadA1</sub></i> de Tn21 clonado en el vector pACYC184, Cl <sup>R</sup>                       | P. Roy y N. Messier |
| pLQ443      | <i>attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub></i> clonado en el vector pACYC184, Cl <sup>R</sup>                | P. Roy y N. Messier |
| pLQ445      | <i>attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub></i> clonado en el vector pACYC184, Cl <sup>R</sup>                | P. Roy y N. Messier |
| pLQ426      | <i>attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub></i> clonado en el vector pACYC184, Cl <sup>R</sup>                | P. Roy y N. Messier |
| PM11-1(AS)2 | <i>intl1-attI1</i> (pLQ3poly:: <i>attI1</i> – <i>attI1</i> S al M13F) clonado en el vector pMAL-c2, Amp <sup>R</sup> | Quiroga MP.         |

Cl<sup>R</sup>: resistencia a cloranfenicol. Amp<sup>R</sup>: resistencia a ampicilina.

Los plásmidos portadores de las estructuras *attI1-aadA1-attC<sub>aadA1</sub>* con el sitio *attI1* y *attC<sub>genecassette-aadA1-attC<sub>aadA1</sub></sub>* cada una con un sitio *attC* diferente *upstream* del gen *cassette aadA1* usados en este estudio (Tabla 12) fueron construidos a partir de cepas con diferente arreglo de genes *cassettes* en la zona variable de integrones y sus estructuras se detallan en la siguiente figura (Figura 21).

## Resultados y discusión



**Figura 21. Estructura de los genes *cassettes* usados para los ensayos de recombinación.** Las flechas naranjas representan al gen *aadA1*, el rectángulo celeste al sitio *attI1*, el rectángulo rosa al *attC<sub>aadA3</sub>*, el rectángulo amarillo al *attC<sub>aacA1</sub>* y el rectángulo rojo al *attC<sub>dhfr1</sub>*. Los rectángulos a la derecha del gen *aadA1* representan al sitio *attC<sub>aadA1</sub>*.

Para evaluar los eventos de escisión del gen *cassette aadA1* se realizaron PCRs con los cebadores PACYC184 5' y PACYC1843'N (Tabla 2 del capítulo 3 de materiales y métodos), a partir del ADN plasmídico extraído luego de la incubación con IPTG en los ensayos de recombinación realizados con los clones mencionados y también a partir de las colonias obtenidas luego de la transformación de dicho ADN en *E. coli* TOP10 para cuantificar los eventos de escisión de cada gen *cassette* (Tabla 13). Adicionalmente, se muestra el esquema general de como ocurre la escisión del gen *cassette* en el plásmido portador del mismo (Figura 22) y los geles de agarosa (Figura 23) representativos de cada una de las PCRs de escisión para cada uno de los plásmidos empleados en este estudio.

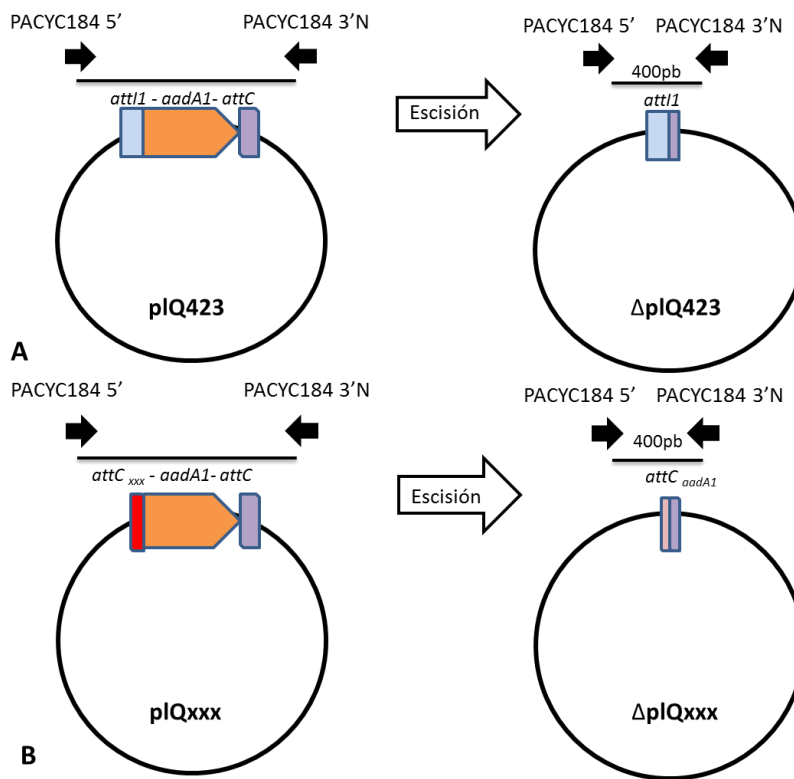
Resultados y discusión

**Tabla 13.** Frecuencias de escisión estimadas para los genes *cassettes* estudiados en distintos ensayos de recombinación.

| Plásmidos | Estructuras analizadas                                 | Escisión (%) |
|-----------|--|--------------|
| pLQ423    | <i>attI1-aadA1-attC<sub>aadA1</sub></i>                | 25,5         |
| pLQ443    | <i>attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub></i> | 98,8         |
| pLQ445    | <i>attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub></i> | 100          |
| pLQ426    | <i>attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub></i> | >3,33%       |

El porcentaje de escisión corresponde al promedio de al menos tres ensayos independientes.

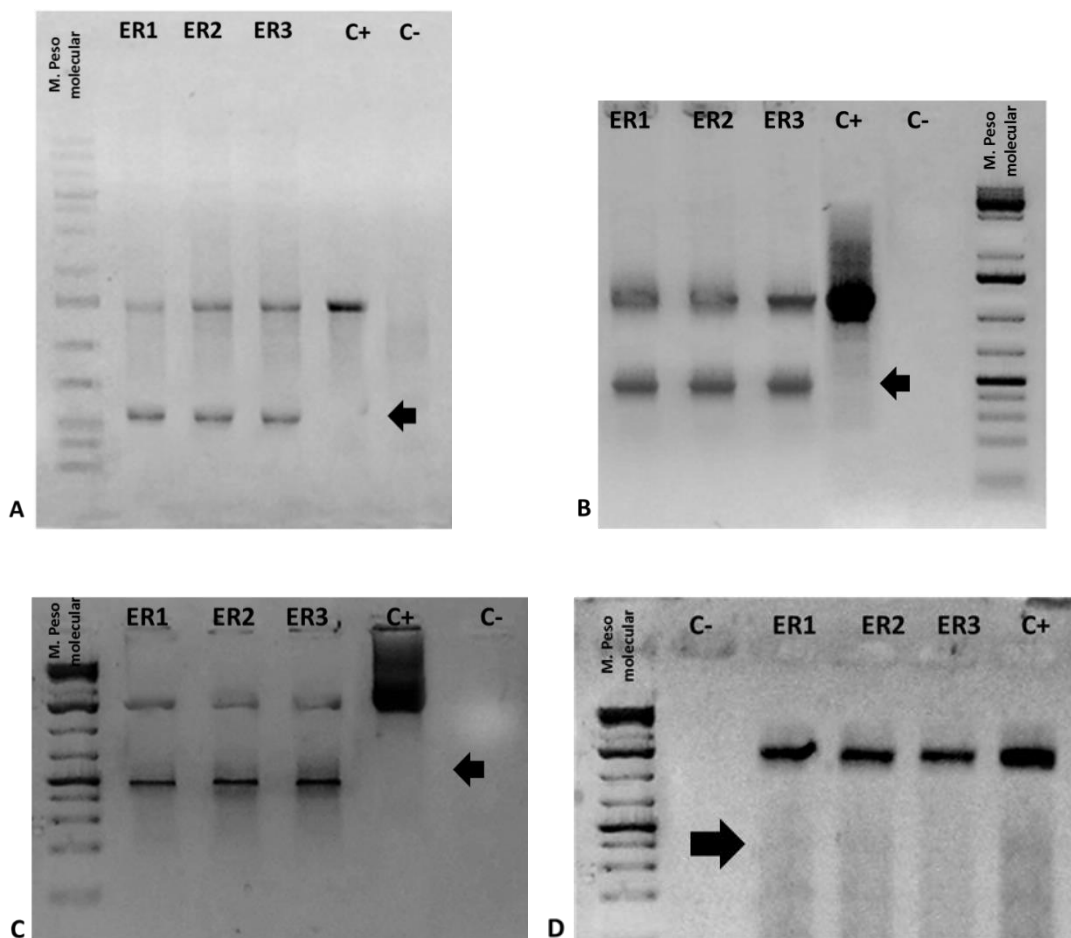
**Esquema general de la escisión del gen *cassette aadA1* en los distintos clones empleados en este estudio**



**Figura 22.** Esquema general del proceso de escisión del gen *cassette aadA1* para los diferentes arreglos construidos en los plásmidos usados en los ensayos de recombinación. A la izquierda se ve al plásmido que contiene al gen *cassette*, y a derecha el plásmido luego de la escisión del gen *cassette*. El rectángulo celeste representa al sitio *attI1*, el rojo a los distintos sitios *attC* y el violeta al *attC<sub>aadA1</sub>*, la flecha naranja representa a *aadA1*, las negras a los cebadores, la línea negra corta representa el amplicón correspondiente a amplificaciones positivas (de aproximadamente 400 pb) y la línea larga al amplicón negativo para la escisión (1200 pb). **A.** Corresponde al clon que contiene al sitio *attI1* *upstream* del gen *cassette aadA1* **B.** Corresponde a los clones que contienen diferentes sitios *attC* *upstream* del gen *cassette aadA1*.

## Resultados y discusión

### Evaluación por PCR de la escisión del gen *cassette aadA1* en los distintos clones empleados en este estudio



**Figura 23. Geles de agarosa donde se muestra el resultado de la evaluación por PCR de las escisiones para los distintos ensayos de recombinación.** La evaluación de la escisión del gen *cassette aadA1* se realizó a través de una PCR con los cebadores PACYC184 5'-PACYC184 3'N. Se muestran los resultados para cada uno de los clones usados en los ensayos de recombinación presentados en este estudio, obtenido luego de la incubación con IPTG. **A.** PCR de escisión para el clon pLQ423 (*attI1-aadA1-attC<sub>aadA1</sub>*) **B.** PCR de escisión para el clon pLQ443 (*attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>*) **C.** PCR de escisión para el clon pLQ445 (*attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>*) **D.** PCR de escisión para el clon pLQ426 (*attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>*). Se muestra con una flecha de color negro la fila donde se encuentra el amplicón producto de la escisión, el cual es de menor intensidad en el último en ensayo (Figura 23, D).

## Resultados y discusión

Como se mencionó anteriormente, la integrasa y otras proteínas de la familia de las tirosina recombinasas realizan intercambio de ADN con el elemento no palindrómico del integrón o sitio *attI* y el elemento de recombinación palindrómico sitio *attC*. La integración o inserción de genes *cassettes* ocurre preferencialmente entre un sitio *attC* de un gen *cassette* circularizado con el sitio *attI1* de un integrón, en el caso de la escisión de un gen *cassette*, la reacción más común ocurre entre dos sitios *attC*. La integrasa de clase 1 se caracteriza por el reconocimiento de una estructura de sitio *attC* plegada de cadena simple, no de cadena doble (Francia y col., 1999). Al observar los resultados, se puede identificar que la reacción de escisión entre el plásmido que contiene a la integrasa y el plásmido que contiene la estructura *attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>* (pLQ426) (Tabla 13) da como resultado el valor más bajo dentro de los porcentajes de escisión para las estructuras analizadas. Este resultado puede deberse a que el sitio *attC<sub>dhfr1</sub>* tiene una particularidad de secuencia y estructural de la cual se describió en detalle en el capítulo de análisis de las estructuras secundarias de los sitios *attC* de los clones usados en este trabajo.

A partir de la comparación de los porcentajes de escisión correspondientes a las 4 estructuras analizadas, observamos que los genes *cassettes attI1-aadA1-attC<sub>aadA1</sub>* (25%), *attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>* (98%) y *attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>* (100%) son blancos favorables para la escisión mediada por IntI1, a diferencia del gen *cassette attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>* donde el valor de frecuencia de escisión fue el más bajo de todos los ensayos (>3,33%).



**El gen *cassette attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>* y el *cassette attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>* son escindidos más eficientemente que los demás genes *cassettes* analizados, con una frecuencia de escisión cercana al 100%.**

Como puede observarse en la Tabla 13, para los ensayos realizados con las 3 primeras estructuras, obtuvimos los resultados positivos de escisión más elevados, lo cual puede deberse a que estructuralmente hablando y por las características que tienen las secuencias de los diferentes sitios *attC upstream* y *downstream* del gen *cassette aadA1*, son blancos favorables para ser reconocidos por la integrasa. Particularmente, como ya hemos descrito a lo largo de este trabajo, la integrasa a diferencia de otras proteínas de la familia de las recombinasas puede intercambiar ADN usando dos sitios con dos características intrínsecas a cada uno de ellos: el sitio *attI* no palindrómico y el sitio *attC* palindrómico. El primero de ellos, durante la recombinación se encuentra en forma de doble cadena y suele ser más conservado con respecto a la integrasa que lo reconoce y para el caso del sitio *attC*, su cadena inferior se pliega en forma de estructura secundaria exponiendo un par de bases nucleotídicas de gran importancia para el reconocimiento de la integrasa: las bases extrahelicoidales. Para este estudio calculamos que las frecuencias de escisión son más altas cuando la recombinación implica la participación de dos sitios *attC*, lo cual coincide con lo reportado por varios autores como la forma más exitosa en la que ocurre la escisión de los diversos genes *cassettes* contenidos en integrones (Collis y Hall, 1992; Collis y col., 1993).

## Resultados y discusión

Como ya se describió en el apartado donde analizamos las estructuras secundarias de los diferentes sitios *attC* que fueron estudiados, describimos que a pesar de ser estructuras distintas, de diferente longitud y secuencia, guardan características en particular que son claves para la recombinación, principalmente la presencia e identidad de bases extrahelicoidales durante el plegamiento de la hebra inferior del sitio *attC* en particular y más específicamente, la distancia entre ellas, que juega un papel fundamental en el reconocimiento de la integrasa.

Se ha descrito en otros estudios, que generalmente distintas integrasas durante el proceso de recombinación tienen preferencia por diferentes sitios *attC* con respecto a otros, principalmente debida al reconocimiento de dichas bases extrahelicoidales (Johansson y col., 2004; Larouche y Roy, 2011). Para el caso de nuestro estudio en particular, en los ensayos realizados con el clon pLQ426, correspondiente a la estructura *attC<sub>dhfr1</sub>-aadA1- attC<sub>aadA1</sub>*, el valor de la frecuencia de escisión del gen *cassette aadA1* fue menor al 3%. Nuestros resultados concuerdan con lo reportado por Roy y Larouche (2011) quienes evaluaron la actividad de escisión del gen *cassette aadA1* y otros genes de resistencia antibiótica flanqueados por distintos sitios *attI* y *attC*, por medio de la actividad recombinasa de IntI1 y otras integrasas (IntI1, IntI2\*179E, IntI3, SonIntIA y CchIntIA), demostrando que hay un efecto pronunciado en el valor de las frecuencias de escisión debido a la identidad y el espaciado de las bases extrahelicoidales de los sitios *attC* implicados en la recombinación, siendo las más comunes y reconocidas una Timina (T) y una guanina (G) separadas por una distancia de 6 nucleótidos. Sin embargo, se demostró que IntI1 continúa escindiendo

## *Resultados y discusión*

también estructuras con una T y C como bases extrahelicoidales, siempre y cuando la distancia de 6 nucleótidos se conserve. Esto demuestra que las identidades de las bases extrahelicoidales no afectan el reconocimiento de las integrasas pero si influyen la reacción de recombinación. En este estudio en particular de 2011, los autores demostraron cualitativamente que los sitios *attI1*, *attC<sub>aadA3</sub>* y *attC<sub>aacA1</sub>* localizados *upstream* del gen *cassette aadA1* son blancos favorables para que la integrasa realice escisión, específicamente para el caso del *attI1 upstream* del gen *aadA1* en una medida que ellos denominan moderada (20 a 75%) y para el caso de los sitios *attC<sub>aadA3</sub>* y de *attC<sub>aacA1</sub>* en una medida que ellos denominan fuerte (>75%). Sin embargo, la metodología que usaron para determinar estos valores es cualitativa pero concuerda con los resultados que obtuvimos para las frecuencias de recombinación, específicamente las frecuencias de escisión de las distintas estructuras en las que se encuentra el gen *cassette aadA1*. Para el caso del sitio *attC* correspondiente al gen *cassette attC<sub>dhfr1</sub>* en la estructura *attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>*, se demuestra que no es un blanco óptimo para que la integrasa realice escisión, disminuyendo la frecuencia de recombinación del gen *cassette* que se encuentra *downstream* del mismo en la zona variable de un integrón. Este tipo de resultados nos podrían acercar a información sobre los genes *cassettes* que tienden a estar en ciertas posiciones dentro de distintos integrones y los factores que dentro de sus características moleculares influyen sobre su dispersión.

En este mismo trabajo (Roy y Larouche, 2011) se describió que el sitio *attC<sub>dfrA1</sub>* o *attC<sub>dhfrA1</sub>* es un sitio de 95pb, que cuando se encuentra ubicado *upstream* de genes *cassettes*

## Resultados y discusión

como *aadA1* y de *sat2* al realizar ensayos de recombinación, presentan una escisión baja (Larouche y Roy, 2011). Dentro de los mismos experimentos se comprobó que otras estructuras que involucraban al sitio *attC<sub>aacA12-orf</sub>* (117 pb) tienen una alta frecuencia de escisión (Larouche y Roy, 2011).

Es importante destacar que a través de estos resultados se demostró no solo el profundo efecto de la identidad y del espaciado de las bases extrahelicoidales en los sitios *attC* y como esto está directamente relacionado con la escisión del gen *cassette*. Todas las integrasas empleadas en este estudio (Int1, Int2\*179E, Int3, SonInt1A y VchInt1A) escindieron todos los genes *cassettes* que estaban flanqueados por sitios *attC* cuyas bases extrahelicoidales eran una G y una T y separadas por 6 nucleótidos, a excepción VchInt1A para la cual no se detectó que escinda a la estructura *attC<sub>blaIMP-1-aac(6')-Ia-orfG + orfH-attC<sub>orfH</sub></sub>* que contiene al sitio *attC<sub>blaIMP-1</sub>* al que esta integrasa no reconoció en ninguna de las otras estructuras evaluadas. La hebra inferior del sitio *attC* plegado del gen *cassette dfrA1* (de 95 pares de bases) tiene las bases extrahelicoidales en las posiciones 71 o 72 y la segunda en la posición 80, separadas por una distancia de 7 u 8 nucleótidos. En este estudio, al modificar la estructura de este sitio *attC* manteniendo las citosinas en las mismas posiciones y la distancia de bases extrahelicoidales, no hubo cambios en las frecuencias de escisión. Al modificar por ejemplo la presencia de una citosina como una segunda base extrahelicoidal, se favoreció la escisión por parte de Int1, demostrando que las integrasas tienen una preferencia específica en elegir las bases extrahelicoidales. Adicionalmente se redujo a través de técnicas moleculares la distancia entre las bases extrahelicoidales del sitio *attC<sub>dfrA1</sub>*,

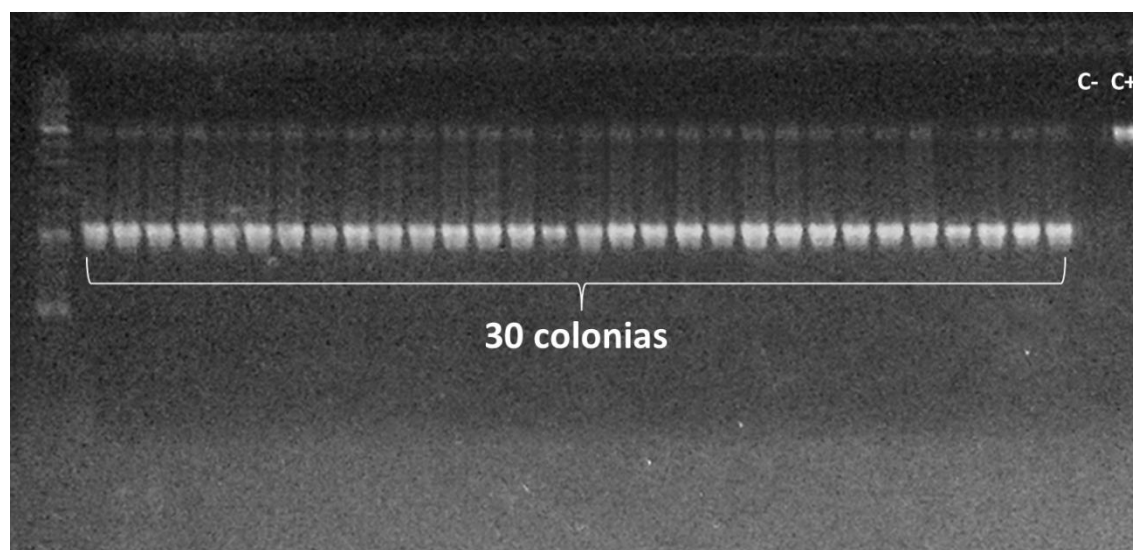
## *Resultados y discusión*

transformándola en 6 en vez de 7 u 8, lo que llevó a un incremento en la escisión por Int1. Esto demuestra que Int1 trabaja más eficientemente cuando la distancia entre las bases extrahelicoidales es más corta. Se comprobó también que no siempre la identidad de las bases afecta la unión de las diferentes integrasas pero si tienen una influencia sobre la reacción de recombinación. Toda esta información concuerda con los resultados obtenidos en este trabajo de tesis, dónde las frecuencias más altas de escisión se observaron en las estructuras con sitios *attC* que contienen en su secuencia bases extrahelicoidales a una distancia de 6 o 7 nucleótidos *attC<sub>aadA3</sub>* y *attC<sub>aacA1</sub>*, a pesar de que las diferencias entre estos dos sitios *attC* son destacables, especialmente por la longitud de su secuencia, siendo el *attC<sub>aacA1</sub>* el más largo de todos, conservan la similitud de sus bases extrahelicoidales.

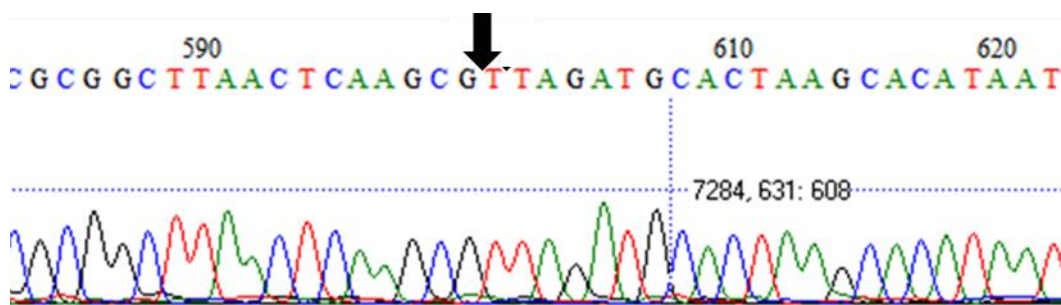
Según lo reportado, los genes *cassettes* se integran preferencialmente en el sitio *attI1* que contiene el sitio GTTRRY (o *core site*) (Stokes y col., 1997; Collis y col., 2001). En la especie bacteriana ambiental *Nitrosomonas europaea* se encontraron 2 integrasas en su genoma, una de ellas ubicada en un integrón cromosómico la cual fue denominada como (IntINeu). Se realizaron una serie de experimentos de ensayos de recombinación donde se ensamblaron distintas estructuras genómicas consistentes en distintos arreglos de genes *cassettes* con diferentes sitios *att* tanto *upstream* como *downstream* del ORF, para nuestro caso de interés se usó para la construcción de varios de los clones el sitio *attC<sub>aadA1</sub>* *upstream* de distintos ORFs de resistencia, los cuales mostraron altos niveles de escisión por IntINeu excepto para aquella estructura que tenía *upstream* el sitio *attC* de *aadA1* seguido del ORF correspondiente al *cassette dfrA1* y su respectivo sitio *attC* (Leon y Roy, 2003).

#### 4.4.1.1- Verificación del evento de escisión del 100%

El evento de recombinación que dio lugar a la escisión del gen *cassette aadA1* para el ensayo realizado con el plásmido pLQ445 (*attC<sub>aacA1</sub>-aadA1- attC<sub>aadA1</sub>*), resultó en valores cercanos al 100%, razón por la cual decidimos comprobar que dicha recombinación fuese de tipo sitio específica. A partir de una de las colonias que resultó positiva para el ensayo de escisión del gen *cassette aadA1* (Figura 22), se analizó la secuencia del producto resultante de la reacción de PCR de escisión y se determinó que la recombinación es sitio específica, mostrando el punto de corte específico para la integrasa, (Figura 23).



**Figura 24.** Gel de agarosa donde se observa la frecuencia de escisión (100%) del ensayo de recombinación con el clon pLQ445 *attC<sub>aacA1</sub>-aadA1- attC<sub>aadA1</sub>*. Se observan las bandas de escisión correspondientes a las 30 colonias, seguidas del control negativo y el control positivo (plásmido pLQ445 sin escindir).



**Figura 25. Secuencia del fragmento de escisión secuenciado.** Se observa marcado con una flecha el lugar específico donde ocurrió la escisión del gen *cassette aadA1* en el ensayo de recombinación con el clon pLQ 445 con la estructura ***attC<sub>aacA1</sub>-aadA1- attC<sub>aadA1</sub>***.

#### 4.4.2- Evaluación de la capacidad de IntI1 para insertar el lado 5' de los genes *cassettes*

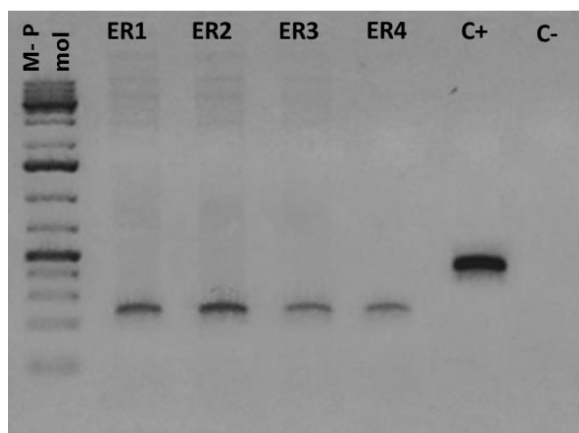
***attI1-aadA1-attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>*, *attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>* y *attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>***

Se evaluó la capacidad de IntI1 de insertar el lado 5' de los genes *cassettes attI1-aadA1-attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>*, *attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>*, *attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>* en el sitio *attI1* del clon PMI1-1 (AS)2. Para evaluar este evento se usaron los cebadores M13Fndel y *aadA1* que generan un amplicón de 278 pb (Tabla 14 y Figura 26). Los valores obtenidos se analizarán en el apartado siguiente.

**Tabla 14.** Frecuencias de inserción en el sitio *attI1*, del lado 5' de los *genes cassettes* analizados, estimadas en los ensayos de recombinación *in vivo*.

| Sitio blanco                      | Estructuras analizadas   | Inserción 5' (%) |
|-----------------------------------|--|------------------|
| PMI1-1(AS)2<br><i>IntI1-attI1</i> | pLQ423: <i>attI1-aadA1-attC<sub>aadA1</sub></i>                | 8,8              |
|                                   | pLQ443: <i>attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub></i> | 2,2              |
|                                   | pLQ445: <i>attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub></i> | 4,4              |
|                                   | pLQ426: <i>attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub></i> | 1,1              |

El porcentaje de inserción 5' corresponde al promedio de al menos tres ensayos independientes.



**Figura 26.** Gel de agarosa donde se muestra el resultado de la evaluación por PCR de las inserciones en el sitio *attI1*, del lado 5' de los *genes cassettes* para los distintos ensayos de recombinación. La evaluación de la inserción del lado 5' del gen *cassette aadA1* se realizó a través de una PCR con los cebadores M13Fndel y *aadA1*. Se muestran los resultados para cada uno de los clones usados en los ensayos de recombinación presentados en este estudio. ER1: PCR de inserción a 5' para el clon pLQ423 (*attI1-aadA1-attC<sub>aadA1</sub>*) ER2: PCR de inserción a 5' para el clon pLQ443 (*attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>*) ER3: PCR de inserción a 5' para el clon pLQ445 (*attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>*) ER4: PCR de inserción a 5' para el clon pLQ426 (*attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>*). Control positivo: pC16Ildc (Quiroga, MP Tesis Doctoral).

#### 4.4.3- Evaluación de la capacidad de Int1 para insertar el lado 3' de los genes *cassettes*

*attI1-aadA1-attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>*, *attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>* y *attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>*

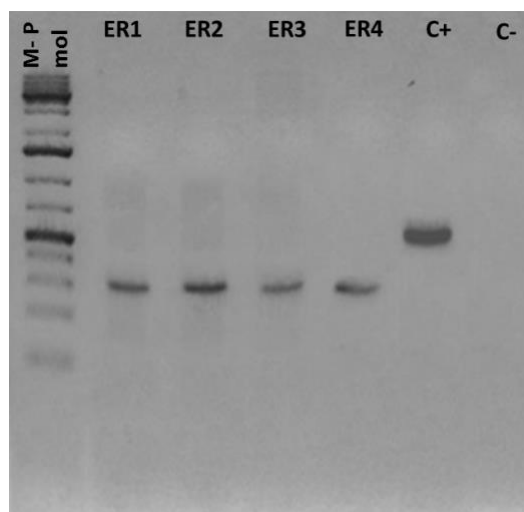
Se evaluó la capacidad de Int1 de insertar a los genes *cassettes attI1-aadA1-attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>*, *attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>*, *attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>* en el sitio *attI1* del clon PMI1-1 (AS)2 del lado 3' del *cassette*. Se usaron los cebadores *aadAR* y M13R que generan un amplicón de 249 pb, (Tabla 15 y Figura 27).



**Tabla 15.** Frecuencias de inserción en el sitio *attI1*, del lado 3' de los *genes cassettes* analizados, estimadas en los ensayos de recombinación.

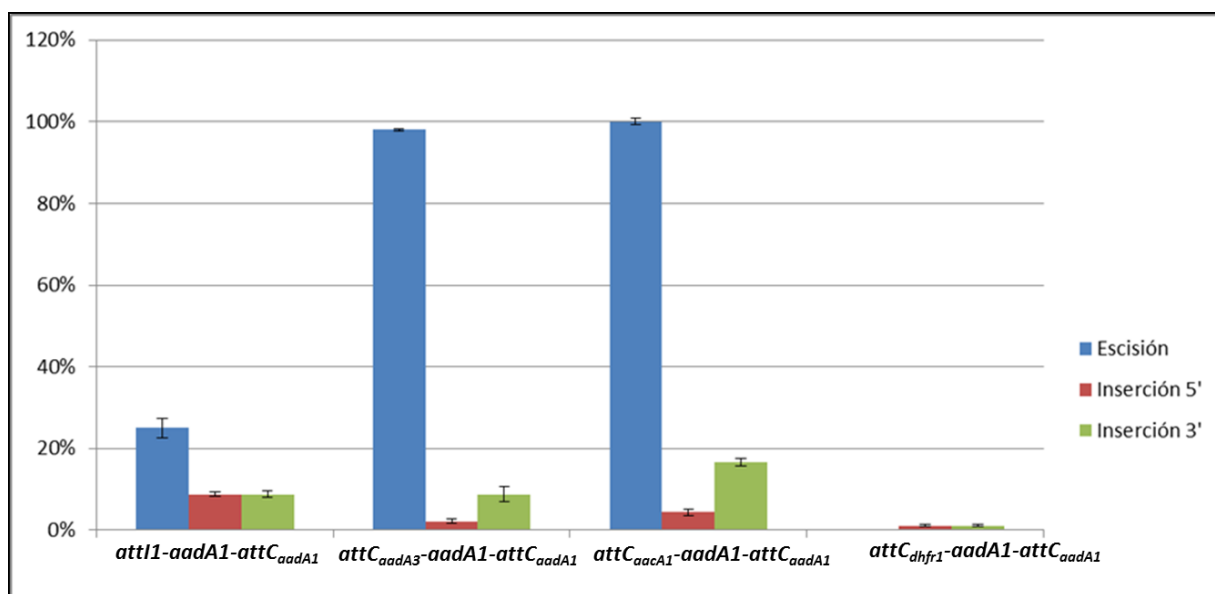
| Sitio blanco                      | Estructuras analizadas   | Inserción 3' (%) |
|-----------------------------------|--|------------------|
| PMI1-1(AS)2<br><i>IntI1-attI1</i> | pLQ423: <i>attI1-aadA1-attC<sub>aadA1</sub></i>                | 8,8              |
|                                   | pLQ443: <i>attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub></i> | 8,8              |
|                                   | pLQ445: <i>attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub></i> | 16,6             |
|                                   | pLQ426: <i>attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub></i> | 1,1              |

El porcentaje de inserción 3' corresponde al promedio de al menos tres ensayos independientes.



**Figura 27.** Gel de agarosa donde se muestra el resultado de la evaluación por PCR de las inserciones en el sitio *attI1*, del lado 3' de los *genes cassettes* para los distintos ensayos de recombinación. La evaluación de la inserción del lado 3' del gen *cassette aadA1* se realizó a través de una PCR con los cebadores *aadAR* y *M13R*. Se muestran los resultados para cada uno de los clones usados en los ensayos de recombinación presentados en este estudio. ER1: PCR de inserción a 3' para el clon pLQ423 (*attI1-aadA1-attC<sub>aadA1</sub>*) ER2: PCR de inserción a 3' para el clon pLQ443 (*attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>*) ER3: PCR de inserción a 3' para el clon pLQ445 (*attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>*) ER4: PCR de inserción a 3' para el clon pLQ426 (*attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>*). Control positivo: pC16Ildc (Quiroga, MP Tesis Doctoral).

Con respecto a los valores de inserción presentados en este trabajo, consiste en un aporte nuevo ya que únicamente en anteriores trabajos se evaluó cuantitativamente la frecuencia de escisión de estos clones en particular. En la figura 28 se resumen en forma de porcentaje las frecuencias de escisión e inserción para cada uno de los ensayos de recombinación.



**Figura 28. Porcentajes de recombinación de los genes *cassettes* estudiados en este trabajo.** Porcentajes de inserción del lado 5', del lado 3', y de escisión de los genes *cassettes* *attI1-aadA1-attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>*, *attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>* y *attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>*. Los porcentajes de inserción y de escisión corresponden al promedio de al menos tres ensayos independientes. Mostramos las barras de error típico.

A partir de los estudios de recombinación realizados por Quiroga, MP (2012) en su tesis doctoral, trabajo en el cual se empleó la misma metodología para realizar los ensayos de recombinación presentados en el presente trabajo, con el fin de evaluar las frecuencias de recombinación de estructuras o plataformas genéticas determinadas de nuestro interés. Se determinó la funcionalidad de genes *cassettes* inusuales (tipo  $\Delta attI$ ) ya que presentó tanto eventos de escisión como de inserción a través de mecanismos de recombinación sitio específica, para los *cassettes* atípicos *attI*-ORF-*attI* específicamente en el caso del gen *cassette aadB* (confiere resistencia a kanamicina, gentamicina, tobramicina) y también se encontró que para lo que se ha denominado genes *cassettes* típicos con una estructura *attI*-

## Resultados y discusión

ORF-*attC* o *attC*-ORF-*attC* tienen un valor alto para sus frecuencias de escisión. La frecuencia mayor se presenta para la estructura inusual *attI1-aadB-attI1* con una frecuencia cercana al 100%. En este estudio se encontró un valor similar de frecuencias de escisión para las estructuras *attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>* y para *attC<sub>aaCA1</sub>-aadA1-attC<sub>aadA1</sub>* lo cual podría demostrarnos que hay ciertos sitios *attC* que pueden ser un mejor blanco para la integrasa para realizar la escisión, o la combinación de ciertos tipos de sitio de recombinación (*attI* o *attCs*) *upstream* y *downstream* del marco de lectura del gen *cassette* podría favorecer la escisión del 100% del gen *cassette* en diversos eventos de recombinación. Estos arreglos en tándem de genes *cassettes* no se observan en la clínica, lo que nos puede estar indicando que no es una plataforma genética estable para el integrón por lo que debe escindir esta estructura en una alta frecuencia para ser insertada en otra región de recombinación. Para analizar estos resultados hay que tener en cuenta que el sitio *attI1* recombina cuando se encuentra en forma de doble cadena y el sitio *attC* únicamente cuando se ha producido el plegamiento de su cadena inferior que forma una estructura secundaria, lo que podría favorecer el aumento de la frecuencia de escisión. Algunas mutaciones que afectan el compromiso del plegamiento de la estructura disminuyen las frecuencias de recombinación, el plegado global de la hebra inferior es esencial para que ocurra la recombinación (Bouvier y col., 2005).

Para el caso de las frecuencias de inserción de este mismo gen *cassette* atípico (*attI1-aadB-attI1*) se encontró un porcentaje de inserción del 7% en las posiciones 5' y 3' del sitio *attI1* indicando que todas fueron inserciones netas, y a través del análisis de sus secuencias

## Resultados y discusión

que se inserta de manera sitio específica (Quiroga, 2012). Este resultado es similar al de este estudio para la estructura *attI1-aadA1-attC<sub>aadA1</sub>* donde se determinó que la frecuencia de inserción a 5' y 3' es del 8,8% para ambas posiciones. Fue el único de los casos para este estudio en que podemos considerar que todas las inserciones fueron netas.

El porcentaje de inserción del gen *cassette attI1-aadB-attC<sub>aadB</sub>* a 5' y 3' fueron de 4 y 9%, respectivamente, mientras que para el gen *cassette attC<sub>sat2</sub>-aadB-attC<sub>aadB</sub>* fueron del 11 y 6% (también a 5' y a 3', respectivamente). Comparamos esos resultados con los valores obtenidos para este estudio, en el que los porcentajes de inserción en el sitio *attI1* para los casos *attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>* fueron del 2,2 y del 8,8% para 5' y 3'; para la estructura *attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>* de 4,4% y de 16,6% para 5' y 3', respectivamente; y para el caso de la estructura *attC<sub>dfrA1</sub>-aadA1-attC<sub>aadA1</sub>* ambos valores fueron menores al 3%. En este estudio se evidenció que se favoreció la recombinación entre el sitio *attI1* y los dos primeros sitios *attC* a través de la vía de formación de cointegrados y evidentemente el *attC<sub>dfrA1</sub>* tiene una influencia negativa tanto para los eventos de escisión como los de inserción en el sitio de recombinación *attI1*. Al comparar todos estos valores se observa que para la estructura *attC<sub>sat2</sub>-aadB-attC<sub>aadB</sub>* el valor es mayor para la inserción a 5' en el sitio *attI*, superando los valores obtenidos para los genes *cassettes* restantes *attI1-aadB-attC<sub>aadB</sub>*, *attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>* y *attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>* que tienen un valor mayor de inserción a 3' del mismo sitio *attI*, por lo cual no podemos determinar si hay una tendencia específica que defina la inserción de estas estructuras genéticas que evaluamos en este trabajo. La estructura que

## Resultados y discusión

incluye el *attC<sub>dfrA1</sub>* como ya se ha mencionado a lo largo de todo este documento tiene una influencia marcada en disminuir los eventos de recombinación (escisión e inserción).

Al comparar los valores se observa que para las estructuras de *attC*-ORF-*attC* el valor es mayor para la inserción a 5' superando los valores obtenidos para el gen *cassette aadB* pero conservando la tendencia a insertarse más en esta posición del sitio *attI1* (sin tener en cuenta la estructura que incluye el *attC<sub>dfrA1</sub>* que como ya se ha mencionado a lo largo de todo este documento tiene una influencia marcada en disminuir los eventos de recombinación).

Al realizar la comparación estructural de los distintos sitios *attCs* trabajados en este estudio (ver en detalle: 4.3- Estructuras secundarias de los sitios *attC* usados en los estudios de recombinación), se ve que no se puede establecer una condición específica bajo la cual la recombinación ya sea para la escisión o para la inserción se vea notablemente influenciada. Simplemente, parece ser que los eventos de recombinación de escisión son de menos costo energético para la integrasa que los eventos de inserción y que definitivamente están marcados por la similitud entre las bases extrahelicoidales y la distancia entre ellas de los sitios *att* que flanquean al gen *cassette* tanto *upstream* como *downstream*, creemos que entre más similitudes tengan los dos sitios la recombinación va a ser más eficiente, específicamente para las características que influyen con la recombinación (bases extrahelicoidales y el espaciado entre ellas).

Según una de las publicaciones más recientes en la cual se construyeron integrones sintéticos para comparar la escisión y el aumento de la resistencia de un gen específico de

## *Resultados y discusión*

resistencia a Cloranfenicol (*cat2*) al estar bajo presión antibiótica y ver si en algún momento bajo dicha presión, dicho gen *cassette* se insertaba en el primer lugar disponible del integrón, se vió que era más frecuente el evento de escisión de los genes *cassettes upstream* del mismo, pero no la escisión de este gen y la inserción en el primer lugar del integrón: un corte y pegado del gen *cassette*. Adicionalmente según estos autores, la escisión del gen *cassette* y la formación de un cointegrado requiere únicamente un evento de recombinación, mientras que la escisión y la integración de un gen *cassette* requeriría dos eventos de recombinación y los eventos simples de reordenamiento son más frecuentes que los de reordenamiento múltiple (Barraud y Ploy, 2015).

El modelo presentado en este trabajo incluye únicamente el gen *cassette aadA1* que es uno de los más importantes en la clínica y sobre todo se presenta con una alta frecuencia en los integrones clínicos. Con respecto a los valores de escisión de la estructura *attI1-aadA1-attC<sub>aadA1</sub>* se sabe que son menores en comparación con las estructuras *attC-ORF-attC* por lo que es difícil ver que el gen *cassette* que esté en primera posición en el integrón se escinda más frecuentemente ya que adicionalmente sería el que más se está expresando su resistencia. Sin embargo, como lo pudimos ver en este estudio y en otros similares, esto también depende de la secuencia y las características de la estructura secundaria del sitio *attC* del que estemos hablando ya que hay algunos que por sus características no son eficientes para realizar eventos de recombinación y por alguna razón deben estar en cierto lugar dentro del rearreglo de la región variable de los integrones. Además, evidentemente depende de la presión antibiótica que se le esté ejerciendo a la bacteria que contenga

### *Resultados y discusión*

integraciones en su genoma ya que Int11 tiene la propiedad intrínseca para capturar sitios *attC* en el momento de su extrusión, estabilizándolos y recombinándolos eficientemente. Con respecto a esto, se demostró a través de experimentos *in vivo* y como ya hemos venido describiendo a lo largo de los resultados, que únicamente la hebra inferior del sitio *attC* es usada como sustrato durante la integración de los genes *cassettes*. Diversos procesos bacterianos que permiten la entrega de ADN en forma de simple cadena como la conjugación y la replicación, favorecen el plegamiento adecuado de los sitios *attC* y que la integrasa del integrón estabiliza el plegamiento de los sitios *attC* de una forma específica (Loot y col., 2014).

# Capítulo 5: Conclusiones

---

- ✓ Identificamos y determinamos 37 alelos *aadA-like* asociados a integrones multirresistentes a través de la búsqueda bioinformática en la base de datos INTEGRALL.
- ✓ El gen *cassette aadA1* es el alelo más frecuente del grupo de los alelos *aadA-like* con el mayor número de registros en la base de datos INTEGRALL.
- ✓ De los 37 alelos *aadA-like*, el gen *cassette aadA1* es el único asociado a integrasas de tipo 1 y 2; los alelos *aadA-like* restantes están asociados a la integrasa de clase 1 o a integrasas de esta misma clase deletadas.
- ✓ El orden de prevalencia de los alelos *aadA-like* fue, en orden descendente, *aadA1*, *aadA2* y *aadA5*.
- ✓ En referencia al gen *cassette aadA1* y su ubicación dentro de la zona variable de los integrones de clase 1, la mayoría corresponden al arreglo *dfrA1-aadA1*, y en segundo lugar se identificó a *aadA1* como único gen *cassette*.
- ✓ En referencia al gen *cassette aadA1* y su ubicación dentro de la zona variable de los integrones de clase 2, la mayoría de corresponden al arreglo *dfrA1-sat2-aadA1* seguido del arreglo *sat2-aadA1*, y a diferencia de lo que se encontró en los integrones de clase 1, no se encontró al gen *cassette aadA1* como único gen *cassette* de la zona variable.
- ✓ El gen *cassette aadA1* en integrones de clase 1 tiende a ubicarse con el mismo valor de probabilidad en la primera, segunda y tercera posición de las zonas variables de los integrones analizados.



## Conclusiones

- ✓ El gen *cassette aadA1* en integrones de clase 2 tiende a ubicarse con el mismo valor de probabilidad en la segunda y tercera posición de las zonas variables de los integrones analizados.
- ✓ En referencia al gen *cassette aadA2*, el arreglo más frecuente de la zona variable de integrones de clase 1 fue *aadA2* como único gen *cassette*, seguido del arreglo *dfrA12-orfA-aadA2*.
- ✓ En referencia al gen *cassette aadA5* el arreglo más frecuente de la zona variable de integrones de clase 1 fue *dfrA17-aadA5* seguido en mucha menor medida por *aadA5* como único gen *cassette* de la zona variable.
- ✓ Otros genes *cassettes* asociados a integrones de clase 1 fueron *aadA6*, *aadA7*, *aadA1a*, *aadA1b*, *aadA11* y *aadA4* con bajo número de registros.
- ✓ A partir del análisis de las secuencias de los ORFs de los 37 alelos *aadA-like*, encontramos la formación de un grupo predominante denominado como grupo *aadA1-like*.
- ✓ A partir del análisis de las secuencias de los sitios *attC* de los 37 alelos *aadA-like*, identificamos 15 variantes de sitios *attC* asociados a los 37 alelos.
- ✓ Determinamos que hay intercambio entre los ORFs y los sitios *attC* de los 37 alelos *aadA-like* para formar distintos genes *cassettes*.
- ✓ La variante 1 de los sitios *attC* correspondiente a los 37 alelos, denominada variante correspondiente al grupo *aadA1-like*, es la más dispersa con el mayor número de secuencias asociadas a los registros contenidos en el GenBank y referidas al gen *cassette aadA1*.
- ✓ Con respecto al análisis de las estructuras secundarias de los sitios *attC* encontramos que de las 4 regiones 1L o *inverse core site* de los 4 sitios *attC* estudiados *attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>*,

## Conclusiones

*attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>*, los sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>* coinciden exactamente con la secuencia consenso de la región 1L, mientras que los sitios *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>* presentan la transversión de un nucleótido cada uno.

✓ A partir de la comparación de los porcentajes de escisión los genes *cassettes attI1-aadA1-attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>* y *attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>* son blancos favorables para la escisión mediada por IntI1, a diferencia del gen *cassette attC<sub>dhfr1</sub>-aadA1- attC<sub>aadA1</sub>* donde el valor de frecuencia de escisión fue el más bajo de todos los ensayos.

✓ El gen *cassette attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>* y el gen *cassette attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>* son escindidos más eficientemente que los demás genes *cassettes* analizados, con una frecuencia de escisión cercana al 100%.

✓ En este estudio se evidenció, que se favoreció la recombinación entre el sitio *attI1* y los dos sitios *attC<sub>aadA3</sub>* y *attC<sub>aacA1</sub>* a través de la vía de cointegrados.

✓ Los valores de inserción presentados en este trabajo, consisten en un aporte nuevo ya que en anteriores trabajos se evaluó solamente la frecuencia de escisión de los clones estudiados aquí.

✓ En este estudio se encontró que para lo que se ha denominado genes *cassettes* típicos con una estructura *attI-ORF-attC* o *attC-ORF-attC* tienen un valor alto para sus frecuencias de escisión, excepto cuando se involucra el sitio *attC<sub>dfrA1</sub>* que tiene características moleculares específicas que causan una influencia marcada en disminuir los eventos tanto de inserción como de escisión.

✓ Según los datos obtenidos de frecuencias de recombinación, los eventos de escisión son de menos costo energético para la integrasa que los eventos de inserción.

# Resumen

---

La Transferencia Horizontal Genética (THG) es una de las estrategias evolutivas más eficientes de las bacterias, ya que la adquisición de material genético foráneo permite aumentar la variabilidad genética de un determinado genoma, de manera casi instantánea. Los integrones son elementos genéticos, claves en la captura, expresión y dispersión de genes que se encuentran bajo forma de genes *cassettes* entre las bacterias gram-negativas. Codifican una recombinasa sitio específica, la integrasa (IntI), la cual reconoce dos sitios de recombinación muy diferentes entre sí, el sitio *attI*, adyacente al gen *intI*, y los sitios *attC*. A su vez, los genes *cassettes* son los elementos móviles de esta plataforma genética, y están formados por un marco de lectura abierto (ORF) y un sitio *attC*, siendo la arquitectura de los genes *cassettes* típicos insertos en la región variable del integrón *attI1*-ORF-*attC* y *attC*-ORF-*attC*. Los integrones de clase 1 y 2 son los más exitosos dentro del ambiente hospitalario, asociados a multirresistencia antibiótica mientras que otros han sido detectados esporádicamente.

La mayoría de los genes de resistencia antibiótica que se encuentran en la clínica actualmente se originaron a partir de bacterias del ambiente, incluyendo especies productoras de antibióticos. Se han descrito dos ramas de genes *cassettes aadA* que confieren resistencia a estreptomicina/espectinomicina asociados a integrones en cepas clínicas. Estas ramas son *aadA1* y *aadA2*. Ambas son modelo de estudio, ya que se ha descrito que al menos el gen *cassette aadA1* habría estado presente previo a la era antibiótica y se habría transmitido a través de eventos de THG a las bacterias clínicas

## Resumen

actuales. Dicho gen *cassette* se encuentra ampliamente distribuido en los aislamientos clínicos aportando a la problemática de la multiresistencia antibiótica, razón por la cual realizamos un estudio molecular y funcional del mismo en este trabajo.

A través de análisis bioinformáticos identificamos un grupo de 37 alelos que denominamos *aadA-like*. Encontramos que únicamente el gen *cassette aadA1* está asociado a integrones de clase 1 y 2, a diferencia de los alelos *aadA-like* restantes, que están asociados exclusivamente a los de clase 1. El alelo *aadA1* es el más abundante y asimismo es el que más arreglos presenta en la zona variable del integrón en el cual está contenido. En segundo y tercer lugar de prevalencia, están los alelos *aadA2* y *aadA5* respectivamente. Con respecto a los genes *cassettes* asociados, encontramos que el gen *cassette aadA1* está estrechamente relacionado con el gen *cassette dfrA1* o se encuentra solo en la zona variable en integrones de clase 1. Adicionalmente, este gen *cassette* se tiende a ubicar en las 3 primeras posiciones de la zona variable del integrón, lo que explica la alta expresión de la resistencia en aislamientos clínicos.

En cuanto al análisis de la dispersión del gen *cassette aadA1*, lo encontramos un total de 86 arreglos de genes *cassettes*. La asociación del gen *cassette aadA1* a integrasas de integrones de clase 1 fue del 84%. Los genes *cassettes aadA1* y *dfrA1* se encontraron juntos tanto en integrones de clase 1 y 2, lo que nos da una idea de que estos dos genes *cassettes* están muy asociados en los integrones clínicos.

Con respecto al estudio de los componentes genéticos de los genes *cassettes aadA-like* a través de análisis de alineamiento y máxima similitud, se observó la formación de

## Resumen

algunos grupos, destacando el grupo que incluye a los que denominamos como alelos *aadA1-like* (7 alelos).

Mediante los análisis utilizando CLUSTALW y máxima similitud se identificaron 15 variantes del sitio *attC* asociadas a los 37 alelos *aadA-like* con una mayor identidad de secuencias entre las bases nucleotídicas involucradas en la recombinación específicamente en las regiones 1L y 1R (RYYYAAC y GTTRRRY, respectivamente). La variante 1 del sitio *attC* proveniente del agrupamiento de máxima similitud contiene a casi la mitad de los alelos (15 de 37). Al igual que para el caso del análisis de los ORFs, la agrupación de la mayoría de los alelos corresponde a los denominados alelos *aadA1-like* lo cual confirma la alta dispersión de este gen *cassette* en particular.

Para el estudio funcional del gen *cassette aadA1* analizamos las características moleculares de los sitios *attC* flanqueantes de este gen *cassette* en los clones en los cuales realizamos ensayos de recombinación. La estructura de los genes *cassettes* clonados en cada uno de estos plásmidos corresponde a estructuras típicas observadas en diferentes arreglos de integrones de la clínica: *attC<sub>aadA1</sub>*, *attC<sub>aadA3</sub>*, *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>*. Analizamos las regiones que están involucradas en el proceso de recombinación, la región *core site* o 1R, la región 1L o *inverse core site* y la presencia y espaciado de bases extrahelicoidales.

Encontramos que el sitio *attC<sub>aadA1</sub>* es el único que se ajusta completamente a la secuencia consenso de la región 1R o *core site* 5'-GTTRRRY-3'. Los tres sitios *attC* restantes (*attC<sub>aadA3</sub>*, *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>*) se caracterizan por tener una transversión de una Purina a Pirimidina en uno de sus nucleótidos, pero no en aquellos nucleótidos donde la integrasa realiza el corte en el proceso de recombinación 5'-G'TT-3' (El punto de corte es entre la G y la

## Resumen

T). Los sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>* tienen un 100% de identidad de secuencia entre sus regiones 1L 5'-GTTTAAC-3' y coinciden exactamente con la secuencia consenso de la región 1L, mientras que los correspondientes a los sitios *attC<sub>aacA1</sub>* y *attC<sub>dfrA1</sub>* presentan la transversión de un nucleótido cada uno.

Al observar los determinantes de reconocimiento estructural de importancia para la recombinación, se pudo determinar que los sitios *attC<sub>aadA1</sub>* y *attC<sub>aadA3</sub>* mantienen una identidad en cuanto a sus bases extrahelicoidales G y T, así como la distancia entre ellas (6 pares de bases), lo cual podría estar contribuyendo a aumentar las frecuencias de escisión durante la recombinación. Con respecto al sitio *attC<sub>aacA1</sub>* las bases extrahelicoidales conservan la misma identidad G y T y el mismo espaciado entre ellas (6 pares de bases) y presentan valores muy similares de escisión cuando este sitio *attC* está en la estructura *attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>*.

Dos sitios de recombinación *attC*, el sitio *attC<sub>aadA3</sub>* de 60 pb y el sitio *attC<sub>aacA1</sub>* de 118 pb, muy diferentes en longitud y en su región espaciadora, presentan frecuencias de recombinación muy parecidas, lo cual indica que la importancia de la identidad de las bases extrahelicoidales y el espaciado entre las mismas, es la característica fundamental que influye sobre la frecuencia de recombinación mientras que el sitio *attC<sub>dfrA1</sub>* cuyas bases extrahelicoidales son C y A con una distancia entre ellas de 7 a 8 pb presenta un marcado descenso en las frecuencias de escisión.

En cuanto a los ensayos de recombinación de genes *cassettes* mediados por *Int1* realizados *in vivo*, se utilizaron en cada caso al clon PMI1-1 (AS)2 (*int11-att11*) que contiene al gen de la integrasa de los integrones de clase 1 cotransformando *E. coli* TOP10 con de los

## Resumen

siguientes clones: pLQ423 (*attI1-aadA1-attC<sub>aadA1</sub>*), pLQ443 (*attC<sub>aadA3</sub>-aadA1-attC<sub>aadA1</sub>*), pLQ445 (*attC<sub>aacA1</sub>-aadA1-attC<sub>aadA1</sub>*) o pLQ426 (*attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>*). Se observó que las tres primeras estructuras son blancos favorables para la escisión mediada por IntI1, a diferencia del gen *cassette attC<sub>dhfr1</sub>-aadA1-attC<sub>aadA1</sub>* donde el valor de frecuencia de escisión fue el más bajo de todos los ensayos. Con respecto a las frecuencias de inserción en el sitio *attI1*, al analizar la inserción neta, se observó que se favoreció la recombinación entre el sitio *attI1* y los sitios *attC<sub>aadA3</sub>* y *attC<sub>aacA1</sub>* a través de la vía de cointegrados.

Al analizar en conjunto los resultados de bioinformática y funcionales observamos que los patrones moleculares que favorecen a la escisión e inserción neta de *cassettes* difieren entre sí y por lo tanto deben seguir siendo estudiados en forma particular para lograr un posible consenso de parámetros que reúna todos los requerimientos de los sitios blanco de la integrasa de integrones de clase 1.

# Referencias bibliográficas

---

Barraud O, Ploy MC. "Diversity of class 1 integron gene *cassette* rearrangements selected under antibiotic pressure". *J Bacteriol.* 197:2171–2178. 2015.

Bennett, P.M. "Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria". *Br. J. Pharmacol.* 153(Suppl. 1): 347-357. 2008.

Biskri, L., M. Bouvier, A. M. Guerout, S. Boissard, and D. Mazel. "Comparative Study of Class 1 Integron and *Vibrio cholerae* Superintegron Integrase Activities." [In eng]. *J Bacteriol* 187, no. 5: 1740-50. 2005.

Bissonnette, L., Roy, P.H. "Characterization of *In0* of *Pseudomonas aeruginosa* plasmid pVS1, an ancestor of integrons of multiresistance plasmids and transposons of gram-negative bacteria". *J. Bacteriol.* 174: 1248-1257. 1992.

Bouvier, M., G. Demarre, and D. Mazel. "Integron *Cassette* Insertion: A Recombination Process Involving a Folded Single Strand Substrate." [In eng]. *EMBO J* 24, no. 24: 4356-67. 2005.

Bouvier, M., M. Ducos-Galand, C. Loot, D. Bikard, and D. Mazel. "Structural Features of Single-Stranded Integron *Cassette attC* Sites and Their Role in Strand Selection." [In eng]. *PLoS Genet* 5, no. 9: e1000632. 2009.

Brown, H.J., Stokes, H.W., Hall, R.M. "The integrons *In0*, *In2*, and *In5* are defective transposon derivatives". *J. Bacteriol.* 178: 4429-4437. 1996.

Bunny K L, Hall R M, Stokes H W. "New mobile gene *cassettes* containing an aminoglycoside resistance gene, *aacA7*, and a chloramphenicol resistance gene, *catB3*, in an integron in pBWH301. *Antimicrob Agents Chemother.* 39: 686–693. 1995.

Cambray, G., A. M. Guerout, and D. Mazel. "Integrons." [In eng]. *Annu Rev Genet* 44: 141-66. 2010.

Cameron, F. H., D. J. Groot Obbink, V. P. Ackerman, and R. M. Hall. "Nucleotide Sequence of the *aad(2'')* Aminoglycoside Adenylyltransferase Determinant *aadB* Evolutionary Relationship of This Region with Those Surrounding *aadA* in R538-1 and *dhfrII* in R388." [In eng]. *Nucleic Acids Res* 14, no. 21: 8625-35. 1986.



## Referencias bibliográficas

Collis, C. M., and R. M. Hall. "Expression of Antibiotic Resistance Genes in the Integrated Cassettes of Integrons." [In eng]. *Antimicrob Agents Chemother* 39, no. 1: 155-62. 1995.

Collis, C. M., and R. M. Hall. "Gene Cassettes from the Insert Region of Integrons Are Excised as Covalently Closed Circles." [In eng]. *Mol Microbiol* 6, no. 19: 2875-85. 1992a.

Collis, C. M., and R. M. Hall. "Site-Specific Deletion and Rearrangement of Integron Insert Genes Catalyzed by the Integron DNA Integrase." [In eng]. *J Bacteriol* 174, no. 5: 1574-85. 1992b.

Collis, C. M., G. D. Recchia, M. J. Kim, H. W. Stokes, and R. M. Hall. "Efficiency of Recombination Reactions Catalyzed by Class 1 Integron Integrase IntI1." [In eng]. *J Bacteriol* 183, no. 8: 2535-2542. 2001.

D'Costa, V.M., McGrann, K.M., Hughes, D.W., and Wright, G.D. "Sampling the Antibiotic Resistome", *Science*. vol. 311, no. 5759: 374–377. 2006.

Davies, J. "Inactivation of Antibiotics and the Dissemination of Resistance Genes". *Science*. vol. 264: 375–382. 1994.

Esiobu, N., Armenta, L., and Ike, J. "Antibiotic Resistance in Soil and Water Environments", *Int. J. Environ. Health Res*. vol. 12: 133–144. 2002.

Fluit, A. C., and F. J. Schmitz. "Class 1 Integrons, Gene Cassettes, Mobility, and Epidemiology." [In eng]. *Eur J Clin Microbiol Infect Dis* 18, no. 11: 761-70. 1999.

Francia, M. V., J. C. Zabala, F. de la Cruz, and J. M. Garcia Lobo. "The IntI1 Integron Integrase Preferentially Binds Single-Stranded DNA of the *attC* Site." [In eng]. *J Bacteriol* 181, no. 21: 6844-9. 1999.

Gravel, A., B. Fournier, and P. H. Roy. "DNA Complexes Obtained with the Integron Integrase IntI1 at the *attI1* Site." [In eng]. *Nucleic Acids Res* 26, no. 19: 4347-55. 1998a.

Gravel, A., N. Messier, and P. H. Roy. "Point Mutations in the Integron Integrase IntI1 That Affect Recombination and/or Substrate Recognition." [In eng]. *J Bacteriol* 180, no. 20: 5437-42. 1998b.

Grindley, N. D., K. L. Whiteson, and P. A. Rice. "Mechanisms of Site-Specific Recombination." [In eng]. *Annu Rev Biochem* 75: 567-605. 2006.

Guerin, E., T. Jove, A. Tabesse, D. Mazel, and M. C. Ploy. "High-Level Gene Cassette Transcription Prevents Integrase Expression in Class 1 Integrons." [In eng]. *J Bacteriol* 193, no. 20: 5675-82. 2011.

## Referencias bibliográficas

Hall, R. M., C. M. Collis, M. J. Kim, S. R. Partridge, G. D. Recchia, and H. W. Stokes. "Mobile Gene Cassettes and Integrons in Evolution." [In eng]. *Ann N Y Acad Sci* 870: 68-80. 1999.

Hall, R. M., D. E. Brookes, and H. W. Stokes. "Site-Specific Insertion of Genes into Integrons: Role of the 59-Base Element and Determination of the Recombination Cross-over Point." [In eng]. *Mol Microbiol* 5, no. 8: 1941-59. 1991.

Hanau-Bercot, B., I. Podglajen, I. Casin, and E. Collatz. "An Intrinsic Control Element for Translational Initiation in Class 1 Integrons." [In eng]. *Mol Microbiol* 44, no. 1: 119-30. 2002.

Harbottle, H., Thakur, S., Zhao, S., White, D.G. "Genetics of antimicrobial resistance". *Anim. Biotechnol.* 17: 111- 124. 2006.

Johansson, C., M. Kamali-Moghaddam, and L. Sundstrom. "Integron Integrase Binds to Bulged Hairpin DNA." [In eng]. *Nucleic Acids Res* 32, no. 13: 4033-43. 2004.

Jove, T., S. Da Re, F. Denis, D. Mazel, and M. C. Ploy. "Inverse Correlation between Promoter Strength and Excision Activity in Class 1 Integrons." [In eng]. *PLoS Genet* 6, no. 1: e1000793. 2010.

Larouche, A., and P. H. Roy. "Effect of *attC* Structure on *Cassette* Excision by Integron Integrases." [In eng]. *Mob DNA* 2, no. 1: 3. 2011.

Leon, G., and P. H. Roy. "Excision and Integration of *Cassettes* by an Integron Integrase of *Nitrosomonas europaea*." [In eng]. *J Bacteriol* 185, no. 6: 2036-41. 2003.

Levesque, C., L. Piche, C. Larose, and P. H. Roy. "PCR Mapping of Integrons Reveals Several Novel Combinations of Resistance Genes." [In eng]. *Antimicrob Agents Chemother* 39, no. 1: 185-91. 1995.

Levesque, C., S. Brassard, J. Lapointe, and P. H. Roy. "Diversity and Relative Strength of Tandem Promoters for the Antibiotic-Resistance Genes of Several Integrons." [In eng]. *Gene* 142, no. 1: 49-54. 1994.

Liebert, C.A., Hall, R.M., Summers, A.O. "Transposon Tn21, flagship of the floating genome". *Microbiol. Mol. Biol. Rev.* 63: 507-522. 1999.

Loot, C., Parissi, V., Escudero, J. A., Amarir-Bouhram, J., Bikard, D., & Mazel, D. The "Integron Integrase Efficiently Prevents the Melting Effect of Escherichia coli Single-Stranded DNA-Binding Protein on Folded *attC* Sites". *Journal of Bacteriology*, 196(4): 762–771. 2014.

MacDonald, D., G. Demarre, M. Bouvier, D. Mazel, and D. N. Gopaul. "Structural Basis for Broad DNA-Specificity in Integron Recombination." [In eng]. *Nature* 440, no. 7088 : 1157-62. 2006.

## Referencias bibliográficas

Mazel, D. "Integrans: Agents of Bacterial Evolution." [In eng]. *Nat Rev Microbiol* 4, no. 8: 608-20. 2006.

Mindlin, S., Petrova, M., Gorlenko, Zh., Soina, V., Khachikian, N., Karaevskaya, E., 2009. "Multidrug Resistant Bacteria in Permafrost: isolation, biodiversity, phenotypic and genotypic analysis". In: New Permafrost and Glacier Research. Nova Science Publishers, Inc, Hauppauge, 89-105. 2009.

Partridge, S. R., G. D. Recchia, C. Scaramuzzi, C. M. Collis, H. W. Stokes, and R. M. Hall. "Definition of the *attI1* Site of Class 1 Integrans." [In eng]. *Microbiology* 146 (Pt 11): 2855-64. 2000.

Partridge, S. R., G. Tsafnat, E. Coiera, and J. R. Iredell. "Gene Cassettes and Cassette Arrays in Mobile Resistance Integrans." [In eng]. *FEMS Microbiol Rev* 33, no. 4: 757-84. 2009.

Petrova, M., Gorlenko, Z., Mindlin, S. "Tn5045, a novel integron-containing antibiotic and chromate resistance". *Res. Microbiol.* 162: 337-345.2011.

Quiroga, C., P. H. Roy, and D. Centron. "The *S.ma*.12 Class C Group II Intron Inserts at Integron *attC* Sites." [In eng]. *Microbiology* 154, no. Pt 5: 1341-53. 2008.

Quiroga, MP. "Estudio molecular y funcional de los sitios de recombinación atípicos de los integrones". Tesis Doctoral. 2012.

Recchia, G. D., and R. M. Hall. "Gene Cassettes: A New Class of Mobile Element." [In eng]. *Microbiology* 141 ( Pt 12): 3015-27. 1995a.

Recchia, G. D., and R. M. Hall. "Origins of the Mobile Gene Cassettes Found in Integrans." [In eng]. *Trends Microbiol* 5, no. 10: 389-94.1997.

Recchia, G. D., and R. M. Hall. "Plasmid Evolution by Acquisition of Mobile Gene Cassettes: Plasmid pIE723 Contains the *aadB* Gene Cassette Precisely Inserted at a Secondary Site in the IncQ Plasmid RSF1010." [In eng]. *Mol Microbiol* 15, no. 1: 179-87. 1995b.

Recchia,G.D. and Sherratt,D.J. "Gene acquisition in bacteria by integron-mediated site-specific recombination". In Craig,N.L., Craigie,R., Gellert,M. and Lambowitz,A.M. (ed.), Mobile DNA II. ASM Press, Washington, DC, pp. 162-176. 2002.

Riccio, M.L., Pallecchi, L., Docquier, J.D., Cresti, S., Catania, M.R., Pagani, L., Lagatolla, C., Cornaglia, G., Fontana, R., Rossolini, G.M. "Clonal relatedness and conserved integron structures in epidemiologically unrelated *Pseudomonas aeruginosa* strains producing the VIM-1 metallo- $\beta$ -lactamase from different Italian hospitals". *Antimicrob. Agents Chemother.* 49: 104-110. 2005.

## Referencias bibliográficas

Riesenfeld, S.C., Goodman, R.M., and Handelsman, J. "Uncultured Soil Bacteria Are a Reservoir of New Antibiotic Resistance Genes". *Environ. Microbiol.* 2004, vol. 6: 981–989. 2004.

Rosser, S. J., and H. K. Young. "Identification and characterization of class 1 integrons in bacteria from an aquatic environment". *J. Antimicrob. Chemother.* 44:11–18. 1999.

Rowe-Magnus DA, Mazel D. "Resistance gene capture". *Curr Opin Microbiol.* 2:483-488. 1999.

Rowe-Magnus, D. A., A. M. Guerout, and D. Mazel. "Super-Integrans." [In eng]. *Res Microbiol* 150, no. 9-10: 641-51. 1999.

Rowe-Magnus, D. A., A. M. Guerout, P. Ploncard, B. Dychinco, J. Davies, and D. Mazel. "The Evolutionary History of Chromosomal Super-Integrans Provides an Ancestry for Multiresistant Integrans." [In eng]. *Proc Natl Acad Sci USA* 98, no. 2: 652-7. 2001.

Rowe-Magnus, D. A., and D. Mazel. "Integrans: Natural Tools for Bacterial Genome Evolution." [In eng]. *Curr Opin Microbiol* 4, no. 5: 565-9. 2001.

Stokes, H. W., and R. M. Hall. "A Novel Family of Potentially Mobile DNA Elements Encoding Site-Specific Gene-Integration Functions: Integrans." [In eng]. *Mol Microbiol* 3, no. 12: 1669-83. 1989.

Stokes, H. W., D. B. O'Gorman, G. D. Recchia, M. Parsekhian, and R. M. Hall. "Structure and Function of 59-Base Element Recombination Sites Associated with Mobile Gene *Cassettes*." [In eng]. *Mol Microbiol* 26, no. 4: 731-45. 1997.

Summers, A.O. "Genetic linkage and horizontal transfer, the roots of the antibiotic multi resistance problem". *Anim. Biotechnol.* 17: 125-135. 2006.

Sundstrom, L., P. Radstrom, G. Swedberg, and O. Skold. "Site-Specific Recombination Promotes Linkage between Trimethoprim- and Sulfonamide Resistance Genes. Sequence Characterization of *dhfrV* and *sull* and a Recombination Active Locus of Tn21." [In eng]. *Mol Gen Genet* 213, no. 2-3: 191-201. 1988.

Walsh, T.R. "Combinatorial genetic evolution of multiresistance". *Curr. Opin. Microbiol.* 9: 476-482. 2006.

Ward, J. M., and J. Grinsted. "Physical and Genetic Analysis of the Inc-W Group Plasmids R388, Sa, and R7K." [In eng]. *Plasmid* 7, no. 3: 239-50. 1982.

White, P. A., C. J. McIver, and W. D. Rawlinson. "Integrans and Gene *Cassettes* in the *Enterobacteriaceae*." [In eng]. *Antimicrob Agents Chemother* 45, no. 9: 2658-61. 2001.

# Anexos

**Anexo 1. Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 1 que incluyen al gen *cassette aadA1***

| Integrasa    | Estructura de la zona variable   | Cantidad |
|--------------|--|----------|
| <i>intI1</i> | <i>dfrA1-aadA1</i>   | 79       |
| <i>intI1</i> | <i>aadA1</i>   | 74       |
| <i>intI1</i> | <i>aacA4-catB8-aadA1</i>   | 14       |
| <i>intI1</i> | <i>bla<sub>OXA-1</sub>-aadA1</i>   | 12       |
| <i>intI1</i> | <i>aadB-aadA1</i>  | 5        |
| <i>intI1</i> | <i>catB8-aadA1</i>   | 3        |
| <i>intI1</i> | <i>estX-aadA1</i>  | 3        |
| <i>intI1</i> | <i>aacA4-aadA1-catB2</i>   | 2        |
| <i>intI1</i> | <i>aadB - aadA1 - cmlA</i>   | 2        |
| <i>intI1</i> | <i>aadB - aadA1 - tnp</i>  | 2        |
| <i>intI1</i> | <i>aadB - aadA1 - cmlA6</i>  | 2        |
| <i>intI1</i> | <i>dfrA15 - aadA1</i>  | 2        |
| <i>intI1</i> | <i>dfrA15 - aadA1 - qacEdelta1</i>   | 2        |
| <i>intI1</i> | <i>dfrB1 - aadA1</i>   | 2        |
| <i>intI1</i> | <i>aacA - aadA1 - qacEdelta1 - sul1 - ISCR1 - bla<sub>CTX-M-2</sub> - orf3/qacEdelta1</i>                  | 1        |
| <i>intI1</i> | <i>aacA17 - aadA1</i>  | 1        |
| <i>intI1</i> | <i>aacA4 - aadA1</i>   | 1        |
| <i>intI1</i> | <i>aacA4 - aadA1 - qacEdelta1 - sul1</i>   | 1        |
| <i>intI1</i> | <i>aacA4 - aadA1 - qacEdelta1 - sul1 - ISCR1 - sapC-like - sapB-like - qnrB10</i>                          | 1        |
| <i>intI1</i> | <i>aacA4 - aadA1 - qacEdelta1 - sul1 - ISCR1 - sapC-like - sapB-like - sapA-like - qnrB10 - qacEdelta1</i> | 1        |
| <i>intI1</i> | <i>aacA4m - aadA1 - qacEdelta1</i>   | 1        |
| <i>intI1</i> | <i>aacA7 - aadA1 - qacEdelta1</i>  | 1        |
| <i>intI1</i> | <i>aacC1 - aadA1</i>   | 1        |
| <i>intI1</i> | <i>aadB - aadA1 - orf</i>  | 1        |
| <i>intI1</i> | <i>aadB - aadA1 - aacA</i>   | 1        |
| <i>intI1</i> | <i>blaIMP-18 - aadA1 - qacEdelta1</i>  | 1        |
| <i>intI1</i> | <i>blaIMP-18 - aadA1 - bla<sub>OXA-2</sub> - aadA1 - qacEdelta1</i>  | 1        |
| <i>intI1</i> | <i>blaP1a - aadA1 - qacEdelta1 - sul1 - orf5 - orf6</i>  | 1        |
| <i>intI1</i> | <i>blaVIM-1 - aadA1 - qacEdelta1 - sul1 - orf5 - tniBdelta3</i>  | 1        |
| <i>intI1</i> | <i>blaVIM-2 - aadA1 - qacEdelta1</i>   | 1        |
| <i>intI1</i> | <i>catB3 - aadA1</i>   | 1        |
| <i>intI1</i> | <i>dfrA1 - aadA1 - qacEdelta1 - sul1 - ISCR1 - bla<sub>CTX-M-2</sub> - orf3 - qacEdelta1 - sul1</i>        | 1        |
| <i>intI1</i> | <i>dfrA1 - aadA1 - qacEdelta1 - sul1 - ISCR1 - bla<sub>CTX-M-2</sub> - orf3 - sul1</i>                     | 1        |
| <i>intI1</i> | <i>dfrA28 - aadA1</i>  | 1        |

Anexos

|              |   |   |
|--------------|---|---|
| <i>intI1</i> | <i>dfrA14 - aadA1</i>   | 1 |
| <i>intI1</i> | <i>dfrB1 - aadA1 - catB2</i>  | 1 |
| <i>intI1</i> | <i>dfrB3 - aadA1</i>  | 1 |
| <i>intI1</i> | <i>dfrB6 - aadA1</i>  | 1 |
| <i>intI1</i> | <i>estX - aadA1 - qacEdelta1 - sul1 - ISCR1 - bla<sub>CTX-M-2</sub> - orf3/qacEdelta1 - sul1 - orf5</i> | 1 |
| <i>intI1</i> | <i>orf5 - sul1 - qacEdelta1 - aadA1 - bla<sub>OXA-1</sub></i>   | 1 |
| <i>intI1</i> | <i>orfI - aadA1</i>   | 1 |
| <i>intI1</i> | <i>orfE-like - aadA1</i>  | 1 |
| <i>intI1</i> | <i>qacH - aadA1</i>   | 1 |
| <i>intI1</i> | <i>sat2 - aadA1</i>   | 1 |
| <i>intI1</i> | <i>dfrA1 - aadA1 - acp-like - bla<sub>TLA-2</sub></i>   | 1 |

**Anexo 2. Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 2 que incluyen al gen *cassette aadA1***

| <b>Integrasa</b> | <b>Estructura de la zona variable</b>  | <b>Cantidad</b> |
|------------------|--|-----------------|
| <i>intI2</i>     | <i>dfrA1 - sat2 - aadA1</i>  | 30              |
| <i>intI2</i>     | <i>sat 2 - aadA1</i>   | 7               |
| <i>intI2</i>     | <i>dfrA1 - sat2 - aadA1 - orfX</i>   | 6               |
| <i>intI2</i>     | <i>estX - sat2 - aadA1</i>   | 6               |
| <i>intI2</i>     | <i>dfrA1 - aadA1</i>   | 3               |
| <i>intI2</i>     | <i>dfrA1 - sat2 - aadA1 - ybeA</i>   | 3               |
| <i>intI2</i>     | <i>dfrA1 - sat2 - aadA1 - ybeA - ybfA - ybfB - ybgA</i>                          | 2               |
| <i>intI2</i>     | <i>aadA1 - orfX</i>  | 1               |
| <i>intI2</i>     | <i>aadA1 - sat2 - dfrA1</i>  | 1               |
| <i>intI2</i>     | <i>dfrA1 - catB2 - sat2 - aadA1</i>  | 1               |
| <i>intI2</i>     | <i>dfrA1 - sat2 - aadA1 - qacEdelta1 - sul1</i>                                  | 1               |
| <i>intI2</i>     | <i>dfrA1 - sat2 - aadA1 - ybeA - ybfB - ybgA</i>                                 | 1               |
| <i>intI2</i>     | <i>dfrA1 - sat2 - aadB - catB2 - dfrA1 - bla<sub>CARB-4</sub> - aadA1 - ybeA</i> | 1               |
| <i>intI2</i>     | <i>dfrA1 - sat2 - IS911 - aadA1 - orfX</i>                                       | 1               |
| <i>intI2</i>     | <i>dfrA1:: - IS911 - ::dfrA1 - sat2 - aadA1</i>                                  | 1               |
| <i>intI2</i>     | <i>insB - insA - dfrA1 - sat2 - aadA1</i>  | 1               |
| <i>intI2</i>     | <i>IS1 - sat2 - aadA1 - orfX</i>   | 1               |
| <i>intI2</i>     | <i>sat2 - aadA1 - orfX</i>   | 1               |
| <i>intI2</i>     | <i>sat2 - aadA1 - ybeA</i>   | 1               |
| <i>intI2</i>     | <i>sat2 - aadB - catB2 - dfrA1 - sat2 - aadA1</i>                                | 1               |
| <i>intI2</i>     | <i>sat2 - ereA1 - aadA1 - orfX</i>   | 1               |

**Anexo 3. Arreglos más frecuentes de genes cassettes de la zona variable de integrones de clase 1 que incluyen al gen cassette *aadA2***

| Integrasa    | Estructura de la zona variable   | Cantidad |
|--------------|--|----------|
| <i>intI1</i> | <i>aadA2</i>   | 70       |
| <i>intI1</i> | <i>dfrA12-orfF-aadA2</i>   | 53       |
| <i>intI1</i> | <i>aadB-aadA2</i>  | 6        |
| <i>intI1</i> | <i>sat-<i>psp</i>-aadA2</i>  | 4        |
| <i>intI1</i> | <i>aacA4-aadA2</i>   | 3        |
| <i>intI1</i> | <i>bla<sub>IMP-9</sub>-aacA4 -bla<sub>OXA-10</sub>-aadA2</i>   | 3        |
| <i>intI1</i> | <i>dfr12 -orfF -aadA2</i>  | 3        |
| <i>intI1</i> | <i>dfrA12 -aadA2</i>   | 3        |
| <i>intI1</i> | <i>dfrA12 -orf -aadA2</i>  | 3        |
| <i>intI1</i> | <i>dfrA12 -orfF -aadA2 -orf -orf</i>   | 3        |
| <i>intI1</i> | <i>estX -<i>psp</i> -aadA2 -tnp1294 -aadA1 -qacH -tnp440 (IS440) -sul3</i>                             | 3        |
| <i>intI1</i> | <i>aacA4 -bla<sub>OXA-4</sub> -aadA2</i>   | 2        |
| <i>intI1</i> | <i>aadA2 -orfF -dfrA12</i>   | 2        |
| <i>intI1</i> | <i>aadA2 -qacEdelta -bla<sub>PSE-1</sub> -qacEdelta1 -sul1</i>   | 2        |
| <i>intI1</i> | <i>aadA2 -qacEdelta1 -sul1</i>   | 2        |
| <i>intI1</i> | <i>bla<sub>PSE-1</sub> -aadA2 -qacEdelta1</i>  | 2        |
| <i>intI1</i> | <i>dfrA12 -aadA2 -qacEdelta1</i>   | 2        |
| <i>intI1</i> | <i>dfrA12 -orfF -aadA2 -qacEdelta1</i>   | 2        |
| <i>intI1</i> | <i>dfrA12 -orfF -aadA2 -qacEdelta1 -sul1 -ISCR1 -bla<sub>CTX-M-2</sub> -orf3/qacEdelta1 -sul1</i>      | 2        |
| <i>intI1</i> | <i>orf -ereB -orf2 -bla<sub>PSE-1</sub> -aadA2 -qacEdelta1 -sul1 -orf5</i>                             | 2        |
| <i>intI1</i> | <i>5'CS   aacA4'-3   bla<sub>IMP-6</sub>   aadA2   tnpA (IS91 family) -3'CS   qacEA1   sul1   orf5</i> | 1        |
| <i>intI1</i> | <i>5'CS   aadB   aadA2 -3'CS</i>   | 1        |
| <i>intI1</i> | <i>5'CS   dfrA12   gcuF   aadA2</i>  | 1        |
| <i>intI1</i> | <i>5'CS   dfrA12   gcuF   aadA2 -3'CS   qacED1   sul1</i>  | 1        |
| <i>intI1</i> | <i>5'CS   dfrA12   gcuF   aadA2 -3'CS   qacED1   sul1   orf5D</i>                                      | 1        |
| <i>intI1</i> | <i>5'CS   dfrA12   gcuF   aadA2 -3'CS   qacED1   sul1   orf5delta</i>                                  | 1        |
| <i>intI1</i> | <i>5'CS   dfrA12   gcuF   aadA2 -3'CS   qacEdelta1   sul1</i>  | 1        |
| <i>intI1</i> | <i>aacA -aadA2 -qacEdelta1 -sul1 -ISCR1 -floR -orf -qacEdelta1 -sul1</i>                               | 1        |
| <i>intI1</i> | <i>aacA27 -aadA2 -qacEdelta1 -sul1 -ISCR1 -qnrA1 -ampR</i>   | 1        |
| <i>intI1</i> | <i>aacA3 -catB2 -aadA2</i>   | 1        |
| <i>intI1</i> | <i>aacA3 -aadA2</i>  | 1        |
| <i>intI1</i> | <i>aacA4 -aacC1 -orf -aadA2 -cmlA1 -qacEdelta1 -sul1</i>   | 1        |
| <i>intI1</i> | <i>aacA4 -bla<sub>OXA-4</sub> -aadA2 -qacEdelta1 -sul1 -ISCR1 -qnr -dfrA3b -orf105 -sul1</i>           | 1        |
| <i>intI1</i> | <i>aacA4 -bla<sub>P1b</sub> -aadA2 -qacEdelta</i>  | 1        |
| <i>intI1</i> | <i>aacA4 -blaP1b -aadA2</i>  | 1        |

Anexos

|              |   |   |
|--------------|---|---|
| <i>int11</i> | <i>aadA1 -cmlA1 -aadA2 -sch_054 -sat -int11</i>   | 1 |
| <i>int11</i> | <i>aadA2 -cmlA -aadA1 -qacEdelta1 -sul1</i>   | 1 |
| <i>int11</i> | <i>aadA2 -cmlA1 -orf4</i>   | 1 |
| <i>int11</i> | <i>aadA2 -dfrA1</i>   | 1 |
| <i>int11</i> | <i>aadA2 -dfrA12</i>  | 1 |
| <i>int11</i> | <i>aadA2 -linF -qacEdelta1 -sul1</i>  | 1 |
| <i>int11</i> | <i>aadA2 -linF -aacA4-CR -qacEdelta1</i>  | 1 |
| <i>int11</i> | <i>aadA2 -linG</i>  | 1 |
| <i>int11</i> | <i>aadA2 -lnuF</i>  | 1 |
| <i>int11</i> | <i>aadA2 -orf1 -dfrA12</i>  | 1 |
| <i>int11</i> | <i>aadA2 -qacEdelta1 -sul1 -ISCR1 -bla<sub>CTX-M-2</sub> -orf3/qacEdelta1</i>               | 1 |
| <i>int11</i> | <i>aadA2 -qacEdelta1 -sul1 -ISCR1 -catA2 -qacEdelta1 -sul1 -orf5 -orf6</i>                  | 1 |
| <i>int11</i> | <i>aadA2 -qacEdelta1 -sul1 -orf341 -bla<sub>DHA-1</sub> -ampR -qacEdelta1 -sul1</i>         | 1 |
| <i>int11</i> | <i>aadA2 -qacEdelta1 -sul1 -ISCR1 -sapA-like -qnrB2 -qacEdelta1 -sul1 -ISCR1 -dfrA19</i>    | 1 |
| <i>int11</i> | <i>aadA2 -qacEdelta1 -sul1 -orf5</i>  | 1 |
| <i>int11</i> | <i>aadA2 -qnrVC1</i>  | 1 |
| <i>int11</i> | <i>aadA2 -sul1</i>  | 1 |
| <i>int11</i> | <i>aadA2 -sat</i>   | 1 |
| <i>int11</i> | <i>aadA2 -sul1 -orf -mphR -mrX -mphA</i>  | 1 |
| <i>int11</i> | <i>aadB -cmlA5 -aadA2 -strA -strB -tnpM -Δint11 -bla<sub>GES-12</sub> -aacA4 -dfrA7</i>     | 1 |
| <i>int11</i> | <i>aadB -aadA2 -qacEdelta1 -sul1 -ISCR1 -ampC -ampR -sul1 -orf5</i>                         | 1 |
| <i>int11</i> | <i>bla<sub>GES-1</sub> -aacA4 -dfrA15b -cmlA4 -aadA2</i>                                    | 1 |
| <i>int11</i> | <i>bla<sub>GES-1</sub> -aacA4 -dfrA15b -cmlA4 -aadA2</i>                                    | 1 |
| <i>int11</i> | <i>bla<sub>OXA-31</sub> -aadA2 -cmlA1</i>   | 1 |
| <i>int11</i> | <i>bla<sub>PSE-1</sub> -aadA2 -cmlA1 -aadA1</i>   | 1 |
| <i>int11</i> | <i>blaP1 -aadA2 -cmlA1 -catB11 -qacEdelta1 -sul1 -ISCR1 -qnrA1 -sul1 -ISCR1 -dfrA19</i>     | 1 |
| <i>int11</i> | <i>blaP1 -aadA2 -cmlA1 -catB11 -qacEdelta1 -sul1 -ISCR1 -qnrA1 -sul1 -ISCR1 -tnpA -tnpA</i> | 1 |
| <i>int11</i> | <i>bla<sub>P1b</sub> -aadA2</i>   | 1 |
| <i>int11</i> | <i>bla<sub>VIM-2</sub> -catB8 -aacA4 -aadA2 -qacEdelta1</i>                                 | 1 |
| <i>int11</i> | <i>bla<sub>VIM-1</sub> -aacA4 -orfPa85 -blaPSE-1 -aadA2 -qacEdelta1 -sul1</i>               | 1 |
| <i>int11</i> | <i>bla<sub>VIM-2</sub> -aacA4 -blaP1b -aadA2</i>  | 1 |
| <i>int11</i> | <i>bla<sub>VIM-2</sub> -aacA4 -aadA2 -qacEdelta1</i>  | 1 |
| <i>int11</i> | <i>catB3 -aadA2</i>   | 1 |
| <i>int11</i> | <i>cmlA1 -aadA2 -qacEdelta1 -sul1 -orf5 -orf6</i>   | 1 |
| <i>int11</i> | <i>dfrA12 -aadA2 -qacEdelta1 -sul1</i>  | 1 |
| <i>int11</i> | <i>dfrA12 -orf -aadA2 -intron -intron</i>   | 1 |
| <i>int11</i> | <i>dfrA12 -orf -aadA2 -qacEdelta1 -sul1 -orf -qnrA1</i>                                     | 1 |
| <i>int11</i> | <i>dfrA12 -orf -aadA2 -qacEdelta1 -sul1 -ISCR1 -bla<sub>CTX-M-14</sub> -tnp</i>             | 1 |



Anexos

|                          |   |   |
|--------------------------|---|---|
| <i>intI1</i>             | <i>dfrA12 -orfF -aadA2 -aadA2 -cmlA1 -aadA1 -qacH -tnp440 (IS440) -sul3</i>                           | 1 |
| <i>intI1</i>             | <i>dfrA12 -orfF -aadA2 -cmlA -aadA1 -qacH -tnp -sul3</i>  | 1 |
| <i>intI1</i>             | <i>dfrA12 -orfF -aadA2 -cmlA1 -aadA1 -qacH -tnp440 (IS440) -sul3</i>                                  | 1 |
| <i>intI1</i>             | <i>dfrA12 -orfF -aadA2 -cmlA1 -aadA1 -qacH -tnp440 (IS440) -tnpA (IS10) -sul3</i>                     | 1 |
| <i>intI1</i>             | <i>dfrA12 -orfF -aadA2 -qacEdelta1 -sul1 -orf5</i>  | 1 |
| <i>intI1</i>             | <i>dfrA12 -orfF -aadA2 -qacEdelta1 -sul1 -orf5 -IS6100 -ISCR1 -bla<sub>CTX-M-14</sub> -IS903</i>      | 1 |
| <i>intI1</i>             | <i>dfrA12 -orfF -aadA21 -qacH -tnp440 (IS440) -sul3</i>   | 1 |
| <i>intI1</i>             | <i>dfrA12 -orfF -aadA2 -qacEdelta1 -sul1 -ISCR1 -bla<sub>CMY-11</sub> -yqgF -gshB -orf -orf -sul1</i> | 1 |
| <i>intI1</i>             | <i>dfrA12 -tnp -aadA2</i>   | 1 |
| <i>intI1</i>             | <i>dfrA15 -aadA2 -qacEdelta1 -sul1 -ISCR1 -qnrA1 -ampR</i>  | 1 |
| <i>intI1</i>             | <i>dfrA16 -aadA2</i>  | 1 |
| <i>intI1</i>             | <i>dfrA15b -cmlA4 -aadA2 -qacEdelta1 -sul1 -ISCR1 -bla<sub>CTX-M-59</sub> -orf3/qacEdelta1</i>        | 1 |
| <i>intI1</i>             | <i>dfrA15b -cmlA4 -delta aadA2</i>  | 1 |
| <i>intI1</i>             | <i>dfrA16 -aadA2 -qacEdelta1 -sul1 -ISCR1 -bla<sub>CTX-M-9</sub> -orf1005 -qacEdelta1 -sul1</i>       | 1 |
| <i>intI1</i>             | <i>dfrA16 -aadA2</i>  | 1 |
| <i>intI1</i>             | <i>dfrA16 -aadA2 -ISCR1 -qnr -dfrA3b -orf105 -sul1</i>  | 1 |
| <i>intI1</i>             | <i>dfrA16 -blaP1b -aadA2 -ereA1</i>   | 1 |
| <i>intI1</i>             | <i>dfrA16 -aadA2 -qacEdelta1 -sul1 -ISCR1 -qnr -ampR -sul1 -orf5 -orf6</i>                            | 1 |
| <i>intI1</i>             | <i>dfrA16 -bla<sub>pSE-1</sub> -aadA2 -ereA2</i>  | 1 |
| <i>intI1</i>             | <i>dfrA27 -aadA2 -qacEdelta1 -sul1 -ISCR1 -qnrA1 -sul1 -ISCR1</i>                                     | 1 |
| <i>intI1</i>             | <i>dfrA27 -aadA2</i>  | 1 |
| <i>intI1</i>             | <i>dfrB7 -aadA2</i>   | 1 |
| <i>intI1</i>             | <i>estX -orf -aadA2</i>   | 1 |
| <i>intI1</i>             | <i>estX -psp -aadA2 -cmlA1 -aadA1 -qacH -tnpA IS26 -sul3</i>  | 1 |
| <i>intI1</i>             | <i>estX -psp -aadA2 -cmlA -aadA1 -qacH -tnp -sul3</i>   | 1 |
| <i>intI1</i>             | <i>msrB -msrA -orf -aadA2 -qacE -sul1 -orf -tniB -delta1</i>  | 1 |
| <i>intI1</i>             | <i>orf1005 -aadA2delta -qacEdelta1</i>  | 1 |
| <i>intI1 (truncated)</i> | <i>aadA2</i>  | 1 |
| <i>intI1</i>             | <i>aadA2</i>  | 1 |
| <i>intI1</i>             | <i>aadA2 - qacEdelta1 - sul1 - ISCR1 - catA2</i>  | 1 |
| <i>intI1</i>             | <i>catB2 - aac - aadA2</i>  | 1 |
| <i>intI1</i>             | <i>orf64 -aadA2 -qacE</i>   | 1 |
| <i>intI1</i>             | <i>qacH -aadB -cmlA10 -aadA2 -qacEdelta1 -sul1</i>  | 1 |
| <i>intI1</i>             | <i>qacH -dfrA32 -ereA -aadA2 -cmlA -aadA1</i>   | 1 |
| <i>intI1</i>             | <i>qacH -dfrB6 -aadA2 -qacEdelta1</i>   | 1 |
| <i>intI1</i>             | <i>sat -psp -aadA2 -sul1 -orf5</i>  | 1 |
| <i>intI1</i>             | <i>sat -psp -aadA2 -cmlA1 -tnpA -qacH -tnp440 (IS440) -sul3</i>                                       | 1 |

Anexos

|                   |  |   |
|-------------------|--|---|
| <i>int11</i>      | <i>sat -psp -aadA2 -cmlA1 -aadA1 -qacH -tnp440 (IS440) -sul3 -orf1 -mefB</i>                                     | 1 |
| <i>int11</i>      | <i>sat -psp -aadA2 -cmlA -aadA1 -qacH -tnp -sul3</i>   | 1 |
| <i>int11</i>      | <i>sat -psp -aadA2 -cmlA1 -aadA2 -qacH -tnp440 (IS440) -sul3</i>   | 1 |
| <i>int11</i>      | <i>sul1 -qacEdelta1 -aadA2</i>   | 1 |
| <i>int11</i>      | <i>dfrA12D6 - gcuFD3 - aadA2</i>   | 1 |
| <i>int11</i>      | <i>-dfrA12 - aadA2</i>   | 1 |
| <i>int11</i>      | <i>-bla<sub>VIM-1</sub>-aacA7 --dfrA1-aadA1-qacEdelta1-sul1 - dfrA12 - orfF - aadA2</i>                          | 1 |
| <i>int11delta</i> | <i>bla<sub>VIM-1</sub>-aacA7 - dfrA1 - aadA1 - qacEdelta1 - sul1 - dfrA12 - orfF - aadA2 - qacEdelta1 - sul1</i> | 1 |

**Anexo 4. Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 1 que incluyen al gen *cassette aadA5***

| Integrasa    | Estructura de la zona variable  | Cantidad |
|--------------|---|----------|
| <i>int11</i> | <i>dfrA17-aadA5</i>   | 100      |
| <i>int11</i> | <i>aadA5</i>  | 9        |
| <i>int11</i> | <i>dfrA1-aadA5</i>  | 3        |
| <i>int11</i> | <i>bla<sub>IMP1</sub>-aadA5</i>   | 2        |
| <i>int11</i> | <i>bla<sub>OXA1</sub>-aadA5</i>   | 1        |
| <i>int11</i> | <i>5'CS - catB8j-aacA4' - aadA5 - 3'CS</i>  | 1        |
| <i>int11</i> | <i>5'CS - bla<sub>IMP-1</sub> - aacA3 - aadA5 - 3'CS</i>  | 1        |
| <i>int11</i> | <i>5'CS - dfrA17g - aadA5 - 3'CS</i>  | 1        |
| <i>int1</i>  | <i>5'CS - dfrA17h - aadA5</i>   | 1        |
| <i>int11</i> | <i>aacA4 - bla<sub>BEL-1</sub> - smr2 - aadA5</i>   | 1        |
| <i>int11</i> | <i>aadA5 - qacE1</i>  | 1        |
| <i>int11</i> | <i>bla<sub>IMP-1</sub> - aacA4 - aadA5 - qacEdelta1</i>   | 1        |
| <i>int11</i> | <i>bla<sub>IMP-8</sub> - aacA3 - aadA5</i>  | 1        |
| <i>int11</i> | <i>bla<sub>OXA-1</sub> - orf1 - aadA5</i>   | 1        |
| <i>int11</i> | <i>bla<sub>OXA-31</sub> - aadA5</i>   | 1        |
| <i>int11</i> | <i>chrA - sul1 - qacEdelta1 - orf5 - aadA5 - dfrA17</i>   | 1        |
| <i>int11</i> | <i>dfrA1 - catB3 - aadA5</i>  | 1        |
| <i>int11</i> | <i>dfrA1 - orf1 - aadA5 - qacEdelta1</i>  | 1        |
| <i>int11</i> | <i>dfrA17 - aadA5 - qacEdelta1 - sul1 - ISCR1 - bla<sub>CTX-M-1</sub> - orf3 - orf1005 - qacEdelta1</i> | 1        |
| <i>int11</i> | <i>dfrA17a - aadA5</i>  | 1        |
| <i>int11</i> | <i>dfrA5 - aadA5 - sul1</i>   | 1        |
| <i>int11</i> | <i>orf - aadA5</i>  | 1        |
| <i>int11</i> | <i>folA - catB3 - aadA5 - sul1 - orf5</i>   | 1        |
| <i>int11</i> | <i>qacF - aadA5 - bla<sub>OXA</sub></i>   | 1        |
| <i>int11</i> | <i>aacC11 - aadA5</i>   | 1        |
| <i>int11</i> | <i>-qacF - aadA5 - bla<sub>OXA</sub></i>  | 1        |

Anexos

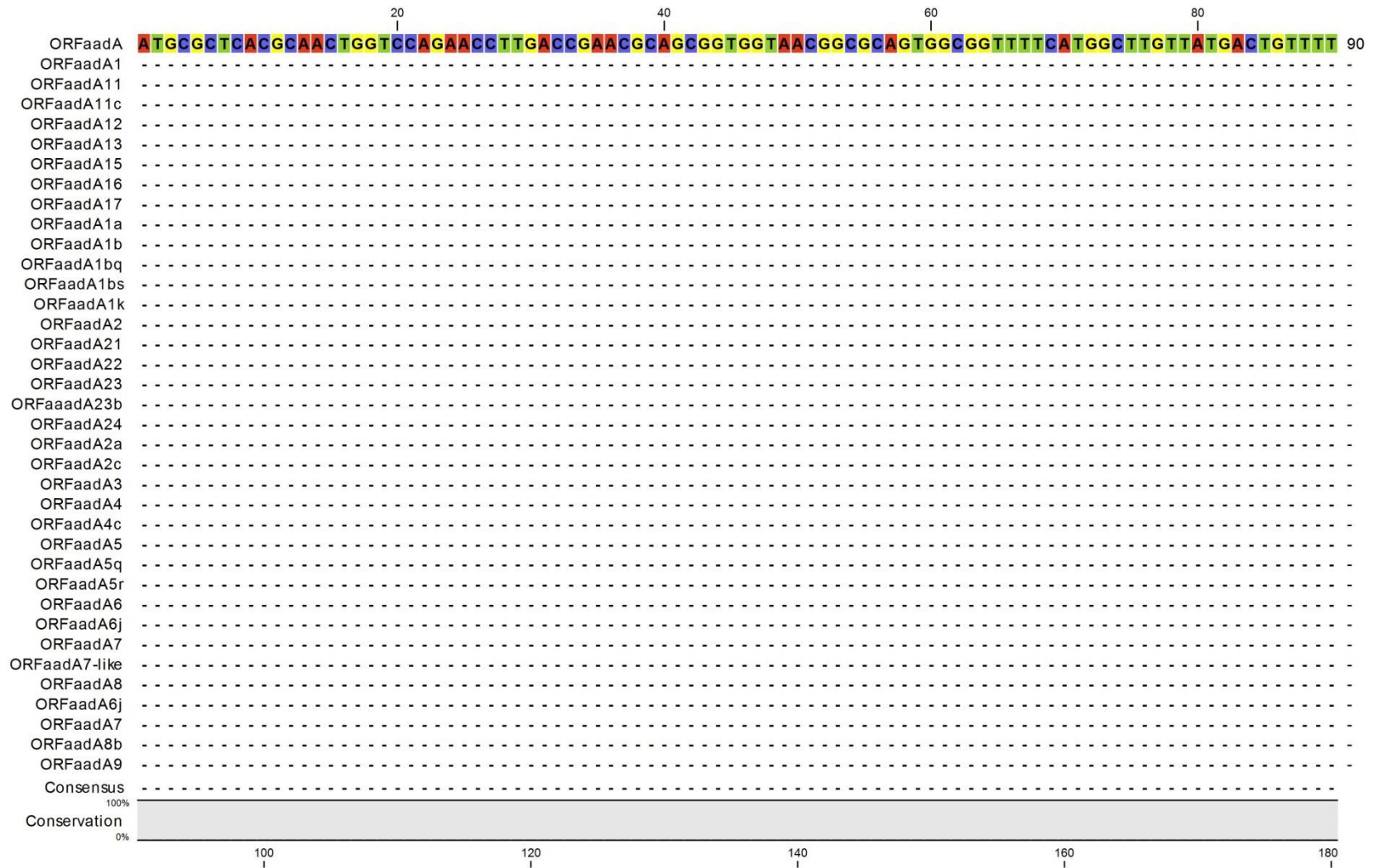
**Anexo 5. Arreglos más frecuentes de genes *cassettes* de la zona variable de integrones de clase 1 que incluyen a los genes *cassettes aadA6, aadA7, aadA1a, aadA1b, aadA11* y *aadA4***

| Integrasa         | Estructura de la zona variable  | Cantidad |
|-------------------|---|----------|
| <i>intI1</i>      | <i>aadA6-orfD</i>   | 7        |
| <i>intI1</i>      | <i>aacA7-aadA6-orfD</i>   | 2        |
| <i>intI1</i>      | <i>aacC5-aadA7</i>  | 6        |
| <i>intI1</i>      | <i>aadA7</i>  | 6        |
| <i>intI1</i>      | <i>aadA1a</i>   | 4        |
| <i>intI1</i>      | <i>estX-aadA1a</i>  | 2        |
| <i>intI1</i>      | <i>aadA1b</i>   | 2        |
| <i>intI1</i>      | <i>aadA11</i>   | 4        |
| <i>intI1</i>      | <i>aadA4</i>  | 5        |
| <i>intI1</i>      | <i>aacA7 - aadA6</i>  | 1        |
| <i>intI1</i>      | <i>aacA7 - aadA6 - orfD</i>   | 1        |
| <i>intI1</i>      | <i>aacA8 - bla<sub>OXA-2</sub> - aacA7 - aadA6 - orfD - qacEdelta1 - sul1 - orf5</i>  | 1        |
| <i>intI1</i>      | <i>aadA6 - gcuD</i>   | 1        |
| <i>intI1</i>      | <i>aadB - bla<sub>VIM-2</sub> - aadA6/aadA10</i>  | 1        |
| <i>intI1</i>      | <i>aadB - aadA6</i>   | 1        |
| <i>intI1</i>      | <i>bla<sub>VIM-4</sub> - aacA7 - aadA6 - orfD</i>   | 1        |
| <i>intI1</i>      | <i>bla<sub>VIM-4</sub> - aadA6 - orfD - qacEdelta1</i>  | 1        |
| <i>intI1</i>      | <i>bla<sub>IMP-15</sub> - aadA6 - orfD - qacEdelta1</i>   | 1        |
| <i>intI1</i>      | <i>sul1 - qacEdelta1 - bla<sub>VEB-1</sub> - aphA6 - ISCR1 - sul1 - qacEdelta1 - orfD - aadA6</i>   | 1        |
| <i>intI1</i>      | <i>orfD - aadA6</i>   | 1        |
| <i>intI1</i>      | <i>aacA4 - bla<sub>OXA-56</sub> - aadA7 - orf1</i>  | 1        |
| <i>intI1</i>      | <i>aacC5 - aadA7 - qacEdelta1 - sul1</i>  | 1        |
| <i>intI1</i>      | <i>aacC5 - aadA7 - qacEdelta1 - sul1delta</i>   | 1        |
| <i>intI1</i>      | <i>aadA7 - aadA7</i>  | 1        |
| <i>intI1</i>      | <i>orf102 - aacA4 - bla<sub>OXA-56</sub> - aadA7 - tnpA - qacEdelta1 - sul1 - ISCR14 - rmtD - orfA - groEL - ISCR14 - sul1 - orf - cmxA</i> | 1        |
| <i>intI1</i>      | <i>aacA4 - aacC1 - orfX - orfX' - aadA1a</i>  | 1        |
| <i>intI1</i>      | <i>aacA4 - aacC1 - orfX - orfX - aadA1a - qacEdelta1 - sul1 - orf5 - orf6</i>   | 1        |
| <i>intI1</i>      | <i>aacA4 - aacC1 - orfX - orfY - aadA1a - qacEdelta1</i>  | 1        |
| <i>intI1</i>      | <i>aacC1 - orfX - orfX - aadA1a</i>   | 1        |
| <i>intI1</i>      | <i>dfrA1 - aadA1a</i>   | 1        |
| <i>intI1</i>      | <i>bla<sub>P1a</sub> - aadA1a - qacEdelta1 - sul1 - orf5 - orf6</i>   | 1        |
| <i>intI1</i>      | <i>sul1 - qacEdelta1 - aadA1a - orfG - aacA1</i>  | 1        |
| <i>intI1</i>      | <i>estX - smr2 - aadA1a</i>   | 1        |
| <i>intI1</i>      | <i>orf1 - cat - aadA1a</i>  | 1        |
| <i>intI1</i>      | <i>5'CS - bla<sub>IMP-18</sub> - aadA1b - bla<sub>OXA-224</sub> - 3'CS - qacED1</i>   | 1        |
| <i>intI1</i>      | <i>dfrA1 - gcu39 - aadA1b - catB2 - catB2</i>   | 1        |
| <i>intI1delta</i> | <i>aadA1bq</i>  | 1        |

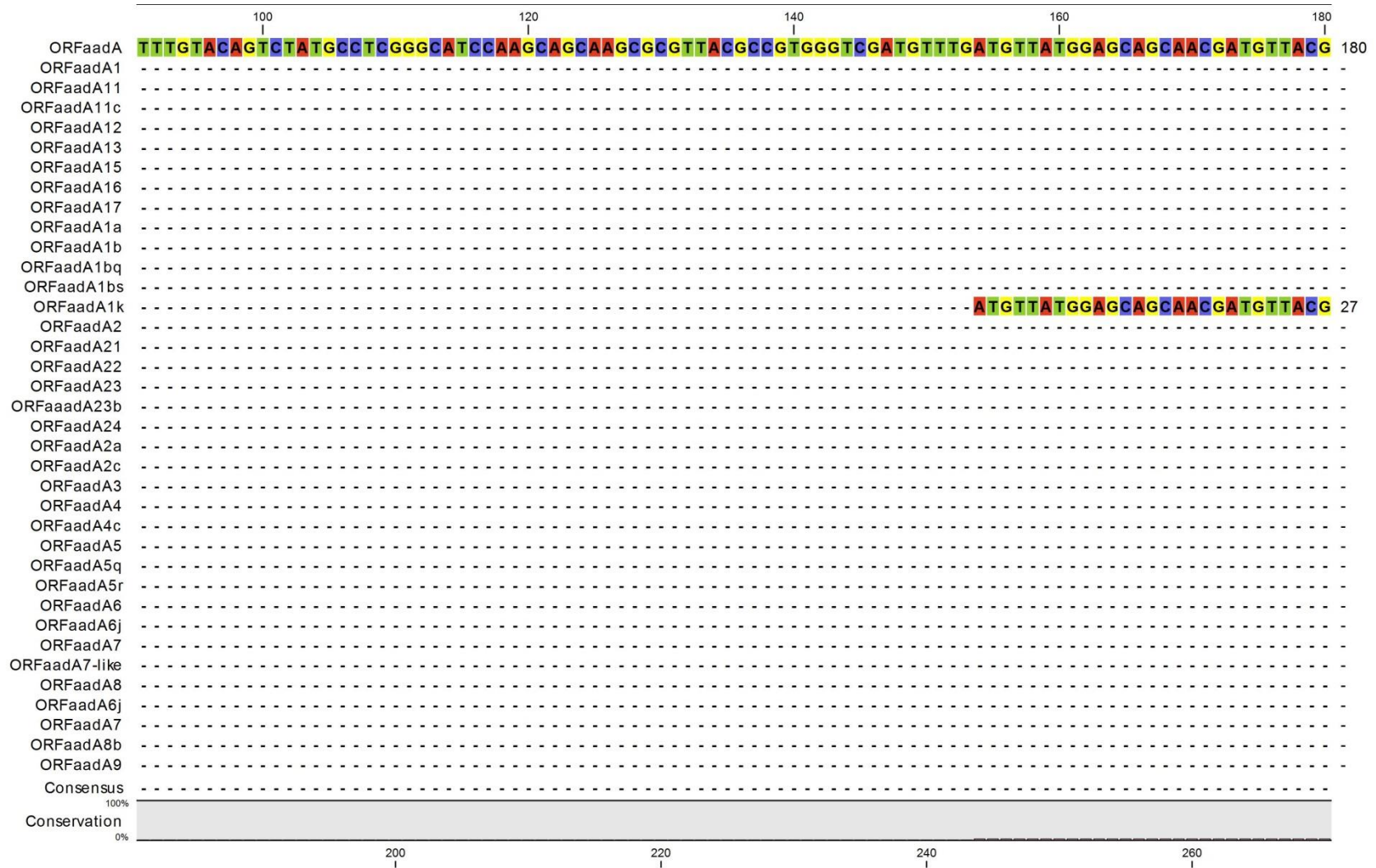
Anexos

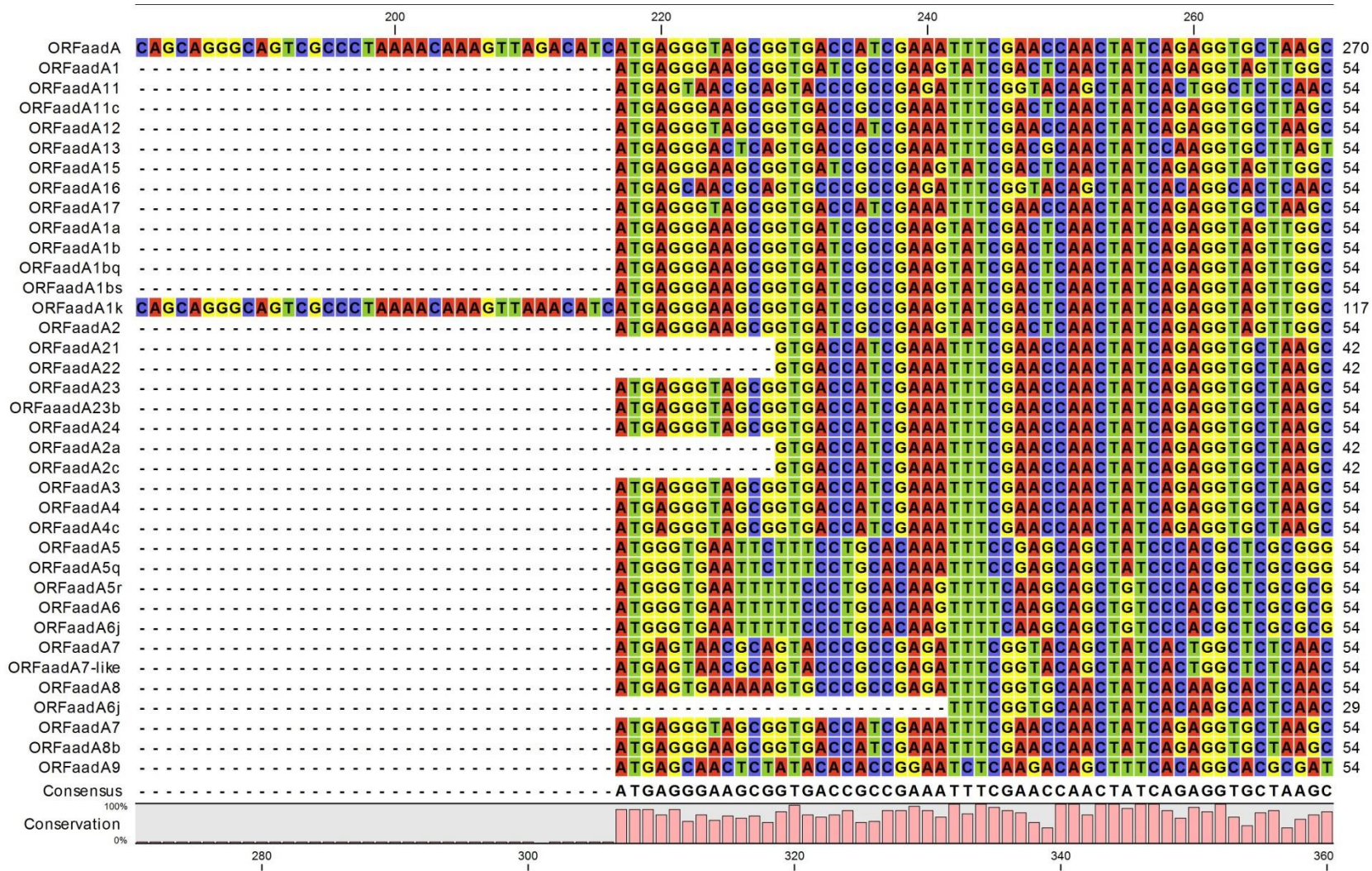
|                    |  |   |
|--------------------|--|---|
| <i>(truncated)</i> |  |   |
| <i>intl1</i>       | <i>aadA11 - qacEdelta1 - sul1</i>  | 1 |
| <i>intl1</i>       | <i>aadB - orf - aadA11 - qacEdelta1 - sul1 - orf5</i>  | 1 |
| <i>intl1</i>       | <i>aadB - orf - aadA11 - qacEdelta1</i>  | 1 |
| <i>intl1</i>       | <i>orf - orf - aadA11 - tniC</i>   | 1 |
| <i>intl1</i>       | <i>dfrB1 - aadA11 - qacEdelta1</i>   | 1 |
| <i>intl1</i>       | <i>dfrA - aadA11 - qacEdelta1 - sul1 - orf5</i>  | 1 |
| <i>intl1</i>       | <i>aadB - orf - aadA11 - qacEdelta1</i>  | 1 |
| <i>intl1</i>       | <i>orf114 - orf114 - aadA11</i>  | 1 |
| <i>intl1</i>       | <i>aadB - aadA4</i>  | 1 |
| <i>intl1</i>       | <i>aacA7 - aadB - aadA4 - orf - bla<sub>OXA-2</sub> - orfD</i>   | 1 |
| <i>intl1</i>       | <i>dfrA13 - aadA4 - tnp</i>  | 1 |
| <i>intl1</i>       | <i>bla<sub>OXA-2</sub> - aadA4 - qacEdelta1 - sul1 - orf5</i>  | 1 |
| <i>intl1</i>       | <i>aadA4 - aacA1 - orfG - orfH - orfI - orfJ - orfK - orfL - orfM - catB2 - qacEdelta1 - sul1 - orf5</i> | 1 |
| <i>intl1</i>       | <i>aadA4 - qacE</i>  | 1 |

**Anexo 6. Alineamiento por ClustalW para los ORFs de los distintos alelos *aadA-like***

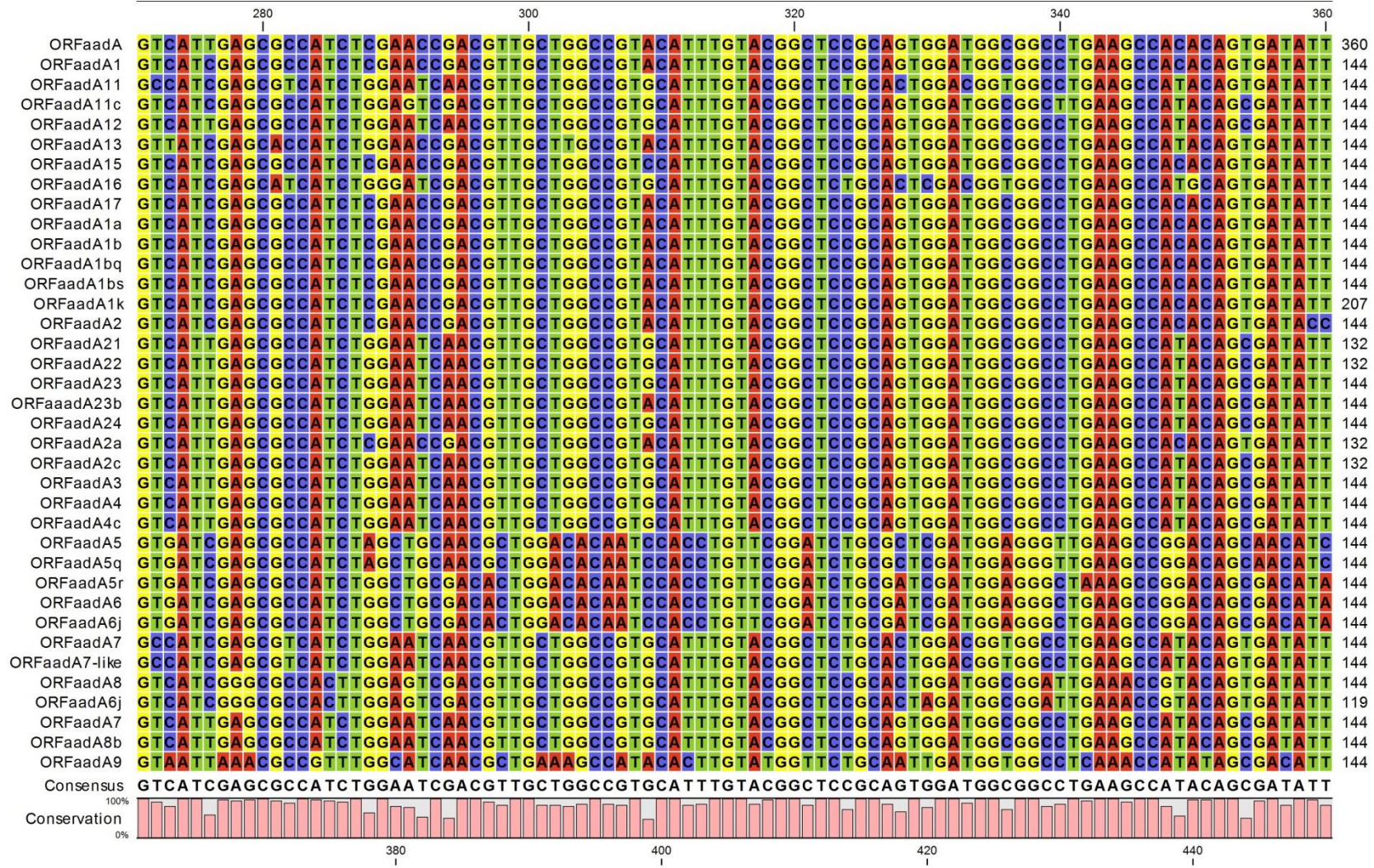


Anexos





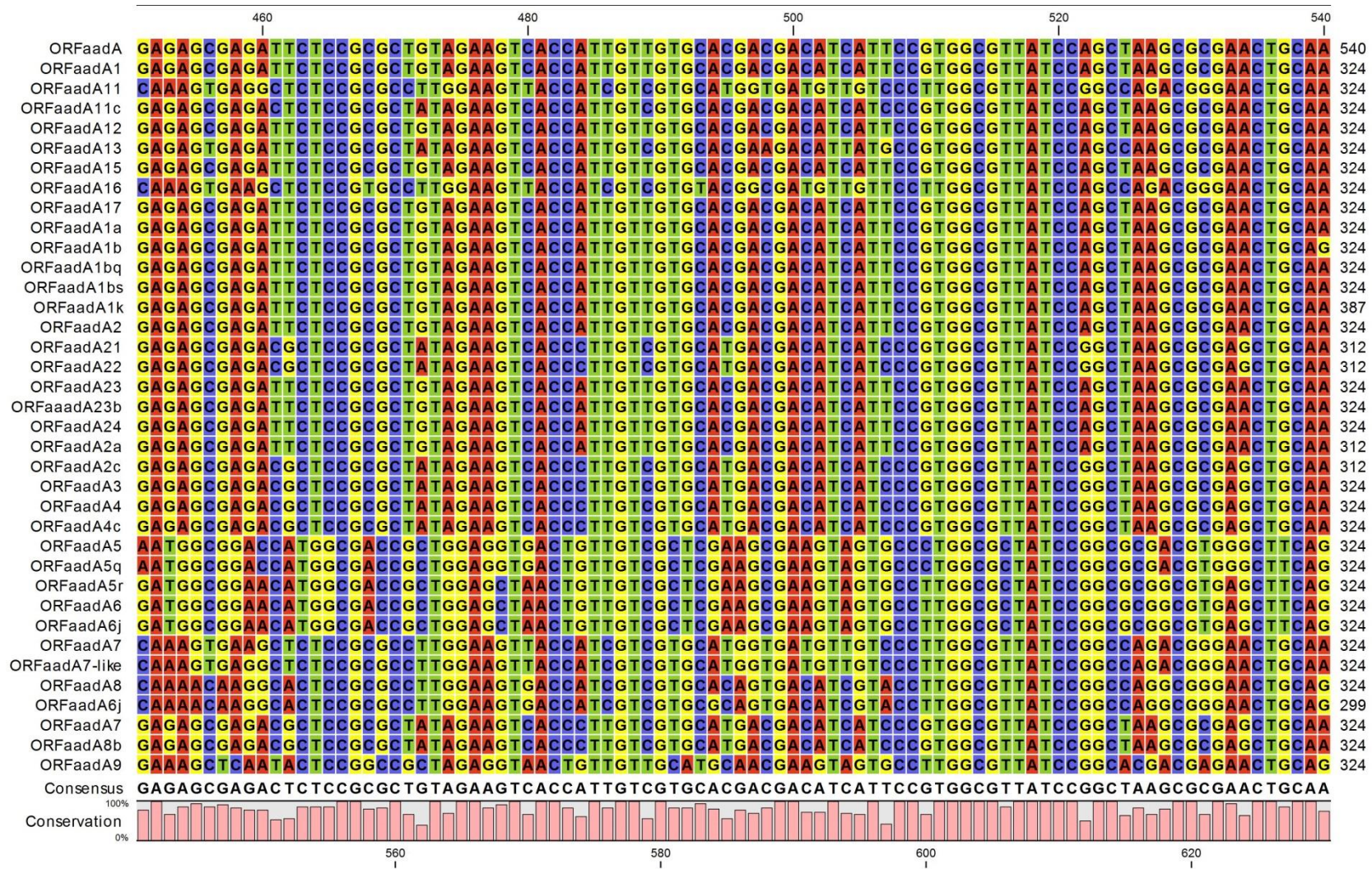
Anexos

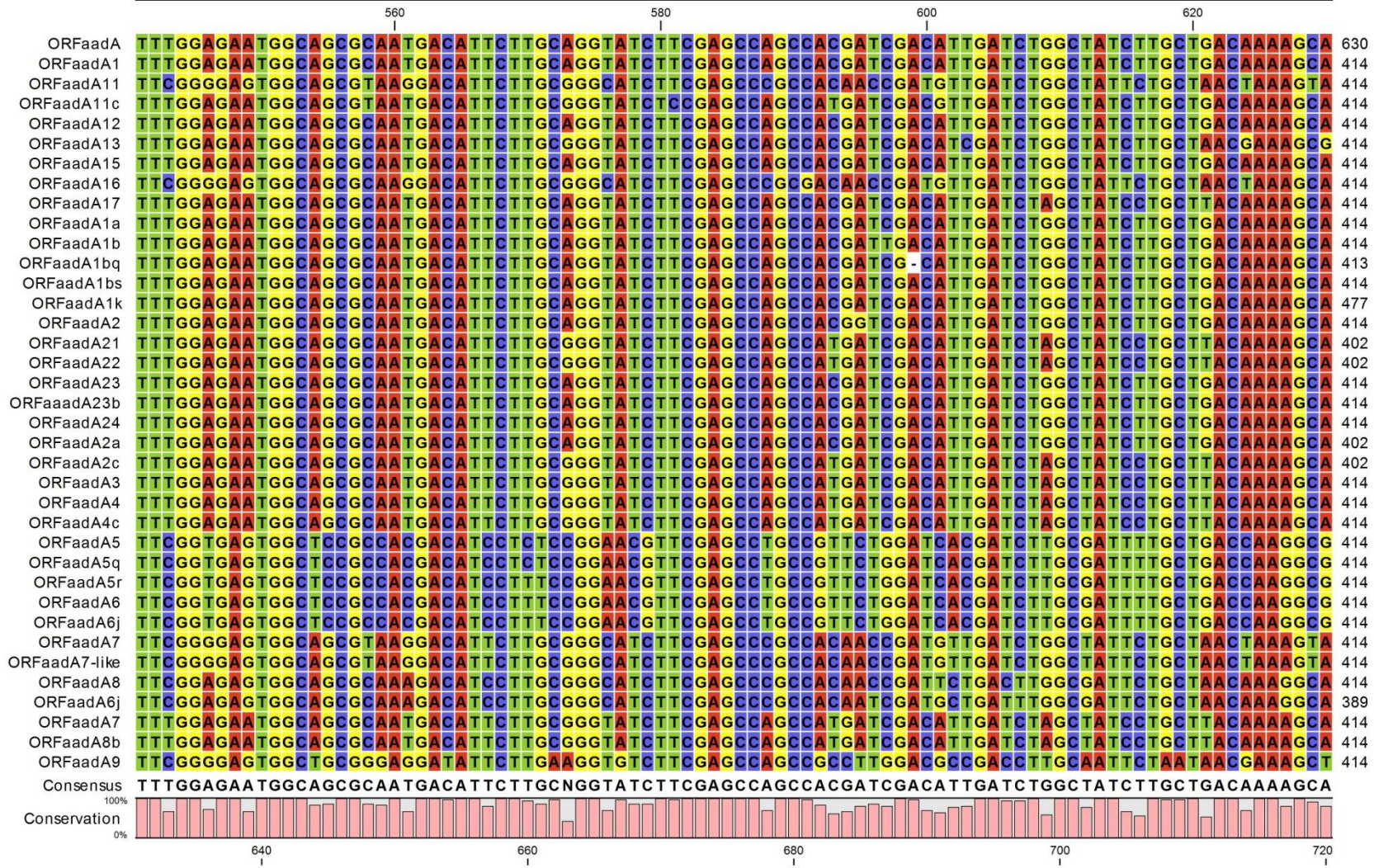


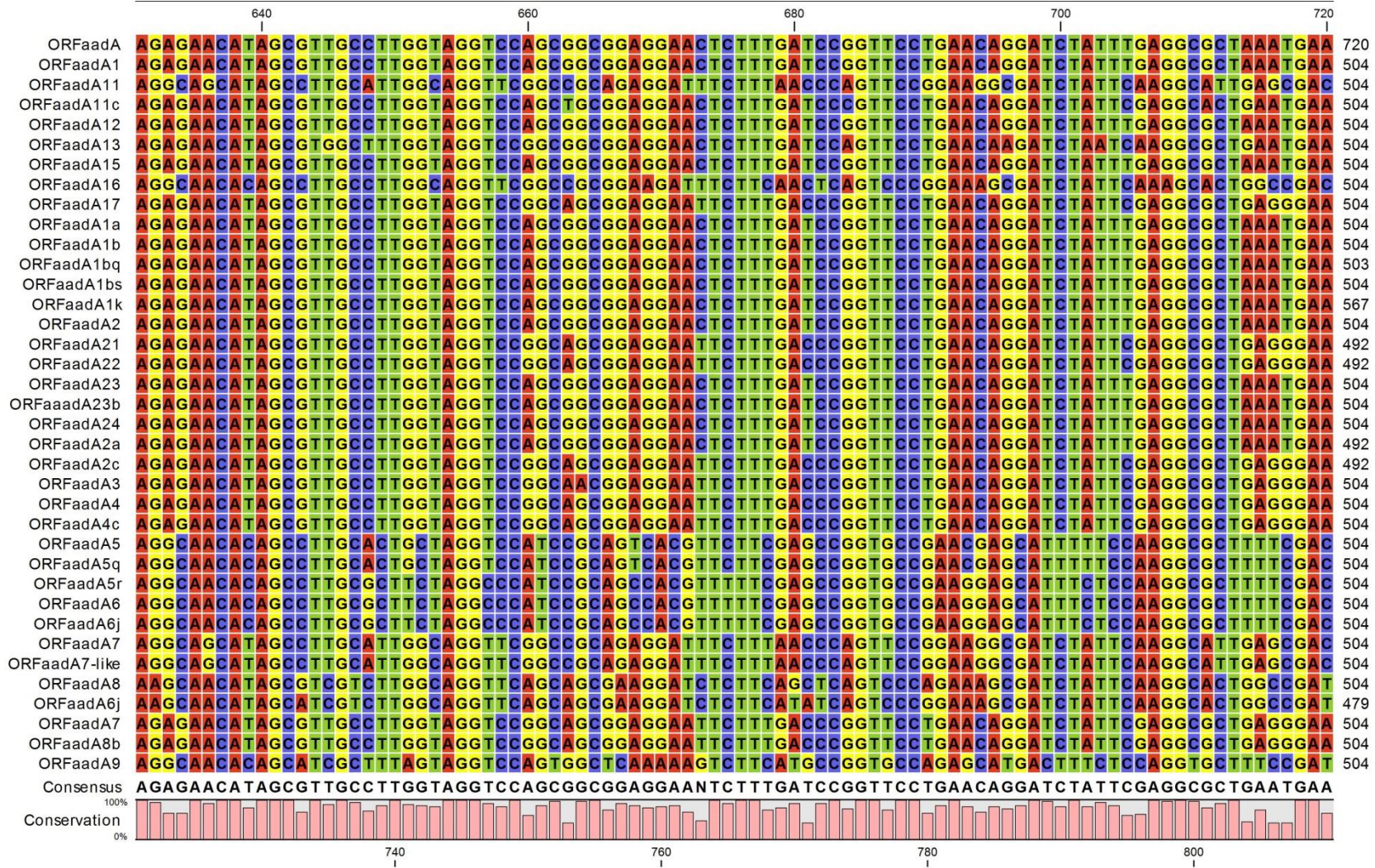




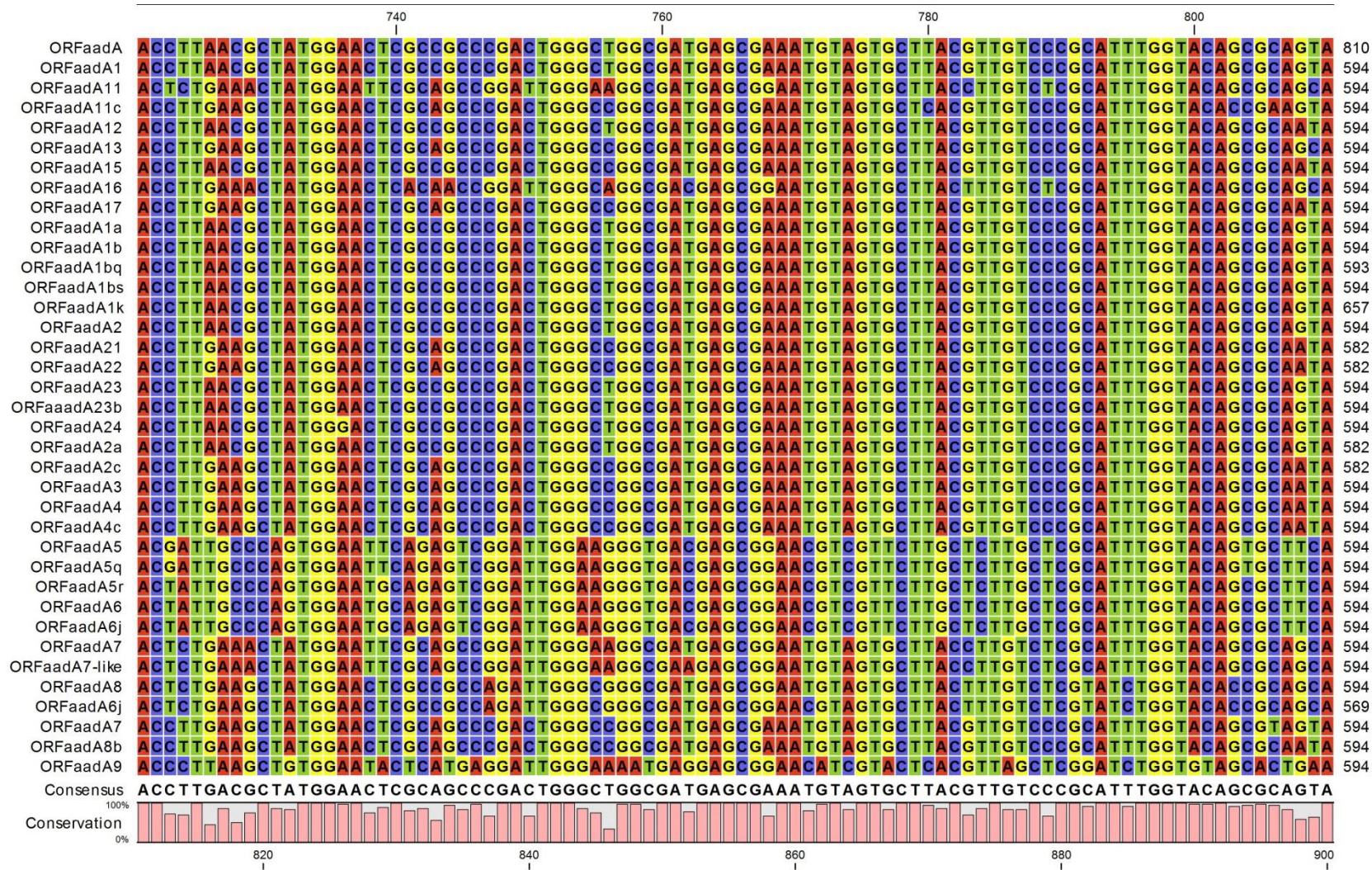
Anexos



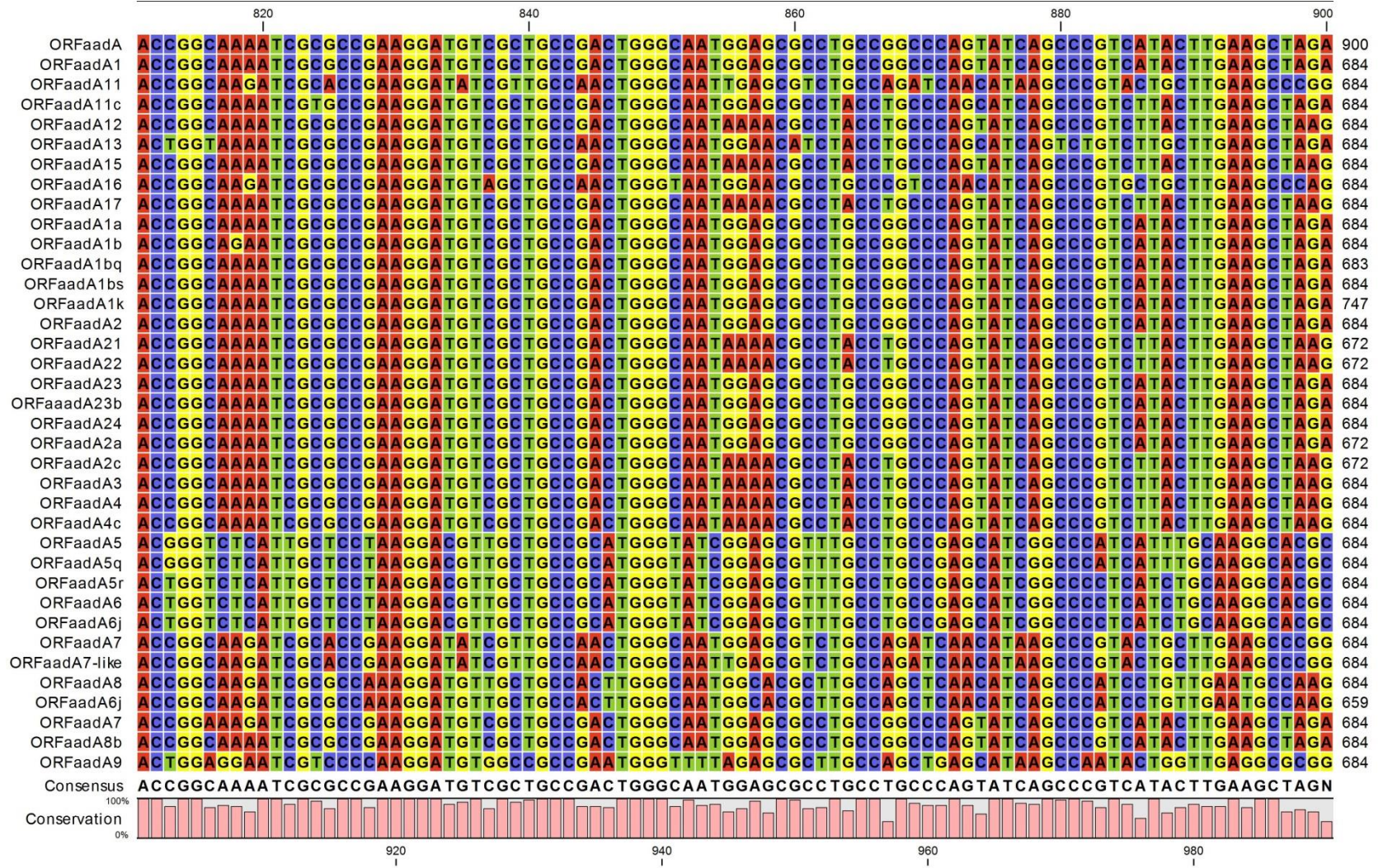




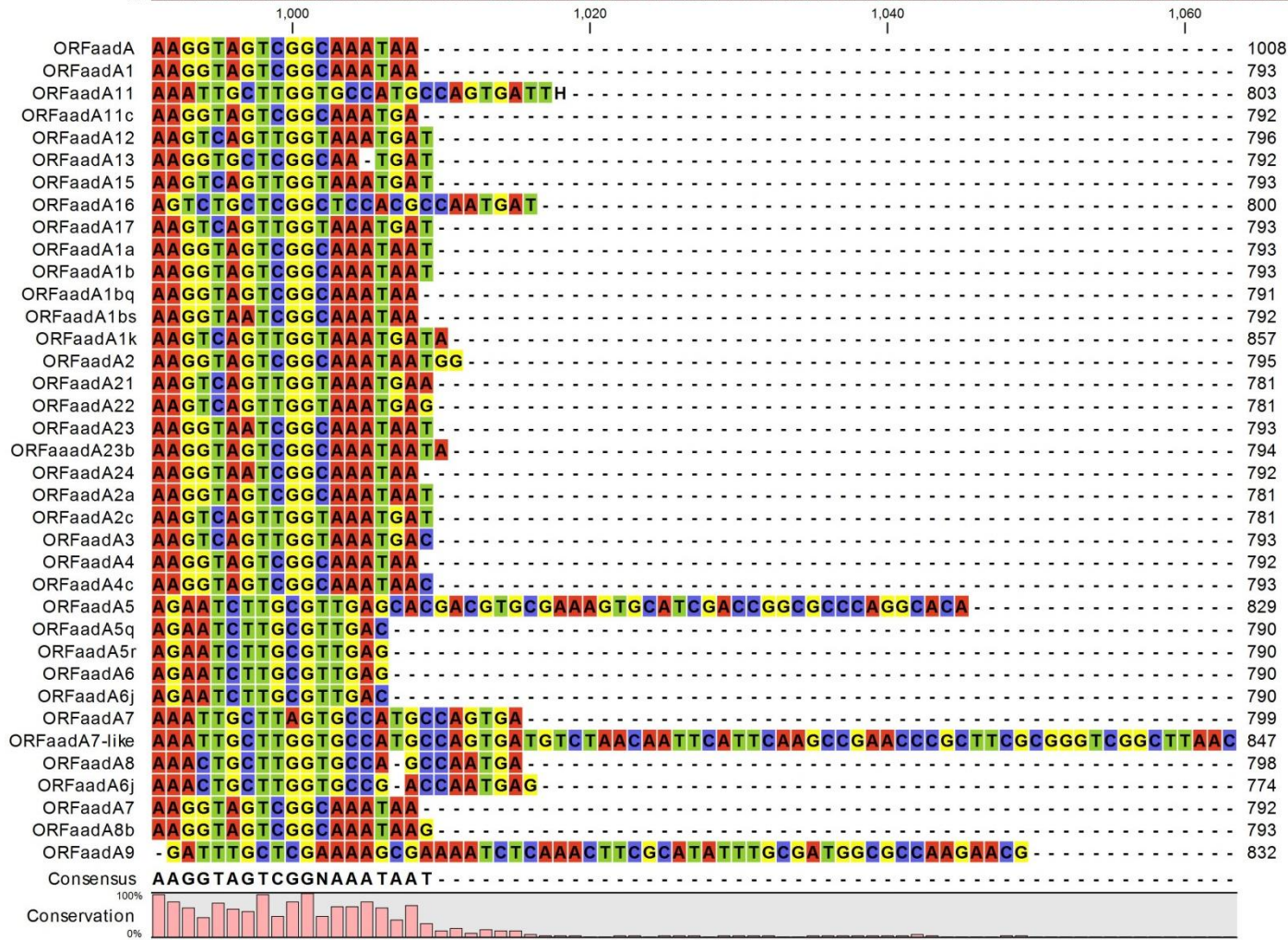
Anexos



Anexos









*Anexos*

Firman este documento de tesis

---

Angélica Viviana Galán  
Tesisista

---

Dra. María Paula Quiroga  
Directora adjunta

---

Dra. Daniela Centrón  
Directora de tesis

Fecha: