



UNIVERSIDAD DE GRANADA  
FACULTAD DE CIENCIAS

**Desarrollo de intrones del grupo II como  
vectores de reconocimiento génico y su  
aplicación en genómica funcional en  
microorganismos y plantas**

*Rafael Nisa Martínez*

Tesis Doctoral  
2011

Editor: Editorial de la Universidad de Granada  
Autor: Rafael Nisa Martínez  
D.L.: GR 4522-2011  
ISBN: 978-84-694-6014-6



**Desarrollo de intrones del grupo II como vectores de reconocimiento génico y su aplicación en genómica funcional en microorganismos y plantas**

Memoria que presenta el licenciado en Bioquímica  
D. Rafael Nisa Martínez  
como aspirante al grado de Doctor

Fdo: *Rafael Nisa Martínez*

Vº Bº  
El Director

Fdo: Nicolás Toro García  
Doctor en Ciencias Biológicas  
Profesor de Investigación del C.S.I.C.



Esta Tesis Doctoral ha sido realizada en el Grupo de Ecología Genética de la Rizosfera, del Departamento de Microbiología del Suelo y Sistemas Simbióticos de la Estación Experimental del Zaidín. (C.S.I.C.)



Parte de los resultados presentados en esta Tesis Doctoral han sido publicados en revistas internacionales o están en preparación:

**Nisa-Martínez, R., Jiménez-Zurdo, J.I., Martínez-Abarca, F., Muñoz-Adelantado, E. y Toro, N.** (2007). Dispersion of the RmInt1 group II intron in the *Sinorhizobium meliloti* genome upon acquisition by conjugative transfer. *Nucleic Acids Res* **35**; 214–222.

**Nisa-Martínez, R., Molina-Sánchez, M.D. y Toro, N.** Relationship between the group II intron RmInt1 and his natural host: *Quid pro quo?*. En preparación.

**Nisa-Martínez, R., Martínez-Abarca, F., y Toro, N.** A new genetic tool based on the group II intron RmInt1. En preparación.

**Nisa-Martínez, R., Jiménez-Zurdo, J.I., Frugier, F., Crespi, M y Toro, N.** Transient expression and subcellular localization of a bacterial group II intron maturase in *Arabidopsis thaliana* protoplasts. En preparación.



*A mis padres  
y a Lola*



# ÍNDICE



<b>INTRODUCCIÓN</b>	3
<b>1. CARACTERÍSTICAS DE LOS INTRONES DEL GRUPO II</b>	5
1.1. DISTRIBUCIÓN EN LA NATURALEZA	5
1.2. ESTRUCTURA	7
<b>1.2.1. Características de cada dominio.</b>	8
1.2.1.1 <i>Dominio I</i>	9
1.2.1.2. <i>Dominio II</i>	11
1.2.1.3. <i>Dominio III</i>	11
1.2.1.4. <i>Dominio IV</i>	12
1.2.1.5. <i>Dominio V</i>	12
1.2.1.6. <i>Dominio VI</i>	14
1.2.1.7. <i>Subdominios inusuales</i>	15
1.2.1.8. <i>Regiones interdominio</i>	16
<b>1.2.2. Otros intrones del grupo II</b>	16
1.2.2.1. <i>Intrones degenerados</i>	17
1.2.2.2. <i>Twintrones</i>	18
1.2.2.3. <i>Intrones del grupo II tipo Trans-splicing</i>	19
1.3. PROTEÍNAS CODIFICADAS POR LOS INTRONES DEL GRUPO II	19
<b>1.3.1. Características generales</b>	19
<b>1.3.2. Modelo tridimensional de una IEP</b>	23
1.4. CLASIFICACIÓN DE LOS INTRONES DEL GRUPO II	24
<b>2. MECANISMOS DE ESCISIÓN EN LOS INTRONES DEL GRUPO II</b>	28
2.1 <i>BRANCHING</i> O FORMACIÓN DE INTRÓN <i>LARIAT</i>	30
2.2 HIDRÓLISIS O FORMACIÓN DE INTRÓN LINEAL	31
2.3 FORMACIÓN DE INTRÓN CIRCULAR	31
2.4. SPLICING ALTERNATIVO	32
2.5. PAPEL DE LAS PROTEÍNAS EN LA ESCISIÓN <i>IN VIVO</i>	33
<b>2.5.1. Las IEPs y su papel en la escisión del intrón.</b>	33
<b>2.5.2. IEPs que participan en el <i>splicing</i> de varios intrones del grupo II</b>	35

<b>2.5.3. Proteínas codificadas por el hospedador</b>	36
<b>3. MOVILIDAD DE LOS INTRONES DEL GRUPO II</b>	40
3.1. MECANISMOS DE <i>RETROHOMING</i>	40
<b>3.1.1. Reconocimiento de la diana</b>	41
<b>3.1.2. Movimiento de intrones mitocondriales de levaduras</b>	44
<b>3.1.3. Retrohoming de intrones bacterianos</b>	49
3.1.3.1 <i>Retrohoming mediado por TPRT</i>	49
3.1.3.2. <i>Movilidad independiente de actividad DNA endonucleasa</i>	50
3.1.3.3. <i>Retrohoming de intrones lineales</i>	52
3.1.3.4. <i>Dependencia del retrohoming de las funciones celulares</i>	53
3.2. MECANISMO DE RETROTRANSPOSICIÓN	53
<b>4. EVOLUCIÓN DE LOS INTRONES DEL GRUPO II</b>	55
<b>5. LOS INTRONES DEL GRUPO II COMO HERRAMIENTAS BIOTECNOLÓGICAS</b>	58
5.1. CARACTERÍSTICAS BIOTECNOLÓGICAS Y APLICACIONES POTENCIALES DE LOS INTRONS DEL GRUPO II	58
5.2. USO DE LOS INTRONES DEL GRUPO II EN PROCARIOTAS	59
5.3. ESTRATEGIAS ACTUALES PARA EL SILENCIAMIENTO GÉNICO Y LA MUTAGÉNESIS INSERCIONAL DIRIGIDA EN EUCARIOTAS	63
5.4. LOS INTRONES DEL GRUPO II COMO ALTERNATIVA PARA LA MUTAGÉNESIS INSERCIONAL DIRIGIDA EN EUCARIOTAS.	66
<b>6. RmInt1, EL INTRÓN DEL GRUPO II DE <i>Sinorhizobium meliloti</i></b>	67
<b>OBJETIVOS</b>	73
<b>MATERIAL Y MÉTODOS</b>	77
<b>1. ESPECIES Y CEPAS BACTERIANAS</b>	77

<b>2. PLÁSMIDOS Y VECTORES DE CLONAJE</b>	77
<b>3. OLIGONUCLEÓTIDOS</b>	79
<b>4. CULTIVOS BACTERIANOS</b>	79
4.1. MEDIOS DE CULTIVO	79
4.2. ANTIBIÓTICOS	80
4.3. CONDICIONES DE CULTIVO	81
4.4. CONSERVACIÓN DE CULTIVOS BACTERIANOS	81
<b>5. AISLAMIENTO DE ÁCIDOS NUCLEICOS</b>	81
5.1. EXTRACCIÓN DE DNA total	81
<b>5.1.1. Lisados de colonia</b>	82
5.2. EXTRACCIÓN DNA PLASMÍDICO	82
<b>5.2.1. Aislamiento de plásmidos por lisis alcalina</b>	82
<b>5.2.2. Extracción de dna plasmídico por precipitación                 con sales de magnesio</b>	83
<b>5.2.3. Aislamiento de DNA plasmídico mediante                 kit comercial</b>	84
<b>5.2.4. Extracción de plásmidos a partir de                 grandes volúmenes de cultivo                 (midipreparaciones)</b>	84
5.3. EXTRACCIÓN DE RNA total	85
<b>5.3.1. Aislamiento de RNA total de <i>S. meliloti</i></b>	85
<b>5.3.2. Extracción de RNA total mediante kit comercial</b>	86
5.4. CUANTIFICACIÓN DE LA CONCENTRACIÓN DE RNA Y DNA	87
<b>6. MANIPULACIÓN Y AMPLIFICACIÓN DE DNA</b>	87
6.1. DIGESTIÓN CON ENZIMAS DE RESTRICCIÓN	87
6.2. REACCIÓN DE DEFOSFORILACIÓN	88
6.3. PCR	88
<b>6.3.1. PCR de colonia</b>	89
<b>6.3.2. RT-qPCR</b>	89
6.3.2.1. <i>Síntesis del cDNA</i>	90
6.3.2.2. <i>qPCR</i>	91

<b>7. ELECTROFORESIS</b>	91
7.1. ELECTROFORESIS EN GELES DE AGAROSA NO DESNATURALIZANTES	92
7.2. ELECTROFORESIS EN GELES DE AGAROSA DESNATURALIZANTES	92
7.3. ELECTROFORESIS EN GELES DE POLIACRILAMIDA DESNATURALIZANTES	92
7.4. MARCADORES DE PESO MOLECULAR	93
<b>8. CLONAJE DE FRAGMENTOS</b>	94
8.1. AISLAMIENTO DE FRAGMENTOS DE DNA DE GEL DE AGAROSA	94
8.2. LIGACIÓN DE FRAGMENTOS DE RESTRICCIÓN	94
<b>8.2.1. Adenilación de amplificadores de PCR</b>	94
8.3. PREPARACIÓN DE CÉLULAS COMPETENTES	94
<b>8.3.1. Células competentes por métodos químicos: RbCl</b>	95
<b>8.3.2. Células electrocompetentes</b>	95
8.4. TRANSFORMACIÓN DE CÉLULAS COMPETENTES	96
<b>8.4.1. Transformación por choque térmico</b>	96
<b>8.4.2. Electrotransformación</b>	96
<b>9. SECUENCIACIÓN Y ANÁLISIS DE SECUENCIAS</b>	96
<b>10. TÉCNICAS DE HIBRIDACIÓN DE ÁCIDOS NUCLEICOS</b>	97
10.1. HIBRIDACIÓN DNA-DNA (SOUTHERN BLOT)	97
<b>10.1.1. Transferencia alcalina por vacío</b>	97
<b>10.1.2. Marcaje de la sonda</b>	98
<b>10.1.3. Hibridación, lavados y revelado</b>	98
10.2. HIBRIDACIÓN EN MICROARRAYS	99
<b>10.2.1. Hibridación genómica comparativa, CGH.</b>	100
10.2.1.1 <i>Marcaje de las muestras.</i>	100
10.2.1.2 <i>Preparación de los microarrays.</i>	102
10.2.1.3 <i>Hibridación y lavados.</i>	102
10.2.1.4 <i>Escaneado y procesamiento de los datos.</i>	102
<b>10.2.2. Hibridación de microarrays</b>	103

10.2.2.1. Marcaje de las muestras.	103
10.2.2.2. Pretratamiento de los microarrays	105
10.2.2.3. Hibridación y lavados.	105
10.2.2.4. Revelado y procesamiento de los datos.	105
<b>11. ENSAYOS DE MOVILIDAD</b>	105
11.1. CONJUGACIONES EN <i>S. meliloti</i>	105
11.2. DETECCIÓN POR HIBRIDACIÓN	106
<b>12. TÉCNICAS CON RADIOISÓTOPOS: EXTENSIÓN DE CEBADOR</b>	107
<b>13. CONSTRUCCIÓN DE MUTANTES</b>	108
<b>14. ENSAYOS CON PROTOPLASTOS DE <i>A. thaliana</i></b>	109
14.1. PREPARACIÓN DEL MEDIO DE CULTIVO JPL 10X CON O SIN AZÚCARES	109
14.2. OBTENCIÓN DE PROTOPLASTOS DE <i>A. thaliana</i> A PARTIR DE CULTIVO CELULAR	110
14.3. TRANSFORMACIÓN DE LOS PROTOPLASTOS	113

## **RESULTADOS Y DISCUSIÓN**

<b>CAPÍTULO I. DISPERSIÓN DEL INTRÓN DEL GRUPO II RmInt1 EN EL GENOMA DE <i>Sinorhizobium meliloti</i> TRAS SU ADQUISICIÓN POR TRANSFERENCIA CONJUGATIVA.</b>	121
I.1. INTRODUCCIÓN	121
I.2. MATERIAL Y MÉTODOS	123
<b>I.2.1. Cepas bacterianas y condiciones de crecimiento</b>	123
<b>I.2.2. Construcciones donadoras de RmInt1 para ensayos         <i>in vivo</i></b>	124
<b>I.2.3. Ensayos de maduración y <i>homing in vivo</i></b>	124
<b>I.2.4. Análisis de la dispersión de RmInt1 en el genoma de <i>S.         meliloti</i> y de la movilidad de los intrones insertados</b>	125
I.3. RESULTADOS	126
<b>I.3.1. Modificación de construcciones donadoras de         RmInt1 para su dispersión</b>	126

<b>I.3.2. Dispersión de RmInt1 en el genoma de <i>S. meliloti</i></b>	130
<b>I.3.3. Capacidad de <i>retrohoming</i> de la copias de intrón insertadas en el genoma y complementación en <i>trans</i> del intrón <math>\Delta</math>ORF con la secuencia codificante de la IEP</b>	133
I.4. DISCUSIÓN	135
<b>CAPÍTULO II. RmInt1 Y <i>Sinorhizobium meliloti</i>: ¿<i>Quid pro quo</i>?</b>	141
II.1. INTRODUCCIÓN	141
II.2. MATERIAL Y MÉTODOS	144
<b>II.2.1. Cepas bacterianas y condiciones de crecimiento</b>	144
<b>II.2.2. Hibridación DNA-DNA</b>	145
<b>II.2.3. Hibridación genómica comparativa</b>	145
<b>II.2.4. Análisis mediante hibridación en microarrays.</b>	146
<b>II.2.5. RT-qPCR</b>	147
<b>II.2.6. Colección de mutantes de <i>S. meliloti</i> 1021 y 2011</b>	148
<b>II.2.7. Construcción del plásmido donador de intrón pGm4</b>	152
<b>II.2.8. Ensayos de movilidad</b>	153
II.3. RESULTADOS	154
<b>II.3.1. Selección de las cepas para los ensayos con microarrays</b>	154
<b>II.3.2. Genes diferencialmente expresados en <i>S. meliloti</i> 1021 en presencia de intrón.</b>	155
<i>II.3.2.1. La presencia de intrones no implica diferencias en el genotipo.</i>	160
<b>II.3.3. Perfil genético de <i>S. meliloti</i> RMO17 comparado con <i>S. meliloti</i> 1021</b>	161
<i>II.3.3.1. Genes diferencialmente expresados en <i>S. meliloti</i> RMO17 respecto a <i>S. meliloti</i> 1021.</i>	163
<b>II.3.4. Genes diferencialmente expresados en <i>S. meliloti</i> RMO17 en presencia de intrón</b>	164

<b>II.3.5. Verificación de los niveles de expresión de RmInt1 mediante RT-qPCR</b>	167
<b>II.3.6. Movilidad de pGm4</b>	168
<b>II.3.7. Ensayos de movilidad en mutantes de <i>S. meliloti</i> 1021 y 2011</b>	169
<i>II.3.7.1. Ligasas</i>	170
<i>II.3.7.2. Helicasas</i>	170
<i>II.3.7.3. Proteínas reparadoras del DNA</i>	171
<i>II.3.7.4. DNAs Polimerasas y RNAs polimerasas</i>	171
<i>II.3.7.5. Nucleasas</i>	173
<i>II.3.7.6. Chaperonas</i>	174
<i>II.3.7.7. Mutantes en otras funciones</i>	174
<b>II.4. DISCUSIÓN</b>	176
<b>II.4.1. Análisis genómico y transcriptómico por microarrays</b>	176
<b>II.4.2. Análisis de la movilidad de RmInt1 en mutantes de <i>S. meliloti</i></b>	179
<i>II.4.2.1 Efectos sobre las etapas iniciales del homing</i>	180
<i>II.4.2.2. Degradación del RNA del intrón tras la transcripción reversa</i>	181
<i>II.4.2.3 Papel de las DNA polimerasas</i>	181
<i>II.4.2.4. Otras proteínas</i>	182

## **CAPÍTULO III. INCORPORACIÓN DE MARCADORES DE RESISTENCIA EN LA SECUENCIA DE RmInt1 PARA SU USO EN MUTAGÉNESIS INSERCIONAL**

187

<b>III.1 INTRODUCCIÓN</b>	187
<b>III.2 MATERIAL Y MÉTODOS</b>	191
<b>III.2.1. Cepas bacterianas y condiciones de crecimiento</b>	191
<b>III.2.2. Construcción de la Unidad de Resistencia a Eritromicina (U.R.E.)</b>	191
<b>III.2.3. Construcción de un marcador de resistencia activable mediante <i>retrohoming</i></b>	193

<b>III.2.4. Ensayos de <i>homing in vivo</i></b>	194
III. 3. RESULTADOS	195
<b>III.3.1. El casete U.R.E. confiere resistencia a eritromicina</b>	195
<b>III.3.2. Construcción de un marcador de resistencia activable por <i>retrohoming</i> y su incorporación a vectores donadores de intrón</b>	195
III.4. DISCUSIÓN	198
<b>CAPÍTULO IV ACTIVIDAD DE RmInt1 EN HOSPEDADORES HETERÓLOGOS BACTERIANOS Y EUCARIOTAS (PLANTAS)</b>	203
IV.1 INTRODUCCIÓN	203
IV.2 MATERIAL Y MÉTODOS	204
<b>IV.2.1. Cepas bacterianas y condiciones de crecimiento</b>	204
<b>IV.2.2. Construcción de plásmidos</b>	204
<b>IV.2.3. Ensayos de <i>homing in vivo</i></b>	206
<b>IV.2.4. Localización de las fusiones IEP-GFP</b>	206
IV.3. RESULTADOS	207
<b>IV.3.1. RmInt1 es móvil en <i>E. coli</i> y <i>P. putida</i>.</b>	207
<b>IV.3.2. Construcción de fusiones IEP-GFP para su uso en bacteria</b>	209
<b>IV.3.3. Localización intrabacteriana de las fusiones IEP-GFP</b>	209
<b>IV.3.4 Localización subcelular en protoplastos de <i>A. thaliana</i></b>	211
IV.4. DISCUSIÓN	213
<b>CONCLUSIONES</b>	219
<b>BIBLIOGRAFÍA</b>	223
<b>ABREVIATURAS Y ANGLICISMOS</b>	267
<b>AGRADECIMIENTOS</b>	273

# **INTRODUCCIÓN**



## INTRODUCCIÓN

La idea general de que, durante el desarrollo de la vida en nuestro planeta, la evolución se ha basado en la síntesis de proteínas precedida por la replicación del RNA se propuso por primera vez hace ya algo más de cuarenta años (Woese, 1967; Crick, 1968; Orgel, 1968). Se ha sugerido que, probablemente, los catalizadores formados exclusivamente por RNA han tenido una gran importancia en los inicios de la evolución de la vida, pero la posibilidad de que estos catalizadores estuviesen aún presentes en los organismos actuales no se tuvo en cuenta. El inesperado descubrimiento de las ribozimas (Kruger *et al.* 1982; Guerrier-Takada *et al.* 1983) fue el comienzo para el debate, que aún hoy continúa, sobre el papel del RNA en los orígenes de la vida (Pace y Marsh, 1985, Sharp, 1985; Lewin, 1986; Noller, 2006).

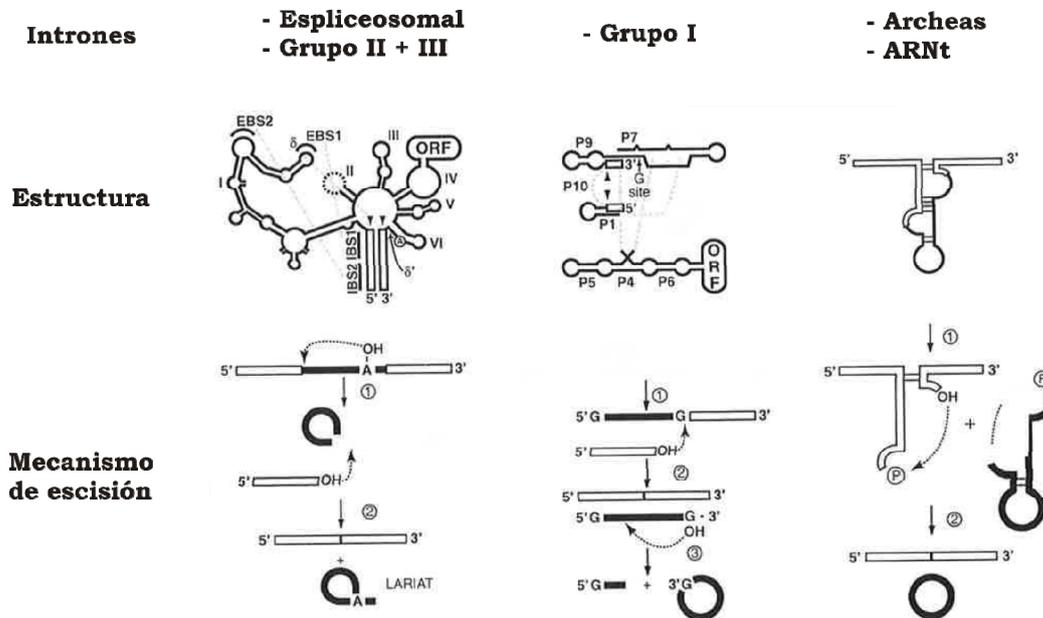
Las ribozimas son moléculas de RNA con actividad catalítica. Desde que fueron descubiertas, la catálisis química del RNA ha sido descrita en muchos procesos biológicos esenciales, como en la maduración y procesamiento del RNA, en la replicación de genomas compuestos por RNA, y en la formación de los enlaces peptídicos durante la traducción (Eckstein y Lilley, 1996; Ban *et al.*, 2000; Muth *et al.*, 2000; Nissen *et al.*, 2000). Las ribozimas naturales, con la única excepción del ribosoma, catalizan la rotura o ligación del esqueleto fosfodiéster del RNA. En contraste con otras ribonucleasas conocidas, las ribozimas catalizan reacciones altamente específicas de secuencia, determinadas por interacciones RNA-RNA entre la ribozima y sus moléculas sustrato.

Las ribozimas pueden ser clasificadas en tres grupos: RNAasa P, pequeños RNAs catalíticos e intrones<sup>1</sup> autocatalíticos (Fig. 1). La RNAasa P es una enzima ubicua que participa en la biosíntesis de los tRNAs (Phizicky y Hopper, 2010). El grupo de los pequeños RNAs catalíticos incluye cuatro motivos catalíticos diferentes que comparten una misma reacción química (Symons, 1992). Finalmente, los intrones autocatalíticos se dividen en dos tipos: intrones del grupo I e intrones del grupo II (Michel *et al.*, 1982; Davies

---

<sup>1</sup>Se han descrito cinco tipos principales de intrones presentes en los organismos vivos en función de diferencias estructurales y de su mecanismo de maduración. Además de los intrones nucleares eucariotas (también conocidos como intrones espliceosómicos) (Burge *et al.*, 1999; Reed, 2000), encontramos un grupo de intrones específicos de arqueobacterias (Lykke-Andersen *et al.*, 1997), un tipo particular de intrones nucleares que se localizan en los ARNt (Abelson *et al.*, 1998) y dos grupos de intrones cuya particularidad es que son capaces de auto-madurar en condiciones de laboratorio: intrones del grupo I y del grupo II (Fig. 1).

*et al.*, 1982; Schmelzer *et al.*, 1982)<sup>2</sup>. A su vez, dentro de los intrones del grupo II se ha hecho una distinción al referirse a intrones del grupo III, que no son más que intrones del grupo II simplificados estructural y funcionalmente que fueron encontrados en el mRNA de los plastidios del protista *Euglena gracilis*<sup>3</sup> (Christopher y Hallick, 1989; Copertino y Hallick, 1993).



**Figura 1.** Tipos de intrones junto con sus estructuras y mecanismos de escisión. En la parte inferior de la figura se observan como cajas blancas los exones y como cajas negras los intrones. (Adaptado de Belfort *et al.* 2002).

Los intrones del grupo II son, como ya hemos dicho, grandes RNAs catalíticos (ribozimas) que pueden codificar proteínas pertenecientes a la familia de las reverso-transcriptasas (Michel y Lang, 1985; Xiong y Eickbush, 1990). Estas proteínas reciben el nombre de IEPs (Intron-Encoded Proteins), aunque alguna como la del intrón Ll.ltrB de *Lactococcus lactis* (*L. lactis*) tiene nombre propio: LtrA. Se sabe que estas proteínas asisten específicamente la maduración de su intrón y participan activamente en otros procesos en los que se ven involucrados estos intrones (Lambowitz, 1989; Lambowitz y Belfort, 1993; Lambowitz *et al.*, 1999).

<sup>2</sup> La distinción entre intrones del grupo I y II se lleva a cabo en 1982 por Michel y colaboradores cuando, al analizar las secuencias completas de diez intrones mitocondriales de levadura y dos incompletas (una de levadura y otra de *Aspegillus nidulans*), encuentran claras diferencias estructurales (Michel *et al.*, 1989).

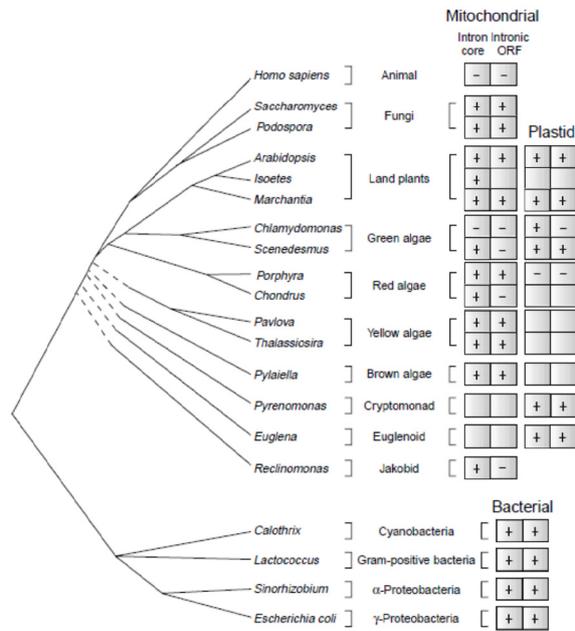
<sup>3</sup> En los intrones del grupo II y III de los plastidios de estos euglenoides protistas también se han encontrado los denominados como “twintrons”, que son intrones dentro de intrones.

Algunos intrones del grupo II son retroelementos genéticos móviles, hecho que se sospechó por su amplia presencia y distribución, (Lambowitz, 1989) pero no fue hasta 1990 cuando se demostró su capacidad para invadir una secuencia específica de DNA libre de intrón (Meunier *et al.*, 1990; Skelly *et al.*, 1991). Sólo se ha podido demostrar la capacidad de movimiento en algunos de ellos: aI1 y aI2 (también conocidos como *coxI-I1* y *coxI-I2*) presentes en *S. cerevisiae* (*S. cerevisiae*) (Meunier *et al.*, 1990; Lazowska *et al.*, 1994); Ll.ltrB de *L. lactis* (Mills *et al.*, 1997), Ec.I5 de *Escherichia coli* (*E. coli*) (Zhuang, 2009) y RmInt1 en *Sinorhizobium meliloti* (*S. meliloti*) (Martínez-Abarca *et al.*, 2000). Recientemente, se ha identificado un grupo de intrones que se comportan de un modo similar, en cuanto a maduración y movilidad, a Ll.ltrB en *E. coli*: Na.1 y Na.2 de *Novosphingobium aromaticivorans* (Karberg, 2005).

## 1. CARACTERÍSTICAS DE LOS INTRONES DEL GRUPO II

### 1.1. DISTRIBUCIÓN EN LA NATURALEZA

Los intrones del grupo II se encontraron por primera vez en el genoma mitocondrial de la levadura *S. cerevisiae* (Michel *et al.*, 1982), pero se encuentran ampliamente distribuidos en la naturaleza (Fig. 2). Debido a la similitud de su mecanismo de maduración y el que presentan los intrones nucleares, son considerados como los antecesores de los intrones espliceosómicos, de los retrotransposones no-LTR y también se piensa que están evolutivamente relacionados con la telomerasa (Padgett *et al.*, 1984; Ruskin *et al.*, 1984; Sharp, 1985; 1991; Cech, 1986; Cavalier-Smith, 1991; Moore *et al.*, 1993; Eickbush, 1994; Koonin, 2006; Chalamcharla *et al.*, 2010; Keating, 2010). Posteriormente se han descrito en bacterias y arqueas (Lambowitz y Zimmerly, 2004; Lehmann y Schmidt, 2003; Toro *et al.*, 2007). No suelen ser muy abundantes en mitocondrias de hongos pero sí en el genoma tanto de mitocondrias (Oda *et al.*, 1992; Knoop y Brennicke, 1993) como de cloroplastos (Ohyama *et al.*, 1986; Shinozaki *et al.*, 1986) de plantas superiores, y están presentes en el genoma de orgánulos de algas (Kück *et al.*, 1987; Kück, 1989; Kono *et al.*, 1991; Manhart y Von der Haar, 1991). En este sentido, es de destacar la especial abundancia de intrones del grupo II (y también del grupo III) en los cloroplastos de *E. gracilis* donde han llegado a identificarse hasta 155 intrones, lo que supone el 40% del genoma del orgánulo (Hallick *et al.*, 1993; Copertino y Hallick, 1993; Doetsch *et al.*, 1998; 2001). También son especialmente abundantes en el genoma del



**Figura 2.** Distribución de los intrones del grupo II en la naturaleza (adaptada de Bonen, 2001). Aunque no está actualizada, aún sirve para ilustrar la amplia distribución filogenética de los intrones del grupo II en orgánulos eucariotas y en bacterias.

cloroplasto de *Arabidopsis thaliana* (*A. thaliana*) (26 intrones; Sato *et al.*, 1999) y en los genomas mitocondriales de ésta (22 intrones; Marienfeld *et al.*, 1999) y de *Marchantia polymorpha* (25 intrones; Oda *et al.*, 1992). Curiosamente, los intrones identificados en orgánulos se encuentran en la mayoría de los casos interrumpiendo genes que codifican proteínas (Open Reading Frames, ORFs) y, en algunos casos, los genes de tRNA y rRNA. Se observan ciertas tendencias en los genomas: en plantas con flor, todos los intrones mitocondriales del grupo II se encuentran interrumpiendo ORFs, mientras que un tercio de los presentes en el

cloroplasto se han insertado en tDNAs (los dos tercios restantes interrumpen ORFs); en cambio, los dos intrones presentes en la mitocondria del alga roja *Porphyra* se localizan en el DNA que codifica los rRNA (Bonen y Vogel, 2001). Llama la atención que aún no se han encontrado intrones del grupo II interrumpiendo genes nucleares eucariotas, aunque se ha descrito que el genoma nuclear de *A. thaliana* (concretamente el cromosoma 2) contiene la inserción de una secuencia de DNA mitocondrial (270 kb) que a su vez incluye un intrón del grupo II (Lin *et al.*, 1999). Recientemente se ha descrito el primer intrón del grupo II en el genoma mitocondrial de un gusano bilateral del género *Nephtys* interrumpiendo el gen *cox1* (Vallès *et al.*, 2008) y también en el genoma de otros animales (Dellaporta *et al.*, 2006; Burger *et al.*, 2009).

A principios de los años 90, Ferat y colaboradores identificaron por primera vez intrones del grupo II en proteobacterias (*Azotobacter*, *Escherichia*) y cianobacterias (*Calothrix*, *Anabaena*, *Nostoc*) (Ferat y Michel, 1993; Ferat *et al.*, 1994), bacterias que se supone que han dado origen a las mitocondrias y los cloroplastos, respectivamente. Más tarde, se identificaría uno de los intrones del grupo II más estudiados y caracterizados, Ll.ltrB en *L. lactis* (Mills *et al.*, 1996; Shearman *et al.*, 1996). Aproximadamente, el 25%

de los genomas bacterianos secuenciados contienen intrones del grupo II<sup>4</sup> y en algunos casos se han descrito hasta 12 copias de estos elementos en el genoma, lo que nos proporciona una idea de la amplia distribución y presencia de este tipo de elementos entre los procariotas (Martínez-Abarca y Toro, 2000a; Dai y Zimmerly, 2002; Lambowitz y Zimmerly, 2004). Al contrario de lo que ocurre con los intrones de orgánulos, los intrones bacterianos suelen localizarse entre genes (no interrumpiendo ORFs) o bien, con mucha frecuencia, dentro de elementos móviles como secuencias de inserción o plásmidos (Ferat *et al.*, 1994; Martínez-Abarca *et al.*, 1998; Mills *et al.*, 1996; Shearman *et al.*, 1996; Klein y Dunny, 2002). Además, normalmente suelen encontrarse como retroelementos activos, no como los de orgánulos, que frecuentemente tienen su estructura de RNA degenerada y carecen de ORF o codifican IEPs degeneradas, por lo que han perdido su movilidad<sup>5</sup> (Michel y Ferat, 1995; Bonen, 2008; Barkan, 2009). Esta localización de los intrones bacterianos del grupo II podría favorecer la dispersión de estos elementos en el mundo procariota (Martínez-Abarca y Toro, 2000a; Lambowitz y Zimmerly, 2004). Mucho más reciente es la descripción de este tipo de elementos en arqueobacterias (Galagan *et al.*, 2002; Deppenmeier *et al.*, 2002), donde se han encontrado únicamente en dos especies de arqueas metanogénicas estrechamente relacionadas, *Methanosarcina acetivorans* y *Methanosarcina mazei*, y se piensa que su presencia podría estar relacionada con transferencia horizontal desde eubacterias (Rest y Mindell, 2003; Dai y Zimmerly, 2003; Toro, 2003).

## 1.2. ESTRUCTURA

La estructuración del RNA de los intrones del grupo II en dominios se infiere de los análisis filogenéticos de su estructura secundaria (Michel *et al.*, 1989). Esta organización, junto con los conocimientos de las numerosas interacciones terciarias y limitaciones estructurales, ha permitido la construcción de modelos tridimensionales hipotéticos de los intrones del grupo II (DeLancastre *et al.*, 2005; Hamill y Pyle, 2006; Dai *et al.*, 2008). A pesar de las nuevas informaciones obtenidas sobre la estructura terciaria funcional (Toor *et al.*, 2008a; Toor *et al.*, 2008b) las designaciones iniciales en los distintos dominios y los distintos modelos, siguen siendo válidos (Fig 3.).

---

<sup>4</sup> Se han encontrado intrones del grupo II en *Proteobacteria*, *Cyanobacteria*, *Acidobacteria*, *Actinobacteria*, *Bacteriodes* y *Firmicutes* (Gram positiva) (base de datos de intrones del grupo II: <http://www.fp.ucalgary.ca/group2introns>).

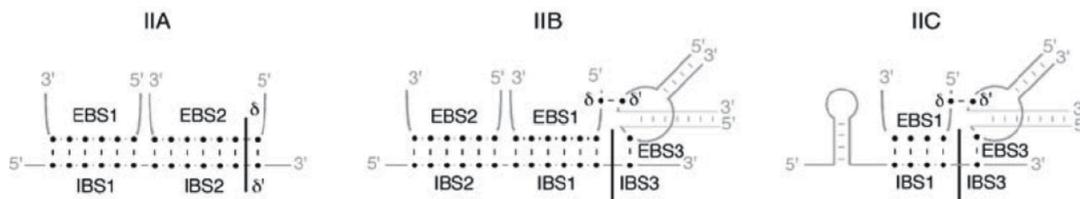
<sup>5</sup> Sólo un tercio de los intrones del grupo II de orgánulos presentan IEP (Toor *et al.*, 2001)



Gracias a estudios en *trans*<sup>6</sup> con los diferentes dominios, hoy podemos asociar funciones concretas a cada uno de ellos

### 1.2.1.1 Dominio I

De los dos exones que flanquean a un intrón del grupo II, el exón 5' muestra una gran complementariedad de bases con secuencias de reconocimiento del DI (las interacciones EBS1-IBS1 e EBS2-IBS2) (Fig. 4) (Jacquier y Michel, 1987; Lambowitz y Zimmerly, 2010), mientras que el reconocimiento del exón 3' se limita a la interacción de un único par de bases (interacción  $\delta$ - $\delta'$  o EBS3-IBS3) (Costa *et al.*, 2000).



**Figura 4.** Interacciones de los exones con el sitio activo en los intrones IIA, IIB y IIC.

El dominio I (DI) es el dominio de mayor tamaño (excluyendo al DIV en el que se encuentra la secuencia que codifica la IEP) y está dividido en varios sectores que controlan el plegamiento, el reconocimiento de los exones y el acoplamiento de los dominios catalíticos. Estudios *in vitro* del plegamiento del RNA indican que el DI es el primero que se pliega y que sirve como andamio para el posterior ensamblaje de los otros dominios (Fedorova y Zinglet, 2007; Pyle *et al.*, 2007, Stenier *et al.*, 2008). La arquitectura global del DI viene dictada por la formación de manera adecuada de lo que se denomina “el cruce de cinco caminos” haciendo referencia a los cinco brazos que componen el DI y que parten de un anillo común (Toor *et al.*, 2008a) y por un conjunto de interacciones internas entre las regiones  $\kappa$ - $\zeta$  del tallo DID con el DV (Waldsich y Pyle, 2007).

Se han determinado tres tipos de interacciones terciarias entre el DI y el DV que son críticas para la catálisis: las ya mencionadas  $\kappa$ - $\kappa'$  y  $\zeta$ - $\zeta'$  que son importantes para el acoplamiento del DV (Costa y Michel, 1995; Boudvillian y Pyle, 1998), y la interacción  $\lambda$ - $\lambda'$ , la cual aproxima el DV al sitio de escisión 5' (5'-ss) (Boudvillian *et al.*, 2000) y participa directamente en la catálisis. Además, se ha demostrado que la interacción intradominio  $\varepsilon$ - $\varepsilon'$

<sup>6</sup> Se trata de estudios *in vitro* en los que los dominios se transcriben como moléculas independientes y, cuando se combinan con otras regiones del intrón y se analizan las reacciones que son capaces de catalizar, podemos deducir las funciones en las que están directamente involucrados.

también es importante para la función catalítica (Jacquier y Michel, 1987, 1990).

La interacción  $\alpha$ - $\alpha'$  facilita la unión del exón 5' (Qin y Pyle, 1998) lo cual juega un papel importante en el plegamiento del intrón (Waldsich y Pyle, 2007). También se ha sugerido que la interacción  $\theta$ - $\theta'$  participa en la estabilización estructural del intrón nativo (Costa *et al.*, 1997b). Esta interacción es también importante para la incorporación del efector catalítico DIII y la unión JII/III (secuencia entre el DII y el DIII) al sitio activo (Fedorova *et al.*, 2003).

También se ha identificado en el DI el sitio de anclaje para el punto de ramificación (*branch point*) dentro de la denominada región de coordinación (*coordination loop*<sup>7</sup>) (Hamill y Pyle, 2006), la cual también contiene residuos que participan en el posicionamiento del sitio 3' de escisión (3'-ss) (interacción EBS3-IBS3; Costa *et al.*, 2000).

En el DI existen elementos estructurales que son críticos para el plegamiento del intrón, si bien, el plegamiento del DI es independiente de otros dominios del intrón (Qin y Pyle, 1997). El plegamiento del DI es el paso limitante en el plegamiento del intrón completo (Su *et al.*, 2005) ya que la mayoría de los grupos funcionales críticos para la compactación y el plegamiento del intrón se encuentran comprendidos en una subestructura relativamente pequeña denominada "elemento de control del plegamiento" (Waldsich y Pyle, 2007). Esta subestructura contiene sitios de anclaje para el DV (elementos  $\kappa$  y  $\zeta$ ) y para el DVI (región de coordinación) (Fig. 3). Para la compactación de la ribozima también son importantes las interacciones  $\alpha$ - $\alpha'$  y  $\beta$ - $\beta'$  (Fedorova y Zingler, 2007), aunque se ha demostrado que el contacto entre  $\beta$  y  $\beta'$  no es indispensable para la función del intrón (Chu *et al.*, 2001).

El primer tallo del DI de algunos intrones del grupo II contiene una inserción que podría ser un posible sitio de unión de proteínas (Adamidi *et al.*, 2003) ya que se ha visto que en el intrón del grupo II de maíz *atpF* es necesaria esta inserción como sitio de unión del cofactor CRS1. También se han identificado en este dominio posibles sitios de unión para la proteína codificada por el intrón bacteriano Ll.ltrB de *L. lactis*.

---

<sup>7</sup> Se denomina región de coordinación (*coordination loop*) porque coordina el acercamiento de todos los factores necesarios para el *splicing*: el sitio de escisión 5' (5'-ss), el sitio de escisión 3' (3'-ss) y el sitio de ramificación (*branch site*) (Costa *et al.*, 2000; Hamill y Pyle, 2006).

### 1.2.1.2. Dominio II

El DII contiene motivos importantes para el ensamblaje estructural del intrón (Costa *et al.*, 1997b), y puede contener inserciones adicionales de RNA que, en algunos casos, forman ORFs que pueden codificar proteínas. Por ello, el DII puede actuar como una vía potencial de codificación de genes adicionales que son transportados a nuevos lugares del DNA cuando el intrón se retrotranspone, aunque los ORFs en el DII son poco comunes (Simon *et al.*, 2008). El DII forma dos interacciones terciarias esenciales tipo tetraloop-receptor<sup>8</sup> con el DI ( $\theta$ - $\theta'$ , vista antes) y con el DIV ( $\eta$ - $\eta'$ ) (Chanfreau y Jacquier, 1996; Costa *et al.*, 1997b). La interacción  $\theta$ - $\theta'$  se mantiene durante todo el proceso de maduración lo que indica que podría estar implicada en la estabilización de la estructura nativa del intrón (Chanfreau y Jacquier, 1996; Costa *et al.*, 1997b) siendo diferente entre los intrones IIA y IIB. Esta interacción también se considera importante para acercar tanto el DIII como la región entre DII/DIII ( $J_{II/III}$ )<sup>9</sup> al centro catalítico (Fedorova *et al.*, 2003).

La interacción con el DIV está estructuralmente conservada entre los intrones IIA y IIB (Costa *et al.*, 1997b). El contacto tetraloop/receptor  $\eta$ - $\eta'$  se establece entre la región central del DII y el bucle terminal del DVI y se ha propuesto que esta interacción podría actuar como un interruptor conformacional entre los dos pasos del *splicing* permitiendo la ligación de los exones (Chanfreau y Jacquier, 1996).

### 1.2.1.3. Dominio III

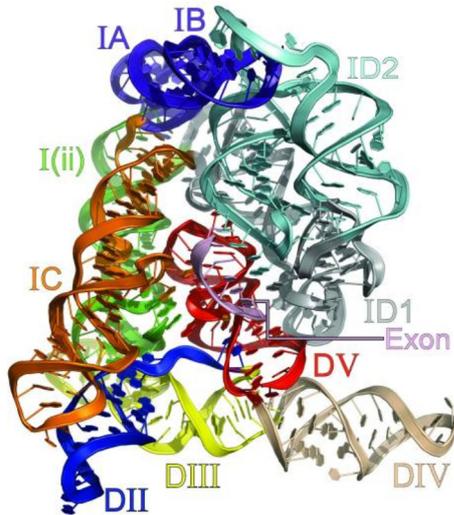
El DIII contiene una región desapareada interna (la región más conservada dentro del DIII) que es importante para el ensamblaje estructural del intrón, formando parte del centro activo (Toor *et al.* 2010). Actúa como efector catalítico (Qin y Pyle, 1998; Pyle y Lambowitz, 2006). No es imprescindible para la catálisis (Koch *et al.*, 1992), pero su presencia incrementa de forma notable la proporción de intrón escindido y la inserción del intrón (Griffin *et al.*, 1995; Xiang *et al.*, 1998; Fedorova *et al.*, 2003). El DIII une el resto del intrón con alta afinidad relativa (Podar *et al.* 1995a). Se cree que el tallo basal y la región desapareada interna podrían estar

<sup>8</sup> Son contactos que suelen estar implicados en la estabilización de la estructura terciaria mediante puentes de hidrógeno. Hay 2 tipos principales descritos en intrones del grupo II: (i) GNRA/receptor G-C, en el que las posiciones 3 y 4 del tetraloop interaccionan con el surco menor del ARN receptor; y (ii) GAAA/receptor motivo 11 nt (CCUAAG...UAUGG) (Costa y Michel, 1995; 1997).

<sup>9</sup> La región entre los dominios II Y III ( $J_{II/III}$ ) es, entre otras, una de las regiones más conservadas del intrón, y juega un papel clave en la construcción del sitio activo (Toor *et al.*, 2008a).

participando en interacciones terciarias con otros dominios (Jestin *et al.*, 1997) mientras que el lazo final interacciona con el DV (interacción  $\eta$ - $\eta'$ ; Fedorova y Pyle; 2005).

#### 1.2.1.4. Dominio IV



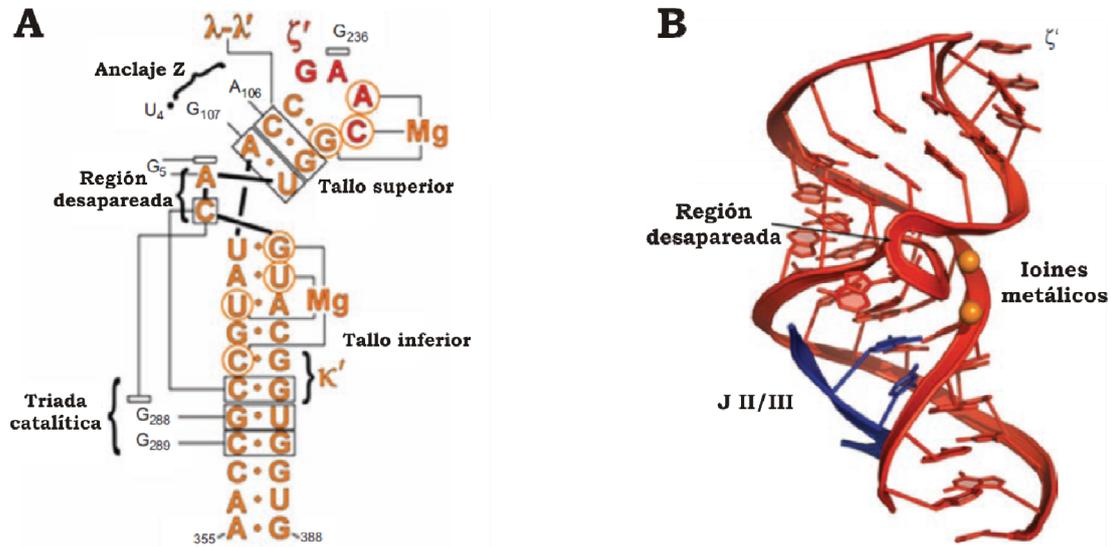
**Figura 5.** Estructura terciaria de un intrón del grupo II. El DV (en rojo) forma el centro catalítico junto con el resto de los dominios, encapsulando este sitio activo. Se observa como el DIV (en gris) se encuentra proyectado hacia el exterior de esta estructura (Toor *et al.*, 2009).

En algunos intrones, el DIV contiene un marco abierto de lectura que codifica una proteína con actividad madurasa, que permite el *splicing* del intrón bajo condiciones fisiológicas y se requiere para la movilidad del intrón (Lambowitz y Zimmerly, 2004). El DIV también contiene el sitio de unión primario para la madurasa (Matsuura *et al.*, 2001; Singh *et al.*, 2002; Watanabe y Lambowitz, 2004) y para otros factores de *splicing* específicos de cada intrón (Ostersetzer *et al.*, 2005). Debido a que se sitúa hacia afuera (Fig. 5) una vez que el intrón se encuentra plegado, puede jugar un papel más general como elemento de unión de proteínas e interactuar así con diferentes cofactores, permitiendo el *splicing* del intrón y su movilidad. Estos y otros aspectos serán tratados más ampliamente en sucesivos apartados.

#### 1.2.1.5. Dominio V

Los intrones del grupo II pueden considerarse como una semilla, dado la capacidad de diseminar información genética y de generar complejidad genómica (Pyle, 2010). El grano de esta semilla sería el pequeño y altamente conservado DV, con estructura de horquilla-bucle que contiene un pequeño codo o desapareamiento asimétrico de dos bases (Schmidt *et al.*, 1996) en el centro del tallo (Fig. 6). El DV es el componente más importante en la reacción de catálisis de los intrones del grupo II (Lehman y Schmidt, 2003; Pyle y Lambowitz, 2006; Toor *et al.*, 2009).

Uno de los motivos más conservados del DV es lo que se ha denominado la triada catalítica situada en la base de este dominio (Fig 6). Cambios en cualquiera de los residuos bloquean completamente el *splicing*, mientras que si se modifican los nucleótidos con los que aparean el efecto es leve (Chanfreau y Jacquier, 1994; Boulanger *et al.*, 1995; Peebles *et al.*, 1995). De los 3, la G es esencial para la maduración del intrón tanto *in vitro*



**Figura 6.** Red de interacciones en las que participa el DV. **(A)** Mapa de las interacciones secundarias entre el DV y el resto del intrón. Los corchetes indican interacciones a distancia y los contactos con los iones metálicos, Las cajas indican las bases apiladas, las etiquetas corresponden con los nombres de las interacciones terciarias. Los nucleótidos en rojo participan en la interacción ζ-ζ'. **(B)** Estructura del DV dentro del centro catalítico de un intrón del grupo IIC. Los nucleótidos del DV se presentan en rojo. Las esferas naranjas corresponden con los iones metálicos. Los nucleótidos de JII/III (en azul) forman una triple hélice con el surco mayor en la base del DV. También se indica la interacción ζ-ζ' con el DI. (Adaptada de Pyle 2010).

como *in vivo* (Boulanger *et al.*, 1995; Peebles *et al.*, 1995) y podría jugar un papel dual: hacia el surco menor, estabiliza la unión del DV al DI, y hacia el surco mayor, participaría activamente en la química de la reacción de *splicing* (Konforti *et al.*, 1998b).

Cuando nos adentramos en la estructura terciaria del intrón, el cual envuelve al DV como un escudo protector (Toor *et al.*, 2008a, 2010; Michel *et al.*, 2009; Keating *et al.*, 2010), se puede apreciar como el DV adopta una conformación que favorece en gran medida su reactividad (Fig. 6). El lazo asimétrico interno gira de forma muy cerrada sobre sí mismo, concentrando el esqueleto de fosfatos en el espacio y creando, por lo tanto, un lugar de potencial electrostático extremadamente negativo. Quizás, como resultado, dos cationes divalentes se unen fuertemente a este sitio. Debido a su posición, y por su localización relativa (Toor *et al.*, 2008b) se cree que estos

dos iones metálicos juegan un papel crítico en el mecanismo de catálisis de los intrones del grupo II. El sitio de unión de los iones metálicos se encuentra sustentado desde abajo por una estructura de triple hélice que resulta de la unión de la región conservada JII/III a una región no variable dentro del surco mayor del tallo inferior del DV (Keating *et al.*, 2010). Esta triple hélice del surco mayor es un nuevo motivo de RNA que parece tener una gran importancia en RNA relevantes biológicamente, incluyendo a la telomerasa (Theimer *et al.*, 2005; Keating *et al.*, 2010).

La subestructura intrónica adyacente parece ser esencial en el mantenimiento de esta conformación catalíticamente reactiva del DV (Toor *et al.*, 2008a). Ésta se logra gracias a la formación de la triple hélice con el JII/III y gracias también a interacciones como  $\lambda$ - $\lambda'$  y la región de anclaje Z del DI (Boudvillain *et al.*, 2000; Toor *et al.*, 2008a), la cual refuerza la única conformación basal dentro de la región desapareada del DV y permite así la carga negativa en el eje.

Además de las interacciones anteriores que ayudan a la reactividad catalítica del intrón, existen muchas otras interacciones entre el DI y la parte de atrás del DV, las cuales son necesarias para presentar y orientar los grupos catalíticos de forma adecuada dentro del centro activo. Se sabe que la horquilla del DV tiene dos “caras”, una catalítica y otra de unión opuesta a la anterior (Abramowitz *et al.*, 1996, Konforti *et al.*, 1998a). Dentro de las interacciones esenciales para el anclaje del DV está la  $\zeta$ - $\zeta'$  que tiene lugar entre el *tetraloop* del ápice del DV y un receptor en el DI (Keating *et al.*, 2008). Una segunda interacción es la  $\kappa$ - $\kappa'$ , la cual soporta la hélice inferior del DV a través de una red de contactos con la base y el esqueleto (Boudvillain y Pyle, 1998; Toor *et al.*, 2010).

#### 1.2.1.6. Dominio VI

Este dominio es importante porque en él encontramos una adenosina desapareada (A) muy conservada a 7-8 nt del final del intrón (Michel *et al.*, 1989). Esta A participa en la primera reacción de *splicing* proporcionando el 2'-OH que actúa como nucleófilo y ataca al primer nucleótido del intrón (5'-ss). Por ser un residuo puntual, conservado e importante en la catálisis, se han desarrollado multitud de estudios relacionados con su mecanismo de acción, el mecanismo por el cual se reconoce, su implicación en la selección del 3'-ss, etc. (revisado en Lehmann y Schmidt, 2003; Molina-Sánchez, 2008.)

Se sabe que el DVI puede entrar y salir del interior del intrón, manteniendo algunos contactos terciarios y formando así una red estable de interacciones (DibHajj *et al.*, 1993; Chin y Pyle, 1995; Chu *et al.*, 1998; 2001). La estructura dinámica de este dominio podría ser importante para su función. Cuando el DVI se localiza dentro del corazón del intrón, participa en la reacción de *branching*, esta posición debe interactuar con la región de coordinación y con otras regiones del intrón (DeLencastre *et al.*, 2005; Hamill y Pyle, 2006).

Cuando el DVI se encuentra totalmente fuera del corazón del intrón interactúa rápidamente con el DII. Esta interacción ( $\eta$ - $\eta'$ ) está muy conservada y une el *tetraloop* del extremo del DVI con una región distal del DII (Costa *et al.*, 1997). Se cree que esta interacción puede jugar un papel regulador clave en la función del DVI, protegiendo a este dominio de la degradación por el centro activo. Es posible que el DVI se encuentre “esperando” en conformación  $\eta$ - $\eta'$  hasta que el 5'-ss se encuentre dentro del centro catalítico y configurado adecuadamente para sufrir el primer paso del *splicing*. Otro papel regulador de esta interacción y de la conformación “en silencio” del DVI está unido al hecho de que los intrones del grupo II pueden escindir-se mediante otros mecanismos diferentes al *branching* (como veremos más adelante), por eso, la interacción  $\eta$ - $\eta'$  puede conferir de forma inherente una flexibilidad mecánica a los intrones del grupo II.

#### 1.2.1.7. Subdominios inusuales

**La inserción I(i):** un pequeño pero interesante conjunto de intrones del grupo II contiene una inserción que se localiza en el tallo inicial del DI (fig. 3) en la hélice I(i), adyacente al bucle interno semiconservado que forma parte de la región de anclaje Z que se observa en los intrones del grupo IIC (Michel *et al.*, 2007) Esta inserción se observó por primera vez al caracterizar el intrón de *Azotobacter vinlandii* (Adamidi *et al.*, 2003; Ferat *et al.*, 2003). Tras esta identificación inicial, se han descrito múltiples variantes clasificándolos en un linaje distinto, dentro de la clase CL1 (Michel *et al.*, 2007). Estos intrones, los cuales contiene esta inserción aproximadamente en la misma posición, tienden a invadir regiones limitantes con genes (Michel *et al.*, 2007; Simón *et al.*, 2008). Se ha propuesto que las secuencias insertadas pudieran contener 5'-ss alternativos y secuencias IBS1, las cuales podrían actuar como reguladores del *splicing* (Ferat *et al.*, 2003; Michel *et al.*, 2007).

**La inserción en 3'-terminal o Dominio VII:** aunque la secuencia nucleotídica de los intrones del grupo II no se encuentra altamente conservada, sí tienen una estructura secundaria distintiva compuesta por seis dominios que irradian desde una unión central (Michel *et al.*, 1989). Fue por lo tanto, sorprendente, cuando se descubrió una nueva clase de intrones del grupo IIB que contenían, aparentemente, un séptimo dominio localizado entre el DVI y el 3'-ss ((Tourasse *et al.*, 2005; Stabell *et al.*, 2007; 2009). Aunque quizás no sea un dominio como tal, esta inserción ha sido identificada en varios intrones del grupo IIB filogenéticamente relacionados. En algunos casos, esta nueva estructura influye en el splicing del intrón (Stabell *et al.*, 2009).

#### 1.2.1.8. Regiones interdominio

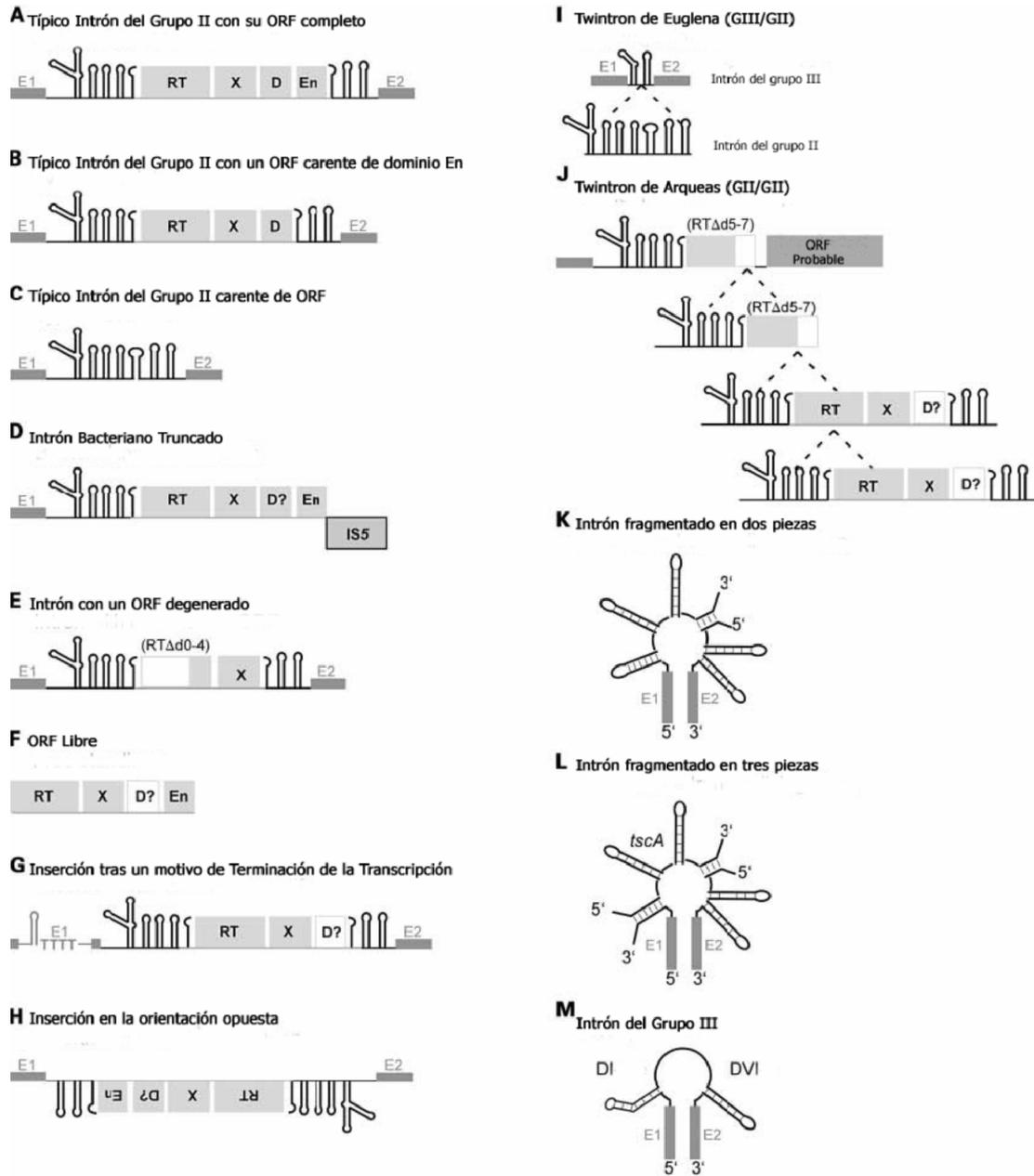
La mayoría de los nucleótidos que conforman la región central de la que irradian los seis dominios de los intrones del grupo II están altamente conservados y juegan un papel importante en la catálisis de estos elementos genéticos. Se sabe que los nucleótidos iniciales del intrón (G3, C4 y G5) participan en interacciones terciarias  $\epsilon$ - $\epsilon'$  y  $\lambda$ - $\lambda'$  (Fig. 3) críticas en la catálisis. Múltiples estudios han demostrado la importancia de la región JII/III entre el DII y el DIII en la catálisis de los intrones del grupo II (Podar *et al.*, 1998b; Mikheeva *et al.*, 2000; Fedorova *et al.*, 2003; DeLencastre *et al.*, 2005; Fedorova y Pyle, 2005). Esta región contiene dos de los nucleótidos más conservados en todo el intrón, los cuales se cree que son importantes para el *splicing*. La interacción  $\gamma$ - $\gamma'$  (Fig. 3) también se da entre esta región y el último nucleótido del intrón, residiendo su importancia en la segunda reacción del *splicing* y en la selección del 3'-ss correcto (Jacquier y Michel, 1990; Jacquier y Jacquesson-Breuleux, 1991). La función catalítica del JII/III depende estrictamente de su conexión física al tallo basal del DII, lo cual implica que el tallo del DII debe facilitar la formación o el posicionamiento de la estructura JII/III (Fedorova *et al.*, 2003).

La presencia de tres nucleótidos que unen el DV y el DVI es uno de los factores determinantes para la selección del sitio de *branching* en los intrones del grupo IIB1 (Chu *et al.*, 2001). Este hecho es compartido por los intrones IIC, pero no se da en los IIA ni en los CL2 (Toor *et al.*, 2001).

#### 1.2.2. Otros intrones del grupo II

1.2.2.1. Intrones degenerados

Muchos intrones de organelas tienen defectos estructurales que perjudican la actividad de la ribozima. En el DNA, tanto mitocondrial como cloroplastídico, de plantas se codifican cerca de veinte intrones del grupo II, ninguno de los cuales tiene la capacidad de sufrir auto-splicing o de ser móvil (Barkan, 2004). Sus variaciones estructurales incluyen desapareamientos, inserciones y deleciones en los DV y DVI, la ausencia de



**Figura 7.** Representación esquemática de la degeneración observada en la naturaleza en los intrones del grupo II (adaptada de Robart y Zimmerly, 2005 por Barrientos-Durán, 2008). Los exones se muestran como cajas de color gris oscuro, el ARN del intrón como líneas de color negro y el marco de lectura (ORF) como cajas de color gris claro. Los dominios I-VI de la estructura secundaria del ARN del intrón están representados como tallos-bucles (stem loops). Las Figuras A-J representan estructuras de DNA de doble cadena, con cajas y tallos posicionados encima o debajo para indicar la orientación de la cadena. Las Figuras K-M representan estructuras secundarias del ARN. **(A)** Típico intrón del grupo II con todos sus motivos (Ejemplos: *L. lactis* Ll.LtrB, intrones mitocondriales de levadura al1, al2) (Belfort et al., 2002) **(B)** Intrón carente de dominio endonucleasa (En)- frecuente en bacterias (Ejemplo: *S. meliloti* RmInt1) (Martínez-Abarca et al., 2000a). **(C)** Intrón carente de ORF (Ejemplo: intrones mitocondriales al5y, cob111) (Michel et al., 1989). **(D)** Intrón bacteriano fragmentado. En esta ocasión el intrón está truncado por la inserción de la IS5 en orientación opuesta (Ejemplo: Intrón *E. coli* I4) (Dai y Zimmerly, 2002). **(E)** Intrón con un ORF degenerado, probablemente sea inmóvil (El ejemplo utilizado muestra al intrón encontrado en el gen *tmK11* (*matK*) en cloroplastos de plantas; también se ha visto en el gen *nad114* (*matR*) en mitocondrias de plantas superiores y en los cloroplastos de *Euglena psbC14* (Mohr et al., 1993; Zimmerly et al., 2001). **(F)** ORF libre, carente de ribozima (Ejemplos: nMat-1a, nMat-1b, nMat-2a, nMat-2b de genomas nucleares de plantas; *matK* en el genoma mitocondrial de *Epiphagus virginianum*) (Mohr y Lambowitz, 2003). **(G)** Inserción tras terminadores rho-independientes (Ejemplos: intrones bacterianos de la clase IIC) (Dai y Zimmerly, 2002). **(H)** Inserción en la orientación reversa (Ejemplos: el intrón X.f.I1 del gen que codifica la DNA metiltransferasa en *Xylella fastidiosa*; el intrón Th.e.I12 del gen que codifica la glicosiltransferasa de *Thermosynechococcus elongatus*) (Dai y Zimmerly, 2002). **(I)** Twintrons del genoma cloroplastidial de *Euglena*. Estos pueden ser resultado de muchas combinaciones posibles de intrones del grupo II y III. (Ejemplo: Twintron del gen *rps3*) (Copertino et al., 1992). **(J)** Twintron en Arqueobacterias. En la figura se muestra un ejemplo de cuatro intrones solapados, estando cada uno de ellos insertados dentro de una secuencia conservada perteneciente al ORF de otro intrón. Dos de las copias están incompletas, por lo que la escisión secuencial no puede eliminar a todos los intrones, y no existe ningún ORF flanqueando a los intrones (Ejemplo: *M.a.I2-F1/M.a.I1-F1/M.a.I1-3/M.a.I5-1* en *Methanosarcina acetivorans*) (Dai y Zimmerly, 2003). **(K)** Intrón codificado en dos localizaciones genómicas distintas y separadas. La escisión *in trans* produce un ARN funcional desde dos exones separados. Frecuentemente se observa en mitocondrias de plantas (Ejemplos: *nad1-14*, en plantas superiores) (Bonen, 1993). **(L)** Intrón separado en tres fragmentos. Similar al anterior, aunque ahora el intrón está codificado por tres piezas (Ejemplo: *Chlamydomonas psaA11*) (Choquet et al., 1988; Goldschmidt-Clermont et al., 1991). **(M)** Intrones del Grupo III. Encontrados únicamente en el genoma cloroplastidial de *Euglena*, estos intrones son de longitud corta, formas degeneradas formadas por porciones de los DI y VI. (Ejemplos: *psbK*) (Doetsch et al., 2001).

la A prominente del dominio VI, y subdominios que son irreconocibles o que directamente están ausentes (Michel *et al.*, 1989; Bonen, 2008). Los ejemplos más extremos se encuentran en el DNA de los cloroplastos de *Euglena*, el cual contiene hasta 155 intrones que carecen de varios dominios (Fig. 7M).

El más pequeño, un intrón del grupo III, comprende unos 100 nucleótidos con dos estructuras que se asemejan a los dominios I y VI, careciendo del DV (Copertino y Hallick 1993). El *splicing* de estos intrones tan degenerados probablemente requiera la participación *in trans* de RNAs o proteínas que compensen la pérdida de su estructura de RNA.

#### 1.2.2.2. Twintrones

A veces, los intrones del grupo II forman derivados conocidos como *twintron*, que no son más que un intrón insertado dentro de otro (Fig. 7I,J),

reflejando su actividad como elementos móviles (Copertino y Hallick 1993). En algunos casos, el intrón invasor interrumpe una región crítica del otro intrón, siendo necesario el splicing del primero para que el segundo pueda funcionar (Hallick *et al.*, 1993), mientras que en otros casos, el intrón que invade no afecta en ninguna medida al intrón invadido (Nakamura *et al.*, 2002). Algunos twintrones forman auténticas “matrioskas” de intrones debido a la inserción sucesiva de varios elementos interrumpiendo al anterior (Dai y Zimmerly, 2003).

### 1.2.2.3. Intrones del grupo II tipo Trans-splicing

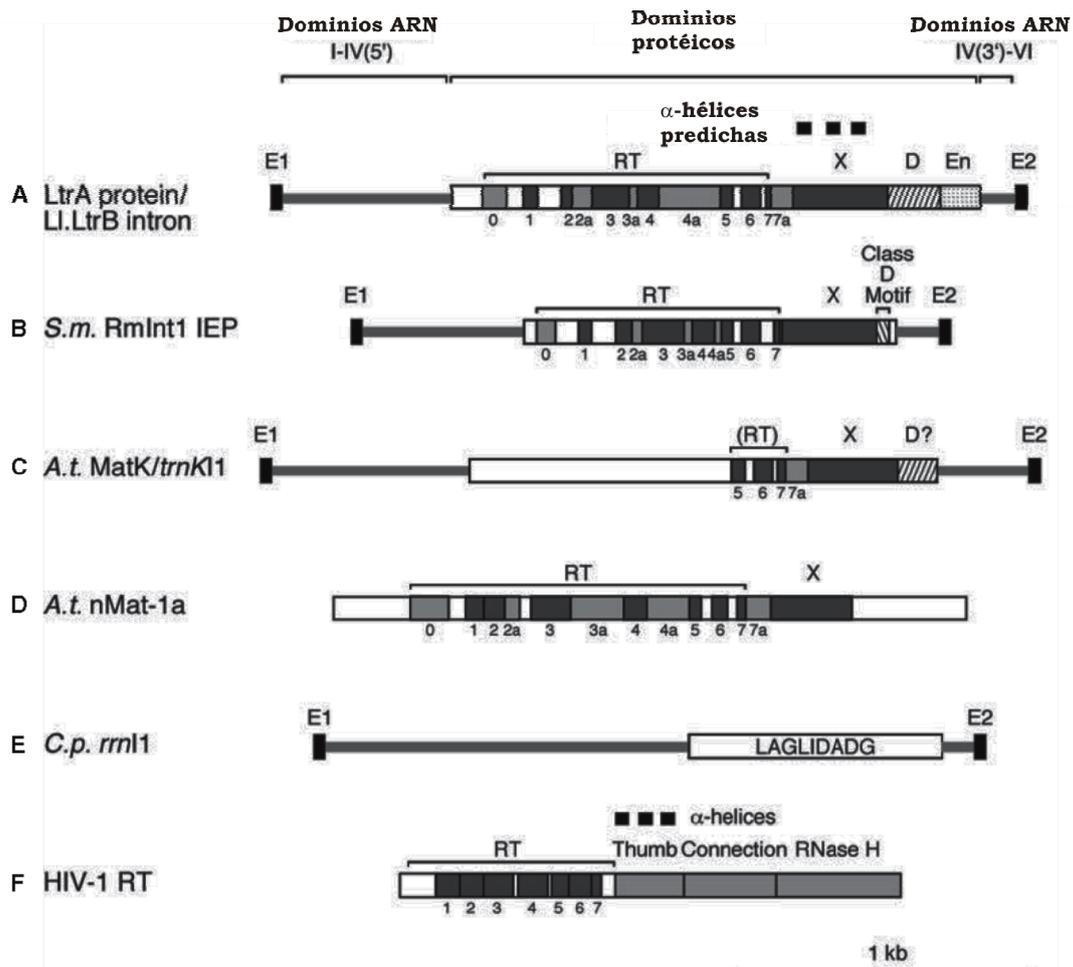
Los intrones del grupo II tipo *trans-splicing* son variantes causadas por reordenamientos genómicos que dividen un intrón en dos o más segmentos que se transcriben de forma separada (Glanz y Kück, 2009). Estos segmentos se reasocian mediante contactos terciarios entre los dominios del intrón del grupo II y actúan en *trans* sobre los exones para producir un RNA mensajero funcional. Estos elementos han surgido de forma natural en los cloroplastos y las mitocondrias de plantas (Qiu y Palmer 2004; Bonen, 2008). Un ejemplo es el intrón *nad1-I4* mitocondrial, que es continuo en algunas plantas con flores, pero en otras sufre *splicing* en *trans*, siendo la única diferencia significativa dos segmentos transcritos de forma independiente (Fig. 7K) (Bonen, 1993). Ejemplos más extremos son el *psaA-I2* del cloroplasto de *Chlamydomonas* spp. y el *nad5-I3* de la mitocondria de algunas angiospermas, los cuales se transcriben en tres segmentos (fig. 7L) (Goldschmidt-Clermont *et al.*, 1991; Knoop *et al.*, 1997)

## 1.3. PROTEÍNAS CODIFICADAS POR LOS INTRONES DEL GRUPO II

### 1.3.1. Características generales

La mayoría de los intrones del grupo II bacterianos y la mitad de los mitocondriales y cloroplastídicos codifican una proteína en su dominio IV, genéricamente llamada IEP (Intron Encoded Protein) y que consta de diferentes dominios y actividades bioquímicas. Las IEPs mejor caracterizadas son las codificadas por los intrones mitocondriales de levaduras *coxI-I1* e *-I2* (Kennell *et al.*, 1993) y la codificada por el intrón de *L. lactis* Ll.LtrB (LtrA) (Matsuura *et al.*, 1997). LtrA (Fig. 8A) consta de cuatro dominios funcionales: reverso transcriptasa (RT); madurasa (X/pulgar); unión a DNA (D) y DNA endonucleasa (En). El dominio RT se puede dividir a su vez en siete bloques

con secuencia aminoacídica conservada (RT1-7) que forman los dedos y la palma de las RTs retrovirales<sup>10</sup>, conteniendo en el RT5 una secuencia altamente conservada (YADD) que forma parte del sitio activo de la RT (Xiong y Eickbush, 1990). Aunque seguramente tengan un plegamiento similar, las RTs de los intrones del grupo II y de los retrotransposones no-LTR son mayores que la de los retrovirus, debido a una extensión aminoacídica terminal (RT0) y a inserciones entre los diferentes bloques (RT2a, 3a, 7a de LtrA; comparar con la RT de HIV-1 en la figura 8F). Algunas de estas inserciones han conservado ciertas características estructurales y pueden ser importantes funcionalmente (Malik *et al.*, 1999; Blocker *et al.*, 2005).



**Figura 8.** IEPs de algunos intrones del grupo II y determinadas proteínas relacionadas. **(A)** Proteína LtrA codificada por el intrón de *L. lactis* L1.LtrB. **(B)** IEP del intrón de *S. meliloti* Rmlnt1, que carece de dominio

<sup>10</sup> La estructura del dominio polimerasa de la RT del HIV-1 se asemeja a una mano derecha humana, en la cual, los distintos subdominios compondrían el pulgar, la palma y los otros 4 dedos (Kohlstaedt *et al.*, 1992).

endonucleasa y que pertenece a la clase bacteriana D (ver más adelante). El “Motivo de la clase D” en el extremo carboxilo es una secuencia conservada que se requiere para el *splicing* y la funcionalidad de los intrones de la clase D (Molina-Sánchez *et al.*, 2010). **(C)** Proteína MatK codificada por el intrón de *A. thaliana trnK1*. Las proteínas MatK conservan la secuencia de los bloques RT5-7 y el dominio X, pero su parte amino terminal ha divergido de las IEPs canónicas de los intrones del grupo II, y carecen del dominio En (Mohr *et al.*, 1993). **(D)** Proteína nMat-1a codificada por un gen nuclear de *A. thaliana*. Las proteínas nMat-1a mantienen los dominios RT y X completos, pero tienen mutaciones que podrían inhibir su actividad RT; las proteínas nMat-2 (no mostradas) también contienen el dominio En, pero también con mutaciones que pueden inhibir su actividad (Mohr y Lambowitz, 2003). **(E)** Proteínas tipo LAGLIDADG codificadas por el intrón de *Cryphonectria parasitica rrm1*. **(F)** RT del HIV-1. Los esquemas de los intrones y de los ORFs están a escala. Las inserciones entre los bloques de la RT se encuentran nombrados como 2a, 3a, 4a y 7a. Existen tres hélices  $\alpha$  predichas, características del dominio “pulgar”, que están representadas encima del dominio X de LtrA (comparar con la RT de HIV-1 en la sección F). (Figura adaptada de Lambowitz y Zimmerly, 2010)

El nombre de dominio madurasa que se le da a veces al dominio X se debe a que ha sido identificado como un sitio en el que determinadas mutaciones afectan al *splicing* del RNA (Mohr *et al.*, 1993; Moran *et al.*, 1994). Se caracteriza por secuencias conservadas y por la predicción de tres hélices  $\alpha$  análogas estructuralmente al dominio “pulgar” de las RTs de los retrovirus (Blocker *et al.*, 2005). Además de en la transcripción inversa, los dominios RT y X/pulgar trabajan juntos en la unión específica al RNA del intrón, promoviendo la formación de la estructura activa de la ribozima y la posición de la proteína para iniciar la síntesis del DNA copia (Saldanha *et al.*, 1999; Wank *et al.*, 1999; Cui *et al.*, 2004).

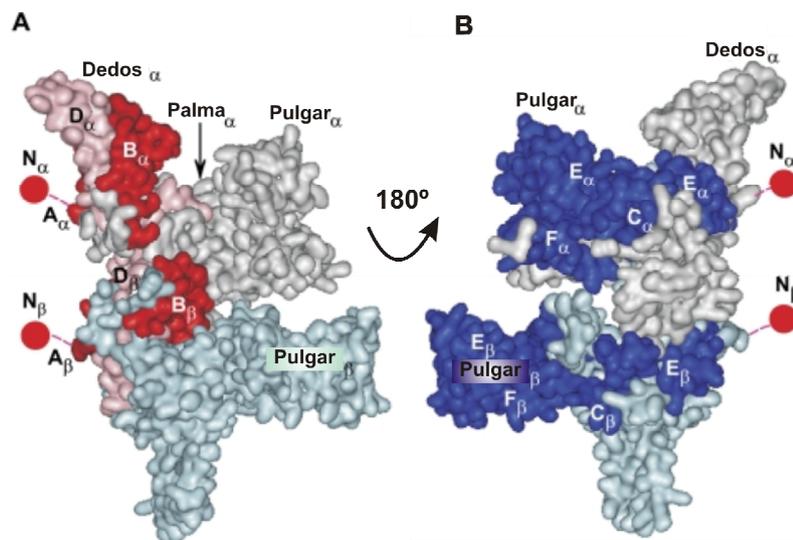
Los dominios D y En interactúan con el DNA diana durante el movimiento del intrón. El dominio D contribuye a la unión al DNA, mientras que el dominio En es una endonucleasa dependiente de  $Mg^{++}$  y corta el DNA diana para generar un cebador de la transcripción inversa (San Filippo y Lambowitz, 2002). Algunos intrones bacterianos del grupo II (el mejor estudiado es RmInt1 de *S. meliloti*) codifican una IEP carente de dominio En y usan un mecanismo diferente para generar el cebador de la transcripción inversa (Fig. 8B) (Estos aspectos de la movilidad serán tratados en posteriores apartados).

Muchas IEPs de intrones del grupo II han perdido secuencias conservadas que se requieren para la actividad RT, pero siguen siendo funcionales durante el *splicing* del RNA, el cual es esencial para la expresión del gen en el que se ha insertado el intrón. Algunas de estas IEPs degeneradas sólo tienen pequeños cambios (mutaciones en el motivo conservado YADD del sitio activo de la RT), mientras que otras se encuentran altamente degeneradas, como les ocurre a algunas IEPs de orgánulos de *A. thaliana* (MatK del cloroplasto, Fig 2C; nMat-1a, -1b, -2a, -2b del núcleo, Fig 2D) (Mohr *et al.*, 1993; Mohr y Lambowitz, 2003).

Por último, existe un pequeño grupo de intrones del grupo II mitocondriales de hongos que codifican IEPs pertenecientes a la familia de

las endonucleasas caracterizadas por el motivo LAGLIDADG (Fig 2E) (Toor y Zimmerly, 2002). Estas enzimas están codificadas en ocasiones por intrones del grupo I, cuya función es promover el *homing* del intrón mediante la rotura de la doble hebra del DNA, promoviendo la recombinación homóloga y, a veces, actuando como maduradoras durante el *splicing* del RNA (Belfort *et al.*, 2002), pero no se conoce aún si su función es similar en los intrones del grupo II.

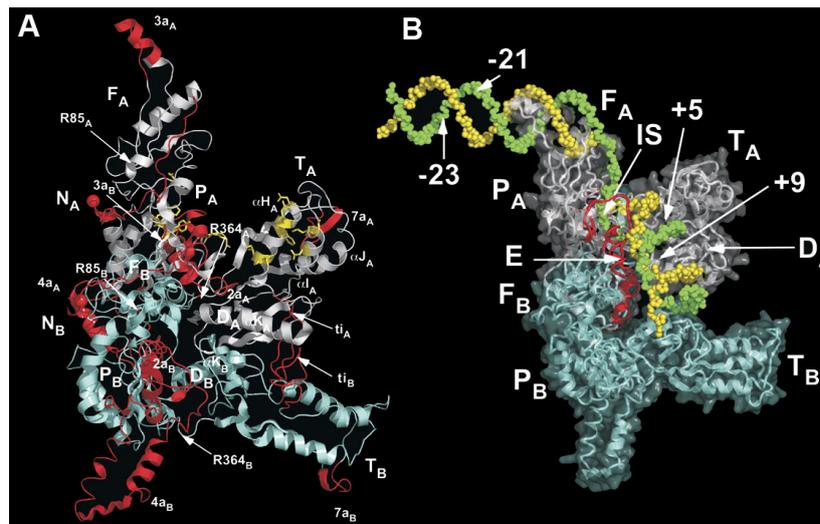
En estudios recientes de la proteína LtrA de *L. lactis* se han identificado determinadas regiones necesarias para la unión de la proteína al DIVa del RNA del intrón (Fig 9): una extensión N-terminal del subdominio RT-0 (región A) y otra región desde RT-0 hasta RT-2 que son esenciales para un *splicing* eficiente del intrón (región B). A la vista de los resultados experimentales se cree que la región que comprende desde RT-3 hasta RT-4 (región D) puede favorecer su unión al DIVa estabilizando la proteína y/o corrigiendo el plegamiento de otras regiones. Finalmente, se han encontrado otras tres regiones (desde RT-2 hasta RT-3 y dos zonas del dominio X; regiones C, E y F) que también son requeridas para el *splicing* de Ll-LtrB (Gu *et al.*, 2010). Existen intrones móviles del grupo II que carecen de algunas de estas regiones, como es el caso de RmInt1 de *S. meliloti*, por lo que deben interactuar con el RNA del intrón de forma diferente (Bloker *et al.*, 2005).



**Figura 9.** Modelo tridimensional de una IEP (ver apartado siguiente) en el que se muestran las regiones requeridas en la unión entre el DIVa y las regiones internas del ARN. (Adaptada de Gu *et al.*, 2010).

### 1.3.2. Modelo tridimensional de una IEP

Debido a que aún no se ha podido cristalizar ninguna IEP de ningún intrón del grupo II, se han tenido que usar otras RTs cristalizadas para conseguir un modelo de IEP. La homología de secuencia entre la RT del HIV-1 y LtrA de *L. lactis* ha permitido construir un modelo tridimensional de LtrA acoplando la secuencia de los dominios RT, X y D en el modelo cristalográfico de rayos X que se dispone de la RT del HIV-1. LtrA es un homodímero (Saldanha *et al.*, 1999), pero hay indicios de que se encuentra como monómero en solución, con dímeros que se ensamblan con el RNA del intrón (Rambo y Doudna, 2004). Por eso, es posible que las dos subunidades adopten estructuras terciarias diferentes dependiendo de las interacciones asimétricas con el DI del RNA del intrón y/o con la otra subunidad. Para la modelización, se asumió que una subunidad de LtrA (subunidad A) tiene una estructura análoga a la subunidad activa p66 de la RT del HIV-1. Se asumió que la segunda subunidad (subunidad B) tenía una estructura similar a la subunidad p51 de la RT del HIV-1. Esta asunción, aunque posiblemente equivocada, era necesaria, ya que la RT del HIV-1 sólo se presenta como dímero.



**Figura 10.** Modelo tridimensional del dímero formado por LtrA obtenido por comparación con la RT de HIV-1 y la endonucleasa VII del fago T4 realizado por Blocker y cols. Cada monómero se representa con un color diferente: la subunidad A en blanco y la B en celeste. Se señalan elementos importantes en la estructura como la palma (P), dedos (F) y pulgar (T), así como el dominio de unión a DNA (D), identificados con el subíndice de cada monómero. (A) En rojo se destacan las principales diferencias respecto de la RT de HIV-1 y en amarillo las coincidencias más fuertes. El extremo N-terminal de ambas subunidades está señalado con una esfera roja y denotado  $N_A$  y  $N_B$ . (B) LtrA unida a su DNA diana. El dominio En está representado en rojo (E) y cada una de las hebras del DNA en amarillo (inferior) y verde (superior). Figura adaptada de Blocker *et al.*, 2005.

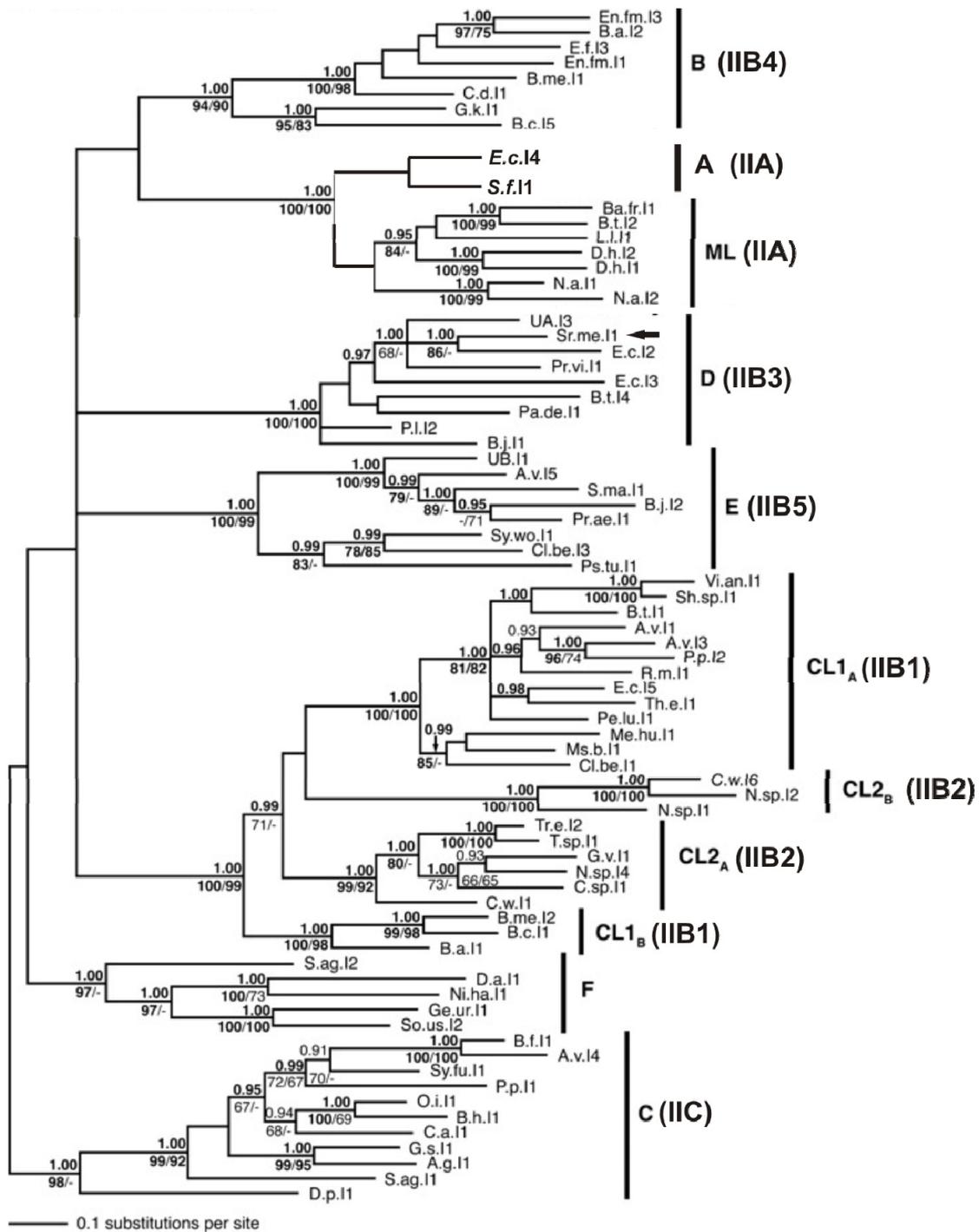
LtrA encaja muy bien en el corazón de la estructura terciaria de la RT del HIV-1, con la subunidad A plegada como los dedos, palma y pulgar (Fig. 9 y 10). La interfase de dimerización entre la palma de p66 y los dedos de p51 de la RT del HIV-1 se mantiene en el modelo para LtrA. Sin embargo, la interfase de dimerización entre los dominios conectores en p66 y p51 son significativamente diferentes, reflejando que la mayoría del dominio de conexión de la RT del HIV-1 es reemplazado por los dominios D y En, aunque su contribución a la dimerización no se conoce. Las hélices  $\alpha$  predichas en el dominio D tienen una gran superficie hidrofóbica que interacciona con regiones distales de la otra subunidad. Pueden existir otras interacciones de dimerización sólo apreciables en presencia del RNA, ya que LtrA dimeriza sólo cuando se une a éste (Rambo y Doudna, 2004).

Este modelo posiciona la mayoría de las inserciones de la RT de LtrA en la espalda de la mano y detrás del pulgar, de forma alejada de la palma y los dedos, por lo que no interferiría en la formación del sitio activo de la RT. El dominio X concuerda bien con el pulgar del modelo de estructura terciaria de la RT del HIV-1, al igual que el dominio D, cuya posición hace pensar en la posibilidad de que participe tanto en la unión a DNA como en la dimerización. Debido a que la identidad entre LtrA y p66 se pierde justo donde comienza el dominio DNA endonucleasa definido en intrones del grupo II (Blocker *et al.*, 2005), a partir de aquí el modelo se ha estructurado en función de la endonucleasa VII del fago T4 con el motivo H-N-H posicionado para producir el corte y la colicina E7 con su sustrato unido. Esta modelización predice que la interacción con el DNA diana va a precisar de la curvatura de este DNA para que la proteína sea capaz de abarcar todas sus posiciones de reconocimiento en ambos exones, hecho que posteriormente se comprueba mediante microscopía de fuerza atómica (Noah *et al.*, 2006).

#### 1.4. CLASIFICACIÓN DE LOS INTRONES DEL GRUPO II

Las IEPs codificadas por cada intrón se unen específicamente a la ribozima del intrón que la codifica, formando una unidad funcional denominada RiboNucleoProteinpartícula (RNP). Como consecuencia, el RNA del intrón y su IEP han co-evolucionado dando lugar a los diferentes linajes de los intrones del grupo II (Fontaine *et al.*, 1997; Toor *et al.*, 2001).

Si atendemos a las diferencias en su estructura secundaria y en secuencia entre los diferentes RNAs de los intrones del grupo II, podemos



**Figura 11.** Relaciones filogenéticas entre intrones bacterianos del grupo II. Esta representación ha sido obtenida a partir del alineamiento de parte de la secuencia nucleotídica del dominio RT de sus IEPs. Con la flecha negra se indica la clasificación de RmInt1. Figura modificada de la que aparece en Simon *et al.*, 2009. Para más detalles ver Figura 2 del artículo original.

distinguir tres clases filogenéticas principales: IIA, IIB y IIC (Michel *et al.*, 1989; Toor *et al.*, 2001, 2002, Zimmerly *et al.*, 2001; Ferat *et al.*, 2003; Toro *et al.*, 2003). Inicialmente, sólo se hizo la distinción entre subgrupos IIA y IIB, subdivididas a su vez en A1, A2, B1 y B2, atendiendo básicamente a detalles estructurales (Michel *et al.*, 1989). Entre esas diferencias se pueden destacar (Fig 3): **(i)** la región desapareada interna donde se encuentran las interacciones  $\epsilon'$  y  $\lambda$  es de  $\sim 14$  nucleótidos (nt) en el caso de los intrones IIA y de 4-5 en los IIB; **(ii)** en los intrones IIA encontramos una adenosina desapareada 7 nt antes del final del intrón (excepcionalmente, 8 nt), mientras que en los IIB suele ser 8 nt; **(iii)** la región donde se encuentran una de las señales de reconocimiento de los exones (EBS1) suele ser asimétrica en el caso de los IIA, sin embargo, los intrones IIB conservan cierta simetría; **(iv)** la región donde se localiza la otra secuencia para el reconocimiento del exón 5' (EBS2) es diferente en los dos casos y está en relación con la longitud de la región distal del DId; **(v)** el tamaño de la región desapareada del DIII varía entre los dos subgrupos, así como la longitud de las regiones interdominio comprendidas entre los dominios I al IV; **(vi)** por último, los intrones IIA terminan con una secuencia consenso YAY mientras que los IIB lo hacen con RAY. La estructura A2 se diferencia de la A1 principalmente porque presenta una inserción entre los dominios I y II, mientras que la B2 difiere de la B1 en una inserción al inicio del intrón, justo en la región opuesta a la inserción que se encuentra en los intrones IIA2 (Michel *et al.*, 1989; Toor *et al.*, 2001).

Posteriormente se observó que estas diferencias se traducen en diferente comportamiento catalítico. En reacciones de auto-maduración llevadas a cabo *in vitro*, se ha determinado que los requerimientos de  $Mg^{2+}$  así como de cationes monovalentes son mayores en el caso de los IIA que, por lo general, suelen ser menos reactivos<sup>11</sup> (Lehmann y Schmidt, 2003); por otro lado, la incidencia de reapertura de los exones<sup>12</sup> una vez ligados es mayor (casi exclusivo) en IIB que en IIA (Schmidt *et al.*, 1990; Hebbbar *et al.*, 1992). Además, estos dos subgrupos tienen en común el mecanismo por el cual determinan el sitio de maduración 5', pero difieren en la selección del sitio de maduración 3' (ver interacciones EBS3-IBS3 y  $\delta$ - $\delta'$  Fig. 4). A esta clasificación se han incorporado los intrones IIC (Grandlund *et al.*, 2001) que

---

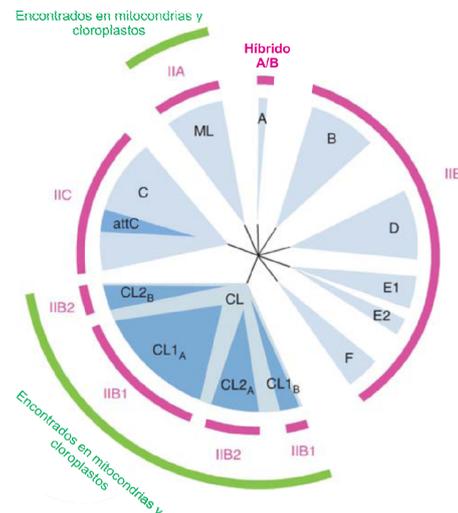
<sup>11</sup> Esta afirmación deriva de comparaciones entre intrones IIB que no poseen ORF ( $\sim 1$  kb) con intrones IIA que sí lo tienen ( $\sim 2,5$  kb). Para conseguir equiparar las eficiencias de auto-maduración, lo que se hizo fue eliminar la mayor parte de esa secuencia, trabajándose con intrones de tamaño similar.

<sup>12</sup> La reapertura de los exones (SER, Spliced Exons Reopening) es una reacción que tiene lugar *in vitro* (aún no se ha determinado *in vivo*), en la cual el propio intrón escindido cataliza el ataque hidrolítico en el punto de ligación de los exones dando lugar a la liberación de los exones (Jarrell *et al.*, 1988; Daniels *et al.*, 1996).

se caracterizan por no presentar EBS2 y tener una secuencia EBS1 de menor extensión que los IIA y IIB (3-4 nt en vez de 6 nt). Además suelen encontrarse insertados aguas abajo de terminadores de transcripción independientes de *rho*, estructuras en horquilla que se piensa podrían conceder parte de la especificidad perdida con el EBS2 para el reconocimiento de una nueva localización donde insertarse (Martínez-Abarca y Toro, 2000a; Grandlund *et al.*, 2001; Toor *et al.*, 2006; Michel *et al.*, 2007). Además se han descrito otra serie de diferencias estructurales: **(i)** el dominio V es más corto que en IIA y IIB, y en lugar de la triada catalítica característica, AGC, presentan los residuos CGC; y, **(ii)** la región donde se describe la interacción  $\varepsilon'$  tiene 11 nt y conserva el motivo AGGAA (ver figura 3).

Existe otra clasificación filogenética considerando la secuencia del dominio RT de la IEP<sup>13</sup>, dividiéndolos en ocho linajes diferentes: clases bacterianas A-F, ML (tipo mitocondrial) y CL (tipo cloroplastídico, que a su vez se subdivide en CL1<sub>A</sub>, CL1<sub>B</sub>, CL2<sub>A</sub> y CL2<sub>B</sub>) (Martínez-Abarca y Toro, 2000a; Toor *et al.*, 2001; Zimmerly *et al.*, 2001; Toro, 2003; Simon *et al.*, 2008; 2009) (Fig 11).

Cada linaje de IEP se asocia con un subgrupo de RNA específico (Fig 12): ML con IIA, la clase bacteriana C con IIC, y los demás con IIB. Las IEPs CL se asocian con RNAs IIB1 y IIB2, mientras que las clases bacterianas A, B, D, E y F se asocian con estructuras IIB típicas (Simon *et al.*, 2009, Lambowitz y Zimmerly, 2010) sugiriendo una co-evolución de ambos componentes (Martínez-Abarca y Toro, 2000a; Toor *et al.*, 2001). Por esta razón algunos autores han propuesto una clasificación que unifique ambos criterios, la estructura de la ribozima y los dominios RT de la IEP (Toro, 2003; Michel *et al.*, 2007): IIA, IIB (subdividido en otros 5 grupos, 1-5) y IIC



**Figura 12** Correlación entre los linajes estructurales (magenta) del ARN de los intrones del grupo II, con los linajes secuenciales de las IEPs (azul). Todos los linajes se encuentran en bacterias, mientras que los encontrados en organelas se muestran en verde. (Modificado de Lambowitz y Zimmerly, 2010).

<sup>13</sup> Los dominios RT 0-7 se encuentran bastante bien conservados entre intrones bacterianos e intrones de mitocondrias, sin embargo el dominio X apenas se conserva entre estos dos subgrupos. Esta divergencia puede ser el resultado de una fuerte presión evolutiva sobre el dominio X en su adaptación para reconocer su intrón de forma específica y asistirlo en sus funciones. En ninguno de los estudios realizados para la clasificación de las IEP se ha considerado el dominio En porque se considera que se ha adquirido en múltiples ocasiones por distintos intrones durante la evolución (Zimmerly *et al.*, 2001).

(Fig 11).

Un dato a resaltar es que en bacterias se dan todos los tipos de familias de intrones del grupo II (Simon *et al.*, 2008), mientras que en mitocondrias y cloroplastos sólo se dan los ML y CL, respectivamente, y no otros linajes. Esta distribución puede indicar que los intrones ML y CL estaban ya presentes en los endosimbiontes bacterianos que colonizaron a los organismos eucariotas (Lambowitz y Zimmerly, 2010).

Casi todos los intrones procariotas descritos hasta nuestros días presentan IEP; sólo algunos intrones identificados en la bacteria termófila *Thermosynechococcus elongatus* y en la arquea *Methanosarcina acetivorans* (*M. acetivorans*) carecen de proteína en su secuencia (Dai y Zimmerly, 2003). En eucariotas inferiores, aproximadamente la mitad de los intrones del grupo II contienen IEP, pero en plantas superiores el porcentaje de intrones con IEP es considerablemente menor. Hay que resaltar que la mayoría de los intrones degenerados se localizan en intrones de cloroplastos.

Este patrón de distribución podría marcar la tendencia de los intrones en su evolución: el origen de los intrones del grupo II que codifican su propia proteína probablemente sea bacteriano y su dispersión hacia organismos con mayor complejidad estructural ha ido asociada a la pérdida de la IEP (Toor *et al.*, 2001). Otra explicación posible es que estos intrones en orgánulos funcionan como verdaderos intrones, mientras que en bacterias se comportan exclusivamente como retroelementos egoístas (Dai y Zimmerly, 2002). La presencia de intrones con dominios RT degenerados en *Euglena*, *Chlamydomonas* y una amplia variedad de plantas terrestres (Copertino y Hallick, 1993; Toor *et al.*, 2001), o de intrones en cuyo dominio IV quedan remanentes de lo que pudo ser una IEP (se han encontrado los codones de inicio y final de traducción además de algunos aminoácidos en *M.a.I6* de *M. acetivorans*; Dai y Zimmerly, 2003), apoyan las teorías anteriores.

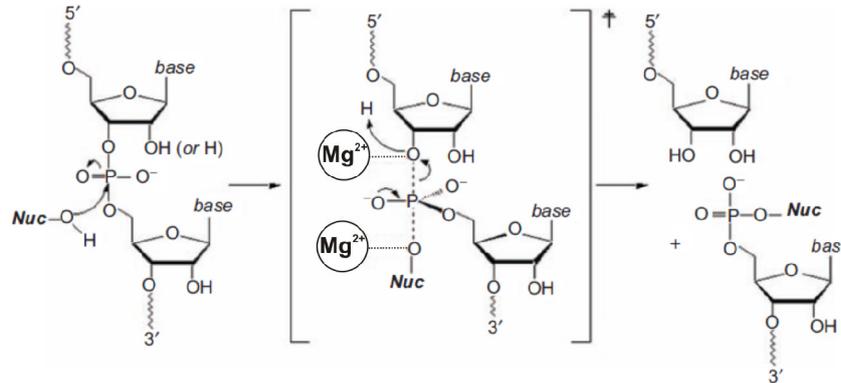
Otra conclusión que puede extraerse del análisis de los árboles filogenéticos es que los fenómenos de transferencia horizontal son muy habituales en procariotas (englobando bacterias y arqueas) (Zimmerly *et al.*, 2001; Dai y Zimmerly, 2002; Belhocine *et al.*, 2004). Los aspectos evolutivos serán tratados más adelante.

## **2. MECANISMOS DE ESCISIÓN EN LOS INTRONES DEL GRUPO II**

El gran impacto de los intrones del grupo II sobre los genomas de organismos vivos es producto de los diferentes mecanismos de auto-*splicing* y retrotransposición. Los intrones del grupo II en general se escinden a partir

de un pre-mRNA como un lazo, estructura adoptada también por los intrones espliceosomales (van der Veen *et al.*, 1986). Esto hace que sean un sistema modelo válido para estudiar el mecanismo de *splicing* en profundidad (Jacquier, 1990). Aunque los estudios *in vivo* han aportado una información valiosa sobre la biología y los mecanismos de los intrones del grupo II, los resultados más importantes provienen del análisis *in vitro* de su actividad catalítica.

Los intrones del grupo II catalizan la rotura de un enlace fosfodiéster y



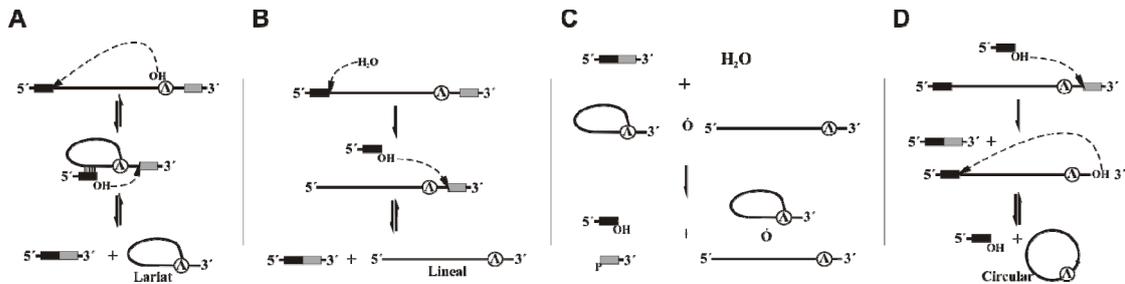
**Figura 13.** Mecanismo químico de catálisis de los intrones del grupo II. Representación de la reacción básica que ocurre durante las reacciones de *splicing* directo y reverso involucrado en la movilidad del intrón. "Nuc" puede representar el grupo 2'-OH del sitio de *branching*, el grupo 3'-OH del exón 5', el final del intrón o una molécula de agua. Los iones metálicos están representados en sus posiciones dentro de la estructura cristalina.

reacciones de ligación tanto sobre moléculas de DNA como de RNA, derivando en varios mecanismos posibles para escindir y migrar a nuevos sitios (Pyle y Lambowitz, 2006). La reacción química básica catalizada por los intrones del grupo II es el ataque nucleofílico de un alcohol o de una molécula de agua sobre un enlace fosfodiéster, resultando en la liberación de un grupo 3'-OH y de un grupo 5'-P o nuevos enlaces fosfodiéster (Fig. 13) (Pyle, 2008). Esto resulta en la inversión de la configuración del fósforo (Podar *et al.*, 1995b). En esta catálisis están implicados dos iones metálicos divalentes que coordinan el ataque nucleofílico y la liberación del grupo 3'-OH (fig 13) (Gordon *et al.*, 2007; Gordon y Piccirilli, 2001; Sontheimer *et al.*, 1999). Por todo ello, desde hace tiempo se creía que los intrones del grupo II son metaloenzimas, con un mecanismo de catálisis parecido a otras enzimas fosfotransferasas. Ésto es consistente con la estructura cristalina de los intrones del grupo II, en la cual, se han observado dos iones metálicos en el sitio activo. (Toor *et al.*, 2008; 2009).

## 2.1 BRANCHING O FORMACIÓN DE INTRÓN LARIAT

Esta es la ruta principal para el auto-*splicing* de los intrones del grupo II que se da a través de dos reacciones consecutivas de transesterificación (Peebles *et al.*, 1987; Jarrel *et al.*, 1988), en la cual se libera el intrón en forma de lazo o *lariat* (Fig. 14A) (Perlman y Butow, 1989; Perlman y Podar, 1996). Esta ruta comienza con el ataque de un grupo 2'-OH de la adenosina desapareada dentro del DVI, residuo que también se conoce como punto/sitio de ramificación o de *branch* (van der Veen *et al.*, 1986; 1987a), al sitio de escisión 5' permaneciendo unidos de forma covalente al primer nucleótido del intrón (enlace 2'-5'; Peebles *et al.*, 1986, Schmelzer y Schweyen, 1986), dando lugar al exón 5' libre y a un intermediario compuesto por el exón 3' unido al lazo. Este paso es seguido por una segunda reacción de transferencia de grupo fosfato que ocurre al atacar el grupo 3'-OH del exón 5' al sitio de escisión 3', produciendo moléculas libres de intrón en forma de lazo y los exones ligados en mayor o menor medida. En el auto-*splicing* de los intrones del grupo II, la primera reacción es la que limita el proceso ya que la segunda reacción es mucho más rápida que la reacción de *branching*, y el intermediario exón 3'-lazo sólo es detectable cuando ciertas mutaciones o condiciones de reacción inhiben el segundo paso de la reacción (Jacquier y Michel, 1987; Jacquier y Jacquesson-Breuleux, 1991; Chanfreau y Jacquier, 1996; Mikheeva *et al.*, 2000). Incluso después de que el enlace entre el intrón y el exón 5' ha sido cortado, éste permanece unido al intrón mediante apareamiento de bases (EBS-IBS) (Jacquier y Michel, 1987; Jacquier y Jacquesson-Breuleux, 1991) lo que le permite estar posicionado para el posterior ataque y unión al exón 3'.

Ambos pasos de esta reacción son fácilmente reversibles (Pyle y Lambowitz, 2006), y moléculas de intrón libres pueden usar lo que se conoce como *splicing* reverso para insertarse en moléculas diana tanto de DNA como de RNA que contengan secuencias similares al exón 5'. Aunque *in vitro* el *splicing* directo y reverso son catalizadas por el RNA del intrón, *in vivo* existen proteínas que juegan un papel fundamental en la función de los intrones del grupo II. La IEP y proteínas codificadas por el hospedador facilitan la formación de estructuras activas de RNA y son requeridas para la mayoría de los pasos de retrotransposición y movilidad (Belfort *et al.*, 2002; Lambowitz y Zimmerly, 2004; Pyle y Lambowitz, 2006; Beauregard *et al.*, 2008; Solem *et al.*, 2009).



**Figura 14.** Mecanismos de escisión descritos en intrones del grupo II. Los exones están representados por recuadros negros (exón 5' o exón 1) y grises (exón 3' o exón 2) y el ARN del intrón se muestra como una línea continua. La flecha en los dos sentidos refleja la reversibilidad de algunas reacciones. Con flechas discontinuas hacemos referencia a los ataques nucleofílicos que tienen lugar en las diferentes rutas. (Modificada de Toro et al., 2007).

## 2.2 HIDRÓLISIS O FORMACIÓN DE INTRÓN LINEAL

Esta ruta también es común en los intrones del grupo II que sufren auto-*splicing*. En ella participa una molécula de agua como nucleófilo durante la primera reacción de *splicing* (van der Veen *et al.*, 1987; Jarrell *et al.*, 1988; Daniels *et al.*, 1996; Podar *et al.*, 1998b; Vogel y Borner, 2002), generando como productos los exones ligados y el intrón en forma lineal (Fig. 14B). Este proceso ocurre sin problemas en la mayoría de los intrones cuya bioquímica ha sido estudiada, representando una ruta alternativa importante a la formación de *lariat*, tanto *in vivo* como *in vitro* (Pyle y Lambowitz, 2006). La primera reacción en el *splicing* reverso es muy eficiente para los intrones lineales (Roitzsch y Pyle, 2009). Aunque el *branching* es la ruta más habitual por la cual los intrones del grupo II sufren auto-*splicing* y se retrotransponen, está claro que el *splicing* hidrolítico representa un mecanismo alternativo que puede conferir determinadas ventajas bajo condiciones ambientales muy específicas. De hecho, existen determinadas familias de intrones del grupo II que sólo tienen *splicing* bajo la ruta hidrolítica, como son aquellos con el sitio de *branching* mutado (Podar *et al.*, 1998a) o los que, directamente, carecen de él (Vogel y Bonen, 2001; Vogel y Borner, 2002).

## 2.3 FORMACIÓN DE INTRÓN CIRCULAR

Se ha visto que varios intrones del grupo IIB tienen una tercera ruta de *splicing*, que conlleva la liberación del intrón en forma circular (Fig 14D) (Jarrell, 1993; Murray *et al.*, 2001; Li-Pook-Than y Bonen, 2006; Molina-Sánchez *et al.*, 2006). El RNA en forma circular (incluyendo el *lariat* de los intrones del grupo II y las ribozimas tipo intrones del grupo I (Nielsen *et al.*,

2005) son muy resistentes a las nucleasas celulares. Esto les proporciona evidentes ventajas a los elementos genéticos móviles mediante RNA, ya que deben permanecer de forma estable dentro del organismo hospedador (Pyle, 2005). Aunque existen varios modelos posibles para la formación de los círculos, se cree que es iniciada por el ataque de exones 5' libres en el ambiente (producto muy abundante proveniente de la reapertura de exones<sup>14</sup> que ya estaban unidos, Fig 14C). Se cree que los exones 5' libres atacan el sitio de *splicing* 3', dejando libre el extremo 3' y uniéndose al exón 3' (Jarrel, 1993; Murray *et al.*, 2001; Molina-Sánchez *et al.*, 2006). Durante el segundo paso en la formación del círculo, el grupo 2'-OH del nucleótido 3' terminal ataca al sitio de *splicing* 5', liberando al intrón circular y regenerando el exón 5' libre. Sin embargo, los intrones circulares pueden provenir también de intrones lineales que hayan sido circularizados por ligasas del hospedador, como ocurre en intrones que carecen de los nucleótidos necesarios en el sitio de circularización (Li-Pook-Than y Bonen, 2006). De forma interesante, se ha observado que cuando aparecen intrones circulares de RNA, también aparecen de DNA (Murray *et al.*, 2001; Molina-Sánchez *et al.*, 2006) lo que nos lleva a una observación que se realizó hace ya 27 años, y es que, en *Podospora anserina*, la presencia de las moléculas de intrón circular de DNA aumentan en células senescentes (Osiewacz y Esser, 1984).

#### 2.4. SPLICING ALTERNATIVO

Estudios *in vitro* han demostrado que el *splicing* de los intrones del grupo II no se lleva a cabo siempre de forma correcta. Por ejemplo, en el intrón de la clase IIB3 RmInt1 de *S. meliloti*, se ha visto que pueden ocurrir interacciones EBS-IBS alternativas (Costa *et al.*, 2006a; 2006b). El intrón B.h.I1 de *Bacillus halodurans* utiliza sitios de escisión 5' aguas arriba de la estructura canónica del intrón. Estos dos ejemplos sólo se han visto *in vitro*, por lo que podría ocurrir que fuesen artefactos experimentales provocados por las condiciones salinas no fisiológicas o por la falta de la IEP. Sin embargo, el intrón B.c.I4 de *Bacillus cereus* sufre *splicing* de forma natural 56 nt aguas abajo del sitio de escisión 3' canónico (Tourasse *et al.*, 2005), por lo que es probable que existan de forma natural reacciones de *splicing* inusuales.

---

<sup>14</sup> La reapertura de los exones o reacción SER (Spliced Exon Reopening) ( Jarrel *et al.*, 1988; Daniels *et al.*, 1996) es catalizada por la forma lineal o lariat del intrón. Los exones son separados hidrolíticamente generando ambos exones libres y reactivos. Esta reacción parece ser específica de los intrones del grupo IIB. Se observa muy rara vez en los intrones IIA, incluso bajo condiciones óptimas para ello (Schmidt *et al.*, 1990; Hebbar *et al.*, 1992).

## 2.5. PAPEL DE LAS PROTEÍNAS EN LA ESCISIÓN *IN VIVO*

Como ya hemos dicho, la mayoría de los experimentos para dilucidar el plegamiento de los intrones del grupo II y los mecanismos de escisión han sido llevados a cabo *in vitro*, bajo condiciones no fisiológicas. Pero se ha visto que *in vivo*, son varias las proteínas que participan en mayor o menor medida en el splicing de los intrones del grupo II. El papel de algunas de ellas es indirecto e inespecífico, como el de las proteínas de membrana mitocondriales Mrs2p y Lpe10p, cuya única contribución a la escisión de los intrones del grupo II mitocondriales es incrementar la concentración de Mg<sup>2+</sup> en el interior de dicho orgánulo (Gregan *et al.*, 2001; Weghuber *et al.*, 2006). Sin embargo, a pesar de la intervención de estas proteínas, las concentraciones de magnesio intramitocondriales son dos órdenes de magnitud inferiores que las usadas en los experimentos *in vitro*, por lo que deben concurrir otros factores que favorezcan el *splicing in vivo*. Estos elementos pueden ser otros iones, pequeñas moléculas u otras proteínas. Se sabe que existe una amplia variedad de proteínas que interaccionan de forma directa (y a veces específica) con el RNA de los intrones del grupo II, promoviendo la escisión al ayudar en el plegamiento y la estabilización de una estructura catalítica activa a bajas concentraciones de magnesio.

### 2.5.1. Las IEPs y su papel en la escisión del intrón.

Las IEPs (ver apartado 1.3) son proteínas que han sido adquiridas por los intrones y que son codificadas normalmente en el DIV de la secuencia intrónica. Estas proteínas multifuncionales se unen específicamente a su propio mRNA, promoviendo la formación de interacciones terciarias en el RNA imprescindibles para que el splicing sea posible (Matsuura *et al.*, 2001). Mientras que los dominios RT y En son fácilmente identificables ya que tienen una secuencia filogenéticamente conservada, las regiones responsables de la función madurasa no son tan claras. Estudios con mutantes (Cui *et al.*, 2004) han demostrado que el dominio X de la IEP, junto con otras regiones de unión al RNA, parecen ser los responsables de la función de maduración (Mohr *et al.*, 1993). Esta funcionalidad durante el *splicing* (actividad madurasa) se observó por primera vez mediante caracterización genética de las IEPs de los intrones mitocondriales de levaduras a11 y a12 (Carignani *et al.*, 1983; Moran *et al.*, 1994). Posteriormente se estudió la actividad bioquímica de la IEP del intrón Ll.LtrB de *L. lactis* gracias a que ha podido ser expresada y purificada, junto con el

RNA del intrón, en el sistema modelo *E. coli*, facilitando así su caracterización bioquímica (Matsuura *et al.*, 1997). La proteína LtrA purificada se une al RNA precursor aún no escindido en relación 2:1 (lo que sugiere que funciona como un dímero) y en una reacción muy rápida, seguida de una etapa más lenta que conlleva un cambio en la conformación del RNA y finalmente al *splicing* de L1.LtrB (Saldanha *et al.*, 1999).

Estudios bioquímicos han demostrado que LtrA tiene un sitio de unión de alta afinidad al DIVa, una estructura tallo-lazo característica en el inicio del ORF de LtrA, participando en la especificidad de intrón (Wank *et al.*, 1999). La unión de LtrA al DIVa se realiza a través del reconocimiento de bases específicas del bucle terminal y en las regiones despareadas (Watanabe y Lambowitz, 2004). Debido a que el DIVa contiene la secuencia Shine-Dalgarno y el codón de iniciación de la IEP, la unión de la propia IEP al DIVa regula su propia traducción reduciéndola (Singh *et al.*, 2002; Cui *et al.* 2004).

Mientras se encuentra unida al DIVa, LtrA tiene contactos adicionales con regiones conservadas del RNA del intrón estabilizando así su estructura activa (Fig 15A) (Wank *et al.*, 1999; Matsuura *et al.*, 2001; Dai *et al.*, 2008). Si el DIVa es deleccionado, LtrA sigue facilitando el *splicing* debido a estas interacciones, aunque la eficiencia se ve muy reducida. La IEP de al2, que es traducida como una proteína de fusión con el exón 5' y luego procesado por proteólisis, también se une al DIVa y a regiones del centro catalítico, sugiriendo unas interacciones que se conservan aún cuando el DIVa no contiene el codón de iniciación del ORF del intrón (Huang *et al.*, 2003).

Una vez identificados los sitios de unión de LtrA al RNA del intrón (Matsuura *et al.*, 2001; Dai *et al.*, 2008) fueron localizados en el modelo tridimensional del RNA del intrón, revelando una gran y continua superficie de contacto con el dominio IVa y las regiones contiguas de los dominios I, II y VI. Esta superficie de unión sugiere que LtrA facilita la reacción de *splicing* al estabilizar las interacciones entre estos dominios así como la estructura plegada del DI. Después del *splicing*, la IEP permanece fuertemente unida al RNA del intrón escindido, aprovechando, probablemente, la mayoría de las interacciones para estabilizar la estructura de la ribozima en el proceso de *splicing* reverso que se da durante la movilidad del intrón. A pesar de las diferencias estructurales entre los intrones IIA, IIB y IIC, los modelos sugieren que la IEP se une y participa de la misma manera en los tres grupos (Dai *et al.*, 2008).

Las regiones de LtrA que son necesarias para el *splicing* fueron identificadas mediante mutagénesis, y localizadas en el modelo

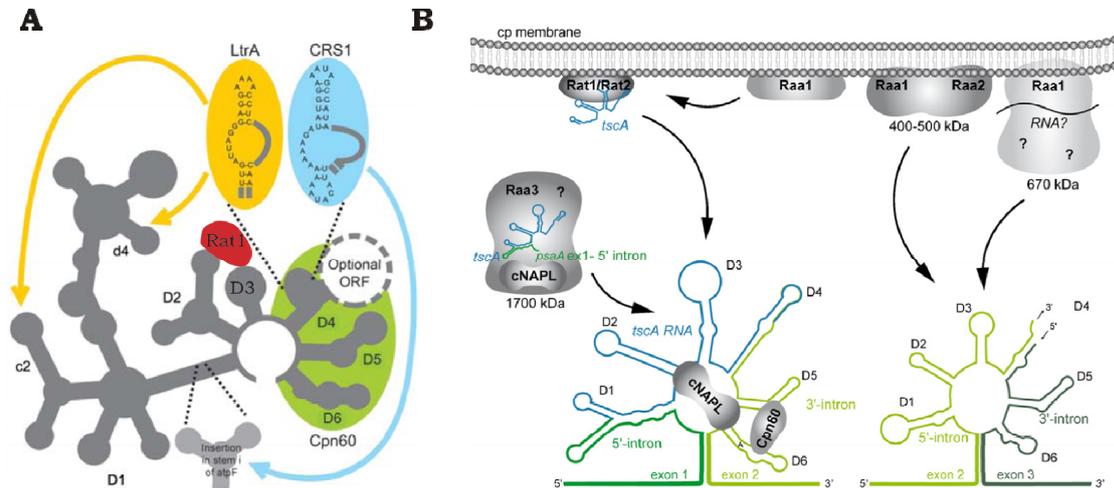
tridimensional de la proteína (Cui *et al.*, 2004; Blocker *et al.*, 2005). Los resultados sugieren que la superficie de unión al RNA se extiende por los dominios RT y X, incluyendo la extensión amino terminal, la región de unión entre el cebador y el molde y determinados puntos en la espalda de la proteína. Experimentos genéticos y bioquímicos han sugerido que la extensión amino terminal de LtrA se une al DIVa, mientras que otras regiones de la RT y del dominio X se unen al “corazón” del intrón (Cui *et al.*, 2004, Gu *et al.*, 2010). A diferencia de las RT de los retrovirus, las RTs de los retroelementos no-LTR se unen de forma específica a sus RNAs molde para iniciar la síntesis del DNA copia durante el splicing reverso. Por lo tanto, las funciones de las RTs de los intrones del grupo II durante la maduración podrían haber evolucionado a partir de interacciones que se usaban inicialmente para el reconocimiento de RNA molde (Kenell *et al.*, 1993).

### **2.5.2. IEPs que participan en el *splicing* de varios intrones del grupo II**

Aunque la mayoría de las IEPs participan de forma exclusiva en el *splicing* del intrón que la codifica, existen algunas que intervienen en el de otros intrones, proporcionando una maquinaria común para la escisión. En el caso más sencillo, una misma IEP puede participar en el *splicing* de varios intrones del grupo II que han proliferado en un mismo genoma (Dai y Zimmerly, 2003; Meng *et al.*, 2005). Otras IEPs, sin embargo, pueden participar en el *splicing* de intrones menos relacionados entre sí. Éste es el caso de la proteína cloroplastídica MatK, codificada por el intrón *trnK11* de plantas terrestres y de forma independiente en algunas plantas no fotosintéticas, que es una IEP altamente modificada y que ha perdido sus funciones en el movimiento de los intrones, pero ha adquirido la habilidad de unirse y participar en el *splicing* de varios intrones del cloroplasto del grupo IIA que carecen de ORF (Fig 8C) (Ems *et al.*, 1995; Vogel *et al.*, 1999; Zoschke *et al.*, 2010). Esta participación en el *splicing* de otros intrones reflejaría la posible pérdida de las interacciones específicas con el intrón, como en las que participa el DIVa, y el incremento de interacciones relacionadas con la estructura compartida por los intrones del grupo IIA.

Un paso más en esta evolución sería el dado por las proteínas nMat (nMat1a, -1b, -2a y -2b) de plantas con flores (Fig 8D; Mohr y Lambowitz; 2003). Estas proteínas, evolucionadas a partir de IEPs de intrones mitocondriales del grupo II, son codificadas por genes del núcleo y son transportadas a los orgánulos para favorecer el *splicing* de intrones del grupo

II que carecen de ORF (Nakawa y Sakurai 2006; Keren *et al.*, 2009). Esta progresión evolutiva en la cual las proteínas han perdido funciones en la movilidad de los intrones del grupo II, lleva a participar en el splicing de múltiples intrones y al final a ser integradas en el genoma del hospedador. Lo cual, limita los posibles daños que pueda producir la movilidad de un intrón y facilita la regulación, por parte del hospedador, de los mecanismos de splicing del RNA y su posible interferencia en otros procesos celulares.



**Figura 15. (A)** Sitios de unión a proteínas en los intrones del grupo II. En gris se ha representado un esquema de la estructura secundaria de un intrón del grupo II. Los óvalos coloreados representan a las proteínas que se asocian con los intrones: LtrA (amarillo), CRS1 (azul), Cpn60 (verde) y Rat1 (rojo). Las flechas indican interacciones con sitios de unión secundarios. **(B)** Estructuras secundarias posibles de los intrones 1 y 2 del ARN de *psaA* y los factores que actúan en *trans* en su splicing (*tscA*, Rat1, Rat2, Raa1, Raa2, Raa3, cNAPL y Cpn60) así como los complejos multiproteicos y ribonucleoproteicos. Figura adaptada de: Fedorova y Zingler, 2007; Jacobs *et al.*, 2010; Solem *et al.*, 2009).

### 2.5.3. Proteínas codificadas por el hospedador

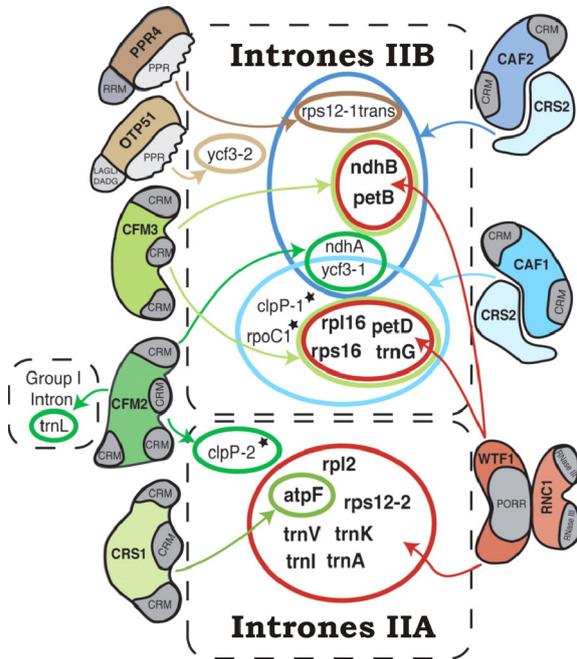
La mayoría de los intrones del grupo II de mitocondrias y cloroplastos no codifican ningún ORF por lo que dependen de otras proteínas codificadas por el hospedador para poder llevar a cabo su *splicing*, ya que no suelen usar IEPs de otros intrones. Esta situación es similar a la de los intrones del grupo I, donde una gran variedad de proteínas del hospedador, como las aminoacil tRNA sintetasas, participan en el *splicing* de distintos intrones (Lambowitz *et al.*, 1999). Los factores del hospedador mejor estudiados son los requeridos por los intrones del grupo II de cloroplastos de plantas (Figs. 15 y 16). Los análisis genéticos, apoyados por estudios *in vivo*, indican que estos intrones utilizan una gran batería de factores de *splicing*, incluyendo proteínas pertenecientes a tres grandes familias de proteínas de unión a

RNA específicas de plantas: CRM (Chloroplast RNA splicing and ribosome Maturation), PORR (reconocimiento del RNA de orgánulos plantas) y PPR (Pentatricopeptide Repeat) (Stern *et al.*, 2010). Estas proteínas funcionan de forma conjunta para facilitar el splicing de diferentes intrones, ya que la mayoría de éstos requieren varios factores de *splicing* que se asocian con el RNA formando grandes complejos (Till *et al.*, 2001; Kroeger *et al.*, 2009).

En el alga verde *Chlamydomonas reinhardtii* se requieren al menos 14 genes nucleares para el splicing en *trans* de sus dos intrones (*psaA-I1* e *-I2*), los cuales parecen funcionar formando grandes complejos, algunos de ellos unidos a la membrana del cloroplasto (Fig 15B) (Goldschmidt-Clermont *et al.*, 1991; Jacobs *et al.*, 2010). Sólo algunos de los componentes han sido identificados y caracterizados, como Rat1 y Rat2, proteínas que muestran homología con polimerasas poli-(ADP-ribosa) (Balczun *et al.*, 2005), Raa2 (antes denominada Maa2) que deriva de una pseudouridín sintasa (Perron *et al.*, 1999), y Raa1, una proteína bifuncional con dos dominios diferentes que es responsable del *splicing* de dos intrones distintos (Merendino *et al.*, 2006). A pesar de su importancia general en el *splicing*, aún no está del todo claro cuales de estos factores están implicados directa o indirectamente. En este alga se ha propuesto que Cpn60 es un factor que participa de forma general en el *splicing* basándose en su unión específica al intrón heterólogo mitocondrial r11 (Bunse *et al.*, 2001). Se asemeja a una proteína de choque térmico con dos dominios ATPasa del tipo Gro-EL y se ha visto que se une al RNA, tanto de intrones homólogos como heterólogos, muy cerca de los dominios DIV y DVI (Fig. 15), aunque su mecanismo de acción aún no se conoce (Balczun *et al.*, 2006). De forma similar se ha visto que cNAPL, una proteína similar a las de ensamblaje del nucleosoma, interacciona con el RNA de intrones del grupo II (Glanz *et al.*, 2006), pero tampoco se conocen todavía ni su sitio de unión ni su función.

CRS2 es una proteína de maíz que deriva de las tRNA peptidil hidrolasas (Jenkins y Barkan, 2001) y co-sedimenta con el RNA de intrones del grupo II. Sin embargo, no presenta una unión específica con intrones, pero sí interacciona con otras dos proteínas, CAF1 y CAF2, que sí interaccionan con intrones individuales (Ostheimer *et al.*, 2003) y pertenecen al grupo de las CRM (Fig. 16) (Barkan *et al.*, 2007). Otra proteína importante perteneciente a las CRM es CRS1, que es un factor de *splicing* que favorece de forma específica la escisión del intrón *atpF* uniendo y compactando su RNA. CRS1 interacciona con una región del DIV que es homóloga al sitio de unión de la madurasa de Ll.LtrB, contactando de forma adicional con una inserción en el DI (Fig 15a y 16). Se sabe que *atpF* no se pliega de forma

correcta *in vitro*, incluso bajo altas concentraciones de magnesio, por lo que además de disminuir los requerimientos de este ión, esta proteína del hospedador es imprescindible para el plegamiento correcto (Ostersetzer *et al.*, 2005).



**Figura 16.** Proteínas codificadas por el núcleo que participan en el splicing de intrones del grupo II de angiospermas. Los intrones encontrados en *Arabidopsis* pero no en maíz están marcados con un asterisco. Los factores de splicing se muestran en el exterior, indicando los dominios conservados. Además de su papel principal sobre *ycf3*, *OTP51* favorece el splicing de los intrones *atpF*, *trnV*, y *trnK* (Falcon de Longevialle *et al.*, 2008). Figura adaptada de Kroeger *et al.*, 2009).

En cloroplastos de angiospermas se ha visto que cada uno de los intrones del grupo II allí presentes necesita de, al menos, una proteína codificada en el núcleo para llevar a cabo su splicing. Las RNP intrónicas más complejas son las pertenecientes a los intrones *ndhB*, *petB*, *petD* y *trnG* que necesitan hasta cinco factores de splicing para cada uno (Fig 16). Cada una de estas proteínas son necesarias para el splicing *in vivo*, y todas ellas co-immunoprecipitan y co-sedimentan con los RNAs de los intrones con los que interacciona. La masa molecular combinada de estas proteínas (unos 300 kDa) supera a la de los intrones con los que se asocian (250 kDa). Por lo tanto, estas RNPs son más proteína que RNA, como les ocurre a los ribosomas y al espliceosoma. Esta

situación podría parecerse a la que se da en los cloroplastos de *C. reinhardtii*, donde sus dos intrones del grupo II requieren de múltiples proteínas para su splicing (Perron *et al.*, 1999; Merendino *et al.*, 2006; Goldschmidt-Clermont *et al.*, 1990). Sin embargo, los intrones de *C. reinhardtii* se transcriben en fragmentos que sufren splicing en *trans* necesitando, seguramente, factores especializados en el ensamblaje del intrón (Fig 15B). Por el contrario, la complejidad de las RNPs de los intrones de los cloroplastos de plantas no tiene precedente ni en bacterias ni en hongos (Fedorova y Zingler, 2007). Las proteínas involucradas en el splicing cloroplastídico están altamente conservadas tanto en monocotiledoneas como en dicotiledoneas (Asakura y

Barkan, 2006; 2007; Asakura *et al.*, 2008), pero son distintas a las de *C reinhardtii*, correlacionándose con el origen diferente de los intrones de cloroplastos de plantas terrestre y de las clorófitas. Por lo tanto, parece que hay varias rutas evolutivas para la adquisición de interacciones con proteínas que conlleven la “degradación” de ribozimas autocatalíticas en enzimas proteín-dependientes.

Además de los factores de *splicing* mencionados anteriormente, un *splicing* eficiente de los intrones del grupo II requiere de la participación de chaperonas de RNA para resolver estructuras estables inactivas o intermediarias que limiten el plegamiento de forma correcta del intrón (Huang *et al.*, 2005; Köhler *et al.*, 2010). Una proteína clave para este papel es Mss116p de mitocondrias de *S. cerevisiae*, un miembro de la familia de las proteínas DEAD-box (helicadas de RNA dependientes de ATP)<sup>15</sup> que participa en reordenamientos estructurales del RNA (Séraphin *et al.*, 1989; Huang *et al.*, 2005). Mutantes disfuncionales en Mss116p son defectivos en el *splicing* de los intrones mitocondriales. Dos de éstos, aI1 y aI2, son intrones móviles de la clase IIA que codifican una IEP con actividad RT/madurasa que estabiliza su estructura activa, mientras que otros dos intrones, aI5 $\gamma$  y bI1, son pequeños intrones de la clase IIB que no codifican proteínas. Mss116p puede ser sustituida en estudios *in vitro* por otras proteínas DEAD-box de *S. cerevisiae* (Ded1p; Del Campo *et al.*, 2009) o de otros organismos: CYT-19 de *Neurospora crassa* (Hunag *et al.*, 2005), SrmB de *E. coli* (Del campo *et al.*, 2009). Por todo esto, Mss116p y las demás proteínas DEAD-box parecen interactuar inespecíficamente con sustratos de RNA o RNP, como es de esperar para las chaperonas de RNA. Los estudios bioquímicos han confirmado estas interacciones no específicas, pero no descartan que tengan otro tipo de capacidad, como la de unión y estabilización estructural del RNA contribuyendo así al *splicing* de los intrones del grupo II (Mohr *et al.*, 2006; Solem *et al.* 2006; Halls *et al.*, 2007; Del Campo *et al.*, 2007, 2009). La participación de las proteínas DEAD-box en el *splicing* de los intrones del grupo II tiene un importante paralelismo con el espliceosoma, en donde estas proteínas y otras helicadas de RNA relacionadas participan en múltiples pasos para acelerar determinadas transiciones estructurales.

---

<sup>15</sup> Las proteínas DEAD-box fueron identificadas al final de los años 80 y se vio que estaban involucradas en el metabolismo del ARN. Contienen nueve motivos conservados, y es en el motivo II donde se encuentra la secuencia de aminoácidos conservada que le da nombre a la familia (D-E-A-D: asp-glu-ala-asp) (Linder *et al.*, 1989). Se ha visto que varios miembros de la familia actúan como helicadas, separando cadenas de ARN en un proceso dependiente de ATP (Pyle, 2008).

### 3. MOVILIDAD DE LOS INTRONES DEL GRUPO II

Además de RNAs catalíticos, algunos intrones del grupo II son elementos genéticos móviles con especificidad de diana. Los intrones pueden insertarse en otros sitios del genoma idénticos a aquellos a partir de los cuales se escindieron, en este caso el proceso se conoce como *homing* (Skelly *et al.*, 1991; Lazowska *et al.*, 1994; Moran *et al.*, 1995); pero también pueden moverse a sitios ectópicos, diferentes a su DNA diana habitual, proceso que se denomina transposición (Mueller *et al.*, 1993b; Sellem *et al.*, 1993). El *homing* se produce con una frecuencia cercana al 100%, es decir, casi siempre que la diana específica del intrón no esté invadida, el intrón va a encargarse de interrumpirla, mientras que la transposición es un proceso mucho menos frecuente (Cousineau *et al.*, 2000; Muñoz-Adelantado *et al.*, 2001). Al nombre de estos procesos se les añadió el prefijo retro- ya que estas rutas conllevan la participación de un intermediario de RNA, y no de DNA como ocurre en el *homing* de los intrones del grupo I, así la denominación pasa a ser *retrohoming* y retrotransposición (Curcio y Belfort, 1996; Ichiyanagi *et al.*, 2002). Esto quedó demostrado mediante la interrupción de un intrón del grupo II con un intrón del grupo I (ambos intrones se escinden tras la transcripción). De este modo, la maduración del intrón del grupo II está condicionada al *splicing* del intrón del grupo I. Si esto ocurre, el proceso de *homing* tiene lugar y el análisis del producto de invasión mostró que el intrón del grupo I había desaparecido (Cousineau *et al.*, 1998; Martínez-Abarca *et al.*, 2004).

Esta capacidad de insertarse en nuevas localizaciones del genoma de su hospedador reside en los complejos conocidos como RiboNucleoProteín-partículas (RNPs). Tras la transcripción del intrón, tiene lugar la traducción de la IEP; ésta interacciona con el RNA transcrito precursor y facilita su maduración. Una vez han ocurrido todos estos procesos, la proteína permanece asociada al RNA escindido formando estos complejos.

#### 3.1. MECANISMOS DE RETROHOMING

Los mecanismos de movilidad usados por los intrones del grupo II comenzaron a descifrarse en la década de los 90 primero con la caracterización de la reacción de corte e inserción en el DNA diana por parte de los intrones mitocondriales de levadura, aI1 y aI2 (Eskes *et al.*, 1997; Yang *et al.*, 1996; Zimmerly *et al.*, 1995a; 1995b), y tras su descubrimiento,

también en bacterias. (Mills *et al.*, 1997; Matsuura *et al.*, 1997; Cousineau *et al.*, 1998).

La principal diferencia entre el *retrohoming* de intrones móviles presentes en mitocondrias de levaduras y los intrones bacterianos es el fenómeno conocido como co-conversión<sup>16</sup> de los exones flanqueantes (Belfort *et al.*, 2002). En levaduras se han descrito dos mecanismos de movilidad en los que, o bien sólo el exón 1 o ambos exones, pueden acompañar al intrón en su movimiento; por el contrario, en bacterias no existe este fenómeno, el intrón se inserta limpiamente en su nueva diana.

Son muy pocos los intrones del grupo II que se haya demostrado que sean móviles a pesar de que muchos de ellos son capaces de madurar (*in vivo* y/o *in vitro*) y formar RNPs. Aún así, cada uno de estos intrones presenta particularidades en el mecanismo de *retrohoming* que lo diferencian de los demás.

### 3.1.1. Reconocimiento de la diana

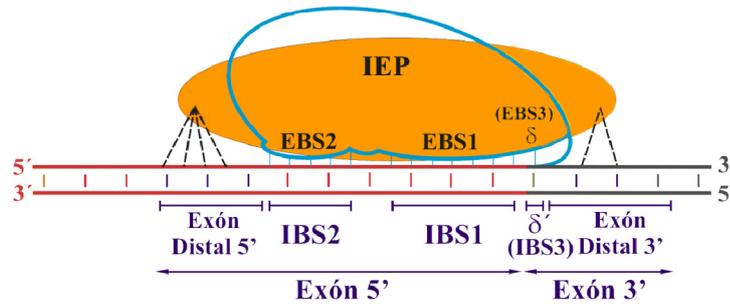
Las secuencias diana de DNA son reconocidas por parte de los intrones del grupo II a través de la IEP y de apareamiento de bases con el RNA del intrón gracias a los complejos de RNPs. En todos los casos hay un apareamiento de bases entre el RNA del intrón y las secuencias de los exones 5' y 3' a través de las interacciones ya mencionadas EBS2/IBS2, EBS1/IBS1 y  $\delta$ - $\delta'$  (intrones IIA) o EBS3/IBS3 (intrones IIB), mientras que la IEP reconoce secuencias adicionales aguas arriba y abajo (Fig 17) (Guo *et al.*, 1997; Yang *et al.*, 1998; Mohr *et al.*, 2000; Singh y Lambowitz, 2002; Jiménez-Zurdo *et al.*, 2003). La importancia de los apareamientos de bases ha sido demostrada gracias a que determinadas mutaciones en la diana podían ser recuperadas por mutaciones complementarias en el RNA del intrón. Las secuencias diana de DNA reconocidas por la IEP varían incluso entre intrones muy relacionados, como *coxI-11* e *-12*, sugiriendo que la IEP puede evolucionar para reconocer dianas diferentes, un hecho que puede ayudar a los intrones del grupo II a establecerse en nuevas localizaciones. Además, son pocas las posiciones reconocidas por la IEP, por lo que la especificidad hacia el DNA diana viene dada principalmente por el RNA del intrón.

Los intrones cuyas IEPs contienen dominios de unión a DNA (D) y endonucleasa (En) (L.LtrB y los intrones mitocondriales de levaduras),

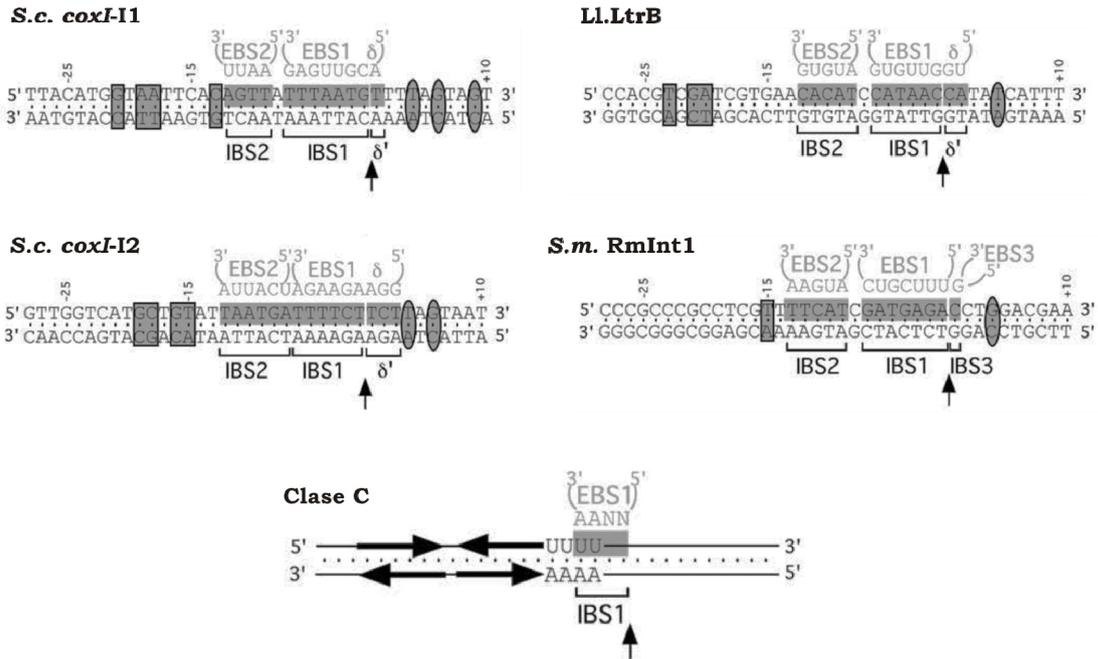
---

<sup>16</sup> La co-conversión entre dos secuencias de DNA es un proceso por el cual se da un intercambio de información genética en dos puntos de forma simultánea.

**A**



**B**



**Figura 17. (A)** Los complejos de RNPs reconocen el sitio diana de DNA de cadena doble o sencilla fundamentalmente por interacciones EBS-IBS ( $\delta$ - $\delta'$  en la subclase IIA), mientras que la IEP también se une específicamente a residuos nucleótidos clave en las regiones distales de los exones 5' (posiciones -23, -21, -20) y 3' (+5 y +10), respectivamente indicados por líneas punteadas en el diagrama. **(B)** Reconocimiento del DNA diana en algunos ejemplos representativos de intrones móviles, así como la propuesta lanzada para intrones de la clase IIC. Con una flecha se indica el punto de inserción del intrón. Las posiciones reconocidas por la IEP en el DNA diana se representan enmarcadas con  si se encuentran en el exón 1 o con  si se localizan en el exón 2. Las posiciones reconocidas por apareamiento de bases están sombreadas en gris (). *coxI-11*, *coxI-12* y *Ll.ltrB* pertenecen a los intrones del grupo IIA; *RmInt1* se encuentra entre los intrones IIB. Figura adaptada de Toro *et al.*, 2007 (A) y Lambowitz *et al.*, 2004 (B).

reconocen dianas relativamente largas (30-35 bp). Para estos tres intrones, mutaciones en nucleótidos clave que son reconocidos por la IEP en la región

distal del exón 5' inhiben tanto el *splicing* reverso como la rotura de la segunda cadena, mientras que mutaciones en el exón 3' sólo inhiben el corte en la segunda cadena del DNA. En el caso de RmInt1, cuya IEP carece de dominio En y de parte del dominio D, los nucleótidos reconocidos por la IEP parecen ser más limitados, donde mutaciones en sólo dos posiciones (-15 y +4) inhiben fuertemente la movilidad (Jiménez-Zurdo *et al.*, 2003). Debido a que RmInt1 no puede llevar a cabo un corte en la segunda cadena, el reconocimiento del exón 3' por la IEP debe ser necesario tanto para el reconocimiento inicial de la diana como para el *splicing* reverso, a diferencia de los otros intrones.

Se han llevado a cabo estudios bioquímicos detallados del reconocimiento del DNA diana por el intrón Ll.LtrB, donde un sistema de expresión eficiente en *E. coli* hace posible la reconstitución de RNPs para estos análisis (Saldanha *et al.*, 1999). Mediante la velocidad de sedimentación, cromatografía de exclusión y crio-microscopía electrónica con RNPs reconstituidas se ha demostrado que su forma precursora tiene una gran estructura pobremente plegada por lo que deben sufrir un cambio conformacional dramático para conseguir el estado catalíticamente activo (Huang *et al.*, 2010). Análisis cinéticos han demostrado que la RNP se une de forma inespecífica al DNA y es entonces cuando busca su secuencia diana mediante mecanismos análogos a aquellos usados por proteínas de unión a sitios específicos de DNA (Aizawa *et al.*, 2003). El reconocimiento inicial parece involucrar a interacciones del surco mayor entre la IEP y bases claves en la región distal del exón 5', incluyendo las bases T -23, G -21 y A -20 en la misma cadena en la cual el intrón realizará su *splicing* reverso (Fig 17) (Singh y Lambowitz, 2002). Estas interacciones reforzadas por contactos entre fosfatos y el esqueleto conllevan el desenrollamiento del DNA, permitiendo que el RNA del intrón se aparee con las secuencias IBS y  $\delta$  para que se de el *splicing* reverso. Experimentos de mutagénesis indican algunas preferencias para determinadas bases desde las posiciones -17 a la -13, donde la IEP cruza el surco menor, pero no está claro si se debe a interacciones directas con bases o indirectas con la estructura del DNA (Perutka *et al.*, 2004). Mutaciones en todas las bases críticas de la región distal del exón 5' inhibe el *splicing* reverso sobre DNA de doble cadena pero tiene poco o ningún efecto sobre el *splicing* reverso en DNA de cadena simple, concluyendo que su reconocimiento se requiere principalmente para el desenrollamiento del DNA (Zhong y Lambowitz, 2003). La rotura de la segunda cadena ocurre en la posición +9 y depende de la misma proteína e interacciones de bases que la posición de la RNP para el *splicing* reverso

(Aizawa *et al.*, 2003) al igual que un pequeño número de interacciones con la IEP con el exón 3', siendo el reconocimiento principal con la T +5 (Fig 17A y B) (Mohr *et al.*, 2000; Singh y Lambowitz, 2001).

El pequeño número de bases reconocidas por la IEP en la región distal del exón 5' nos hace preguntarnos cómo las RNPs reconocen los sitios que pueden aparearse con el RNA del intrón sin desenrollar de forma ineficiente un gran número de sitios no reconocidos. La respuesta parece ser que el desenrollamiento del DNA requiere interacciones determinadas de la IEP y entre bases del RNA del intrón, el cual puede llevar a completar el proceso. La unión de la RNP a los exones 5' y 3' dobla el DNA diana, con el ángulo de torsión aumentando cuando el extremo 3' es reposicionado desde el dominio endonucleasa al sitio activo de la RT para iniciar la transcripción inversa (Noah *et al.*, 2006). Otros intrones del grupo II usan el mismo mecanismo de reconocimiento del DNA diana, pero las secuencias de los exones 5' y 3' reconocidas por la IEP difieren incluso de intrones muy relacionados, sugiriendo una rápida evolución de la especificidad de la IEP (Zhuang *et al.*, 2009).

Los intrones bacterianos de la clase C, cuyas IEPs carecen de dominio En, tienen una especificidad de diana muy definida. Estos intrones se insertan aguas abajo de terminadores de transcripción palindrómicos independientes de *rho* en sitios que tienen un potencial IBS1 pero no un IBS2 (Fig 17B). En estos intrones se ha especulado que las RNPs podrían reconocer específicamente estas estructuras además de sus IBSs para el *retrohoming* (Grandlund *et al.*, 2001; Toor *et al.*, 2006; Michel *et al.*, 2007).

### **3.1.2. Movimiento de intrones mitocondriales de levaduras**

La movilidad de los intrones del grupo II fue demostrada por primera vez para el intrón mitocondrial de *S. cerevisiae* *coxI-I1* (a11) (Meunier *et al.*, 1990) y un año después para *coxI-I1* de *Kluyveromyces lactis* (Skelly *et al.* 1991). Estas investigaciones se llevaron a cabo mediante cruces entre linajes haploides de levadura que contenían diferentes combinaciones de intrones mitocondriales. Durante los cruces, las mitocondrias se fusionaban, permitiendo la recombinación entre sus DNAs. Ambos intrones del grupo II habían ocupado alelos sin intrón con una alta frecuencia, interrumpiendo el 90% de estos alelos. Además, se observó que en *S. cerevisiae*, el *homing* era bloqueado si se mutaban posiciones en la IEP o en el RNA del intrón que inhibiesen el *splicing*. Estudios posteriores han permitido desentrañar los procesos que median la movilidad de estos intrones (Lazowska *et al.*, 1994;

Moran *et al.*, 1995; Zimmerly *et al.*, 1995a; 1995b; Yang *et al.*, 1996; Eskes *et al.*, 1997; 2000).

En el primer estudio detallado sobre la movilidad de los intrones del grupo II (Lazowska *et al.*, 1994) se observó que los intrones aI1 y aI2 eran móviles de forma independiente, pero en algunos cruces, la movilidad de aI1 era bloqueada por algunas pequeñas diferencias en secuencias específicas del alelo en el sitio diana del DNA. Analizando en conjunto la progenie, se vio que la inserción de aI1 en su diana estaba acompañada de una co-conversión asimétrica de las secuencias exónicas flanqueantes, extendiéndose más de 50 bp en el exón 5' pero menos de 25 en el exón 3'. La observación previa de que las RNPs comienzan la síntesis del DNA copia (cDNA) muy cerca del extremo 5' del exón 3' cuando usan como molde un RNA precursor no maduro *in vitro* (Kennell *et al.*, 1993), podría explicar este patrón de co-conversión: el cDNA comienza en el exón 3', continúa a través del intrón hasta bastante avanzado el exón 5' y luego se inserta por recombinación homóloga en la nueva diana.

Un trabajo posterior apunta a la existencia de dos posibles rutas para el *homing in vivo* de estos intrones en levaduras (Moran *et al.*, 1995). Se observó que mutaciones en el centro activo de la actividad RT reducen la movilidad sólo al 40% en aI2, pero que mutaciones que afectan la actividad DNA endonucleasa inhiben por completo el *homing*. Estos eventos de movilidad en ausencia de actividad RT deben producirse a nivel de DNA, probablemente tengan lugar por recombinación mediada por corte producido por la actividad endonucleasa de la IEP. En este caso, el *homing* va acompañado de co-conversión en ambos exones. El proceso dependiente de RT se asemejaría al descrito para el elemento R2 de *Bombyx mori* (Luan *et al.*, 1993): el corte en una de las cadenas (en el exón 3', pero muy próximo al sitio de inserción) permitiría a la RT copiar el intrón usando el 3'OH libre como cebador y el RNA precursor no maduro como molde. Así se explicaría que se produzca co-conversión en el exón 5' pero no en el exón 3'.

Sin embargo, los detalles del proceso se obtuvieron mediante estudios bioquímicos *in vitro* usando preparaciones enriquecidas en RNPs (Zimmerly *et al.*, 1995a; 1995b; Yang *et al.*, 1996). De acuerdo con lo observado por Moran y colaboradores, una de las rutas de movilidad se asemeja al mecanismo TPRT (Target-primed Reverse Transcription) descrito para retrotransposones no-LTR (R2Bm) (Zimmerly *et al.*, 1995a). Estos trabajos demuestran que el *homing* está mediado por corte en las dos cadenas: por un lado, el RNA del intrón se inserta en una de las cadenas (en adelante, la cadena de arriba, superior o sentido por analogía con los esquemas) gracias

al proceso que se ha definido como *splicing* reverso; mientras, la actividad DNA endonucleasa de la IEP produce un corte en la otra hebra ligeramente desplazado aguas abajo del sitio de inserción del intrón (Zimmerly *et al.*, 1995b).

En el *splicing*, las dos transesterificaciones que tienen lugar durante la maduración de los intrones del grupo II por la ruta de *branching* son reversibles (Chin y Pyle, 1995). De hecho, varios autores han sido capaces de demostrar *in vitro* la capacidad del RNA *lariat* de insertarse en otra molécula de RNA que represente los exones unidos (o que al menos contenga secuencias complementarias a los EBSs del intrón; Augustin *et al.*, 1990; Mörl y Schmelzer, 1990a); además, esta inserción tiene lugar justo en el punto de unión de los exones. Esta reacción no está limitada a sustratos RNA, sino que se ha visto, *in vitro*, que estos intrones catalizan de forma relativamente eficiente su inserción sobre DNA (Mörl *et al.*, 1992; Griffin *et al.*, 1995; Yang *et al.*, 1996). En cualquier caso, parece ser que la reacción directa (*splicing*) está favorecida sobre la inversa (Daniels *et al.*, 1996), pero en determinadas condiciones (por ejemplo, que existan otras reacciones acopladas) se puede fomentar la reacción inversa (Aizawa *et al.*, 2003). Tras la reacción de reversión de la maduración, el RNA del intrón queda en forma lineal entre los dos exones, entonces decimos que el *splicing* reverso es completo. Pero también puede ser parcial: el intrón rompe la unión entre los dos exones pero permanece unido en forma de lazo al exón 3'. En levaduras se observa un comportamiento diferencial *in vitro* respecto a inserción completa o parcial: mientras que aI2 se inserta mayoritariamente por *splicing* reverso parcial, aI1 se suele insertar completo en el ~50% de los casos (Zimmerly *et al.*, 1995b; Yang *et al.*, 1996; Eskes *et al.*, 1997; 2000).

El corte generado por la DNA endonucleasa de la IEP de estos intrones de levadura tiene lugar en la posición +10 dentro del exón 3' pero en la cadena opuesta (cadena de abajo, inferior o antisentido) a la que se ha invadido por el RNA del intrón (Yang *et al.*, 1996; Zimmerly *et al.*, 1995a; 1995b). Yang y colaboradores demostraron *in vitro* con RNPs de aI1 que la proporción de *splicing* reverso completo se encuentra entre 38-63% y en más de un 95% de estas moléculas también se había producido el corte de la cadena de abajo (Yang *et al.*, 1996). El extremo 3'OH libre que se ha generado sirve de cebador a la reverso transcriptasa para la síntesis del cDNA, de hecho, se ha propuesto que la reverso-transcripción a través del 3'ss bloquea la escisión y desplaza el equilibrio hacia una inserción completa del intrón (Aizawa *et al.*, 2003).

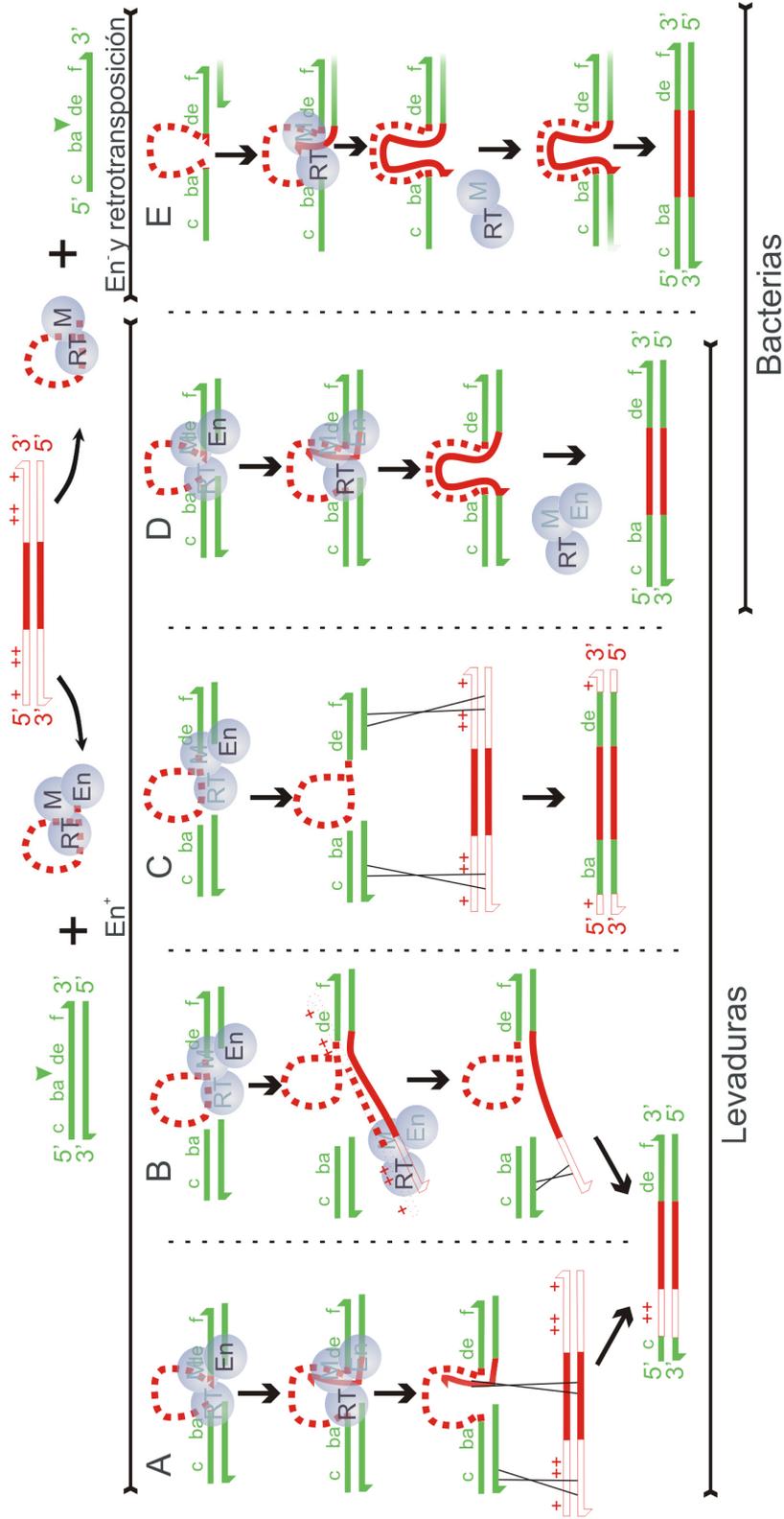
Finalmente, la confirmación definitiva de los mecanismos de movilidad que operan en levaduras deriva del análisis genético de un gran número de eventos de *homing in vivo*, atendiendo a los patrones de co-conversión (Eskes *et al.*, 1997; 2000). Así pues se describen tres rutas para la movilidad de estos intrones:

\* Una ruta minoritaria por la cual, una pequeña proporción de los eventos de movilidad del intrón silvestre se dan sin que exista co-conversión de exones, seguramente mediante la síntesis de un cDNA del intrón completo, que se integra por reparación del DNA, de forma similar a Ll.LtrB (ver más adelante). Un dato interesante es que, esta baja proporción de eventos de movilidad puede verse aumentada hasta un 43% si se mutan determinadas posiciones del DNA diana (Eskes *et al.*, 2000).

\* Una ruta de *homing* (~40% de los eventos analizados) independiente de actividad RT. La integración del intrón tiene lugar por recombinación homóloga a nivel de DNA: la doble rotura generada por las RNPs en el DNA diana se repara mediante el sistema DSB<sub>R</sub> (Double Strand Break Reparir). En estos casos, el intrón se integra acompañado de parte de ambos exones (Fig. 18, mecanismo C).

\* La ruta mayoritaria de *retrohoming* (~60%) tiene como resultado la cotransformación sólo del exón 5' e implica la cooperación de varios mecanismos: por un lado, TPRT para la síntesis del cDNA a partir del intrón y, por otro, recombinación homóloga (DSBR) para la integración definitiva del cDNA del intrón en la nueva diana (Fig. 18, mecanismos A y B).

Existe una enorme controversia al tratar de explicar la resolución definitiva del proceso, ya que en levaduras se dan dos situaciones: intrón insertado parcialmente y la inserción completa. En las dianas en que la inserción del RNA ha sido completa (A), todos los autores coinciden en que el intrón insertado es el molde y que en algún momento, probablemente cuando la RT haya concluido, se produce la reparación de esta región mediante recombinación (muy activa en levaduras). Cuando la inserción ha tenido lugar por *splicing* reverso parcial (B) la RT iniciaría la síntesis usando intrón insertado al menos hasta el enlace modificado 2'-5' presente en el *lariat* (por eso no existe co-conversión del exón 3'), el problema es explicar que sucede después. Hay autores que consideran que podría producirse un cambio de molde al RNA precursor no maduro (Zimmerly *et al.*, 1995b), pero ninguno descarta la posibilidad de que la RT atravesase el enlace y continúe con la síntesis a partir del intrón insertado, resolviendo



**Figura 18.** Mecanismos de movilidad propuestos para intrones del grupo II. Los tres mecanismos de la izquierda se han descrito únicamente en levaduras; en todos los casos dependen de recombinación homóloga. Se observa co-transformación sólo del exón 5' (A y B) o de ambos exones (C). A (mayoritariamente asignado a al1) y/o B (al2 presenta un mayor porcentaje de invasión del ARN por *splicing* reverso parcial) se corresponden con la ruta mayoritaria (60%) que se describe en levaduras, mientras C representa la ruta minoritaria (40%) independiente de RT. El diagrama D se corresponde con la ruta mayoritaria descrita para Li.ltrB y, en determinadas circunstancias, también se manifiesta en levaduras. Se trata de un mecanismo iniciado por TPRT que se resuelve gracias a la maquinaria de reparación de recombinación. En E se representa el mecanismo de *retrohoming* para intrones que carecen de dominio DNA endonucleasa y la hipótesis más aceptada para la retrotransposición. Se produce invasión sobre DNA de cadena simple y probablemente está ligado a replicación.

finalmente mediante recombinación, aunque la síntesis del cDNA no se haya completado (Eskes *et al.*, 1997; 2000).

Se han propuesto muchos factores que determinan la elección de la ruta de movilidad en levaduras: la proporción de integración parcial/completa, la velocidad de iniciación de la síntesis del cDNA, la disociación de la RT, la degradación del RNA, el corte en la cadena de abajo, los sistemas de recombinación/reparación de la mitocondria, e incluso, la naturaleza del DNA diana y del propio intrón (Eskes *et al.*, 2000).

### 3.1.3. Retrohoming de intrones bacterianos

De todos los intrones bacterianos del grupo II caracterizados hasta la fecha como capaces de sufrir splicing *in vivo*, los más estudiados han sido Ll.LtrB de *L. lactis* y RmInt1 de *S. meliloti* mostrándose como unos eficientes elementos genéticos móviles. Por esto se han usado como sistemas modelos experimentales para determinar las rutas de movilidad y sus mecanismos (Lambowitz y Zimmerly, 2004).

El descubrimiento de los intrones del grupo II en bacterias ha implicado el desarrollo de sistemas adecuados para poder llevar a cabo estudios bioquímicos y genéticos detallados (Ferat y Michel, 1993). Sin embargo, la mayoría de los intrones son inmóviles o se expresan a un bajo nivel en el sistema modelo *E. coli*.

#### 3.1.3.1 Retrohoming mediado por TPRT

El mecanismo de movilidad de los intrones del grupo II bacterianos se investigó por primera vez en Ll.LtrB, mediante ensayos genéticos *in vivo* basados en plásmidos, tanto en *L. lactis* como en *E. coli*. Fue posteriormente analizado caracterizando las actividades bioquímicas de RNPs que eran reconstituídas usando la IEP (LtrA) purificada desde su expresión en *E. coli*, y el RNA *lariat* de Ll.LtrB escindido a partir de un precursor obtenido mediante transcripción *in vitro* (Matsuura *et al.*, 1997; Cousineau *et al.*, 1998; Saldanha *et al.*, 1999; Aizawa *et al.*, 2003). Ninguno de los dos elementos presentaba por sí solo las actividades de *splicing* reverso o endonucleasa de la RNP (Saldanha *et al.*, 1999). Las RNPs de los intrones del grupo II se unen inicialmente al DNA de forma inespecífica y es entonces cuando buscan el sitio diana en el DNA (Aizawa *et al.*, 2003). Como ya hemos dicho antes, las RNPs de Ll.LtrB reconocen una región diana en el DNA relativamente larga (Fig 16B), que va desde la posición -25 hasta la +9 siendo

el punto 0 el sitio de inserción del intrón (Guo *et al.*, 2000; Mohr *et al.*, 2000; Karberg *et al.*, 2001; Singh y Lambowitz, 2001). Entonces se produce el retrohoming mediante el mecanismo denominado TPRT<sup>17</sup> (Fig 18D). El RNA de la RNP rompe la cadena de arriba o sentido, justo en la unión de los exones en la diana de DNA de doble cadena, y mediante la reacción de *splicing* reverso el intrón se integra en la diana. Mientras, LtrA usa su capacidad endonucleasa para cortar la cadena de abajo o antisentido en la posición +9. El extremo 3' de la cadena antisentido cortada es puesto en contacto con el centro activo de la reverso transcriptasa. La IEP, mediante su dominio RT, utiliza este extremo 3' como cebador para la transcripción reversa del RNA del intrón que se ha insertado. Se genera así un cDNA copia del intrón que se integra en su nueva localización mediante mecanismos de reparación independientes de recombinación homóloga (Cousineau *et al.*, 1998), a diferencia de lo que ocurre en levaduras, donde se da co-conversión de los exones. En *E. coli*, estos mecanismos de reparación parece que conllevan la síntesis de un cDNA del intrón completo, eliminación de la cadena de RNA por la RNAasa H del hospedador, la síntesis de la segunda cadena de DNA por la DNA polimerasa III y el uso de DNA ligasas para sellar las muescas (Smith *et al.*, 2005).

### 3.1.3.2. Movilidad independiente de actividad DNA endonucleasa

Ninguno de los mecanismos descritos anteriormente son susceptibles de ser llevados a cabo por intrones cuya IEP carezca de dominio DNA endonucleasa (En-) como consecuencia de su manifiesta incapacidad para generar el cebador que permita iniciar el *retrohoming* por TPRT. A pesar de ello, son muchos los intrones bacterianos en los que este dominio está ausente, pero esto no impide que algunos se muevan eficientemente a nuevas localizaciones (figura 18E).

Para tratar de caracterizar el mecanismo por el que estos intrones son capaces de moverse, Lambowitz y colaboradores construyeron un mutante de Ll.ltrB que, sin tener afectadas otras actividades, era incapaz de producir corte en la cadena de abajo (San Filippo y Lambowitz, 2002; Zhong y Lambowitz, 2003). En este mutante, YRT, se reduce considerablemente el *retrohoming* y a ello hay que sumar la particularidad de que pasa a depender de la replicación celular. Además muestra una clara tendencia a insertarse en la cadena que sirve como molde para la síntesis de la hebra conductora

---

<sup>17</sup> TPRP viene de Target DNA-Primed Reverse Transcription en la cual se da una transcripción inversa del ARN del intrón que se ha insertado usando como cebador el DNA diana.

durante la replicación, lo que sugiere que esta cadena naciente podría ser utilizada por la RT como cebador para iniciar la síntesis del cDNA. También existe movilidad sobre dianas que se encuentran en la otra hebra, pero es residual. En cualquier caso, parece ser que la DNA polimerasa debería desengancharse al encontrarse con la RNP, dejando la cadena naciente en posición para permitir la reverso-transcripción a la IEP (el *retrohoming* es absolutamente dependiente de RT), para después volver a la horquilla de replicación y completar el proceso (Zhong y Lambowitz, 2003).

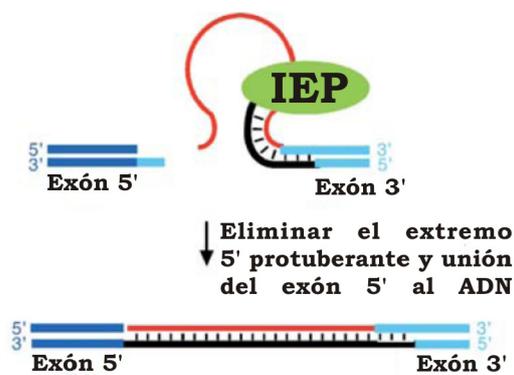
Sin embargo, un estudio de la localización respecto del origen de replicación de los intrones del grupo II que carecen de forma natural de dominio En, muestra que la mayoría de estos intrones se encuentran insertados en la hebra que sirve de molde para la síntesis de la cadena retardada durante la replicación (justo la orientación opuesta a la predicha con el mutante YRT de Ll.ltrB; Ichihyanagi *et al.*, 2002). La explicación para esta observación, es que los intrones En<sup>-</sup> son capaces de insertarse sólo cuando el DNA se encuentra como cadena sencilla, una vez que la horquilla de replicación ha atravesado el sitio de inserción, y que usan como cebador la cadena retardada naciente (no se ha determinado si usarían los cortos fragmentos de RNA sintetizados por la primasa o los propios fragmentos de Okazaki; Martínez-Abarca *et al.*, 2004). Se ha especulado que podrían existir algún tipo de interacciones con la maquinaria de replicación del hospedador que faciliten el *retrohoming* de estos intrones, pero aún no se ha encontrado más relación que ser uno de los momentos en el que parte del genoma se encuentra como DNA de cadena simple. Se han sugerido otros mecanismos que permitan iniciar la síntesis del cDNA como, por ejemplo, muescas al azar en la cadena de abajo (intrón *cob-I1* de *Schizosaccharomyces pombe*, Schäfer *et al.*, 2003) o iniciación de novo (RT codificada por el plásmido Mauriceville en mitocondria de *Neurospora*, Wang y Lambowitz, 1993).

El caso más estudiado de este tipo de intrones es RmInt1, localizado en una secuencia de inserción en el genoma de *Sinorhizobium meliloti* (Martínez-Abarca *et al.*, 1998). Su frecuencia de *retrohoming* es muy alta, similar a la del intrón de *Lactococcus* o a los de levadura (Martínez-Abarca *et al.*, 2000). Aunque se ha demostrado *in vitro* que es capaz de insertarse en DNA de cadena simple y doble (con mucha menor eficiencia en el segundo caso, y tanto menos cuanto más largo es el sustrato), se ha comprobado que no es capaz de cortar la cadena de abajo (Muñoz-Adelantado *et al.*, 2003). Este intrón también muestra una clara tendencia a insertarse en la cadena molde para la síntesis de la hebra retardada, aunque en menor frecuencia

también se inserta en la otra cadena (Martínez-Abarca *et al.*, 2004; Nisa-Martínez *et al.*, 2007).

Otro caso no menos curioso es el de los intrones de la clase IIC, que también carecen de dominio DNA endonucleasa aunque conservan parcialmente el dominio de unión a DNA. El reconocimiento tan particular de su diana ha despertado el interés de los investigadores que han llegado a proponer un posible mecanismo de movilidad relacionado no sólo con la replicación sino con la naturaleza funcional de los terminadores bajo los que se localizan (figura 17B; Granlund *et al.*, 2001; Michel *et al.*, 2007). Se ha visto que el DNA de cadena simple facilita la formación de la estructura en tallo-lazo de un terminador de transcripción o de la región attC que es reconocido por la RNP (Robart *et al.*, 2007; León y Roy, 2009).

### 3.1.3.3. Retrohoming de intrones lineales



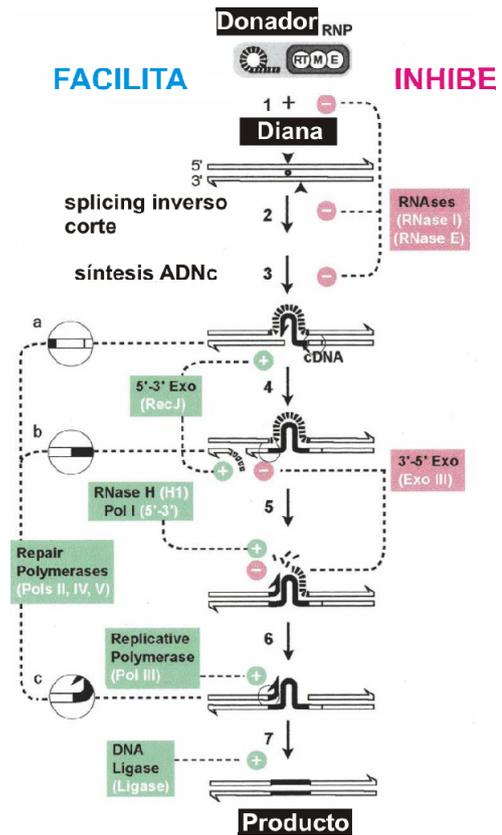
**Figura 18.** Retrohoming de los intrones lineales mediante la primera reacción del splicing reverso, rotura de la cadena de abajo, transcripción inversa y unión del cDNA al exón 5' probablemente por NHEJ (Zhuang *et al.*, 2009b).

El RNA lineal de determinados intrones del grupo II (que se puede generar mediante la ruta hidrolítica de splicing o por la rotura del enlace 2'-5' del *lariat*) usa otra vía de *retrohoming* (Fig. 19), que fue descubierta al inyectar RNPs de Ll.LtrB compuestas por RNA lineal del intrón, en el núcleo de oocitos de *Xenopus laevis* o en embriones de *Drosophila melanogaster* (Zhuang *et al.*, 2009b). En este caso, el RNA lineal del intrón lleva a cabo la primera reacción del *splicing* reverso sobre el DNA diana, de ese modo el

RNA se une al exón 3' pero no al 5'. Se sintetiza entonces el cDNA por la IEP, y esta molécula es la que se une al exón 5', probablemente por unión de extremos no homólogos (NHEJ), un proceso de reparación de DNA muy estudiado. Aunque ineficiente, este mecanismo puede ser usado en el *retrohoming* de los intrones lineales, no sólo en eucariotas, sino también en muchos procariontes que tienen mecanismos análogos al NHEJ (Bowater y Doherty, 2006).

## 3.1.3.4. Dependencia del retrohoming de las funciones celulares

Los intrones del grupo II son elementos genéticos con determinadas características moleculares que les permiten mantenerse y dispersarse en el genoma. Sin embargo, su movilidad depende, en diferente grado, del fondo genético del hospedador por lo que el *retrohoming* estaría supeditado a las funciones celulares para verse completado (Karberg *et al.*, 2001; Toro *et al.*, 2003). Además de requerir de la maquinaria de replicación en las primeras etapas, el *retrohoming* independiente de endonucleasa requiere, en pasos posteriores, de las funciones de reparación. Recientes estudios con el intrón Ll.LtrB en el fondo genético de diferentes mutantes de *E. coli*, deficientes en funciones de reparación de DNA/RNA, han permitido crear un modelo del proceso de *retrohoming* (Fig. 20) que involucra a exonucleasas (RecJ, MutD, y PolI) y RNAasas (RNase H) para el corte de extremos de DNA y eliminación del molde de RNA después de la síntesis del cDNA, junto a los complejos de las DNAs polimerasas y polimerasas de reparación (PolII, PolIII, PolIV y PolV) para la síntesis correcta de la segunda cadena del cDNA (Fig. 20; Smith *et al.*, 2005). También se ha visto que la movilidad de Ll.LtrB está influenciada por interacciones celulares, otros factores del hospedador y de respuesta a estrés (Beauregard *et al.*, 2008; Zhao *et al.*, 2008; Coros *et al.*, 2009).



**Figura 20.** Funciones celulares que intervienen en el *retrohoming*. En azul las que favorecen, y en magenta las inhibitorias. (Smith *et al.*, 2005)

## 3.2. MECANISMO DE RETROTRANSPOSICIÓN

A pesar de las múltiples rutas de *homing* propuestas en los diferentes organismos, distintos estudios han llevado a la conclusión de que, en la mayoría los casos, existe una única vía para la transposición ectópica de los

intrones del grupo II (Dickson *et al.*, 2001; Ichiyanagi *et al.*, 2002; 2003; ver figura 17E). La identificación de los primeros eventos de retrotransposición en intrones mitocondriales se realizó por PCR y una de las observaciones más importantes fue que, a diferencia del *homing*, no existía co-conversión de los exones flanqueantes (Mueller *et al.*, 1993b; Sellem *et al.*, 1993). Así, el mecanismo inicialmente propuesto para este fenómeno dependiente de RT fue la integración del RNA del intrón por *splicing* reverso sobre un RNA que contenga los exones ligados, seguido de la síntesis del cDNA por parte de la RT y la resolución por recombinación homóloga. Los primeros estudios en Ll.ltrB parecían refrendar esta hipótesis (Cousineau *et al.*, 2000). Sin embargo, trabajos posteriores proporcionan una base sólida para afirmar que, también en la retrotransposición, la inserción del RNA tiene lugar a nivel de DNA (Yang *et al.*, 1998; Cousineau *et al.*, 2001 erratum; Dickson *et al.*, 2001; Ichiyanagi *et al.*, 2002). Además, al menos en intrones bacterianos, se ha demostrado que es un proceso independiente de recombinación homóloga (Martínez-Abarca y Toro, 2000b; Ichiyanagi *et al.*, 2002). Pronto se dieron cuenta de que, en la mayoría de estos eventos, la única secuencia de reconocimiento que se observa es la interacción EBS1-IBS1; la conservación del EBS2 y de las posiciones distales en los exones 5' y 3' es mínima en la mayoría de los casos (Yang *et al.*, 1998; Cousineau *et al.*, 2000). En este punto debemos recordar que la interacción de la proteína con las regiones distales de ambos exones permite la separación de la doble hélice y el corte en la cadena de abajo por parte del dominio En (Singh y Lambowitz, 2001), por lo que el dato anterior nos indicaría que para la transposición ectópica se precisa de DNA de cadena simple. Una observación adicional es que este proceso se produce en ausencia de actividad DNA endonucleasa (Cousineau *et al.*, 2000). Si a todos los resultados anteriores le sumamos que existe una clara tendencia a la invasión de sitios no específicos situados en la cadena que sirve como molde para la síntesis de la hebra retardada durante la replicación (Ichiyanagi *et al.*, 2002; 2003), la conclusión es clara: el mecanismo principal por el cual tiene lugar la retrotransposición en todos los intrones hasta ahora caracterizados se asemeja al *retrohoming* de los intrones que carecen de dominio En y se le ha llamado mecanismo de *retrohoming* ilegítimo (ver apartado 3.1.3.2.) (Martínez-Abarca y Toro, 2000b; Dickson *et al.*, 2001; Ichiyanagi *et al.*, 2002; 2003; Belfort *et al.*, 2002).

Finalmente, estudios recientes apuntan a que el mecanismo de retrotransposición depende del organismo hospedador ya que se ha observado un comportamiento diferencial de Ll.ltrB en *L. lactis* y en *E. coli* (Coros *et al.*, 2005). Mientras en *Lactococcus* sigue la vía descrita

previamente, en *E. coli* parece producirse la inserción del intrón también sobre DNA de doble cadena. La explicación de esta observación es que a diferencia de lo descrito anteriormente, la conservación tanto de los EBSs como de las posiciones más relevantes para el reconocimiento por parte de la IEP es mucho mayor que la descrita en *Lactococcus*. Se describe una preferencia de ruta en *E. coli* ligada a las condiciones de crecimiento de las células. Cuando se replican lentamente, prefieren integrarse en DNA de doble cadena mostrando un curioso patrón, cerca del origen y el final de la replicación. Este patrón se ha justificado en virtud de la localización polar de la IEP en este tipo de bacterias y de la posición de estas regiones durante la replicación (Zhao y Lambowitz, 2005; Beauregard *et al.*, 2006). Si crecen rápidamente, el mecanismo de invasión sobre DNA de cadena simple es más frecuente y las inserciones se localizan dispersas a lo largo de todo el genoma; en esta ocasión, replicación y división celular ocurren simultáneamente y el cromosoma está mucho menos compacto (Coros *et al.*, 2005).

#### 4. EVOLUCIÓN DE LOS INTRONES DEL GRUPO II

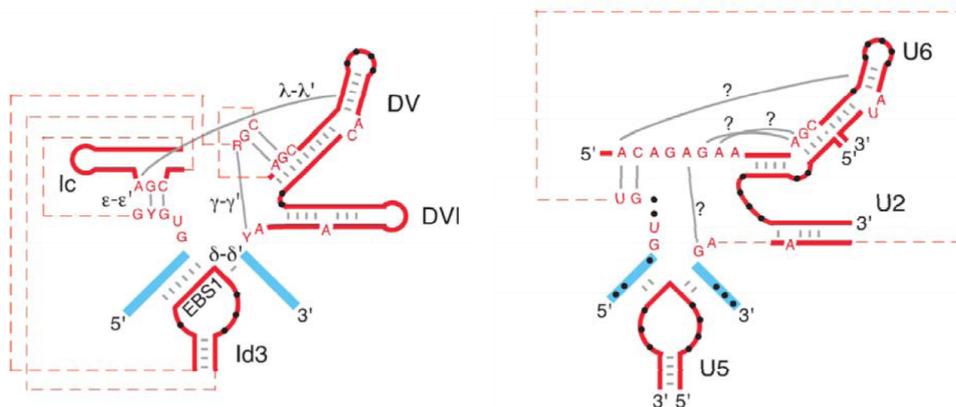
Se piensa que los intrones del grupo II han jugado un papel muy importante en la evolución de los genomas eucariotas como antecesores de los intrones espliceosómicos y de los retrotransposones no-LTR (Sharp, 1985; Cech, 1986; Zimmerly *et al.*, 1995b). Esta relación evolutiva entre los intrones del grupo II y los espliceosómicos se debe a la similitud de sus mecanismos de *splicing*, unión de secuencias semejante y a similitudes estructurales entre regiones clave de dominios de los intrones del grupo II y los RNAs de los espliceosomas (Madhani y Guthrie, 1992; Shukla y Padgett, 2002; Keating *et al.*, 2010). Las similitudes estructurales incluyen: sitios de unión a iones metálicos divalentes en el DV y en el snRNA U6 (RNA nuclear pequeño) respectivamente, que participa en la catálisis; características similares del sitio de ramificación del DVI y de la unión intrón-U2, en el cual, una adenosina equivalente actúa como sitio de *branching* en la primera reacción del *splicing*; similitud entre la interacción  $\epsilon$ - $\epsilon'$  de los intrones del grupo II y el apareamiento de secuencia entre el U6 y la región 5' del intrón; y el reconocimiento de los exones 5' y 3', en los cuales los motivos IBS1-EBS1 y  $\delta$ - $\delta'$  de los intrones de la clase IIA son análogos a la región tallo-lazo del snRNA U5 (Fig 21). La existencia de estas similitudes ha sido reforzada por la reciente cristalización de la estructura de un intrón de la clase IIC (Toor *et al.*, 2008a) y por las cada vez mayores evidencias de que los snRNAs

pueden catalizar reacciones relacionadas con el *splicing* del RNA (Valadkhan *et al.*, 2009). La creencia de relaciones evolutivas entre los intrones del grupo II y los retrotransposones no-LTR viene dada por la similitud de secuencia entre sus RTs (Xiong y Eickbush, 1990; Blocker *et al.*, 2005) y los mecanismos TPRT, en los cuales se usa una rotura en el sitio diana como cebador de la transcripción inversa de un RNA molde unido específicamente (Luan *et al.*, 1993; Zimmerly *et al.* 1995).

La distribución filogenética de los intrones del grupo II, siendo común su existencia en eubacterias y orgánulo eucariotas pero rara en arqueobacterias, sugiere un escenario en el cual los intrones móviles del grupo II se originaron en la eubacterias y luego fueron transmitidos a los eucariotas, seguramente mediante una bacteria endosimbiótica que daría lugar a posteriori a las mitocondrias y los cloroplastos (Cavalier-Smith, 1991;

**Intrones del grupo IIA**

**Intrones espliceosómicos y snRNAs**



**Figura 21.** Similitudes entre el sitio activo de los intrones del grupo II y el del espliceosoma. El ARN de los intrones del grupo II y el snRNA del espliceosoma están dibujados en rojo, y los exones en azul. Las interacciones entre las bases que son similares entre los intrones del grupo II y los intrones espliceosoma aparecen en barras grises, y las bases desapareadas en posiciones similares se muestran como puntos negros. Las líneas discontinuas hacen referencia a conexiones entre secuencias de longitud no especificada. Los interrogantes indican interacciones hipotéticas que pueden darse en el espliceosoma, basadas en interacciones encontradas en los intrones del grupo II (Boudvillain *et al.*, 2000; Toor *et al.*, 2008a). El snRNA U5 es semejante al DId3 de los intrones IIA, mientras que las semejanzas entre ε-ε' y DV/U6 se dan con los intrones IIA y IIB. (Figura adaptada de Lambowitz y Zimmerly, 2010).

Palmer y Logsdon, 1991). Los intrones del grupo II presentes en el ancestro alfa-proteobacteriano de la mitocondria, invadieron al hospedador arquea durante la eucariogénesis (Kooning, 2006). Recientemente se ha descubierto que, lejos de desaparecer del genoma plastidial, algún intrón del grupo II ha adquirido, incluso, alguna función que hace que su presencia sea indispensable para el *splicing* de otros intrones del grupo II presentes (Petersen *et al.*, 2011). La idea de que los intrones del grupo II se originaron como retroelementos en bacterias (hipótesis del retroelemento como

ancestro/”retroelement ancestor hypothesis”) es apoyada por el hecho de que los intrones bacterianos del grupo II incluyen a todos los linajes conocidos y generalmente se comportan como retroelementos, mientras que los intrones de orgánulos pertenecen sólo a los linajes MI y CL (Fig 12), y de forma frecuente carecen de un ORF y/o tienen degeneradas las estructuras de la IEP o del RNA (Toor *et al.*, 2001). El retroelemento ancestral debe haber aparecido por la adquisición por una ribozima que podía hacer auto-*splicing* (Wang *et al.*, 1999) o a partir de un retroelemento que evolucionó a RNA auto-madurante para así minimizar los “daños al hospedador” (Curcio y Belfort, 1996). Aunque los intrones IIC parecen ser la rama más antigua del árbol filogenético (la clase más semejante al último de los ancestros comunes), las pruebas estadísticas son débiles (Rest y Mindell, 2003; Simon *et al.*; 2009).

Se cree que en eucariotas los intrones del grupo II habrían invadido el núcleo y proliferado en el genoma, después de lo cual la estructura ribozímica degeneró fragmentándose en los diferentes snRNA que funcionarían en *trans* formando parte de un aparato de *splicing* común (Sharp, 1991; Chalamcharla *et al.*, 2010). Hipótesis recientes sugieren que la introducción de los intrones del grupo II por un endosimbionte bacteriano fue un paso fundamental en la evolución de los eucariotas, ya que la formación de la membrana nuclear, la cual separa la transcripción de la traducción, ayuda a prevenir la traducción de RNAs que no han madurado de forma completa (Martin y Kooning, 2006). Sin tener en cuenta su origen, la separación entre la transcripción y la traducción impide el acceso inmediato de la IEP al RNA del intrón y hace necesaria la evolución de factores de *splicing* que puedan funcionar en *trans*. En las plantas angiospermas, los factores de *splicing* codificados por el hospedador que actúan sobre los intrones del grupo II del cloroplasto han evolucionado gracias a la expansión y diversificación de diferentes familias de proteínas de unión a RNA, siendo diferentes las proteínas que ayudan a distintos intrones. Tal solución sería inviable para el gran número de intrones de los genomas nucleares de eucariotas, siendo la solución la evolución en los snRNAs derivados de los intrones del grupo II para crear la maquinaria común del *splicing*. Estos snRNAs continúan reconociendo a los intrones mediante secuencias conservadas en 5' y 3', además del nucleótido en el sitio de *branching* similar al de los intrones del grupo II y catalizando el *splicing* mediante las mismas reacciones de transesterificación. Esta maquinaria espliceosomal basada en RNA es, posiblemente, la evidencia más fuerte de que el aparato de *splicing* de los eucariotas ha evolucionado a partir de los

intrones del grupo II o de “parientes” cercanos más que provenir de una síntesis *de novo*.

## **5. LOS INTRONES DEL GRUPO II COMO HERRAMIENTAS BIOTECNOLÓGICAS**

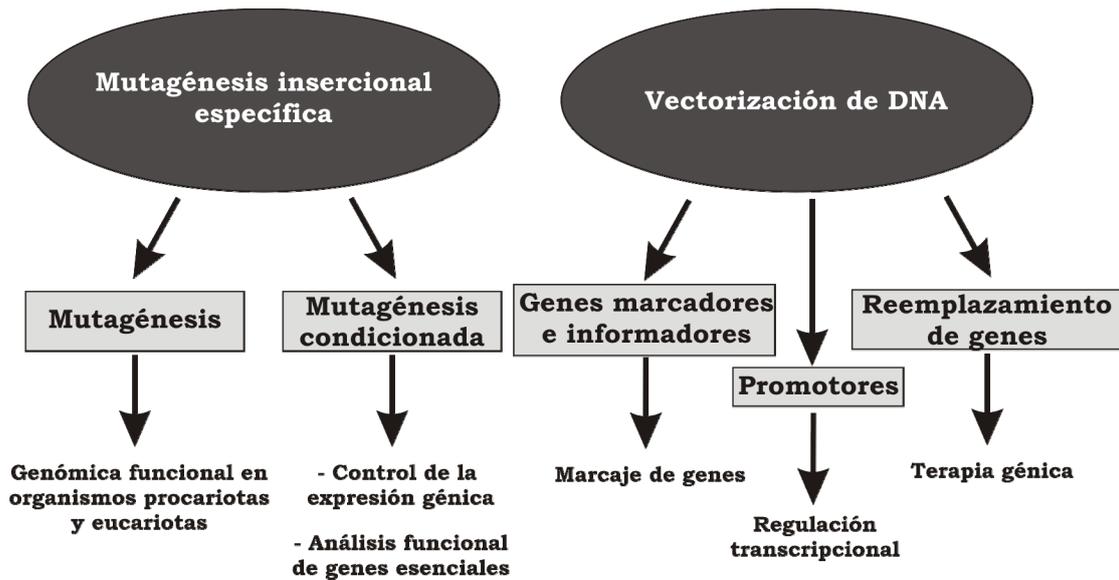
### **5.1 CARACTERÍSTICAS BIOTECNOLÓGICAS Y APLICACIONES POTENCIALES DE LOS INTRONS DEL GRUPO II**

Como en el caso de muchas otras ribozimas (Mills, 2001; Lewin y Hauswirth, 2001; Burt, 2003; Puerta-Fernández *et al.*, 2003; Long *et al.*, 2003), los intrones del grupo II se están desarrollando con el fin de ser usados como herramientas biotecnológicas eficientes, de aplicabilidad tanto en procariotas como en eucariotas ya que el uso de estos retroelementos proporciona una serie de ventajas sobre otras tecnologías ya desarrolladas para la manipulación del DNA (Cui y Davies, 2007). Estas ventajas están en gran medida relacionadas con las características particulares de estos intrones (Belfort *et al.*, 2002; Lambowitz y Zimmerly, 2004; Robart y Zimmerly, 2005; Pyle y Lambowitz, 2006; Plante y Cousineau, 2006; Toro *et al.*, 2007). En primer lugar, son elementos móviles capaces de integrarse específicamente en su DNA diana con alta eficiencia mediante un proceso independiente de recombinación homóloga. A este hecho podemos sumar el que la inserción es muy estable, de manera que se mantiene en sucesivas generaciones. Por otro lado, el reconocimiento de la diana se conoce relativamente bien; se produce principalmente por apareamiento de bases<sup>18</sup>, por lo que la especificidad de la diana puede ser modificada al cambiar las secuencias EBSs dentro del RNA del intrón. Además, en las versiones modificadas más recientes (aquellas en que la IEP es expresada en *trans* fuera del RNA del intrón, versiones  $\Delta$ ORF), el intrón puede portar información dentro de determinadas regiones del dominio IV de la ribozima. Por último, aunque es muy probable que la finalización del proceso dependa de funciones celulares, estos requerimientos son mínimos y probablemente impliquen a genes del hospedador esenciales en otros procesos.

Gracias a estas características, los intrones del grupo II tienen un amplio rango de aplicaciones en ingeniería genética, como herramientas de mutagénesis insercional específica o como vehículos transportadores de información genética (vectorización de DNA) (Fig.22).

---

<sup>18</sup> La proteína reconoce determinadas posiciones para facilitar el desenrollamiento del ADN y para producir el corte en la cadena de abajo, pero el escaso número de interacciones que establece permite encontrar potenciales dianas en casi todos los genes (Perutka *et al.*, 2004).



**Figura 22.** Aplicaciones biotecnológicas de los intrones del grupo II.

## 5.2. USO DE LOS INTRONES DEL GRUPO II EN PROCARIOTAS.

Históricamente se han venido usando determinados métodos para obtener mutantes genéticos en procariotas como los transposones de DNA que han sido una herramienta muy útil de mutagénesis en bacterias. El principal beneficio de trabajar con ellos es que las secuencias insertadas pueden usarse como base a partir de la cual se puede explorar el ambiente genómico próximo mediante PCR y secuenciación. Mediante este procedimiento se han mapeado varias bacterias (Salama *et al.*, 2004; Liberati *et al.*, 2006; Benton *et al.*, 2004). Sin embargo, los sistemas más comunes de mutagénesis, basados en transposones, presentan varias limitaciones importantes: 1) Frecuencias de inserción relativamente bajas; 2) Preferencia de la mayoría de ellos por zonas concretas del genoma (*hot spots*); 3) Intensa búsqueda de la interrupción génica específica. Existen también sistemas específicos de mutagénesis basados en recombinación homóloga que consiste, esencialmente, en la clonación de un fragmento del gen a mutagenizar en un plásmido (p.e., no replicativo o suicida), estando éste interrumpido por un marcador de resistencia a antibióticos. El gen correspondiente en el genoma es interrumpido por inserción del plásmido completo (simple entrecruzamiento) o por intercambio del correspondiente locus con el correspondiente alelo clonado en el plásmido (doble entrecruzamiento). Este sistema de generación específica de mutaciones presenta a su vez varias limitaciones: 1) Dependencia fundamental de un

sistema eficiente de recombinación en la célula, 2) Generación de genes compuestos como consecuencia del evento de cointegración; 3) La resolución del cointegrado ocurre a baja frecuencia, con lo que en la práctica hay reemplazamientos que no se consiguen; y 4) En organismos con gran número de secuencias repetidas, como son las secuencias de inserción, el incremento de la recombinación homóloga global puede causar inestabilidad genómica. Todo ello, hace que este sistema de mutagénesis pueda ser relativamente costoso en tiempo y recursos.

Desde que la movilidad de los intrones del grupo II se demostró utilizando los intrones *coxI-I1* de levaduras se ha trabajado mucho para llegar a entender los procesos por los cuales se lleva a cabo (Meunier *et al.*, 1990; Skelly *et al.*, 1991). Una vez que los mecanismos de reconocimiento de la diana y de integración fueron definidos, comenzó una nueva era para el uso de los intrones del grupo II como herramientas biotecnológicas en la mutagénesis insercional.

Los primeros intentos de modificar los intrones del grupo II fueron dirigidos a alterar su especificidad de diana, cambiando las interacciones de bases entre el RNA y el DNA de los intrones mitocondriales de levaduras aI1 y aI2, y del intrón Ll.LtrB de *L. lactis*. (Guo *et al.*, 1997; Eskes *et al.*, 1997; Matsuura *et al.*, 1997). El mayor éxito se logró usando el intrón Ll.LtrB debido al desarrollo de un sistema de expresión eficiente para este intrón en *E. coli* (Matsuura *et al.*, 1997), el cual ha llegado a ser el intrón mejor caracterizado y el más usado en experimentos de ingeniería genética.

Poco tiempo después de que se llevasen a cabo los primeros experimentos en los cuales se habían modificado las interacciones entre bases, se establecieron los requerimientos para el reconocimiento de la diana por parte de Ll.LtrB (Mohr *et al.*, 2000; Guo *et al.*, 2000; Zhong *et al.*, 2003). Se observó que los intrones del grupo II podían ser redirigidos para que se insertasen de forma eficiente en dianas de DNA deseadas, modificando las secuencias EBS del RNA del intrón (Guo *et al.*, 2000; Karberg *et al.*, 2001). La primera demostración de que la modificación dirigida de los EBSs de Ll.ltrB conduce a la inserción del intrón en el gen que nos interese se realizó con un sistema de plásmido a plásmido utilizando como marcador de selección el gen de resistencia a kanamicina (Km<sup>R</sup>) portado por el intrón en su DIV (Cousineau *et al.*, 1998; Mohr *et al.*, 2000). El proceso fue bastante ineficiente por lo que se modificó el sistema para mejorar el rendimiento del proceso. En muchos casos, se buscó mejorar el límite de detección de los eventos de *homing* (selección de colonias resistentes a tetraciclina, Tc<sup>R</sup>, Guo *et al.*, 2000), en otros se trató de rentabilizar el tiempo para llevar a cabo la

modificación de los intrones (Karberg *et al.*, 2001; Zhong *et al.*, 2003) y en muchos casos se prescindió de un sistema de selección (mapeo por PCR con cebadores específicos, Frazier *et al.*, 2003; Yao *et al.*, 2005).

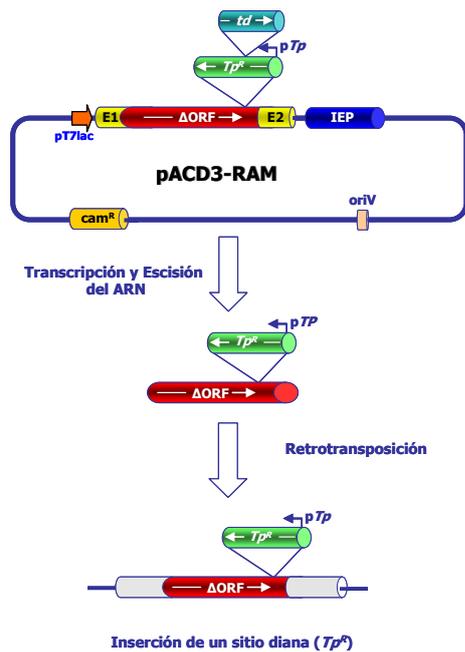
El avance más importante se consiguió cuando se usó una colección de intrones modificados en sus EBSs, posición  $\delta$  y los IBSs del exón 5' (para no interferir en el *splicing*) de forma aleatoria (Karberg *et al.*, 2001) y que contenían un marcador genético activado mediante retrotransposición (RAM) para generar una base de datos de intrones y sitios diana junto con sus eficiencias de integración (Zhong *et al.*, 2003). Originalmente se trataba de un sistema de doble plásmido basado en la activación de la resistencia a tetraciclina ( $Tc^R$ ) cuando el intrón se inserta en la diana (Guo *et al.*, 2000). El plásmido receptor contiene una diana (el gen que deseamos interrumpir) y, a continuación, se encuentra la secuencia de un gen de resistencia a tetraciclina carente de promotor<sup>19</sup>. Por su parte, el plásmido donador expresa el intrón bajo el promotor inducible T7lac; este intrón (en versión  $\Delta ORF^{20}$  con EBSs,  $\delta$  e IBSs modificados al azar y la IEP expresada aguas abajo del exón 2) porta en su DIV el promotor de la RNA polimerasa del fago T7 de manera que cuando se produce el salto del intrón sobre la diana se activa la resistencia a tetraciclina. De este modo, la eficiencia de *homing* se puede evaluar directamente a partir de la razón entre el número de colonias  $Tc^R Amp^R$  (plásmido receptor invadido) y el número de colonias  $Amp^R$  (plásmido receptor sin invadir). A partir de las formas invadidas, por secuenciación, se pueden seleccionar las combinaciones de intrones modificados más interesantes, complementar, si fuese necesario, las EBSs con las IBSs del exón 1 para optimizar el *splicing* y, finalmente, utilizarlos para interrumpir la secuencia deseada en el genoma de la célula. Este sistema es muy eficiente, consiguiéndose frecuencias de invasión entre el 1 y el 80%, sin embargo conlleva demasiado tiempo encontrar el intrón óptimo.

La información obtenida con estos estudios fue usada para generar un algoritmo que permitiese diseñar intrones redirigidos hacia un determinado gen con alta eficiencia y así reducir el tiempo en el que obtener el mejor intrón. Nació así una nueva herramienta para la mutagénesis genética en procariontes (Perutka *et al.*, 2004). Por otro lado, la introducción de un sistema de selección ligado al salto del intrón y su inserción cromosómica

<sup>19</sup> Para disminuir el ruido de fondo, rodeando a estas secuencias se encuentran distintos terminadores de transcripción: la diana se encuentra flanqueada aguas arriba por el doble terminador de la ARN polimerasa de *E. coli* y la ARN polimerasa del fago T7 (T1) y aguas abajo por tres copias del terminador de la ARN polimerasa de *E. coli* (T2); mientras, el gen de resistencia está seguido del terminador  $\Phi$  (que termina la transcripción a partir de la ARN polimerasa del fago T7 pero no de la de *E. coli*) (Karberg *et al.*, 2001).

<sup>20</sup> Versiones en las que la IEP es expresada en *trans* fuera del ARN del intrón ya que ha sido deleccionada parcialmente del dominio IV (En siguientes capítulos desarrollaré esta propuesta).

aumenta la frecuencia de detectar la interrupción de un gen al 100%. Con este fin, Zhong y colaboradores insertaron un gen de resistencia a trimetoprima,  $Tp^R$ , en el dIV del intrón en orientación opuesta a la transcripción del  $\Delta ORF$ ; este gen, a su vez, se encuentra interrumpido por un intrón del grupo I ( $tdI$ ) dispuesto en la misma dirección de la transcripción (RAM) (Fig 23). Así, la  $Tp^R$  no se activa hasta que el intrón se ha insertado en el cromosoma tras la maduración del  $tdI$  (Zhong *et al.*, 2003). Posteriormente, el RAM fue flanqueado por sitios FRT para la recombinasa *F1p*, de manera que podrían hacerse múltiples rondas de inserciones usando el mismo gen marcador para la selección. Esto es, tras una primera inserción se expresa en *trans* la recombinasa que va a eliminar el gen de resistencia; ahora tenemos células sensibles al antibiótico que



**Figura 23.** Representación de la inserción de una construcción derivada del intrón silvestre, que porta como marcador de selección de la retrotransposición (RAM) al gen de resistencia a la trimetoprima clonado en orientación opuesta al componente ribozimico ( $\Delta ORF$ ) del intrón L.LtrB. A su vez, dicho marcador está interrumpido por el intrón autocatalítico del grupo I clonado en la misma orientación que el componente ribozimico  $\Delta ORF$ . Figura adaptada de Zhong *et al.* 2003

(Chen *et al.*, 2005; Yao *et al.*, 2006; Zarschler *et al.*, 2009) y Gram-negativas (*E. coli*, *Shigella flexneri*, *Salmonella enterica*

pueden ser sometidas a nuevas inserciones y seleccionar los eventos de salto mediante  $Tp^R$ , repitiendo el proceso cuantas veces consideremos oportuno (Zhong *et al.*, 2003; Cui y Davis, 2007). Este sistema ha sido trasladado a un vector de amplio rango de hospedador que puede ser potencialmente usado en todas las bacterias en las que sea capaz de replicarse (Yao y Lambowitz, 2007).

Esta tecnología permitió el desarrollo de un sistema de vectores bacterianos programables para la mutación de genes, llamado "Targetron", comercializado por SigmaAldrich. Recientemente otros intrones del grupo II han sido adaptados en esta dirección, expandiendo el rango de dianas accesibles y ofreciendo otras propiedades útiles (Zhuang *et al.*, 2009a; García-Rodríguez *et al.*, 2011). La tecnología Targetron y sus derivadas se han usado con éxito para interrumpir e introducir genes en una gran variedad de cepas bacterianas Gram-positivas (*L. lactis*, *Clostridium perfringens*, *Staphylococcus aureus*, y *Paenibacillus alvei*) (Chen *et al.*, 2005; Yao *et al.*, 2006; Zarschler *et al.*, 2009) y Gram-negativas (*E. coli*, *Shigella flexneri*, *Salmonella enterica*

*serovar Typhimurium*, *Franciscella tulariensis*, *Pseudomonas aeruginosa* y *Agrobacterium tumefaciens*) (Karberg *et al.*, 2001; Yao y Lambowitz, 2007). Al introducir intrones redirigidos en una cepa bacteriana mediante un plásmido donador que expresa el intrón y la IEP, se puede generar inserciones condicionales y no condicionales, dependiendo de la orientación de la inserción del intrón y de la expresión de la IEP (Guo *et al.*, 2000; Karberg *et al.*, 2001; Frazier *et al.*, 2003; Yao *et al.*, 2006). Claramente, los intrones del grupo II pueden ofrecer una serie de herramientas muy útiles en la ingeniería genética de procariotas.

### 5.3. ESTRATEGIAS ACTUALES PARA EL SILENCIAMIENTO GÉNICO Y LA MUTAGÉNESIS INSERCIONAL DIRIGIDA EN EUCARIOTAS”

En muchos casos, los métodos actuales de mutación genética en eucariotas son específicos de organismos y en ocasiones limitados por desventajas inherentes. En el organismo modelo de mamíferos, el ratón, se ha venido usando el RNA interferente como método de silenciamiento génico, que genera pequeñas moléculas de RNA de doble cadenas siRNAs (Paddison *et al.*, 2002). Una ribonucleasa citoplasmática llamada DICER procesa estos siRNAs a partir de precursores más largos. La cadena antisentido del siRNA permanece asociada como efector del complejo de silenciamiento inducido por RNA (RISC) para reconocer y cortar un RNA mensajero complementario, conduciéndolo a su degradación (Meister *et al.*, 2004). La aplicación de este sistema ha sido efectivo, en varios sentidos, para determinar el papel de genes específicos en el genoma de ratones, pero la interpretación de los resultados obtenidos mediante estos estudios es complicada debido al hecho de que este método produce una disminución de los mensajeros, no un silenciamiento total (Mittal, 2004). Además, los oligonucleótidos de siRNA usados en estos experimentos son inestables y deben ser repuestos constantemente para prolongar el silenciamiento génico (Sandy *et al.* 2005). Concentraciones intracelulares altas de siRNAs pueden ser mantenidas expresándolos desde plásmidos estables o desde vectores virales integrados. Este tipo de sistemas de expresión pueden correr el riesgo de integrarse de forma incontrolada en el genoma con consecuencias no deseadas. En los últimos años, la necesidad de desarrollar sistemas eficientes de diseño, identificación y liberación efectiva de los siRNAs han cobrado más importancia.

Aunque los siRNAs son ampliamente usados en ratones, el método más común en el silenciamiento génico es la integración específica de sitio de

DNA exógeno mediante la recombinación homóloga (RH). Las ventajas de la RH sobre los siRNAs es que la RH crea mutaciones en la línea germinal que se mantiene en el cultivo celular y es transferible a los animales. En general, esto conlleva la introducción en los cultivos de células germinales de DNA de doble cadena extraño que es modificado para que contenga regiones de secuencias homólogas al cromosoma (Wang y Zhou, 2003). La molécula de DNA donadora suele incluir un marcador seleccionable de forma que interrumpa el gen diana después de la integración. La liberación del DNA donador en el núcleo por microinyección, electroporación o por transfección lipídica requiere que la maquinaria de recombinación alinee las regiones de homología con el cromosoma y que incorpore el DNA exógeno en el sitio diana. Las células mutadas son seleccionadas usando los marcadores que se expresan desde el DNA integrado. Debido a que la RH es un proceso muy eficiente ( $10^6$  integraciones por célula), los métodos de escrutinio pueden llegar a ser una labor intensiva y demasiado costosa. Además, la necesidad de crear e introducir material transgénico en la célula siempre puede conllevar efectos no deseados.

Una alternativa al uso de plásmidos de doble cadena como donador de DNA en la RH son los virus adenoasociados. El genoma del AAV, un parvovirus humano con un amplio tropismo celular, es encapsidado como DNA de cadena simple. Los AAV recombinantes (AAVr) pueden ser modificados intercambiando los genes virales situados entre las secuencias repetidas terminales invertidas (ITRs) (Porteus *et al.*, 2003, Rusell e Hirata, 1998, Samulski *et al.*, 1989). Los AAVr son más eficientes en tres o cuatro órdenes de magnitud que los plásmidos de doble cadena. Otra de las ventajas de los AAVr es que las enzimas de reparación de DNA se unen de forma más eficiente a los ITRs (Yan *et al.*, 2009). Aunque la alta eficiencia de los AAVrs es interesante, la necesidad de una alta multiplicidad de infección para obtener estos resultados limita su uso a procesos no terapéuticos (Porteus *et al.*, 2003).

En otros organismos modelo, como es el caso de pez cebra (*Danio rerio*), el método más usado para el estudio de la función génica se basa en oligonucleótidos de RNA modificados y dirigidos a un gen denominados morfolidos (Nasevicius y Ekker, 2000). Pero como ocurre con los siRNAs sólo conlleva una bajada en la expresión génica, además sólo es efectiva en un 26% de las ocasiones (Iversen y Newbry, 2005). Los efectos no específicos asociados a los morfolidos no se han caracterizado completamente, pero se sabe que tiene un pequeño impacto sobre los genes que se expresan en embriones mayores de dos días (Eisen y Smith, 2008; Iversen y Newbry,

2005). Un método prometedor en el pez cebra es el conocido como TILLING (Targeting Induced Local Lesions in Genomes). El tilling es un método híbrido entre la genética clásica y la inversa en el cual se usa una colección de mutantes obtenidos mediante modificación química con el agente N-etil N-nitrosourea (ENU) para detectar mutaciones raras en genes de interés. Pero al ser un método basado en la PCR el escrutinio puede dar falsos positivos al existir una alta frecuencia de polimorfismo de un único nucleotido (SNPs) de forma natural.

Últimamente se están usando las nucleasas de dedo de zinc (ZFN) (Porteus y Carroll, 2005). En éstas, un dominio endonucleasa de restricción (*FokI*) es fusionado a un dominio de unión de DNA, que en este caso es dominio de unión a los llamados dedos de zinc de DNA (Chandrasegaran y Smith, 1999). Debido a que la especificidad de reconocimiento de los dedos de zinc puede ser manipulada, la nucleasa quimera puede ser modificada para reconocer secuencias específicas del genoma. Al usar este tipo de endonucleasas redirigibles, se pueden aprovechar las ventajas que ofrecen los mecanismos de reparación celular para llevar a cabo modificaciones en el genoma. Cuando se produce una rotura en la doble cadena del DNA entran en funcionamiento muchos mecanismos celulares de reparación, edición o mutación del genoma. Por ejemplo, la reparación mediante NHEJ (unión de extremos no homólogos) puede interrumpir un gen que ha sufrido una rotura de forma natural o no (van Gent *et al.*, 2001). Además, la introducción de una rotura de doble cadena en un sitio específico de un gen puede estimular los mecanismo de recombinación homóloga que repara este tipo de roturas usando la cromátida hermana sin daños como molde (van Gent *et al.*, 2001). De forma alternativa, la introducción de un DNA exógeno que sea lo suficientemente homólogo a la región donde está la rotura puede hacer que en la reparación se incorpore el DNA extraño al sitio de corte. Esta última estrategia se ha usado con las ZFNs (Cathomen y Joung, 2008, Porteus y Carroll, 2005). La combinación entre ZFNs y RH ha obtenido eficiencias de hasta un 39% en determinada líneas celulares (Cathomen y Joung, 2008), por lo que ha sido usado en varios organismos modelo, con el ratón *Xenopus*, pez cebra y *Drosophila*. Aunque tiene sus ventajas, como todos los métodos también tiene sus inconvenientes, que radican fundamentalmente en: 1) el requerimiento de ingeniería proteica para redirigir las nucleasas viene limitada por la proporción y extensión en las cuales estas enzimas pueden ser modificadas; 2) el ensamblaje modular es complicado obteniendo menos de un 6% de ZFNs funcionales, los que puede causar baja actividad y alta toxicidad (Ramirez *et al.*, 2008; Bae *et al.*, 2003);

3) la capacidad de obtener construcciones con gran actividad y baja toxicidad en células humanas requiere de grandes colecciones muy difíciles de manejar si no se tiene la experiencia y pericia adecuadas (Cornu *et al.*, 2008; Pruet-Miller *et al.*, 2008).

#### 5.4. LOS INTRONES DEL GRUPO II COMO ALTERNATIVA PARA LA MUTAGÉNESIS INSERCIONAL DIRIGIDA EN EUCARIOTAS.

Debido a los claros inconvenientes mostrados en los métodos descritos anteriormente, parecía necesario el desarrollo de nuevas tecnologías que soslayan estos problemas, y los intrones del grupo II ofrecen características muy atractivas para su potencial aplicación en transgénesis y mutación génica. 1) Redirigir los intrones del grupo II es mucho más sencillo que en el caso de los ZFNs y ofrece una versatilidad mayor que aquellos métodos que se basan en eventos al azar. El mecanismo de interrupción génica basado en los intrones del grupo II ofrece también un amplio rango de actividades bioquímicas que pueden ser explotadas, como la interrupción génica por integración directa del intrón en el sitio diana, lo que permite crear, como hemos dicho antes silenciamiento génico condicionado o no, (si está presente la IEP después de la inserción el intrón es eliminado del transcrito permitiendo una transcripción y traducción silvestre del gen); 2) los intrones del grupo II pueden crear cortes en cadena doble de forma específica (Karberg *et al.*, 2001, Mastroianni *et al.*, 2008), y así generar la interrupción de un gen mediante mecanismos de NHEJ o recombinación homóloga; 3) los intrones del grupo II ofrecen la capacidad de transportar genes a una diana determinada mediante la inserción de información genética en el DIV del intrón (Karberg *et al.*, 2001; Frazier *et al.*, 2003, Martínez-Abarca *et al.*, 2004; García-Rodríguez *et al.*, 2011). Un vector, como un intrón del grupo II, que pueda transportar un transgén a una localización determinada de un genoma es una herramienta muy útil ya que, no sólo facilitaría la expresión del transgén en un determinado lugar, sino que a la vez podría interrumpir un gen en concreto.

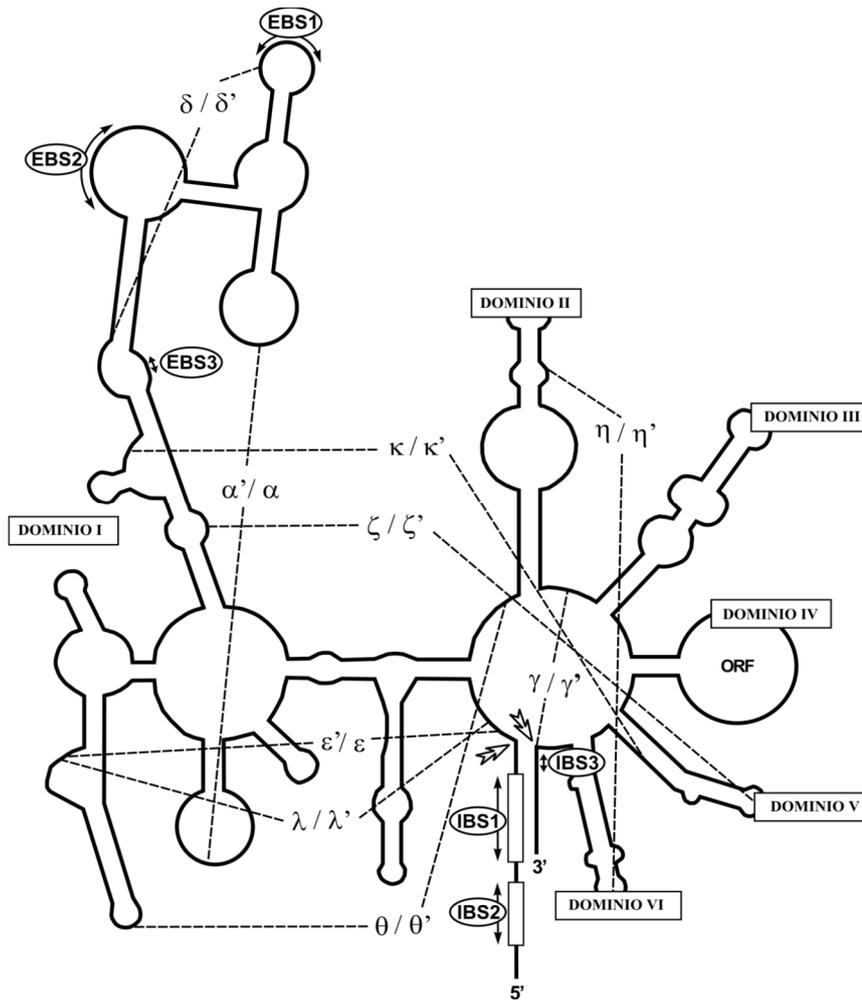
Gracias a todas las características comentadas sobre los intrones del grupo II, se han llevado a cabo experimentos con estos intrones en algunos eucariotas, donde los obstáculos a salvar han sido la accesibilidad de las RNPs al núcleo y las concentraciones subóptimas de magnesio (Mastroianni *et al.*, 2008). Se ha visto que RNPs microinyectadas en el núcleo de oocitos de *Xenopus laevis*, en embriones de *Drosophila melanogaster* y del pez cebra tienen un *retrohoming* eficiente si se suplementa con magnesio (Mastroianni

*et al.*, 2008; Vernon, 2009; Zhuang *et al.*, 2009b). Una solución más general sería identificar y seleccionar intrones que funcionen a concentraciones menores de magnesio. Los intrones del grupo II no sólo pueden ser usados para integrarse en un sitio específico del DNA, sino que también se puede generar una rotura en la doble cadena que estimule la integración dirigida por recombinación homóloga. La rotura de la doble hebra viene dada por un splicing inverso parcial y por la rotura de la segunda cadena por parte de las RNPs de los intrones del grupo II (Karber *et al.*, 2001; Mastroianni *et al.*, 2008).

## **6. RmInt1, EL INTRÓN DEL GRUPO II DE *Sinorhizobium meliloti***

RmInt1 es el primer intrón descrito en bacterias del orden Rhizobiales. Se identificó al analizar la secuencia de una región del plásmido críptico pRmeGR4b presente en la cepa GR4 de la bacteria *Sinorhizobium meliloti*, un microorganismo capaz de fijar nitrógeno en simbiosis con alfalfa (Martínez-Abarca *et al.*, 1998). Siguiendo el comportamiento de los intrones bacterianos (presumiblemente para facilitar su dispersión), se suele encontrar asociado a la secuencia de inserción *ISRm2011-2*, aunque también se puede encontrar invadiendo otras ISs relacionadas (*ISRm10-1* e *ISRm10-2*, Martínez-Abarca y Toro, 2000b; *ISRm10-3*; Muñoz-Adelantado, 2003; Biondi *et al.*, 2011) e incluso se ha localizado otro sitio de transposición ectópica en el gen *oxi1* (Martínez-Abarca y Toro, 2000b). Es interesante destacar que RmInt1 se ha identificado en el 90% de los aislados de suelo de *S. meliloti* y generalmente suele encontrarse en alto número de copias por genoma (de 1 a 11 copias) (Martínez-Abarca *et al.*, 1998). Además se encuentra ampliamente distribuido entre otras especies de la misma familia de acuerdo a patrones de herencia vertical y transferencia horizontal, aunque se observa una tendencia a la inactivación por pérdida del extremo 3' del intrón (Fernández-López *et al.*, 2005).

La estructura secundaria de la ribozima se ajusta a la descrita para intrones del grupo IIB3 (figura 23; Martínez-Abarca *et al.*, 1998). Se ha puesto de manifiesto su capacidad para madurar *in vitro* en ausencia de la proteína dando lugar a moléculas en forma *lariat* pero también en forma lineal, dependiendo de las condiciones de reacción (Costa *et al.*, 2006a). *In vivo*, se observó escisión del intrón (Muñoz-Adelantado *et al.*, 2003) y posteriormente se determinó que las formas de escisión podían ser *lariat* o circular (Molina-Sánchez *et al.*, 2006), mientras que se ha demostrado que la unión de los exones para este intrón es muy ineficiente (Chillón *et al.*, 2011).

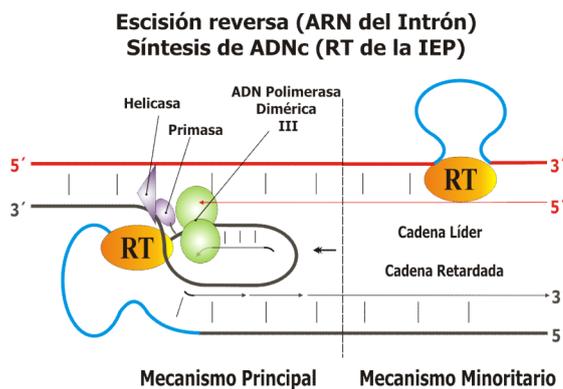


**Figura 24.** Representación esquemática del intrón del grupo II RmInt1. Este intrón puede ser plegado en una estructura secundaria típica de intrones del grupo II. En el esquema se indican los dominios I a VI, señalando el ORF contenido en el dominio IV. Las secuencias de unión a exones (EBS1, EBS2 y EBS3) y las secuencias de unión al intrón (IBS1, IBS2 e IBS3) se indican con flechas de doble punta. Las flechas blancas indican los sitios de corte para la inserción del intrón RmInt1 entre los exones 5' y 3'. Las líneas punteadas muestran las interacciones terciarias.

Por otro lado, la proteína codificada por RmInt1 presenta los 7 dominios característicos de este tipo de RT junto con el dominio X (madurasa), pero carece de los dos dominios presentes en el C-terminal de intrones IIA: dominio de unión a DNA y dominio DNA endonucleasa (Fig. 23; Martínez-Abarca *et al.*, 2000). En su lugar presenta una corta extensión de 20 aminoácidos que participa en la maduración de RmInt1 y es su inserción en al DNA (San Filippo y Lambowitz, 2002, Molina-Sánchez *et al.*, 2010). En función de la distribución de dominios RT, RmInt1 se enmarca dentro de los intrones bacterianos de la clase D (Toro *et al.*, 2002).

RmInt1 es uno de los pocos intrones en los que se ha demostrado la capacidad de movimiento en el genoma de su hospedador (Martínez-Abarca *et al.*, 2000). Las primeras pruebas de este fenómeno se obtuvieron en sistemas de salto desde el genoma a un plásmido que contenían su secuencia diana (Martínez-Abarca *et al.*, 2000; Jiménez-Zurdo *et al.*, 2003). El siguiente paso fue el diseño de un sistema sencillo de *homing* de plásmido (donador) a plásmido (receptor) (Martínez-Abarca *et al.*, 2000; Jiménez-Zurdo *et al.*, 2003), y cuya mejora se describe en posteriores capítulos de esta tesis (Nisa-Martínez *et al.*, 2007). En todos los casos, se comprobó que se trataba de un mecanismo independiente de RecA (Martínez-Abarca *et al.*, 2000; Martínez-Abarca y Toro, 2000b) y mediado por un intermediario de RNA (Martínez-Abarca *et al.*, 2004).

En nuestro laboratorio se han desarrollado numerosos estudios para profundizar en los mecanismos moleculares implicados en la maduración y el movimiento del intrón, tanto *in vivo* (Jiménez-Zurdo *et al.*, 2003) como *in vitro* (Muñoz-Adelantado *et al.*, 2003). Los estudios *in vitro* muestran que las RNPs de este intrón tienen actividad reverso transcriptasa usando un sustrato exógeno poli(rA)/Oligo(dT), aunque la mera adición de cebadores es suficiente para que se produzca polimerización (Muñoz-Adelantado *et al.*, 2003). Se ha comprobado que son capaces de insertarse en sustratos de DNA de cadena doble y de cadena sencilla mediante una reacción de *splicing*



**Figura 25.** Mecanismos principales de invasión propuestos para RmInt1. (Martínez-Abarca *et al.*, 2004)

reverso, pero, como cabría esperar, no producen corte de la cadena de abajo (Muñoz-Adelantado *et al.*, 2003). En cuanto a los requerimientos para el *retrohoming* *in vivo*, se ha determinado que la diana mínima necesaria para unos niveles adecuados de movilidad se extiende desde la posición -20 en el exón 5' hasta la posición +5 en el exón 3'. La mutación de los IBSs en esas dianas ha permitido comprobar que el IBS1 es menos restrictivo para el reconocimiento durante el *homing* que el IBS2 o que el IBS3 (Jiménez-Zurdo *et al.*, 2003). Por el contrario, los estudios desarrollados en los últimos años en relación con los requerimientos de secuencia para el *splicing* *in vivo* muestran justamente la especificidad inversa: la eficiencia de maduración de RmInt1 no está tan condicionada por

la identidad del IBS2 o el IBS3 como por la secuencia IBS1 (Barrientos-Durán, 2008; 2011).

Estudios basados en la pauta de invasión de plásmidos con replicación unidireccional tipo  $\theta$  (Martínez-Abarca *et al.*, 2004) así como en la dispersión del intrón en el genoma de la bacteria tras haber sido transformada con un plásmido donador (Nisa-Martínez *et al.*, 2007), muestran una clara tendencia para la invasión de la hebra que sirve como molde en la síntesis de la cadena retardada durante la replicación (Fig 24). También es capaz de insertarse en la otra cadena pero con mucha menor eficiencia (Fig 24). Por tanto, RmInt1 es el principal representante de los intrones cuya ruta de *retrohoming* es independiente de actividad En (Pyle y Lambowitz, 2006; Toro *et al.*, 2007, Lambowitz y Zimmerly, 2010). Recientemente se ha desarrollado un sistema de mutagénesis dirigida basado en RmInt1, por lo que este intrón también puede ser usado como herramienta biotecnológica (García-Rodríguez *et al.*, 2011).

# **OBJETIVOS**



## OBJETIVOS

Desde su descubrimiento a finales de los 90, nuestro grupo de investigación está trabajando en la caracterización del intrón del grupo II de *Sinorhizobium meliloti*, RmInt1. Los diferentes trabajos se han abordado desde muy diferentes perspectivas: molecular (*in vitro*), funcional (*in vivo*) y evolutiva. En los últimos años, además, se está trabajando en el diseño de herramientas de manipulación genética basadas en las propiedades de este intrón. Para llegar a comprender todos los procesos en los que RmInt1 se ve involucrado y ser capaz de modificarlos, es preciso caracterizar bioquímica y molecularmente las actividades que este intrón lleva a cabo. Así pues, el objetivo general de este trabajo es profundizar en la biología básica de RmInt1 particularizando en los siguientes objetivos:

1. Modificación de la estructura de RmInt1 y análisis de su capacidad de dispersión en *S. meliloti*.
2. Estudio de la interrelación entre RmInt1 y su hospedador natural *S. meliloti*. Determinación de los factores del hospedador que intervienen en la movilidad del intrón y análisis de cómo afecta la presencia de RmInt1 al transcriptoma general de *S. meliloti*
3. Construcción de un marcador de resistencia que pueda ser transportado por RmInt1. Uso del casete para la detección de eventos de movilidad.
4. Análisis de la movilidad de RmInt1 en otras especies bacterianas. Localización subcelular de la proteína codificada por el intrón en bacterias y células eucariotas (plantas).



# **MATERIAL Y MÉTODOS**





## 1. ESPECIES Y CEPAS BACTERIANAS

Todas las especies y cepas utilizadas en los ensayos de laboratorio se incluyen en la siguiente tabla:

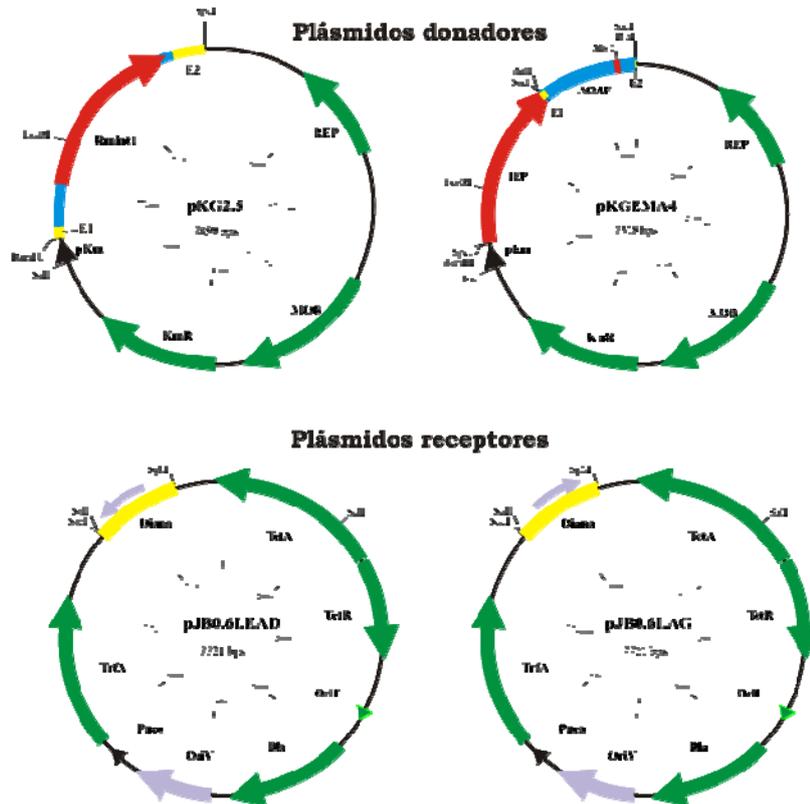
CEPA BACTERIANA	CARACTERÍSTICAS	REFERENCIA
<i>Sinorhizobium meliloti</i> GR4	Aislado de suelo de la EEZ; Nod+, Fix+	Casadesús y Olivares, 1979
<i>S. meliloti</i> RMO17	Aislado de nódulo: Nod+, Fix+	Villadas <i>et al.</i> , 1995
<i>S. meliloti</i> 1021	Nod+, Fix+, Sm <sup>r</sup> ; Genoma secuenciado y anotado	Galibert <i>et al.</i> , 2001
<i>S. meliloti</i> 2011	Nod+, Fix+, Sm <sup>r</sup> ; Derivado de SU47	Casse <i>et al.</i> , 1979
<i>Escherichia coli</i> DH5a	F <sup>-</sup> , ø80dlacZΔM15, Δ(lacZYA-argF)U169, deoR, recA1, endA1, hsdR17(rK <sup>-</sup> , mK <sup>+</sup> ), phoA, supE44, λ <sup>-</sup> , thi-1, gyrA96, relA1	Bethesda Research Lab.
<i>E. coli</i> HB101	supE44, Δ(mcrC-mrr), recA13, ara-14, proA2, lacY1, galK2, rpsL20, xyl-5, mtl-1, leuB6, thi-1	Promega
<i>E. coli</i> ES1301 mutS	lacZ53, mutS201::Tn5, thyA36, rha-5, metB1, deoC, IN(rrnD-rnE)	Promega
<i>E. coli</i> HMS174 (DE3)	F <sup>-</sup> recA1 hsdR(r <sub>K12</sub> <sup>-</sup> m <sub>K12</sub> <sup>+</sup> ) (DE3) (Rif <sup>R</sup> )	Novagen
<i>Pseudomonas Putida</i> KT2440	Derivado de <i>P. putida</i> mt-2; pTOL <sup>-</sup>	(Bagdasarian <i>et al.</i> , 1981)

## 2. PLÁSMIDOS Y VECTORES DE CLONAJE

Se utilizaron una serie de vectores para el clonaje de fragmentos dependiendo de la estrategia diseñada en cada momento. Por lo general, se utilizaron vectores de alto número de copia y que permiten la selección de colonias blancas como pGEM-T (Promega) o pKSBluescript (Stratagene).

Abajo se incluyen los mapas a escala de los plásmidos que se han utilizado en este trabajo, así como una tabla donde se señalan las características más relevantes a las que se hace mención a lo largo de este trabajo.

NOMBRE	USO	REFERENCIA
pKG2.5	Construcción donadora de intrón en la que RmInt1 se encuentra flanqueado por exones -174/+466. Resistencia a Km.	Martínez-Abarca <i>et al.</i> , 2000
pKGEMA4	Construcción donadora de intrón en la que la ribozima de RmInt1 ( $\Delta$ ORF) se encuentra flanqueado por exones -20/+5. Km <sup>R</sup>	Nisa-Martínez <i>et al.</i> , 2007
pJB0.6 LAG/LEAD	Plásmido receptor derivado de pJB3Tc19 que contiene 640 bp de la ISRm2011-2 clonada en una u otra orientación respecto del <i>orV</i> . Se ha usado en ensayos de <i>homing</i> por hibridación. Ap <sup>R</sup>	Martínez-Abarca <i>et al.</i> , 2000 ; 2004





### 3. OLIGONUCLEÓTIDOS

Los oligonucleótidos empleados en este trabajo se han sintetizado con *DNA Synthesizer® 3400* (Applied Biosystem) en el servicio del Instituto de Parasitología y Biomedicina “López-Neyra” (C.S.I.C., Granada). También se han usado oligonucleótidos sintetizados por Sigma.

En cada capítulo de resultados se presentan tablas que contienen los oligonucleótidos utilizados.

### 4. CULTIVOS BACTERIANOS

#### 4.1. MEDIOS DE CULTIVO

El medio completo utilizado para el crecimiento rutinario de *Sinorhizobium meliloti* fue TY (Beringer, 1974) preparado en agua desionizada:

CaCl <sub>2</sub> x 2H <sub>2</sub> O	0,9 g/l
Tryptona (DIFCO)	5 g/l
Extracto de levadura (DIFCO)	3 g/l

Para la preparación del medio sólido se añadió agar (PANREAC) al 1,6%. La esterilización se realizó en autoclave durante 20 minutos a 120°C.

Como medio rico para el cultivo de las diferentes cepas de *Escherichia coli* se utilizó el medio de Luria-Bertani (LB) (Sambrook *et al.*, 1989) preparado con agua desionizada:

NaCl	5 g/l
Tryptona (DIFCO)	10 g/l
Extracto de levadura	5 g/l

Al igual que en el caso del TY, para solidificar el medio se añadió agar (PANREAC) al 1,6%. El medio se esterilizó a 120°C durante 20 minutos.

Como medio mínimo (MM) que permite el crecimiento de *S. meliloti* limitando el de *E. coli* se utilizó la siguiente modificación del medio descrito por Robertsen *et al.* (1981) preparado en agua desionizada:

K <sub>2</sub> HPO <sub>4</sub>	0,3 g/l
---------------------------------	---------

KH <sub>2</sub> PO <sub>4</sub>	0,3 g/l
MgSO <sub>4</sub> .7H <sub>2</sub> O	0,15 g/l
CaCl <sub>2</sub>	0,05 g/l
FeCl <sub>3</sub>	0,006 g/l
NaCl	0,05 g/l
Glutamato sódico	1,1 g/l
Manitol	10 g/l
Biotina	0,0002 g/l
Pantotenato cálcico	0,0001 g/l

Ajustar el pH entre 6,8 y 7,0 con KOH si fuera necesario.

Para solidificar el medio se añadió agar purificado (OXOID) al 1,6% y se esteriliza durante 20 minutos a 120°C.

Como medio de cultivo restrictivo para coliformes se utilizó el Endo Agar™ LES (DIFCO). Este medio se utiliza específicamente para verificar que las colonias de *S. meliloti* seleccionadas de MM tras las conjugaciones están limpias de *E. coli*. Se prepara disolviendo 51 g por litro de agua desionizada, calentando hasta disolución completa y luego autoclavado a 120°C durante 20 minutos.

#### 4.2. ANTIBIÓTICOS

Las concentraciones finales de los antibióticos precisados en los diferentes protocolos fueron:

La adición de antibióticos a los medios de cultivo se hizo a partir de

ANTIBIÓTICO	<i>E. coli</i>	<i>S. meliloti</i>	<i>P. putida</i>
Tetraciclina (Tc, Sigma)	10 mg/l	10 mg/l	20 mg/ml
Kanamicina (Km, Roche)	50 mg/l	180 mg/l	50 mg/l
Ampicilina (Ap, Sigma)	200 mg/l	200 mg/l	
Estreptomicina (Sm, Sigma)	50 mg/l	250 mg/l	
Espectinomicina (Spc, Roche)	100mg/l	100 mg/l	
Gentamicina (Gm, Sigma)	10 mg/l	45 mg/l	
Cloranfenicol (Cm Sigma)	50 mg/l	50 mg/l	



soluciones 100 veces concentradas preparadas en agua desionizada (Km, Ap) o en agua-etanol al 50% (Tc). Las soluciones preparadas en agua se esterilizaron por filtración utilizando unidades Minisart® NML (Sartorius) de 0,2  $\mu\text{m}$  de tamaño de poro.

#### 4.3. CONDICIONES DE CULTIVO

Las distintas cepas de *S. meliloti* se crecieron a 28°C, siendo su tiempo de generación en medio líquido de aproximadamente 2,5 h a 190 r.p.m. en agitador orbital.

*E. coli* se incubó a 37°C y su tiempo de generación en medio líquido ronda los 30 min a 190 r.p.m. en agitador orbital.

#### 4.4. CONSERVACIÓN DE CULTIVOS BACTERIANOS

Para la conservación prolongada de los cultivos bacterianos se utilizó la congelación. Este método se basa en la paralización del metabolismo celular al disminuir la disponibilidad de agua. Para que la viabilidad celular se mantenga durante el periodo de conservación se precisa de lo que se denominan crioprotectores que evitan el daño que los cristales de agua podrían ocasionar en las paredes celulares; en nuestro caso, el crioprotector elegido fue glicerol. Se utiliza a una concentración de aprox. 20% (v/v); de manera que, a 250  $\mu\text{l}$  de glicerol 87% estéril dispuesto en criotubos, se le adicionaron 750  $\mu\text{l}$  de cultivo bacteriano crecido hasta fase logarítmica tardía-fase exponencial. Los viales se congelaron rápidamente y se almacenaron a una temperatura de -80°C.

### 5. AISLAMIENTO DE ÁCIDOS NUCLEICOS

En este apartado vamos a describir los protocolos de extracción de los diferentes ácidos nucleicos utilizados en posteriores ensayos.

#### 5.1. EXTRACCIÓN DE DNA total

Para este propósito se utilizó el kit comercial *AquaPure genomic DNA isolation* (Bio-Rad). Recogimos 1ml de cultivo de *Rhizobium* y lo lavamos con 0,2 ml de Sarcosil al 0,1% en TE. La lisis se lleva a cabo por la adición de 0,6 ml de “*Lysis Solution*” incubando 10 min a 80°C. Se deja enfriar a temperatura ambiente y añadimos 0,003 ml de “*Solución RNase*”. Se mezcla invirtiendo 25 veces e incubamos una hora a 37°C. Pasado este tiempo se adicionan 0,2 ml de “*Protein Precipitation*” y agitamos vigorosamente durante 20 seg. Incubar en hielo 10 min. La muestra se centrifuga a máximas revoluciones durante 5 min y recogemos el sobrenadante en un tubo nuevo.

El DNA se precipita añadiendo 0,6 ml de isopropanol frío invirtiendo hasta la mezcla total durante 10 min a -20°C. Centrifugamos 10 min y lavamos el sedimento con 0,2 ml de etanol 70%. Se resuspende con 0,04 ml de agua bidestilada.

### **5.1.1. Lisados de colonia**

La siguiente técnica consiste en la obtención rápida de DNA con suficiente calidad y cantidad para protocolos como PCR pero insuficiente para otras técnicas (Villadas *et al.*, 1995). Consiste en resuspender en 0,1 ml de agua desionizada<sup>1</sup> estéril una colonia picada de placa en condiciones de esterilidad. Se sedimentan las células por centrifugación durante 2 min a 13000 r.p.m. Se elimina el sobrenadante y se lava 2 veces más con 0,1 ml de agua. Tras el último lavado, las células se resuspenden en 0,1 ml de agua y se incuban en agua hirviendo durante 10 min. Centrifugamos 2 min y el sobrenadante resultante (concretamente 2-4 µl) directamente se usa de molde para el PCR.

## **5.2. EXTRACCIÓN DNA PLASMÍDICO**

Dependiendo de las necesidades de cada ensayo, se utilizó uno u otro protocolo. En líneas generales, el aislamiento mediante kit comercial se restringió al envío de muestras para secuenciación. El procedimiento de lisis seguida de precipitación con sales de magnesio funciona muy bien en células de *E. coli*; sin embargo, el procedimiento rutinario para la extracción plasmídica en *S. meliloti* fue la lisis alcalina. Cuando trabajamos con plásmidos de bajo número de copia, con frecuencia se precisa partir de un elevado volumen de cultivo y recurrimos a las midipreparaciones de plásmidos mediante kit comercial.

### **5.2.1. Aislamiento de plásmidos por lisis alcalina**

El aislamiento de DNA plasmídico de bajo peso molecular se realizó siguiendo el protocolo modificado por Sambrook y colaboradores (Birnboim y Doly, 1979; Ish-Horowicz y Burke, 1981; Sambrook *et al.*, 1989). Se recogen 1,5 ml de cultivo bacteriano crecido hasta fase exponencial y se sedimentan a 13000 r.p.m. durante 2 min. Lavamos las células con 500 µl de sarcosil 0,1% en TE y volvemos a centrifugar, retirando el sobrenadante. Resuspendemos las células con 100 µl de lisozima 4 mg/ml disuelta en solución I (glucosa 50 mM, Tris-HCl 25 mM pH 8, EDTA 10 mM; conservada

---

<sup>1</sup> Debemos tener presente que si estamos lisando células de rizobios este agua debe contener sarcosil 0,1% (p/v).



a 4°C). Se incuba 5min a temperatura ambiente, tras los cuales se añaden 200 µl de solución II (NaOH 0,2 M, 1% SDS). Se homogeniza la preparación agitando el tubo por inversión varias veces y se mantiene en hielo durante 5 min. Entonces se adicionan 150 µl de solución III (acetato potásico 5 M pH 4,8) que ayudará a la neutralización de la solución y la precipitación de proteínas. Se mezcla invirtiendo el tubo y se incuba en hielo 5 min para minimizar el posible daño con DNAsas celulares. Centrifugamos a 13.000 r.p.m. durante 5 para sedimentar los restos celulares y recogemos el sobrenadante en un tubo nuevo. Añadimos un volumen (~400 µl) de ΦCIA y mezclamos vigorosamente. Nuestra muestra va a separarse en dos fases tras una centrifugación de 5 min a 13.000 r.p.m. Recogemos la fase superior acuosa y la mezclamos con un volumen de CIA (elimina posibles restos de fenol en las preparaciones). Volvemos a mezclar y centrifugamos durante 5 min. Pasamos la fase acuosa a un tubo nuevo y precipitamos el DNA con 2,5 volúmenes de etanol 100% frío. Podemos optar por precipitar 30 min a -20°C o 15 min a -70°C, pero lo más habitual para evitar la precipitación de DNA genómico son 5 min a temperatura ambiente. Centrifugamos el DNA precipitado y lavamos el sedimento con 200 µl de etanol al 70%. Dejamos secar y resuspendemos en 25 µl de agua desionizada en la que habremos diluido una solución 100 veces concentrada de RNAsa A<sup>2</sup> (10 µg/ml de concentración final).

### **5.2.2. Extracción de DNA plasmídico por precipitación con sales de magnesio**

Esta técnica está basada en el método de Studier (1991), permite el aislamiento rápido de plásmidos de *E. coli*. Recogemos 1,5 ml de cultivo bacteriano por centrifugación y, tras retirar el sobrenadante, se resuspenden las células en 100 µl de agua desionizada. La lisis tiene lugar por la adición de 100 µl de una solución compuesta por NaOH 0,1 M, EDTA 10 mM y 2% SDS. Se homogeneiza por agitación y se hierven durante 2 min. El DNA lineal, normalmente cromosómico, se precipita por la adición de 50 µl de MgCl<sub>2</sub> 1 M. Tras homogeneizar completamente (aparición de precipitado blanco), se incuba en hielo de 2-5 min. Centrifugamos 1 min para sedimentar el precipitado en el fondo del tubo y añadimos 50 µl de acetato potásico 5 M para eliminar proteínas. El tubo invertido se mezcla con vórtex

---

<sup>2</sup> La solución RNAsa 10 mg/ml se prepara diluyendo, en las cantidades indicadas, la enzima (liofilizada, Sigma) en una solución con 10 mM Tris-HCl pH 7.5 y 15 mM NaCl. Hervir durante 15' y dejar enfriar a RT. Se conserva a -20°C. Debemos distribuirla en alícuotas porque la congelación-descongelación hace que pierda actividad (Sambrook *et al.*, 1989).

evitando que se levante el precipitado blanco del fondo y se incubaba 5 min en hielo. Centrifugamos durante 5 min y recogemos el sobrenadante a un tubo nuevo donde precipitaremos el DNA plasmídico con 0,6 ml de etanol 100% guardado a -20°C. Se deja 5 min a temperatura ambiente y se centrifuga 5 min. Lavamos el sedimento con 200 µl de etanol al 70% y dejamos secar a vacío. Resuspender en el volumen adecuado dependiendo de la concentración deseada o necesaria para las diferentes aplicaciones; lo más habitual para un plásmido de alto número de copias es resuspender en 25 µl agua desionizada adicionada de RNAsa 10 µg/ml.

### **5.2.3. Aislamiento de DNA plasmídico mediante kit comercial**

Durante este periodo hemos usado distintos kits comerciales para el aislamiento de DNA plasmídico en todos los casos se han seguido las indicaciones de uso del proveedor: *QIAprep Spin Miniprep Kit* (Qiagen) y *PerfectPrep® plasmid mini* (Eppendorf). En todos los casos consistió en la lisis celular seguida de la unión del DNA en disolución a una resina/membrana y elución con agua. La ventaja fundamental es la limpieza del DNA y la ausencia de RNAsa en los pasos finales.

### **5.2.4. Extracción de plásmidos a partir de grandes volúmenes de cultivo (midipreparaciones)**

Con este fin también se usó un kit comercial, *QIAGEN Plasmid Midi Kit* (Qiagen). Es un protocolo rápido y sencillo de aislamiento de DNA plasmídico ultrapuro superenrollado con alto rendimiento. Está basado en protocolos de purificación plasmídica por el procedimiento de lisis alcalina modificada, seguido de la unión del plásmido a una resina de intercambio aniónico en condiciones de baja sal y pH apropiado. Todo tipo de posibles contaminantes son eliminados mediante el lavado con soluciones media sal y posteriormente el DNA es eluido en tampones con alta sal. Tras lo cual precisa ser precipitado. Partimos de 100 ml de cultivo que recogemos por centrifugación a 7000 r.p.m. (rotor GSA en centrífuga *Sorvall*) durante 15 min a 4°C. Se siguieron las instrucciones del proveedor hasta eluir nuestro DNA. La precipitación se llevó a cabo en tubos de 1,5 ml en alícuotas de 1 ml a las que se añadieron 700 µl de isopropanol frío. Centrifugamos a máximas revoluciones durante 30 min y lavamos el exceso de sales con 200 µl de etanol 70% frío. Dejamos secar y el DNA contenido en cada tubo se resuspende en 20 µl de agua bidestilada (De 100 ml de cultivo de partida se obtienen ~100 µl de plásmido ultrapuro).



### 5.3. EXTRACCIÓN DE RNA total

#### 5.3.1. Aislamiento de RNA total de *S. meliloti*

Este protocolo presenta algunas modificaciones sobre el descrito por Cabanes *et al.* (2000). Se lleva a cabo la inoculación de 0.5 ml. del cultivo bacteriano saturado en un tubo que contiene 9.5 ml. de medio con el correspondiente antibiótico. Se crece hasta el final de su fase exponencial, esto es  $\sim 0,6$  U de  $A_{600}$ . Recogemos los 10 ml de cultivo en dos tubos de 1,5 ml gracias a varias centrifugaciones de 1-2 min a 14000 r.p.m. Resuspendemos las células sedimentadas con 300  $\mu$ l de solución compuesta por 1.4% SDS, 4mM EDTA y 42 mg/ml proteinasa K. Se incuba durante 10 min a 65°C, agitando de manera esporádica (cada 3 min). Pasado este tiempo, los tubos se ponen en hielo y se adicionan 150  $\mu$ l de NaCl 5M a 4°C, se mezclan vigorosamente y se mantienen en hielo durante 10 min. Centrifugamos 15 min a 4°C y 13000 r.p.m. Se pone el sobrenadante en un tubo nuevo y se precipita con 1 ml de etanol 100% frío (en este punto podemos dejarlo varios días). Se centrifuga el ácido nucleico precipitado durante 30 min a la máxima velocidad de la centrifuga y 4°C. Eliminamos muy bien todos los restos de etanol y dejamos secar ligeramente a temperatura ambiente. Resuspendemos los precipitados correspondientes a los 10 ml de cultivo inicial en 85  $\mu$ l de agua bidestilada mezclando los duplos para el tratamiento con DNasa I. La digestión del DNA en nuestras preparaciones tiene lugar por la incubación durante una hora a 37°C de estos 85  $\mu$ l con 50 U de DNasa I libre de RNAsas (Roche) en presencia de 60 U de inhibidor de RNAsas (GE healthcare) en el tampón adecuado (50 mM Tris-HCl pH 7.5, 1 mM  $MgCl_2$ ). Podemos agitar de vez en cuando. Transcurrido este tiempo, se eliminan las proteínas mediante dos fenolizaciones sucesivas: en la primera se añade 1V (100  $\mu$ l) de  $\Phi$ CIA<sup>3</sup>, se toma la fase superior y se mezcla con 1V de CIA. En el primer caso se centrifuga durante 10 min y en el segundo es suficiente con 5 min a 4°C y máxima velocidad. La fase acuosa superior se precipita con 600  $\mu$ l de etanol 100%/75 mM acetato sódico 3 M pH 5,2 frío. Se mantiene al menos una hora a -80°C. Sedimentar por centrifugación a 4°C durante 30 min y lavar con 0,5 ml de etanol 70% frío. Secar la muestra sin temperatura y resuspender en 20  $\mu$ l de agua bidestilada.

---

<sup>3</sup> Al trabajar con RNA es conveniente que utilicemos fenol a pH ácido (amresco) porque este tipo de fenol arrastra con más facilidad al DNA, dejando en la fase acuosa el ARN (Sambrook *et al.*, 1989).

### 5.3.2. Extracción de RNA total mediante kit comercial

Para el conseguir marcar con éxito los RNAs utilizados para la hibridación de los DNA microarrays, ha sido precisa la extracción de RNA total mediante el uso de kit comercial, concretamente *RNeasy Protect Bacteria Mini Kit* (Qiagen).

Con objeto de paralizar lo más inmediatamente posible cualquier alteración que la bacteria pudiera sufrir a nivel de transcriptoma, la recogida y centrifugación de las muestras se lleva a cabo de manera rápida (1 min). Una vez sedimentadas las células, retiramos el sobrenadante y rápidamente congelamos en N<sub>2</sub> líquido. La descongelación se lleva a cabo en hielo.

*S. meliloti* es una bacteria bastante mucosa por lo que, previo a lisis, usamos 100 µl de lisozima 0,4 mg/ml (por cada tubo correspondiente a 1,5 ml de cultivo) que ayuda a digerir la pared y mejora la lisis celular. Se incuban 5 min a temperatura ambiente (RT). Para proceder con la lisis se añaden 350 µl de tampón RLT (al que añadimos en el momento β-mercaptoetanol a la concentración indicada por la casa comercial), agitamos vigorosamente y se centrifuga 2 min para eliminar las envueltas celulares. Se pasó el sobrenadante a un tubo nuevo al que añadimos 250 µl de Etanol 100% mezclando con pipeta. Se procede a fijar el ácido nucleico sobre las columnas (por cada columna se pasa el sobrenadante correspondiente a 3 ml de cultivo) centrifugando 30 seg a max revoluciones. Se lavaron las columnas con 700 µl de tampón RW1. Se trataron dos veces sucesivas con 500 µl de RPE y se llevó a cabo una última centrifugación de 2 min para eliminar cualquier resto de etanol de las columnas que reduce la eficiencia de elución. Finalmente se eluyó en dos pasos adicionando 30 µl de agua sobre la resina e incubando un min a RT para posteriormente centrifugar 1 min.

Por lo general, la tecnología en la que se basan la columna minimiza la cantidad de DNA que es capaz de permanecer pegado a la resina. Sin embargo, este protocolo precisa de un RNA completamente libre de DNA (que puede interferir en la eficiencia de hibridación por inhibición competitiva). Para ello, llevamos a cabo la digestión del RNA resultante con DNAsa (Qiagen). Reunimos 120 µl de RNA y se añadieron 14 µl de tampón RDD y 6 µl de DNAsa. Se digirió durante una hora a 28°C (la incubación podría alargarse media hora más). A los 140 µl de reacción se añadieron 490 µl de tampón RLT (con β-mer) y 350 µl de etanol 100%. Mezclamos pipeteando, se dispuso sobre una columna y centrifugamos 30 seg. Se lava en dos ocasiones: primero con 350 µl de RW1 y luego, con 500 µl de RPE. Al igual que se describe anteriormente, eliminamos cualquier resto de soluciones de



lavado mediante una centrifugación adicional de dos min. Se eluye en dos pasos de manera que obtengamos un volumen final de 50  $\mu$ l.

Por lo general, el rendimiento de los kits comerciales suele estar ligado a dilución de la muestra; así, llevamos a cabo un último paso adicional de concentración del RNA. Con este fin se utilizaron columnas Microcon 30 (Millipore) que además de concentrar nuestra muestra también suponen una limpieza del RNA que, en ocasiones, se traduce en una mayor eficiencia de marcaje. Llevamos la muestra hasta un volumen final de 500  $\mu$ l con agua, ponemos sobre la parte gris de la columna y mezclamos pipeteando. Centrifugamos hasta reducir el volumen a  $\sim$ 35  $\mu$ l (generalmente, se utilizaron en torno a 8 min pero depende de la muestra). Damos la vuelta a la columna sobre un tubo nuevo y centrifugamos 2 min.

Aunque el RNA es estable a  $-20^{\circ}\text{C}$  es aconsejable conservarlo a  $-80^{\circ}\text{C}$ .

#### 5.4. CUANTIFICACIÓN DE LA CONCENTRACIÓN DE RNA Y DNA

El cálculo de la concentración de ácidos nucleicos se estimó siguiendo el método espectrofotométrico descrito en Sambrook *et al.* (1989). Se midió la absorbancia de las muestras en solución acuosa a longitudes de onda 260 y 280 nm utilizando cubetas de cuarzo. La concentración se determinó en función de los valores estándar: 1 U a  $\text{D.O.}_{260\text{nm}}$  se corresponde con 50  $\mu\text{g/ml}$  de DNA de cadena doble, 40  $\mu\text{g/ml}$  de DNA de cadena sencilla y RNA, y 33  $\mu\text{g/ml}$  de oligonucleótidos. La relación  $\text{D.O.}_{260\text{nm}} / \text{D.O.}_{280\text{nm}}$  se usó como estimador de la pureza de la preparación, considerándose valores por debajo de 1,8 como indicadores de contaminación por proteínas o fenol.

Más recientemente se ha utilizado *NanoDrop*<sup>®</sup> *ND-1000*, un espectrofotómetro que permite la determinación de la concentración a partir de pequeños volúmenes de muestra (1  $\mu$ l).

## 6. MANIPULACIÓN Y AMPLIFICACIÓN DE DNA

### 6.1. DIGESTIÓN CON ENZIMAS DE RESTRICCIÓN

La digestión completa del DNA se realizó siguiendo las indicaciones de temperatura y tampón óptimos recomendadas por los proveedores (Roche y New England Biolabs). Lo más común es utilizar 5 U de enzima/ $\mu\text{g}$  DNA a digerir aunque, en el caso de muestras complejas como DNA total, esta cantidad debe elevarse al doble. A esto debemos sumarle una prolongación del tiempo de digestión de manera que en ocasiones la digestión se ha realizado durante la noche, aunque lo habitual es una hora. En el caso de

querer realizar digestiones con 2 enzimas de restricción simultáneamente hay que tener precaución con el tampón elegido. Debemos tratar de buscar un tampón compatible para ambas (instrucciones del proveedor) y, en caso de no ser posible, realizar primero la digestión con la enzima que precise el tampón con menor fuerza iónica. Si aún así no fuese factible, el DNA se limpiaría de sales y enzima tras la primera digestión, y se procedería con la segunda, así sucesivamente. Hay dos alternativas para eliminar las sales: la primera es el uso de kits comerciales como *QIAquick PCR Purification Kit* (Qiagen), pero también podemos optar por fenolizar y precipitar nuestro DNA.

## 6.2. REACCIÓN DE DEFOSFORILACIÓN

En ocasiones es necesaria la eliminación de los grupos fosfato presentes en los extremos 5' de vectores linearizados para evitar que vuelvan a ligarse. Esto nos evita el análisis de muchas colonias que sean vector re-ligado en vez de contener el inserto que nos interesa. La defosforilación de 5 µg de plásmido digerido precisa 1 U de fosfatasa alcalina (CIP, *Calf Intestine Phosphatase*; Roche) en tampón correspondiente durante 1 h a 37°C. La reacción separa añadiendo 55 µl de una solución compuesta por tampón STE 1x [1 mM Tris-HCl pH8, 100 mM NaCl y 1 mM EDTA] y SDS 1% e incubando 15 min a 68°C. Para obtener mejor rendimiento en la ligación es conveniente fenolizar y precipitar.

## 6.3. PCR

La reacción en cadena de la polimerasa (Polymerase Chain Reaction, PCR) consiste en la amplificación en sucesivos ciclos de un material de partida gracias al uso de una pareja de cebadores y una polimerasa termoestable (*Taq* polimerasa). Puede tener distintas finalidades y en nuestro caso la hemos utilizado para amplificar insertos que después necesitamos clonar añadiendo determinadas enzimas en los extremos hasta para la construcción de mutantes, pasando por comprobaciones rutinaria de la presencia de un determinado inserto. Consiste en 3 pasos fundamentales repartidos en 3 ciclos: desnaturalización, anillamiento y extensión. La desnaturalización suele realizarse a 94°C 30 seg aunque el tiempo depende del tipo de molde; al comienzo del programa de PCR suele incluirse un primer ciclo de desnaturalización de 2 min para disponer de todo el DNA como cadena simple. La temperatura de anillamiento depende de la composición de los cebadores pero suele rondar los 55°C y el tiempo estándar es 30 seg. La extensión no es más que la temperatura óptima para



la actividad de la enzima; el tiempo de este ciclo suele estar en función del tamaño del amplificado y del tipo de enzima usado (por ejemplo, Taq polimerasa 1Kb/30 seg). Además el programa de PCR suele terminarse con un ciclo prolongado de extensión de 5-10 min para completar los fragmentos. Depende de la finalidad perseguida se utilizaron distintas polimerasas: *TripleMaster*<sup>®</sup> (Eppendorf), *High Fidelity*<sup>®</sup> (Roche) o *Pfu* polimerasa (Promega), si era necesaria una baja tasa de errores en el producto final, mientras que para comprobaciones de rutina se usó una *Taq* polimerasa purificada en el laboratorio.

Las reacciones se llevaron a cabo en un volumen final de 25  $\mu$ l conteniendo 10 a 100 ng del DNA molde, 25 pmol de cada uno de los cebadores específicos, 0,1 mM de deoxinucleósidos trifosfato (1 M si se trata de DNA total o de lisados celulares), el tampón especificado para cada aplicación y 2 unidades de DNA polimerasa. La amplificación se realizó en un termociclador *Eppendorf Mastercycler*<sup>®</sup> y el programa empleado varió dependiendo del objetivo.

La limpieza de este tipo de productos (eliminación de cebadores y sales) se realizó mediante columnas *MicroSpin*<sup>™</sup> *S-300HR* (GE Healthcare) o con el uso del *QIAquick PCR Purification Kit* (Qiagen).

### **6.3.1. PCR de colonia**

Es una estrategia muy rápida para la comprobación de la presencia de determinados insertos a partir de las colonias en placa. Es más sucia que utilizar lisados de colonias pero de efectividad similar. Se trata de repartir los 25  $\mu$ l de la mezcla de reactivos de PCR y, en lugar de poner el DNA o el lisado, picar directamente una colonia de la placa con palillo y agitar dentro de la mezcla.

### **6.3.2. RT-qPCR**

Se trata de una variante del PCR convencional que permite la cuantificación de la expresión génica. A diferencia del PCR, la inclusión en la reacción de un intercalante y la tecnología adecuada para la detección de la intensidad de fluorescencia en cada ciclo, la convierte en una herramienta cuantitativa y no sólo cualitativa. La síntesis de CDNA que precede la qPCR nos capacita para evaluar los niveles de un determinado RNA mensajero presente en nuestras muestras. En nuestro caso, esta técnica se utilizó para verificar los resultados derivados de los estudios de hibridación de DNA microarrays, por lo cual los RNAs utilizados fueron exactamente los mismos que se describen para esta técnica (RNA extraído con kit Qiagen).

Existen diferentes enzimas y diferentes procedimientos que acoplan ambos procesos en un sólo paso o que los dividen en dos procedimientos independientes. En los últimos años, las casas comerciales han desarrollado nuevas mejoras en el campo de la PCR a tiempo real que aumentan su eficiencia y reproducibilidad. Nosotros utilizamos AMV reverso transcriptasa (Roche) para la síntesis del cDNA y Taq platinum polimerasa (Invitrogen) para la amplificación. AMV RT es una enzima muy eficiente que además presenta un dominio RNasa H<sup>4</sup>; la importancia de esta actividad radica en que no amplifica el cDNA sintetizado, sino que se limita a copiar el RNA presente en la muestra. Por otro lado, la particularidad de la polimerasa es que se encuentra acomplejada con un anticuerpo que impide su funcionalidad; tan pronto como tiene lugar el primer ciclo de desnaturalización la enzima se libera del complejo y comienza a funcionar. Este hecho reduce el fondo en la reacción, factor determinante a la hora de la cuantificación.

Finalmente decir que, aunque existen diversas formas de expresar los datos resultantes, en nuestro caso decidimos hacer una cuantificación relativa, esto es, usando un control interno a nuestra propia muestra.

#### 6.3.2.1. Síntesis del cDNA.

Partimos de 10 µg de RNA total procedente de las distintas cepas de *S.meliloti*. Se anilla con 50 ng de hexámeros de secuencia aleatoria en presencia de 50 mM Tris-HCl, 8 mM MgCl<sub>2</sub>, 30 mM KCl, 1 mM dithiothreitol, pH 8.5 y en un volumen final de 8,25 µl. Para asegurarnos que el RNA está completamente desnaturalizado, se calienta a 90°C durante 2 min; seguidamente se incuba a periodos de 15 seg a temperaturas decrecientes en 5°C, esto es, 85°C 15 seg/ 80°C 15 seg...hasta alcanzar los 25°C. En este momento, se le añade la mezcla de extensión que contiene dNTPs (0,625 mM cada uno), 15 U de inhibidor de RNAsas y 50 U de AMV RT hasta alcanzar un volumen final de 10 µl y se incuba 1 h a 42°C. Pasado este tiempo, se vuelven a añadir 50 U de enzima y esperamos 2 h más. La inactivación de la reverso transcriptasa tiene lugar a 85°C durante 15 min.

Es recomendable usarlo fresco, pero también es posible mantenerlo a 4°C o a -20°C.

---

<sup>4</sup> El sustrato de la RNasa H es la molécula de ARN que se encuentra formando parte de un híbrido con DNA.



### 6.3.2.1. qPCR.

Se utilizaron diferentes controles y a distintos niveles. Por un lado, los controles de agua para comprobar el nivel de fluorescencia basal, si es que lo hubiera. También se incluyeron muestras no sometidas a reverso transcripción que serían indicativo, en caso de amplificación, de contaminación con DNA en nuestras muestras. Finalmente, aunque como se ha descrito previamente la cuantificación llevada a cabo fue relativa, se utilizaron curvas patrón para la determinación de los parámetros de amplificación correspondientes a cada pareja de oligonucleótidos.

En la reacción de PCR están presentes todos los componentes que habitualmente se encuentran, pero además se incluye un agente intercalante fluorescente que en nuestro caso fue Sybr Green. La amplificación se lleva a cabo en iCycler iQ system (Bio Rad) controlada por el programa iQ5 de la misma casa comercial. Se mezclaron 0,2 mM de la pareja de cebadores correspondiente con una dilución 1:10 del CDNA sintetizado previamente (o en su defecto la cantidad equivalente de muestra no sometida a reverso transcripción, agua o DNA de concentración conocida) en presencia del tampón suministrado por la casa comercial, 1,5 mM de MgCl<sub>2</sub>, 0,1 mM dNTPs, una dilución 1:10000 de Sybr Green<sup>5</sup> y 0,5 U de Platinum Taq polimerasa. El programa utilizado para la amplificación fue: 95°C 3 min, seguido de 40 ciclos 95°C 15s/68°C 30s/72°C 15s, con un ciclo final de extensión a 72°C 3 min.

Las muestras se disponen al menos en duplos aunque lo ideal son triplos. El número de moléculas presentes en los estándares de DNA oscila entre 10 y 10<sup>6</sup>.

Se utilizan placas de 96 pocillos y es importante que no tengan burbujas y que todo el contenido se encuentre en el fondo porque de lo contrario obtendremos lecturas anómalas de la fluorescencia.

## 7. ELECTROFORESIS

Consiste en la separación de macromoléculas en función de su tamaño y carga. Los ácidos nucleicos, con carga negativa, migran hacia el polo positivo.

---

<sup>5</sup> Se prepara mezclando 1µl Sybr Green + 1µl fluoresceína + 18µl DMSO. Esta mezcla puede conservarse a 4°C durante una semana. Posteriormente, se realiza la dilución de 3 µl de la mezcla anterior en 500 µl agua y esta dilución es la que se usa en el PCR. La dilución se aconseja prepararla fresca.

### 7.1. ELECTROFORESIS EN GELES DE AGAROSA NO DESNATURALIZANTES

El análisis de productos de PCR, de DNA total o plasmídico o de fragmentos de restricción se llevó a cabo mediante electroforesis horizontal en geles de agarosa (agarosa *SeaKem® LE*, Cambrex/Iberlabo) al 1% en TAE 1x: [TrisHCl 40 mM, EDTA Na<sub>2</sub> 2 mM, ácido acético glacial 0,1142%(v/v)]. El voltaje habitual de trabajo fue de 120 V aunque determinados procedimientos requirieron velocidades inferiores. Como tampón de carga se utilizó una solución 6x (1µl por cada 5 de muestra) de 0,25% (p/v) de azul de bromofenol y glicerol 45% (v/v).

El revelado de los geles de agarosa se realizó por inmersión de los mismos en una solución de bromuro de etidio 1 mg/ml en agua durante 15 a 20 min. Este intercalante (cancerígeno) es visible en un transiluminador que emite luz UV de 365 nm de longitud de onda. Las fotografías de los geles se realizaron con *Gelprinter* y *Gelstation* de TDI para grabar y procesar las imágenes.

### 7.2. ELECTROFORESIS EN GELES DE AGAROSA DESNATURALIZANTES

Esta técnica se ha utilizado para comprobar la integridad del RNA total extraído. Consiste en la adición de formaldehído durante la preparación del gel de agarosa como agente desnaturalizante. El porcentaje del gel fue 1,4% (p/v). Se prepara disolviendo (calentando) la agarosa (Roche) en agua; cuando se disipe el calor se añade MOPS<sup>6</sup> [4X: MOPS 80 mM, acetato sódico 20 mM y EDTA 4 mM; ajustar a pH 7 con NaOH] y formaldehído hasta concentración final 1X y 1,875% (v/v), respectivamente. El tampón de electroforesis es MOPS 1X. Para evitar la posible degradación que pudiera ocurrir al sumergir el gel en la solución de bromuro de etidio, éste se adiciona directamente en el tampón de carga 6X [glicerol 45% (v/v), azul de bromofenol 0,09% (p/v) y bromuro de etidio 0,25 mg/ml].

### 7.3. ELECTROFORESIS EN GELES DE POLIACRILAMIDA DESNATURALIZANTES

Para técnicas que precisan de una buena resolución de las bandas recurrimos a electroforesis vertical en geles de poliacrilamida desnaturalizantes por la adición de urea 8 M (Roche). El tampón de electroforesis fue TBE 1X. Se preparan secuencialmente, esto es, primero se

---

<sup>6</sup> El MOPS es sensible al calor y a la luz, perdiendo parte de su actividad.



procede a la disolución completa de la urea aplicando calor pero impidiendo que hierva; después se mezclan el resto de componentes en las proporciones adecuadas para finalmente añadir el agente iniciador y el catalizador. Como ejemplo práctico: 25,2 g de urea, 6 ml TBE 10x, 9 ml de Acrid/Bis 40% completando con agua hasta un volumen final de 60 ml; 20  $\mu$ l de TEMED y 600  $\mu$ l de APS 10%.

A la hora de cargar este tipo de geles hay que tener dos precauciones: que el gel haya precorrido durante 1h a 1500 V (~45W) y haya adquirido temperatura (en torno a los 50°C) y que las muestras antes de cargarlas estén desnaturalizadas. Para la desnaturalización se requiere, además de estar resuspendidas en un tampón desnaturalizante [Tampón 2x: 0,3% Azul de bromofenol, 0,3% Xilencianol, 10 mM EDTA pH7,5 y 97,5% formamida desionizada], el calentamiento de las muestras durante 5 min a 95°C y luego se pasan a hielo. El gel se corre a 1700 V (~45W); la migración de los colorantes es un indicativo del tiempo de electroforesis. Para el revelado de los geles es preciso secar los geles a 80°C y vacío (*Gel Dryer*® 583, Bio-Rad). A continuación se exponen a una pantalla de exposición (*Imagin Plate 2040*, Fujifilm) y se revelan tras 12-48 h (*Personal Molecular Imager*® FX dependiente de *Quantity One*®, ambos de Bio-Rad).

#### 7.4. MARCADORES DE PESO MOLECULAR

Los marcadores de peso molecular fueron en todos los casos comerciales:

- Marcador II: DNA del fago  $\lambda$  digerido con la enzima HindIII. Está compuesto por 8 fragmentos que abarcan desde los cientos de nucleótidos hasta más de 20Kbp (Roche; Universidad Autónoma de Madrid). También se dispone marcado con digoxigenina que marcan el peso molecular en las hibridaciones DNA-DNA.
- Marcador III: DNA del fago  $\lambda$  digerido con las enzimas HindIII y EcoRI. En este caso el patrón de digestión comprende 13 fragmentos que abarcan desde los cientos de nucleótidos hasta más de 20Kbp (Roche). También marcado con digoxigenina.
- Marcador pGEM: DNA del plásmido pGEM-T digerido con las enzimas HinfI y EcoRI generando 15 fragmentos que se encuentran entre las pocos nt hasta las ~2500 bp (Promega).
- Marcador  $\Phi$ 29: DNA del fago con el mismo nombre digerido con HindIII. El patrón de digestión se compone de 14 fragmentos entre las 72 y la 4370 bp (UAM).

## 8. CLONAJE DE FRAGMENTOS

En este apartado se incluyen distintos protocolos utilizados habitualmente para el clonaje de insertos de interés en vectores.

### 8.1. AISLAMIENTO DE FRAGMENTOS DE DNA DE GEL DE AGAROSA

Para el aislamiento de fragmentos de restricción separados por electroforesis horizontal en gel de agarosa se utilizaron kits comerciales de acuerdo con las directrices proporcionadas por los proveedores: *PerfectPrep Gel Cleanup*® (Eppendorf) y/o *QIAquick Gel Extraction Kit* (Qiagen). El gel se tiñe por inmersión en bromuro de etidio, de manera que podamos cortar la banda de interés sirviéndonos del transiluminador UV.

### 8.2. LIGACIÓN DE FRAGMENTOS DE RESTRICCIÓN

La unión covalente de dos moléculas de DNA lineal se lleva a cabo gracias a la T4 DNA ligasa (Roche). La proporción molar más habitual vector:inserto suele ser 1:3 aunque puede variar dependiendo del tamaño de los fragmentos a clonar. Las reacciones se llevaron a cabo en un volumen final de 10-15  $\mu$ l a 14-16°C durante toda la noche.

#### 8.2.1. Adenilación de amplificadores de PCR

En ocasiones el producto que se precisa ligar no es un fragmento de restricción sino un producto de PCR. En el mercado se encuentran disponibles vectores que permiten la clonación de amplificadores con T protuberantes (extremos que deja la *Taq* polimerasa), pero algunas polimerasas con actividad exonucleasa 3'-5' dejan extremos romos. Por eso es un procedimiento habitual la adición de adeninas en los extremos de estos fragmentos; este procedimiento está catalizado por la *Taq* polimerasa (5 U) en presencia de 0,2 mM de dATP (70°C durante 30 min). Este producto puede ser utilizado directamente en la reacción de ligación.

### 8.3. PREPARACIÓN DE CÉLULAS COMPETENTES

Una vez que tenemos el DNA circularizado necesitamos que se replique en la bacteria. Con este fin se utilizan un tipo células que está preparado para captar DNA exógeno, las células competentes. Su preparación puede estar mediada por métodos químicos o físicos que alteran la pared celular de la bacteria. En nuestro caso sólo se han preparado células competentes de distintas cepas de *E. coli* y observamos una mayor eficiencia de transformación en células electrocompetentes.



### 8.3.1. Células competentes por métodos químicos: RbCl

En general, este tipo de protocolos deben llevarse a cabo en condiciones de esterilidad; es aconsejable que el proceso sea rápido y en frío para que el metabolismo celular se enlentezca en el mayor grado posible y consigamos el mayor número de células viables para la transformación.

Se cultivan las células de *E. coli* en 100 ml de medio LB hasta que se alcance una  $A_{600}$  de 0,4 (fase logarítmica). Entonces se detiene el crecimiento incubando el cultivo en hielo durante 15 min. Transcurrido este tiempo, las bacterias se sedimentan por centrifugación (6000 r.p.m., 10 minutos, 4 °C). Con cuidado, se resuspende el sedimento con 32 ml de solución RF1 estéril y preenfriada a 4 °C [RF1 (por 100 ml): 1,2 g de RbCl; 0,99 g de  $MnCl_2 \times 4 H_2O$ ; 0,294 g de acetato potásico; 0,15 g de  $CaCl_2 \times H_2O$ ; 11,9 ml de glicerol; agua destilada hasta 100 ml. Ajustar el pH a 5,8 con una solución de ácido acético 0,2 M. Esterilizar por filtración y conservar a 4°C]. Se incubaron en hielo durante 15 minutos y después se centrifugaron 6.000 r.p.m. durante 10 minutos. Se retira el sobrenadante y las células se resuspenden en 4 ml de solución RF2 preenfriada a 4 °C [RF2 (por 50 ml): 0,1046 g de ácido morfolino propanosulfónico (MOPS); 0,06 g de RbCl; 0,55 g de  $CaCl_2 \times H_2O$ ; 5,95 ml de glicerol; agua destilada hasta 50 ml. Ajustar el pH a 6,8 con una solución diluida de NaOH. Esterilizar por filtración y conservar a 4°C]. Se reparten en alícuotas de 100  $\mu$ l en tubos preenfriados a 4°C y se congelan inmediatamente con nitrógeno líquido. Se conservan por un periodo limitado a -80°C.

### 8.3.2. Células electrocompetentes

Para este protocolo se precisa la limpieza completa de las sales del medio de cultivo de manera que no interfieran con la electroporación, el verdadero causante de la desorganización de la pared celular. A su vez, se añade un agente crioprotector que va a permitir la viabilidad de las células conservadas a -80°C, glicerol. Consiste en sucesivos lavados con volúmenes cada vez más reducidos de glicerol 10%; partimos de 500 ml de cultivo a una  $A_{600}$  de 0,5 cuyo metabolismo se retrasa por incubación en hielo durante 20 min. Se recogen a 6.000 r.p.m. durante 15 minutos y 4°C y se resuspenden en 100 ml de glicerol 10% frío, volviéndose a centrifugar. En el siguiente paso se resuspenden en 20 ml de glicerol para acabar, tras este segundo lavado, con las bacterias resuspendidas en 2 ml de glicerol 10%. Se reparten en alícuotas de 50  $\mu$ l que inmediatamente se congelan en  $N_2$  líquido. Se conservan a -70°C durante un periodo limitado.

#### 8.4. TRANSFORMACIÓN DE CÉLULAS COMPETENTES

Se pueden usar diferentes maneras de inducir la desestabilización de las membranas celulares. En células cuya pared ha sido dañada por algún agente químico se suele usar choque térmico para abrir poros en la membrana; en cambio, las células electrocompetentes, cuya pared celular se encuentra intacta, precisan de una descarga eléctrica que cambie el potencial de membrana y permitir la captación del DNA.

##### 8.4.1. Transformación por choque térmico

Se sacaron las células competentes a hielo hasta su descongelación total (~20 min). Se añadió en condiciones de esterilidad el DNA, procedente de ligación o no, y se dejó incubando con las células durante otros 20-25 min. Las células son sometidas a un choque térmico de 42°C durante 90 seg, para seguidamente devolverlas a hielo durante otros 5 min en los que se permitirá la captación del DNA. Ahora debemos permitir la recuperación de las células por la adición de 900 µl de LB fresco durante 1 h a 37°C (en agitación a bajas revoluciones). Transcurrido este tiempo, se siembra parte o todo el cultivo en medio LB sólido, dependiendo de la eficiencia de transformación que esperemos.

##### 8.4.2. Electrotransformación

En este protocolo también debimos descongelar las células en hielo tras lo cual añadimos el DNA con el que deseamos transformar. Transferimos la suspensión celular a una cubeta de electroporación que se somete a un pulso eléctrico según las recomendaciones hechas por el fabricante. En nuestro caso usamos un electroporador *Eppendorf 2510* que recomienda un pulso de 1800 V durante 3-5 milisegundos. Como en el caso anterior, se recompone la estabilidad celular por la adición de 950 µl de LB fresco. A continuación se procede como en el apartado anterior.

#### 9. SECUENCIACIÓN Y ANÁLISIS DE SECUENCIAS

Para la secuenciación de las construcciones presentadas en este trabajo hemos recurrido al servicio de secuenciación de DNA/Genómica del Instituto de Parasitología y Biomedicina “López-Neyra”. Este servicio amplifica las muestras con un termociclador *PE-9600* (Perkin-Elmer) y las secuencia mediante el secuenciador *ABI 373 XL Stretch* (Perkin-Elmer) usando el kit *ABI PRISM Big Dye™ Terminator Cycle Sequencing Ready Reaction* (Perkin-Elmer).



La visualización de los picos de secuencia se realizó mediante el programa *Chromas Lite 2.01*. El trabajo rutinario en las secuencias (búsqueda de dianas de restricción, diseño de cebadores, comparativa de secuencias, búsqueda de ORFs, etc.) se llevó a cabo con el programa *Clone Manager Professional Suite®*.

Para la comparativa e identificación de nuestras secuencias, así como para la búsqueda de nuevas secuencias en las bases de datos, se utilizaron distintos algoritmos BLAST disponibles en la página web del NCBI (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>) (Pearson y Lipman, 1988; Altschul *et al.*, 1990).

Para la predicción de la estructura secundaria del RNA se utilizó programa *mfold* accesible en la siguiente dirección web: <http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form> (Zuker, 2003).

## 10. TÉCNICAS DE HIBRIDACIÓN DE ÁCIDOS NUCLEICOS

### 10.1. HIBRIDACIÓN DNA-DNA (*SOUTHERN BLOT*)

#### 10.1.1. Transferencia alcalina por vacío

Una vez que hemos separado el DNA total o plasmídico en geles de agarosa tenemos que transferir y fijar este material a una membrana de *biodine®* cargada positivamente. Para ello nos servimos de un sistema de transferencia alcalina (NaOH 1 M) por vacío (sistema *VacuGene™ XL* de Pharmacia). El proceso se ve facilitado por la rotura del DNA provocada por la exposición prolongada (15-20 min) a los rayos UV en el transiluminador. La membrana debe ser de un tamaño superior al de la ventana de transferencia y ésta, a su vez, más pequeña que el gel. Para evitar el depósito de sales, la membrana se embebió en agua desionizada para luego equilibrarla con 20xSSC [1xSSC: NaCl 150 mM, citrato sódico 15 mM, pH 7,0] durante 5 min. Se prepara el sistema de transferencia según las instrucciones de los proveedores con cuidado de que no se formen burbujas de aire entre la membrana y el gel. Se conecta el vacío y, una vez comprobada la estanqueidad del sistema, se cubre con NaOH 1M. Debemos vigilar que el sistema se mantenga constante en torno a 55 mbar durante 1-2 h. El siguiente paso fue lavar la membrana con una solución de 2xSSC durante 10 minutos en suave agitación y dejar secar a temperatura ambiente. La fijación del DNA a la membrana se llevó a cabo por calor a vacío envuelta en papel Whatman 3MM (120°C 35 min, 70 cm Hg). Se puede conservar a temperatura ambiente.

### 10.1.2. Marcaje de la sonda

En este tipo de experimentos el marcaje se lleva a cabo con digoxigenina acoplada a un dNTP que en nuestro caso fue dUTP. La sonda se sintetizó por PCR y su tamaño osciló entre los 200-600 bp. Un procedimiento que aumenta en gran medida la especificidad de la sonda es utilizar como molde un DNA purificado de gel amplificado con la misma pareja de cebadores con la que queremos obtener nuestra sonda marcada. Con este molde se preparó una reacción de PCR que contenía los componentes habituales de un PCR pero además se adicionó digoxigenina-11-dUTP (Roche) en una concentración final de 0,1 mM. Se completan 30 ciclos del PCR y el amplicón se limpia con columnas *MicroSpin™ S-300HR* (GE healthcare). Se ponen en paralelo controles sin DIG para comprobar el retraso en la migración en gel como consecuencia de la incorporación de la marca.

Esta sonda se diluye a 60 ng del DNA /ml de solución de prehibridación. Entre 5-10 ml son suficientes para cubrir completamente la membrana. Estas soluciones de hibridación se conservan a -20°C y pueden ser reutilizadas en sucesivas ocasiones.

### 10.1.3. Hibridación, lavados y revelado

Este proceso se llevó a cabo en un horno de hibridación *Hybridiser® HB-1D* (Techne). Se incubó el filtro en una solución de prehibridación (20 ml/100 cm<sup>2</sup>) compuesta por 5xSSC, sarcosil 0,1%, SDS 0,02%, formamida 50% y solución de bloqueo (Roche) al 2% durante 2 horas a 42°C. Entretanto, habremos desnaturalizado la sonda mediante ebullición durante 10 min y posterior paso a hielo (5-10 min). Se desecha la solución de prehibridación y se añade la solución de hibridación (ver apartado anterior) con la que se incubaba a 42°C toda la noche.

Se retiró la sonda y se procedió con los lavados y el revelado de la hibridación. Se realizaron dos lavados de 5 min a temperatura ambiente con 50 ml de 2xSSC/SDS al 0,1% (p/v), seguido de 2 lavados durante 15 min a 68°C con 0,1xSSC/SDS 0,1% (p/v). Estos lavados han servido para eliminar la sonda pegada inespecíficamente. Se disminuye la temperatura del rulo y la del horno de hibridación que trabajará en adelante a temperatura ambiente. Se hace un enjuague con 50 ml de tampón de lavado [0,3% de Tween-20 en tampón 1 (ácido málico 0,1 M, 0,15 M NaCl, pH 7,5)] para eliminar los restos de detergente antes de proceder con el bloqueo de la membrana.



La detección de la sonda se basa en el uso de anticuerpos anti-digoxigenina conjugados con la enzima fosfatasa alcalina. Esta enzima va a encargarse de defosforilar sustratos quimioluminiscentes como es el CSPD que da lugar a un intermediario que emite luz. Para que estos anticuerpos no se fijen a lo largo de toda la membrana es preciso evitar el reconocimiento inespecífico mediante el bloqueo de la superficie del filtro. Con este fin se añaden 30 ml de una solución de bloqueo [Blocking Reagent (Roche) 1% (p/v) en tampón 1] con la que se incuba la membrana durante 30 min. Pasado este tiempo se retira esta solución y se añaden 20 ml de solución de bloqueo en los que se habrá diluido 1:10000 el anticuerpo Anti-Digoxigenin-AP (Roche) con el que se incubará otros 30 min. El exceso de anticuerpo no unido a la sonda se retira mediante dos lavados sucesivos con tampón de lavado. A continuación la membrana se preparó para la adición del CSPD por incubación con 50 ml de tampón 2 [Tris-HCl 0,1 M, NaCl 0,1 M, MgCl<sub>2</sub> 50 mM, pH 9,5] durante 5 min. Tras retirar esta solución se añaden 5 ml de una dilución 1:100 del CSPD (Roche) en tampón 3 y se incuba nuevamente 5 min. Se deja secar perfectamente la membrana, se envuelve en plástico transparente y se incuba a 37°C durante 15 min para desencadenar la reacción de la enzima.

La luz emitida por el sustrato es captada por una película Kodak X Omat y se realizan dos tipos de exposiciones: una de corta duración que capta los momentos de mayor emisión del sustrato (4-6 h) y otra que permite la detección de señales débiles (24 h). El revelado de este tipo de películas se realiza del modo tradicional con una solución de revelado y una solución de fijación diluidas según las indicaciones del fabricante.

Sobre estas membranas podrían realizarse hasta cuatro hibridaciones pero como el sistema de detección es el mismo, se precisó de eliminar la sonda de una hibridación a la siguiente mediante dos lavados a 37°C de 15 min en NaOH 0,2 M, SDS 0,1% (p/v). La limpieza de restos de la solución anterior se realiza con un lavado de 5 min con 2xSSC. En este punto podría secarse y guardar el filtro, o comenzar añadiendo la solución de prehibridación.

## 10.2. HIBRIDACIÓN EN MICROARRAYS

La universidad de Bielefeld los microarrays de DNA Sm14kOLI, que contienen 14976 puntos de hibridación duplicados fijados sobre una resina epoxi, distribuidos en 48 gradillas de 26 filas y 24 columnas. Estos puntos de hibridación representan: 6208 oligonucleótidos de 70 nt de longitud

complementarios a los distintos genes codificados en el genoma de *S. meliloti*; a su vez, también contienen 8080 oligonucleótidos de 50 a 70 nt de longitud complementarios a 2881 regiones intergénicas, algunas de las cuales están representadas por más de un oligonucleótido complementario a ambas hebras del DNA. Los microarrays presentan una serie de controles que se corresponden con 17 oligonucleótidos de 70 nt que determinan la especificidad de hibridación (la identidad de sus secuencias varía en un rango de 70-90%); el control negativo son 5 oligonucleótidos 70-mer de secuencias no presentes en *S. meliloti*; así mismo, también se incluyen posiciones vacías y algunas que contienen el tampón de para la fijación de los oligonucleótidos al soporte. Finalmente, también se incluyen 4 oligonucleótidos 70-mer complementarios a distintos transgenes (*gusA*, *lacZ*, *npfII*, *aacC1*).

Se han utilizado dos tipos de técnicas basadas en la hibridación sobre microarrays: Hibridación Genómica Comparativa (también conocida por sus siglas en inglés, CGH) que nos permite observar diferencias genéticas entre dos cepas; y lo que generalmente se conoce como hibridación en microarrays, que permite observar diferencias en los niveles de expresión de los genes que componen el genoma bacteriano. Ninguna de ellas es puramente cuantitativa, pero permiten confeccionar una idea global sobre el proceso que se estudia.

### 10.2.1. Hibridación genómica comparativa, CGH.

Para llevar a cabo este protocolo se usó el kit BioPrime® Total Genomic Labeling System adaptado de Giuntini *et al.*, 2005.

#### 10.2.1.1 Marcaje de las muestras.

Digerimos 10 µg de DNA total (DNAt; previamente extraído con el kit Real Pure) en un volumen final de 100 µl. Las enzimas seleccionadas en esta ocasión fueron *RsaI* y *TaqI* (aunque se podrían haber usado enzimas como *Sau3AI*). Partimos además de dos muestras que queremos comparar en nuestro caso 1021 y RMO17.

	<u><i>RsaI</i></u>	<u><i>TaqI</i></u>
H <sub>2</sub> O	-	-
Tp 10X	10µl	10µl
BSA 100X		1µl
DNAt	-	-



Enzima	2µl	2µl
	4h. 37°C	4h 65°C
+ 1µl de enzima		
	o.n. 37°C	o.n. 65°C

Comprobamos que está bien digerido cargando en gel 5 µl y limpiamos la digestión con el kit de purificación de productos de PCR (Qiagen), resuspendiendo en un volumen final de 50 µl TE 1x (en dos pasos de elución 30µl+20µl). Es importante que sea TE 1x, en caso de eluir en agua debemos añadir EDTA 5mM según las indicaciones del proveedor.

Comprobamos la cantidad recuperada tanto con Nanodrop como en gel cargando 2 µl (lo habitual es añadir el volumen siguiendo las indicaciones del kit).

Descongelamos las mezclas de reacción 2x de los distintos colorantes a RT y protegidos de la luz. Mezclar bien. (IMPORTANTE: proteger de la luz).

Preparamos la siguiente reacción en tubos de PCR de 0.5 ml:

<u>Componente</u>	<u>Tubo 1</u>	<u>Tubo 2</u>
Alexa Fluor® 3 2X Reaction Mix	25µl	-
Alexa Fluor® 5 2X Reaction Mix	-	25µl
Digestión DNAt 1021	22µl	-
Digestión DNAt RMO17	-	<u>22µl</u>
	47µl	47µl

Incubamos 5 min a 95°C para desnaturalizar las muestras y pasamos directamente a hielo, 5 min. Añadimos 3µl de Exo-Klenow Fragment a cada tubo hasta conseguir un volumen final de 50µl. Mezclamos muy bien e incubamos a 37°C durante 2h.

Para la eliminación de material no polimerizado que podría generarnos mucho ruido sobre el array, se realizó la limpieza de las muestras añadiendo 200µl de Binding Buffer B2 a cada tubo y mezclamos muy bien (preparado adecuadamente con isopropanol). Ponemos sobre la columna Pure Link™ Spin y se centrifuga. Se lava con 650µl de Wash Buffer W1 (preparado adecuadamente con etanol) y se realiza una centrifugación final para eliminar restos de tampón. Añadimos 55 µl de Elution Buffer E1 al centro de la columna e incubar 1 min a RT. Centrifugar a velocidad máxima durante 2 min para recuperar las muestras marcadas.

Mezclamos las muestras marcadas con los distintos colorantes obtenidas en el paso anterior (tubos 1 y 2). Añadimos a esta mezcla 30 µl de Cot-1 DNA (1 mg/ml), 17 µl de t-RNA de levadura (6 mg/ml) y 450 µl de TE 1x. Ponemos esta mezcla sobre el lado gris de una columna Microcon 30

(Millipore) y reducimos el volumen hasta aprox. 35  $\mu$ l. Añadimos a cada sonda combinada 8.5  $\mu$ l SSC 20x y 0.74  $\mu$ l SDS10%.

#### *10.2.1.2. Preparación de los microarrays.*

Previo a la hibridación debemos tratar los arrays. Se preparan las soluciones frescas con cada hibridación. Las incubaciones se realizan en oscuridad y en continua agitación.

Incubamos durante 5 min a RT en Solución 1 (Triton X-100 0,001%).

Incubamos dos veces 2 min a RT en Solución 2 (HCl 0,0032%).

Incubamos durante 10 min a RT en Solución 3 (KCl 100 mM).

Lavamos 1 min a RT en agua miliQ.

Mantenemos 15 min en solución de bloqueo (1x QMT Blocking solution (Nexterion) + 0,0075% HCl) a 50°C previamente atemperada.

#### *10.2.1.3. Hibridación y lavados.*

Las muestras marcadas se desnaturalizaron durante 90 min a 100°C, para posteriormente incubar durante 30 min a 37°C. Disponemos sobre la parte central del porta (previamente caliente 50°C) evitando formar burbujas. Al poner el cubre la solución debe distribuirse homogéneamente entre ambos cristales. Colocamos en el interior de la cámara de hibridación y se incuba en un baño de agua a 65°C durante al menos 16h.

Se realiza un primer lavado rápido que nos permite retirar el cubre en 2xSSC/0.2%SDS a 65°C pudiendo mantener la agitación hasta 30 seg. Posteriormente se incuban en agitación en una solución 2xSSC/0.2%SDS a 65°C durante 5 min. Se hicieron lavados sucesivos de 5 min en una solución 0.2xSSC/0.1%SDS a 65°C (agitando arriba-abajo) y dos veces más en 0.2xSSC a RT; cada lavado de 2min. Inmediatamente se procede a secarlos centrifugando 5 min a 1200 rpm con rotor A-2-DWP.

#### *10.2.1.4. Escaneado y procesamiento de los datos.*

La adquisición de la imagen con los consiguientes valores de intensidad media de los fluoróforos para cada señal de hibridación se realizó con un escáner Axon 4100A y el programa de cuantificación GenePix Pro 5.1. Aunque este programa permite la normalización de los datos, hemos preferido utilizar la plataforma on-line diseñada en el centro de biotecnología (CeBiTec) de la universidad de Bielefeld expresamente para la interpretación



de los datos derivados de los arrays utilizados, EMMA 2.8.2 (<https://www.cebitec.unibielefeld.de/groups/brf/software/emma/cgi-bin/emma2.cgi>; Dondrup *et al.*, 2003). Este programa permite, además de la normalización de los datos por distintos métodos, su tratamiento estadístico.

Los dos parámetros más significativos a tener en cuenta durante el análisis de los datos son el valor de A, que determina la intensidad media correspondiente a un punto de hibridación calculado en función de las diferentes replicas, y el valor de M, que refleja el balance entre las dos muestras fluorescentes. El cálculo de A se realiza de acuerdo a la siguiente fórmula:  $A = \log_2(R_i \times G_i)^{0.5}$ , donde  $R_i$  se corresponde con la intensidad de la señal menos la intensidad inespecífica del fondo para la muestra problema y  $G_i$  es igualmente la resta de intensidades, pero en la muestra control. M se calcula aplicando el logaritmo en base 2 al cociente entre  $R_i$  y  $G_i$ .

### 10.2.2. Hibridación de microarrays.

En este caso se trata de una mezcla de cDNAs marcados con diferentes fluoróforos que, por hibridación competitiva, van a dar lugar a una serie de colores que nos van a marcar el nivel de expresión de los diferentes RNAs mensajeros presentes en las muestras. La muestra control se suele marcar con Cy5 mientras que el cDNA problema se marca con Cy3, aunque cada vez es más común el intercambio de marcajes.

El protocolo utilizado es una adaptación del descrito por Becker & Rüberg en el Manual de *S. meliloti* Sm6K Oligo Microarray.

#### 10.2.2.1. Marcaje de las muestras.

Se procede con el anillamiento 10-20  $\mu$ g RNA con 10 ng de random hexámeros en un volumen final de 16  $\mu$ l (15  $\mu$ g) durante 10 min a 70°C. Se incubó en hielo 5 min y añadimos la siguiente mezcla de reacción:

5x first strand buffer	6 $\mu$ l	
0.1 M DTT	3 $\mu$ l	
50x dNTPs (con aa-dUTP) <sup>7</sup>		0.6 $\mu$ l

---

<sup>7</sup> 50x dNTPs

Solución madre 100 mM aa-dUTP: 1 mg aa-dUTP+17  $\mu$ l H<sub>2</sub>O+0.68  $\mu$ l 1 N NaOH. Tomar una gota y medir el pH hasta ajustar a pH 7 (si es necesario se pueden seguir añadiendo 0.1-0.2  $\mu$ l 1 N NaOH). Guardar a -20°C.

4:1 aa-dUTP/dTTP mix:

RNase inhibitor (40 U/ $\mu$ l)	0.5 $\mu$ l
SuperScript II (200 U/ $\mu$ l)	1.5 $\mu$ l

La enzima sintetiza el cDNA durante 1 h a 42°C, momento en que se añade un nuevo  $\mu$ l de SuperScript II manteniendo la reacción a 42°C durante 1 h más. En este punto se podrían guardar a -20°C

Se procede a eliminar el RNA mediante la adición de 15  $\mu$ l NaOH 0.2 M (mezclar pipeteando) manteniendo a 70°C durante 10 min. Posteriormente la solución se neutralizó añadiendo 15  $\mu$ l HCl 0.2 M. Rápidamente se purifica el cDNA mediante la utilización de CyScribe GFX Purification kit, GE: se añadieron 450  $\mu$ l capture buffer al tubo y verter sobre la columna CyScribe GFX; lavamos la columna 3 veces con 600  $\mu$ l EtOH 80% y el cDNA se eluyó en 60  $\mu$ l NaHCO<sub>3</sub> 0.1 M (pH 9.0).

Como se puede apreciar, se trata de una técnica de marcaje indirecto del cDNA por lo que debemos acoplar el fluoróforo correspondiente en los grupos aminoalil presentes en nuestro cDNA. Se resuspende el fluoróforo<sup>8</sup> en los 60  $\mu$ l del paso anterior y se incuba entre 1 y 2 h en oscuridad a RT. El bloqueo de los grupos reactivos no acoplados se lleva a cabo mediante el tratamiento durante 15 min en oscuridad a RT con 4.5  $\mu$ l hydroxylamina 4 M (mezclar pipeteando). Se hace una limpieza de los fluoróforos no incorporados mediante CyScribe GFX Purification kit, GE: se añadieron 600  $\mu$ l capture buffer a la muestra marcada con Cy5, se mezcla pipeteando y se pasa esta mezcla a la muestra marcada con Cy3 para acabar pasando por la columna CyScribe GFX; se lavó 3 veces con 600  $\mu$ l washing buffer (kit) y se eluye en 60  $\mu$ l elution buffer (kit) al centro del filtro. Incubar 5 min en posición vertical en oscuridad a RT.

Para reducir el volumen de la sonda se deseca en speed-vac hasta aprox. 5  $\mu$ l. Añadimos 45  $\mu$ l de DIG Easy Hyb solution+1  $\mu$ l de DNA de esperma de salmón (5  $\mu$ g/ $\mu$ l).

---

100 mM dATP	19.4 $\mu$ l (concentración final 25 mM).
100 mM dCTP	19.4 $\mu$ l (concentración final 25 mM).
100 mM dGTP	19.4 $\mu$ l (concentración final 25 mM).
100 mM dTTP	3.9 $\mu$ l (concentración final 5 mM).
100 mM aa-dUTP	15.5 $\mu$ l (concentración final 20 mM).

Preparar alícuotas de 4  $\mu$ l y guardar a -20°C.

<sup>8</sup> Los fluoróforos se preparan diluyendo la solución proporcionada por la casa comercial. Añadir 10  $\mu$ l de DMSO (libre de agua), mezclar muy bien (cuidado porque no se disuelven). Repartir en alícuotas de 1  $\mu$ l en tubos opacos y desecar durante una hora en el speed-vac. Se guardan en pareja Cy3/Cy5 en bolsas con perlititas desecantes y a -20°C. También se pueden usar directamente con el DMSO.



#### 10.2.2.2. Pretratamiento de los microarrays

El procedimiento es exactamente igual al descrito en el protocolo de CGH, ver apartado 10.2.1.2.

#### 10.2.2.3. Hibridación y lavados.

Estos procedimientos son muy similares a los descritos en CGH sólo que:

-La desnaturalización de las muestras marcadas se realizó a 65°C durante 5-10 min.

-La temperatura de hibridación es 42°C durante al menos 16h.

-Los lavados iniciales con 2xSSC/0.2%SDS se realizaron a 42°C y los sucesivos lavados fueron de 1 min terminando con un lavado adicional de 1 min con 0.1xSSC frío (18°C).

#### 10.2.2.4. Revelado y procesamiento de los datos.

Ver apartado equivalente en CGH.

### 11. ENSAYOS DE MOVILIDAD

Independientemente de cuál fue el sistema de detección de los eventos de retrohoming producidos por RmInt1, la capacidad de transposición de este intrón fue evaluada en sistemas de plásmido a plásmido en la cepa *S. meliloti* en las cepas 1021 y RMO17 y puesta de manifiesto mediante hibridación DNA-DNA. El ensayo consiste en la introducción de dos plásmidos en una misma bacteria: un plásmido donador de intrón (pKG2,5 o pKGEMA4) y un plásmido receptor que contiene 640 bp de la IS*Rm2011-2*, entre las que se encuentra el sitio de inserción de RmInt1 (pJB0,6LEAD o pJB0,6LAG). Si el intrón invade su diana se observa la aparición de un plásmido recombinante derivado del plásmido receptor que contiene el intrón interrumpiendo su secuencia diana. Esto se traduce en la aparición de una banda nueva, de mayor tamaño, en la hibridación.

#### 11.1. CONJUGACIONES EN *S. meliloti*

La incorporación de los plásmidos donador y receptor en *S. meliloti* se realiza por conjugación triparental (Ditta *et al.*, 1980). En este tipo de conjugación, *S. meliloti* RMO17 va a recibir el plásmido que le aporta *E. coli*

DH5a gracias a los genes de movilización presentes en el plásmido RK2013 (*E. coli* HB101:pRK2013). Son necesarias dos rondas de conjugación para la introducción de ambos plásmidos por lo que una forma de simplificar el proceso fue utilizar como cepa receptora *S. meliloti* RMO17 con el plásmido receptor ya introducido, así sólo precisamos de una conjugación.

Estas conjugaciones se llevaron a cabo durante 12-16 h en medio rico TY (se mezclaron cantidades similares de cada tipo bacteriano) incubadas a 28°C. Tras lo cual se hizo una extensión en placas de MM con los correspondientes antibióticos para el aislamiento de colonias de *S. meliloti* libres de *E. coli* (ver resistencias en apartado 2; debemos seleccionar aquellas que contengan los dos plásmidos, por tanto, Km<sup>R</sup> Ap<sup>R</sup>). Se incuban a 28°C hasta la aparición de colonias (2-3 días). Volvemos a realizar una nueva selección en MM pero ya de colonias aisladas, comprobando en paralelo la posible contaminación con coliformes mediante el uso de Endo-Agar sólido.

### 11.2. DETECCIÓN POR HIBRIDACIÓN

El conjunto plásmidos aislados por lisis alcalina de *S. meliloti* RMO17 se digiere con la endonucleasa de restricción *SaI*I y los fragmentos resultantes se separan mediante gel de agarosa 1% en TAE. Esta enzima lineariza el plásmido donador pero tiene dos dianas en el plásmido receptor que flanquean el sitio de inserción del intrón (ver esquema en figura III.8 y apartado 2 de este material y métodos). En el caso de que se haya producido invasión del intrón el patrón de digestión *SaI*I nos va a generar 4 bandas: una correspondiente con el donador (7.854 bp para pKG2,5 o 7318 bp para pKGEMA4); otra que deriva del plásmido receptor (5.360 bp); una tercera que se corresponde con el fragmento que contiene la diana no invadida (2.380 bp); y, finalmente, la diana invadida por el intrón (4.264 bp o 3120 bp, dependiendo de si se invade por el intrón completo o por la versión  $\Delta$ ORF).

Para la detección de los eventos de *homing* se utilizaron dos sondas diferentes, una específica del intrón que nos va a detectar el plásmido donador y la banda correspondiente a la invasión del plásmido receptor, y la otra en el exón 2, que nos permitirá detectar tanto la banda derivada de la invasión del intrón como el fragmento correspondiente a la fracción de plásmido receptor no invadido. Con la sonda en el exón 2 podremos evaluar de la eficiencia de invasión según la fórmula:

$$\text{Eficiencia de Homing} = \frac{\text{P. receptor invadido}}{\text{P. receptor invadido} + \text{P. receptor no invadido}} \times 100$$



<b>OLIGO</b>	<b>SECUENCIA</b>	<b>APLICACIÓN</b>	<b>REF</b>
ε1	5'-GTGAGCGTCGGATGAAAC-3'	Amplificación de la sonda específica de la ribozima de RmInt1 (posiciones 147 a 594 de la secuencia del intrón)	Muñoz-Adelantado, 2003.
18R.0	5'-ACGTTTCTCAATTGAAACG-3'		
2011-b1	5'-TGGACGAAGACGAACATGG- 3	Amplificación de la sonda específica del exón 2 (ISRm2011-2) (posiciones +3 a +466)	Martínez-Abarca <i>et al.</i> , 1998
2011-b2	5'-TTGAAGTAGGCTGCGCATT-3'		

El análisis y tratamiento de imágenes, así como la cuantificación de las bandas se realizó mediante el programa *Quantity One*® 4.6.2. (BioRad).

## 12. TÉCNICAS CON RADIOISÓTOPOS: EXTENSIÓN DE CEBADOR

Esta técnica se emplea de forma rutinaria para detectar sitios de inicios de transcripción o sitios de procesamiento de formas precursoras (Boorstein y Craig, 1989). Consiste en mezclar un RNA problema con un exceso de oligonucleótido complementario a una región cercana al que se supone su extremo 5'; este oligonucleótido marcado radiactivamente será utilizado como cebador por una RT que va a producir un CDNA cuya longitud debe corresponder con la localización del extremo de la molécula. A esto debemos sumar que la mayoría de las RTs descritas no son capaces de superar enlaces modificados 2'-5' (Lorsch *et al.*, 1995); precisamente este tipo de enlaces son los que unen el extremo 5' de los intrones maduros a la A desapareada del dominio VI de la ribozima. Así, esta técnica se plantea como un sistema adecuado para la detección y cuantificación del *splicing* en intrones del grupo II.

Desde el punto de vista experimental se siguieron las directrices marcadas por Marqués *et al.* (1993). El cebador utilizado fue el oligonucleótido P (5'-TGAAAGCCGATCCCGGAG-3') que hibrida entre las posiciones 97 a 80 del RNA de RmInt1 en dirección hacia la región 5'. Marcamos 10 pmol del cebador P con 8 U de T4 PNKinasa (New England

Biolabs) en presencia de 10  $\mu$ Ci de [ $\gamma$ - $^{32}$ P]ATP. Aproximadamente 0,2 pmol del oligonucleótido P marcado (300.000 CPM) se hibridaron con 15  $\mu$ g de RNA total en tampón PIPES 10 mM pH 7,5/ NaCl 400 mM; tras 5 min a 85°C, se hace caer de forma rápida la temperatura a 60°C para después dejar enfriar lentamente hasta 44°C temperatura a la cual se añade la mezcla de extensión que contiene tampón AMV RT 1X, dNTPs 1mM, actinomicina D (SIGMA) 60  $\mu$ g/ml, 2 U de inhibidor de RNAsa (ROCHE) y 7 U de AMV-RT (Roche); se incuba durante a 42°C. La reacción se para añadiendo 150  $\mu$ l de etanol 100% preenfriado + 5  $\mu$ l de acetato sódico 3 M pH 5,2. Los productos se analizaron en gel desnaturalizante (8 M urea) de poliacrilamida al 6%.

El análisis y tratamiento de imágenes, así como la cuantificación de las bandas se realizó mediante el programa *Quantity One*® 4.6.2. (BioRad).

### 13. CONSTRUCCIÓN DE MUTANTES

Se utilizó el sistema de mutagénesis puntual basado en *Altered Sites*® II in vitro *Mutagenesis Systems* (Promega). El fundamento del protocolo es la introducción de mutaciones mediante hibridación con oligonucleótidos que contienen cambios respecto a la secuencia original. Deberemos tener clonada nuestra secuencia a mutar en el vector pALTER-1® (Promega); en el presente trabajo se usó pAL2.5 (que contiene el intrón flanqueado por exones - 174/+466; Muñoz-Adelantado *et al.*, 2003). Se anillan tres oligonucleótidos a la vez: uno de ellos fue nuestro oligonucleótido mutagénico y los otros dos van a servir para cambiar la resistencia que el plásmido confiere a la bacteria (además de darnos una idea de la probabilidad de que se haya introducido nuestra mutación). Originalmente el plásmido es resistente a tetraciclina pero sensible a ampicilina; si ambos oligonucleótidos anillan en su lugar correspondiente, tras la síntesis y reparación de las muescas, deberemos obtener un plásmido Ap<sup>R</sup> pero Tc<sup>R</sup>.

Una vez que hemos comprobado por secuenciación que la mutación es correcta los insertos se introducen por digestión en el vector definitivo.



## 14. ENSAYOS CON PROTOPLASTOS DE *A. thaliana*

### 14.1. PREPARACIÓN DEL MEDIO DE CULTIVO JPL 10X CON O SIN AZÚCARES

<b>Sustancia</b>	<b>1 litro (10X)</b>	<b>[JPL 1X]</b>
KNO <sub>3</sub>	19,65 gr	19,43 mM
CaCl <sub>2</sub> -2H <sub>2</sub> O	1,32 gr	0,9 mM
MgSO <sub>4</sub> -7H <sub>2</sub> O	1,11 gr	0,45 mM
Fe-EDTA 10mM	5 ml	5 μM
Microelementos	3 ml	ver tabla μelem.
Caseína hidrolizada	1 gr	0,1gr/l
Mioinositol	1 gr	550 μM
Glicina	20 mgr	26,6 μM
Sol. Vitaminas	10 ml	ver tabla vit.
Hormona ANA	2 ml	1,07 μM
KH <sub>2</sub> PO <sub>4</sub> 0,125M	30 ml	0,375 mM
Solución A (NaH <sub>2</sub> PO <sub>4</sub> )	1,95 ml	0,03 mM
Solución B (Na <sub>2</sub> HPO <sub>4</sub> )	3,05 ml	0,06 mM
Sacarosa	150 gr	43,8 mM

Repartir en alícuotas de 100 ml y congelar a -20°C (sin azúcar)

Cuando se vaya a usar, diluir hasta 1 litro y añadir el azúcar. Autoclavar 30 min a 115°C.

#### **SOLUCIONES MADRE**

##### **Fe-EDTA 10mM**

FeSO <sub>4</sub> -7H <sub>2</sub> O	2,78 gr
Na <sub>2</sub> EDTA	3,72 gr
Agua Bidest.	1 l

##### **Microelementos** (Murashig y Scoog)

<b>Sustancia</b>	<b>Cant.</b>	<b>[final]</b>	<b>[final en JPL1X]</b>
H <sub>3</sub> B <sub>3</sub>	3,1 g	100mM	30 μM
KI	415 mg	5mM	1,5 μM
MnSO <sub>4</sub> -1 H <sub>2</sub> O	8,45 g	100 mM	30 μM
ZnSO <sub>4</sub> -7 H <sub>2</sub> O	4,3 g	30mM	90 μM
Na <sub>2</sub> MoO <sub>4</sub> -2H <sub>2</sub> O	121 mg	1mM	0,3 μM

CuSO <sub>4</sub> -5H <sub>2</sub> O	12,5 mg	0,1 mM	0,03 μM
CoCL <sub>2</sub> -6H <sub>2</sub> O	11,9 mg	0,1 mM	0,03 μM.
Agua QSP	500 ml		

**Sol. Vitaminas**

[ Sol. Vit.]		[JPL IX]	
Thiamina HCl	40mg	1,18 mM	1,18 μM
Ac. nicotínico	50 mg	4,06 mM	4,06 μM
Pyridexine HCl	50mg	2,43 mM	2,43 μM
Agua QSP	100 ml		

**Hormona ANA**

Disolver 100 mg en Alcohol absoluto 50% (congelar)

**KH<sub>2</sub>PO<sub>4</sub> 0,125M**

Disolver 17 gr en 1 litro de agua QSP.

**Solución A**

Disolver 5,56 gr de NaH<sub>2</sub>PO<sub>4</sub>-2H<sub>2</sub>O en 200 ml de agua QSP.

**Solución B**

Disolver 14,34 gr de Na<sub>2</sub>HPO<sub>4</sub>-12H<sub>2</sub>O en 200 ml de agua QSP.

La solución A y B forman e un tampón fosfato 0,2 M

14.2. OBTENCIÓN DE PROTOPLASTOS DE *A. thaliana* A PARTIR DE CULTIVO CELULAR

Tres días antes de la preparación de los protoplastos, poner 100ml de un cultivo de células, repicadas al 33% en medio JPL (en un erlenmeyer de 250ml). Por ejemplo el viernes para preparar los protoplastos el lunes. Este repicado se hace a partir de un cultivo que se mantiene repicando al 20% una vez por semana en medio JPL.

Preparación de los protoplastos: (con células de tres días)

Preparar 25ml de la solución enzimática en agitación durante 30 min. aproximadamente.



---

Preparar (en esterilidad) un Falcon de 50ml de medio JPL 0,34M Glucosa Manitol.

Transferir (Esterilidad) 40ml del cultivo celular a un Falcon de 50ml.  
Centrifugar 5 min x 1500 rpm x RT y “Brake Low” (sin freno)  
Quitar el sobrenadante con una pipeta de 25ml en esterilidad.

Añadir despacio al pellet los 25 ml de la solución enzimática (desde cerca del borde, por la pared del Falcon inclinado), completar hasta 50ml con el medio JPL 0,34 Gluc Man, añadiéndolo despacio con la pipeta sobre el borde del Falcon. Resuspender suavemente el pellet (sin vortex ni golpes) y verter los 50ml en un Erlenmeyer estéril de 250 ml. Enjuagar el Falcon con 25ml de medio JPL 0,34M Gluc Man (con la pipeta) y echarlos al Erlenmeyer.

Poner en agitación en la cámara de cultivo entre 3h y 3h-30min (Marcar con “Protos y tu nombre” para no confundir con un cultivo contaminado).

Después del tiempo de incubación, verificar que la pared de las células está bien digerida. Tomar, con una punta truncada estéril, 20 $\mu$ l de la suspensión y ponerlos entre porta y cubre y observar al microscopio las bonitas esferas bien redondas.

Todas las CENTRIFUGACIONES que se van a realizar a partir de ahora se harán (a poder ser) en centrífuga con rotor basculante:

800 rpm x 5 min, ¡¡¡sin freno!!!

Atención, la parada sin freno puede suponer que tarde unos 5 min. en parar.

Transferir la suspensión celular suavemente (directamente desde el Erlenmeyer) a dos Falcon de 50 ml. (aproximadamente 37,5 ml por Falcon) y centrifugar.

Quitar el sobrenadante (en esterilidad, despacito y con pipeta). Añadir a cada pellet 25 ml de JPL 0,34M GlucMan, con la pipeta y suavemente por la pared del Falcon. Homogeneizar suavemente haciendo girar los tubos. Centrifugar.

Quitar el sobrenadante. Añadir, a cada pellet, 4,5 ml de JPL 0,28M Sacarosa (permite a los protoplastos flotar). Homogeneizar suavemente y juntar los dos tubos en un Falcon de 14 ml. Centrifugar.

Veremos tres fases: Un pellet de desechos, una fase clara y encima, flotando, los protoplastos. Recuperar los protoplastos con una pipeta estéril y transferir a un tubo de cultivo Sarstedt de 13 ml (de tapón con dos posiciones). Atención, porque cuanto más apuremos hasta la fase intermedia, más diluidos tendremos los protoplastos. Una vez recuperados tomar 10  $\mu$ l (con una punta truncada) para el recuento, poner el tubo en el frigorífico durante 1 hora antes de la transformación, con el tapón en la posición alta (para que el aire pase).

A partir de ahora todas las tomas de protoplastos se harán con una punta truncada estéril bajo campana, y todas las diluciones se harán con medio JPL 0,34M GlucMan.

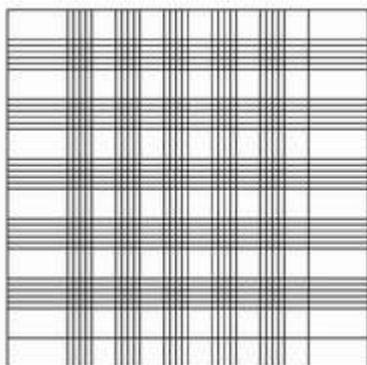
Recuento de los protoplastos en célula Malassez.

Preparar una solución de protoplastos, dilución 50. En un eppendorf de 1,5ml poner 490  $\mu$ l de JPL 0,34 GlucMan, y añadir los 10 $\mu$ l de la suspensión de protoplastos, homogeneizar suavemente.

Preparar la célula Malassez, colocar en el centro el cubre embebido en saliva. Tomar 40 ml de la dilución de protoplastos y añadirlos por un borde de la célula, la suspensión se extiende por tensión superficial bajo el cubre. Ya podemos contar los protoplastos en el microscopio. Vemos sobre la célula de Malassez cuatro tipos de cuadrados (con líneas horizontales y verticales, con los dos y los huecos). Todos son del mismo tamaño. Contar el número de protoplastos por cuadrado en 18 cuadrados aproximadamente, luego calcular la media por cuadrado.

$$\text{N}^\circ \text{ de protos} / \mu\text{l} = \text{Media por cuadrado} \times 100 \times 50 \text{ (dilución)}$$

En general, tendremos aproximadamente 6 protos por cuadrado, lo que da  $3 \cdot 10^4$  proto /  $\mu$ l, por lo tanto,  $1,5 \cdot 10^6$  protos en 50 $\mu$ l (volumen mínimo necesario para hacer un choque osmótico, pero podemos aumentar el volumen para tener más protoplastos).



Por Ejemplo: si contamos (sólo se cuenta en los cuadrados compuestos por líneas verticales y horizontales) 600 protos en la cuadrícula entera, corresponde con los protos en un  $\mu\text{l}$  de la dilución. Ajustamos el factor de dilución (multiplicando por 50) y obtenemos en n° de protos / $\mu\text{l}$  en la solución sin diluir, que serán  $3 \cdot 10^4$ . Esto quiere decir que en 50  $\mu\text{l}$  sin diluir habrá  $15 \cdot 10^5$  protos. Como nosotros necesitamos  $2-4 \cdot 10^5$  protos en 50  $\mu\text{l}$  tendremos que hacer ( $15/3=5$ ) una dilución 1/5 de los protos. concentrados.

### 14.3. TRANSFORMACIÓN DE LOS PROTOPLASTOS

Adaptado del método de transformación de protoplastos por PoliEtilenGlicol (PEG) descrito por Mathur y Koncz.

No hace falta trabajar bajo esterilidad, pero es recomendable.

Los tubos utilizados para la transformación son los tubos de cultivo Sarstedt de 13 ml, con tapón de doble posición. Son resistentes a la congelación con  $\text{N}_2$  líquido que se realiza después del choque osmótico.

En general, para una transformación se usan 15  $\mu\text{l}$  de solución plasmídica, 50  $\mu\text{l}$  de protos. y 150  $\mu\text{l}$  de PEG. Si se aumenta el volumen de protos, se debe aumentar el volumen de las otras soluciones de forma proporcional.

Descongelar el número de alícuotas de PEG necesarias, en hielo (descongelar hasta que quede líquido pero blanco, no se debe poner a RT porque si no se queda transparente, cosa que no deseamos).

Preparar los tubos, tantos como construcciones. Debemos acordarnos de usar los controles adecuados...

Usar puntas cortadas estériles, para los protos y para el PEG.

Poner en todos los tubos (que deben estar en hielo hasta el paso de añadir el PEG):

- 5  $\mu\text{g}$  de plásmido (15 $\mu\text{g}$  max.) en 15  $\mu\text{l}$  de agua

Tratar cada tubo de forma individual hasta la adición del PEG:

- Añadir de  $2-4 \cdot 10^5$  protos en 50 $\mu$ l (añadir por la pared del tubo justo por encima de la solución plasmídica)
- mezclar suavemente.
- añadir 150  $\mu$ l de PEG, justo encima de los protoplastos por la pared del tubo mientras lo giramos
- cerrar con el tapón en posición alta para permitir el paso de aire.
- mezclar suavemente girando el tubo y dejar 20 min en hielo
- poner en oscuridad (RT 23°C) durante otros 20 min

Lavar el PEG añadiendo 500  $\mu$ l de  $\text{Ca}(\text{NO}_3)_2$ . Mezclar suavemente girando el tubo. Añadir otros 500  $\mu$ l de  $\text{Ca}(\text{NO}_3)_2$ . Mezclar suavemente y centrifugar (5 min x 800 rpm x sin freno)

Quitar el sobrenadante con un una punta de 1ml

Resuspender cada pellet suavemente con 500  $\mu$ l de JPL 0,34 M Gluc Man y poner los tubos con el tapón a medio cerrar (en la gradilla) en posición casi tumbado durante unas 15h en oscuridad, para que se produzca la expresión.

## **SOLUCIONES PARA LA TRANSFORMACIÓN**

### **Solución Enzimática:**

Preparar en el último momento, en un Erlen de 100 ml estéril.

Añadir, a 25 ml de JPL 0,34 Gluc Man (por cada Falcon de 50 ml de células):

250 mg (1%) Cellulase R10 (Yakult)

50 mg (0,2%) Macerozyme R10 (Yakult)

Agitar suavemente en agitador magnético durante aproximadamente 30 min.



### **Solución PEG:**

Las alícuotas se deben conservar a  $-20^{\circ}\text{C}$ .

Para 100 ml.:

25 gr de PEG 6000 (25%)

8,2 gr de Manitol (0,45 M)

2,36 gr  $\text{Ca}(\text{NO}_3)_2$  (0,1 M)

Agitar ON

Ajustar a pH 9 con KOH

Alicuotear en 150  $\mu\text{l}$  y conservar a  $-20^{\circ}\text{C}$

Para 50 ml:

12,5 gr de PEG 6000 (25%)

4,1 gr de Manitol (0,45 M)

1,18 gr  $\text{Ca}(\text{NO}_3)_2$  (0,1 M)

Las siguientes soluciones están estériles y se deberá tomar lo que se vaya a usar un Falcon bajo esterilidad.

### **JPL 0,34 M Glucosa Manitol (410 a 430 Osm)**

A 1 L de medio JPL sin azúcar (aprox. 25 Osm) se le añade:

30,5 gr de Glucosa (= 0,1674M, PM= 180,2)

30,5 gr de Manitol (= 0,1692M, PM= 182,2)

Ajustas pH a 5,5 con KOH

### **JPL 0,28 M Sacarosa**

A 1 L de medio JPL sin azúcar se le añade:

96 gr de Sacarosa (=0,28M, PM= 342,3)

Ajusta a pH 5,5 con KOH

### **$\text{Ca}(\text{NO}_3)_2$ 0,275M**

Para 1L de solución

64,94 gr de  $\text{Ca}(\text{NO}_3)_2 \cdot 4\text{H}_2\text{O}$



**RESULTADOS**

**Y**

**DISCUSIÓN**



# **CAPÍTULO I**



## I. DISPERSIÓN DEL INTRÓN DEL GRUPO II *RmInt1* EN EL GENOMA DE *Sinorhizobium meliloti* TRAS SU ADQUISICIÓN POR TRANSFERENCIA CONJUGATIVA.

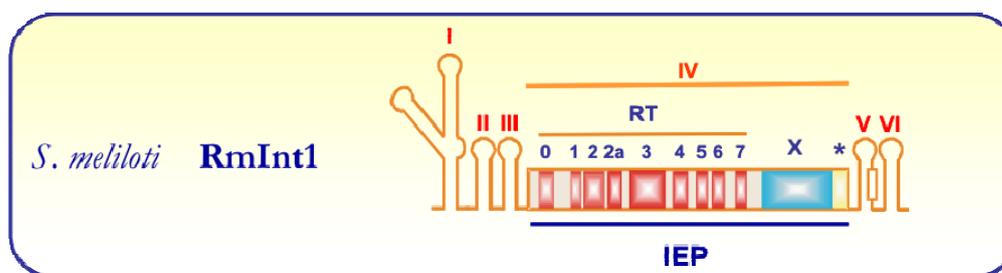
### I.1. INTRODUCCIÓN

Los intrones del grupo II son RNAs catalíticos y, a su vez, retroelementos móviles que se describieron inicialmente en orgánulos de plantas y eucariotas inferiores, aunque posteriormente fueron identificados en bacterias y algunos genomas de arqueas (Michel y Ferat, 1995; Toro, 2003; Lambowitz y Zimmerly, 2004). Un intrón típico del grupo II consta de un RNA con una estructura terciaria dispuesta en seis dominios (DI al DVI) y una proteína con actividad reverso-transcriptasa (RT) y madurasa codificada dentro del DIV. Esta proteína es requerida *in vivo* para el plegamiento del RNA del intrón en su estructura catalíticamente activa que permite su maduración y movilidad (Carignani *et al.*, 1983; Michel *et al.*, 1989; Moran *et al.*, 1994; Saldanha *et al.*, 1999; Matsuura *et al.*, 2001). La movilidad del intrón puede producirse tanto hacia alelos sin intrón (lo que se conoce con el nombre de *retrohoming*) o a sitios ectópicos (recibiendo el nombre de retrotransposición), mecanismo que está mediado por complejos de ribonucleoproteínas (RNPs) formados por la proteína codificada por el intrón (Intron Encoded Protein, IEP) y el RNA escindido en forma de lazo o *lariat*.

Se han definido tres clases filogenéticas de intrones del grupo II (IIA, IIB y IIC) basándose en la identidad de la secuencia de la IEP así como en el grado de conservación de las estructuras en el RNA catalítico (Michel *et al.*, 1989; Toor *et al.*, 2001; Zimmerly *et al.*, 2001; Toro *et al.*, 2002; Ferat *et al.*, 2003; Toro, 2003). Los intrones mejor caracterizados, el intrón bacteriano de *L. lactis*, Ll.LtrB, y los intrones de mitocondrias de levadura aI1 y aI2, pertenecen a la clase IIA, en la cual las IEPs contienen un dominio RT en su región N-terminal que se asemeja a las RTs retrovirales, seguido de un dominio X, asociado con la maduración del RNA y la actividad madurasa, y en su extremo C-terminal presentan un dominio de unión a DNA (D) y un dominio DNA endonucleasa (En) (Mohr *et al.*, 1993; Gorbalenya 1994; Shub *et al.*, 1994; San Filippo y Lambowitz, 2002).

La retrotransposición de los intrones del grupo II ocurre con baja frecuencia y se considera uno de los mecanismos para la dispersión del intrón de forma natural. Como el *retrohoming*, la retrotransposición también ocurre a través de un intermediario de RNA que se inserta en dianas de DNA (Martínez-Abarca y Toro, 2000; Dickson *et al.*, 2001; Ichiyanagi *et al.*, 2002).

La retrotransposición del intrón Ll.LtrB en *L. lactis* ocurre preferentemente sobre la cadena que se usa como molde para la síntesis de la hebra retardada durante la replicación y se supone independiente de actividad endonucleasa, lo que sugiere que el intrón se inserta en DNA que se encuentra momentáneamente como cadena simple en la horquilla de replicación (Ichiyanagi *et al.*, 2002; 2003). En *E. coli*, la retrotransposición de Ll.LtrB ocurre con mayor frecuencia sobre DNA de cadena doble por un mecanismo dependiente de DNA endonucleasa similar al que tiene lugar en el *retrohoming*. Así pues, Ll.LtrB se puede retrotransponer gracias a dos mecanismo distintos y parece que el hospedador podría influir en la elección de la ruta de integración (Coros *et al.*, 2005). Recientemente, se ha descrito transferencia horizontal inter- e intraespecies mediada por conjugación en Ll.LtrB (Belhocine *et al.*, 2004; 2005), demostrando que una vez adquirido este intrón es capaz de dispersarse en el cromosoma por retrotransposición, invadiendo sitios no homólogos. Sin embargo, el bajo nivel de transcritos de Ll.LtrB y su relativa ineficiencia de maduración puede limitar la movilidad y diseminación de Ll.LtrB en la naturaleza (Chen *et al.*, 2005).



**Figura I.1.** Esquema que representa al intrón de *S. meliloti* RmInt1 y que muestra como esta clase de intrones ha perdido tanto el dominio endonucleasa como el dominio de unión a ADN, También se muestra un dominio en la porción carboxi-terminal (\*) de la secuencia codificante de la IEP que también participa en la maduración de RmInt1.

Sorprendentemente, la mitad de las IEPs bacterianas del grupo II anotadas hasta la fecha pierden el dominio Endonucleasa. Muchas otras, como las pertenecientes a la subclase IIB3, (Toro, 2003) también pierden el dominio D (Fig. I.1). Se ha propuesto que tales intrones del grupo II son ancestrales y cuya IEP presenta mínimos requerimientos de secuencia en el DNA diana (Martínez-Abarca y Toro, 2000; Belfort *et al.*, 2002; Ichiyanagi *et al.*, 2002; Lambowitz y Zimmerly, 2004). El intrón mejor estudiado de esta subclase es RmInt1 descrito en *Sinorhizobium meliloti*, un elemento móvil muy eficiente (Martínez-Abarca *et al.*, 2000). Se ha determinado que RmInt1 se mueve gracias a dos rutas de *retrohoming*, con preferencia por una vía en

la que el RNA del intrón se inserta en DNA de cadena simple en la horquilla de replicación y usa como cebador para la reverso transcripción la cadena retardada que se está formando (Martínez-Abarca *et al.*, 2004). Por tanto el mecanismo de *retrohoming* de RmInt1 se asemeja a la transposición ectópica de Ll.LtrB en *L. lactis*.

La mayoría de los intrones del grupo II se localizan dentro de otros elementos móviles y fuera de genes (Toro, 2003). El sitio de *homing* natural de RmInt1 es la secuencia de inserción (IS) ISRm2011-2 (Martínez-Abarca *et al.*, 1998), la cual está presente, generalmente en alto número de copias, en el genoma de casi todos los aislados de *S. meliloti* (Martínez-Abarca *et al.*, 1998; Selbitschka *et al.*, 1995; Fernández-López *et al.*, 2005). RmInt1 se encuentra en el 90% de los aislados de *S. meliloti* (Muñoz *et al.*, 2001; Biondi *et al.*, 2011) y también en otras especies de *Sinorhizobium* y *Rhizobium*, donde los sitios de *homing* se corresponden con elementos IS del grupo ISRm2011-2 (Fernández-López *et al.*, 2005). Así pues, existe la posibilidad de que este intrón haya sido adquirido en las distintas especies de *Sinorhizobium* y *Rhizobium* mediante herencia vertical pero también por transferencia horizontal (Fernández-López *et al.*, 2005). La ausencia de RmInt1 en el 10% de los aislados de *S. meliloti* resulta interesante ya que en estos aislados, en principio, no hay restricción para la movilidad del intrón. RmInt1 es, por tanto, un sistema modelo para el estudio de la supervivencia, evolución y diseminación de los intrones del grupo II en la naturaleza.

En este capítulo se profundiza en la capacidad de *retrohoming* y en la ruta de dispersión de RmInt1 en el genoma de *S. meliloti* una vez que ha adquirido el intrón por transferencia conjugativa. Se han usado construcciones donadoras salvajes y construcciones modificadas basadas en derivados  $\Delta$ ORF con una eficiencia de *retrohoming* mayor. Nuestros datos muestran la extraordinaria capacidad de RmInt1 para dispersarse en la naturaleza por *retrohoming* y revelan la ruta seguida por el intrón para colonizar los sitios diana naturales en este genoma bacteriano.

## I.2. MATERIAL Y MÉTODOS

### I.2.1. Cepas bacterianas y condiciones de crecimiento

Las cepas de *Sinorhizobium meliloti* usadas en este trabajo fueron 1021 (cepa cuyo genoma ha sido secuenciado) (Galibert *et al.*, 2001) y RMO17, cepa que carece de copias de RmInt1 (Villadas *et al.*, 1995). *Escherichia coli* DH5 $\alpha$  se usó para la clonación y mantenimiento de

plásmidos. Las cepas de *S. meliloti* se crecieron a 28°C en medio completo (TY) o en medio mínimo definido (MM) (Robertsen *et al.*, 1981) y *E. coli* DH5 $\alpha$  a 37°C en medio Luria-Bertani (LB). Los medios fueron suplementados, según los requerimientos de las cepas bacterianas, con antibióticos a las siguientes concentraciones: kanamicina, 200  $\mu$ g/ml para rizobios y 50  $\mu$ g/ml para *E. coli*; tetraciclina, 10  $\mu$ g/ml, y ampicilina, 200  $\mu$ g/ml.

### **I.2.2. Construcciones donadoras de RmInt1 para ensayos *in vivo***

En las figuras I.2 e I.3A se muestran esquemáticamente todas las construcciones descritas a continuación. Todos los plásmidos que expresan el intrón RmInt1 salvaje fueron derivados de pBBR1MCS-2 (Km<sup>r</sup>). pKG2 se generó por acortamiento de las secuencias de los exones de pKG2.5 (Martínez-Abarca *et al.*, 2000) a -50/+146 (Costa *et al.*, 2006a) mientras que pBB2.5, pBB20/5H y pBB15/10H fueron generados por invasión *in vivo* de los correspondientes plásmidos receptores que contienen la diana de RmInt1 (Martínez-Abarca *et al.*, 2000; Jiménez-Zurdo *et al.*, 2003). Para construir los derivados RmInt1- $\Delta$ ORF se seleccionó a partir de pKG2 un fragmento del intrón comprendido entre los nucleótidos 611 y 1759 mediante PCR inverso (Costa *et al.*, 2006a), que posteriormente fue clonado en el sitio *Bam*HI (Muñoz-Adelantado *et al.*, 2003) justo delante de la secuencia codificante de la IEP dando lugar a pKG $\Delta$ ORFIEP. Otros variantes  $\Delta$ ORF flanqueados por exones -20/+5 se generaron por amplificación mediante PCR del fragmento *Bam*H1 que contiene el  $\Delta$ ORF; usando como molde el plásmido pKG $\Delta$ ORFIEP y los cebadores apropiados se obtuvieron amplicones que se clonaron por encima (pKG $\Delta$ ORFIEP20/5) o por debajo (pKGIEP $\Delta$ ORF20/5) de la secuencia codificante de la IEP en pKGIEP como fragmentos *Bam*HI o *Sac*I, respectivamente. Otro plásmido donador de RmInt1 en *S. meliloti* RMO17 y 1021 fue pKGEMA4, generado por la inserción de los sitios de restricción *Pm*II y *Bln*I internos a los sitio *Sac*I en pKGIEP $\Delta$ ORF20/5.

### **I.2.3. Ensayos de maduración y homing *in vivo***

La maduración de RmInt1 *in vivo* fue evaluada mediante extensión a partir de cebador de RNA total extraído de células de RMO17 que expresan el intrón salvaje desde los plásmidos pBB2.5, pBB20/5H y pBB15/10H (Muñoz-Adelantado *et al.*, 2003). La cuantificación de las bandas de cDNA resultantes del ensayo anterior se llevó a cabo mediante el programa Quantity One (Biorad). La maduración del intrón fue calculada según la

siguiente fórmula:  $100[S/(S+Pr)]$ , donde S es el RNA del intrón maduro y Pr son las formas precursoras no maduras. La eficiencia de homing *in vivo* del intrón salvaje y sus derivados  $\Delta$ ORF se determinó mediante un ensayo de doble plásmido en *S. meliloti* RMO17 usando como plásmidos receptores pJB0.6LEAD o pJB0.6LAG. La eficiencia de homing se calculó mediante la fórmula  $(H/[H+R]) \times 100$  donde H hace referencia al producto de homing y R al plásmido receptor sin invadir. Los datos son la media de al menos cuatro determinaciones independientes con su correspondiente desviación estándar (Jiménez-Zurdo *et al.*, 2003; Muñoz-Adelantado *et al.*, 2003; Martínez-Abarca *et al.*, 2004).

#### **I.2.4. Análisis de la dispersión de RmInt1 en el genoma de *S. meliloti* y de la movilidad de los intrones insertados**

Los plásmidos donadores pKG2.5 y pKGEMA4 fueron movilizados independientemente a las cepas RMO17 y 1021 por conjugación triparental usando como plásmido auxiliar pRK2013. Los eventos de retrohoming fueron evaluados en transconjugantes individuales procedentes de cada cruce (6-7 colonias) mediante Southern Blot; el DNA genómico digerido con *Bam*HI o *Sal*I (AquaPure Genomic DNA Isolation Kit, BioRad) fue hibridado con sondas específicas de *IS*Rm2011-2 o bien de la secuencia codificante de la ribozima marcadas con digoxigenina (Muñoz-Adelantado *et al.*, 2001). pKG2.5x (Martínez-Abarca *et al.* 2000) y pKGEMA4x se usaron como controles negativos en estos ensayos ya que la secuencia codificante de la IEP en ambos plásmidos presenta un cambio en la pauta de lectura originado por eliminación de un sitio interno *Eco*RI, que da lugar a una IEP truncada.

Para la evaluación de la eficiencia de homing *in vivo* de los intrones recién adquiridos, dos clones individuales de RMO17 que contienen inserciones genómicas del intrón salvaje (4 inserciones) o del intrón  $\Delta$ ORF (5 inserciones) fueron curadas de los plásmidos donadores pKG2.5 o pKGEMA4 sometiendo a las bacterias a sucesivas rondas de crecimiento en medio mínimo sin el antibiótico kanamicina, dando lugar a las cepas RMO17 $\Delta$ 2.5c2 y RMO17 $\Delta$ Ic50. El plásmido receptor pJB0.6LAG que contiene la diana de RmInt1 (Martínez-Abarca *et al.*, 2004) fue movilizado a las cepas curadas. La eficiencia de homing se estableció como se ha descrito en el apartado anterior.

Por otro lado, pKGIEP (Muñoz-Adelantado *et al.*, 2003) que expresa una IEP funcional y pKG4DV que expresa una ribozima no funcional también fueron movilizados independientemente por conjugación a un clon

de RMO17 $\Delta$ Ic50 que contiene el plásmido receptor pJB0.6LAG. pKG4DV es un derivado de pKGEMA4 en el cual los nucleótidos GTT presentes en el dominio V de la ribozima  $\Delta$ ORF (posiciones 1840-1842) se sustituyeron por CGA usando el sistema de mutagénesis Altered Sites II *in vitro* Mutagenesis System (Promega). La eficiencia de *homing* de las variantes  $\Delta$ ORF complementadas se determinó sobre ADN plasmídico de los transconjugantes como ya se ha explicado. Como control negativo de la complementación se usó el plásmido pKGIEPYAHH, un derivado de pKGIEP que expresa una IEP no funcional pues contiene una mutación en el dominio reversotranscriptasa (Muñoz-Adelantado *et al.*, 2003).

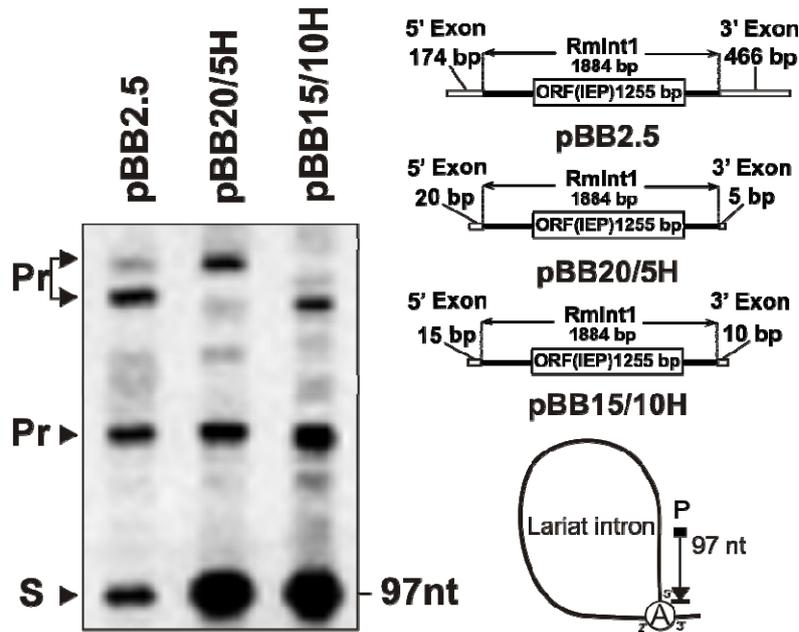
### I.3. RESULTADOS

#### **I.3.1. Modificación de construcciones donadoras de RmInt1 para su dispersión**

Recientemente, nuestro grupo ha demostrado que RmInt1 madura *in vitro*, reacción que se ve favorecida por la eliminación de la mayoría de la secuencia codificante de la proteína que se encuentra en el bucle terminal del DIV y por el acortamiento del exon 5' hasta sus últimos 20 y 15 nucleótidos (Costa *et al.*, 2006a, b). Posteriormente, en un trabajo previo, se ha descrito que el sitio diana comprendido entre las posiciones -20 a +1 del exón 5' y +1 a +5 del exón 3' eran suficientes para asegurar un *homing* eficiente de RmInt1 (Jiménez-Zurdo *et al.*, 2003).

Para conseguir construcciones donadoras de RmInt1 altamente eficientes que mejoren su dispersión, lo primero que hicimos fue analizar la influencia de la longitud de los exones en la maduración del intrón *in vivo*. La reacción de maduración fue caracterizada por extensión a partir de cebador usando un oligonucleótido (P) complementario a la secuencia localizada entre las posiciones 80 y 97 del extremo 5' del intrón (Fig. 4). En estos experimentos, el RNA del intrón escindido se detecta como un producto de extensión de 97 nucleótidos, además de otras bandas de mayor tamaño que derivan de RNA precursor no maduro (Pr) (Muñoz-Adelantado *et al.*, 2003). Como se muestra en la figura. 4 el producto de extensión de 97 nt. Se detectaba en todas la muestras de RNA procedentes de células que contienen construcciones que expresan el intrón flanqueado por parte de la secuencia de los exones originales bien -174/+466 o -20/+5 y -15/+10. la cantidad relativa estimada de intrón maduro fue aproximadamente del 32% cuando éste se transcribía a partir de construcciones con exones largos (pBB2.5),

ésta se vio incrementada al doble cuando se usaron construcciones con exones cortos (-20/+5 y -15/+10 en pBB20/5H y pBB15/10H,

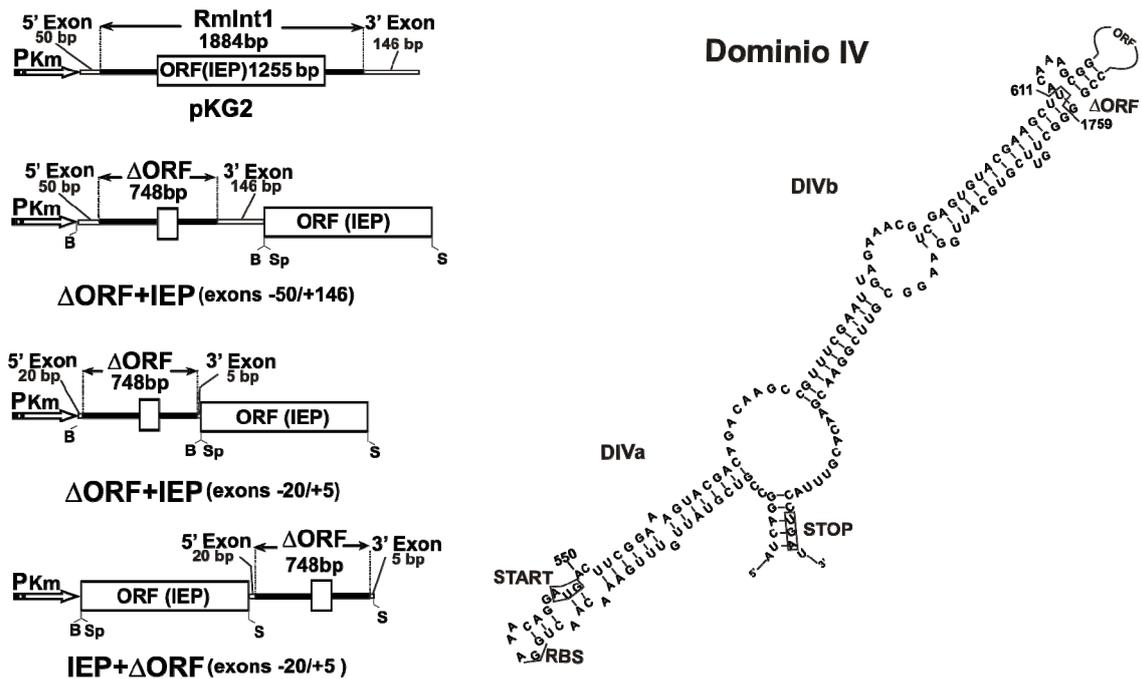


**Figura I.2.** Efecto de la longitud de los exones sobre la maduración de *RmInt1* *in vivo*. El análisis por extensión a partir de cebador se llevó a cabo sobre RNA total (20  $\mu$ g) de células de RMO17 que contenían el plásmido donador pBB2.5, pBB20/H o pBB15/10H. El producto de cDNA de 97-nt corresponde al RNA escindido del intrón (S) mientras que los productos de mayor tamaño son derivados de RNAs precursores no madurados (Pr). En la parte derecha se muestran los esquemas de los plásmidos donadores y del ensayo de extensión a partir de cebador.

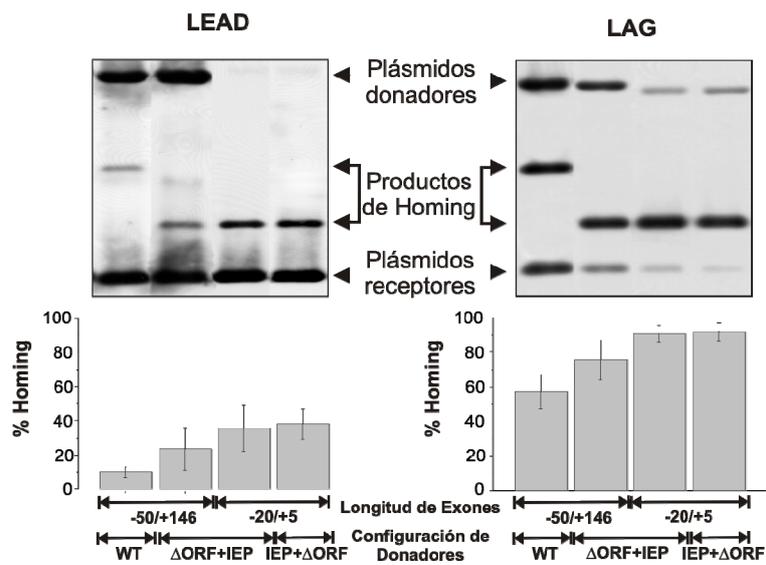
respectivamente) (Fig. I.2). Estos resultados indican que el acortamiento de los exones favorece la maduración de *RmInt1* *in vivo* como ya habíamos demostrado *in vitro*.

Con anterioridad se puso de manifiesto que las frecuencias de movilidad del intrón del grupo II de *L. lactis* Ll.LtrB aumentaban significativamente cuando se usaban construcciones donadoras que expresan la proteína LtrA por debajo del exón 3', y en las cuales se había deletado una gran parte del ORF que codifica LtrA dentro del DIV de la ribozima (Guo *et al.*, 2000). Para demostrar el efecto de una delección similar en *RmInt1*, la eficiencia de *homing* (fracción de dianas invadidas por el intrón) se determinó con un sistema de doble plásmido donde se usaron una batería de plásmidos donadores de *RmInt1* salvaje y  $\Delta$ ORF, todos los

A



B

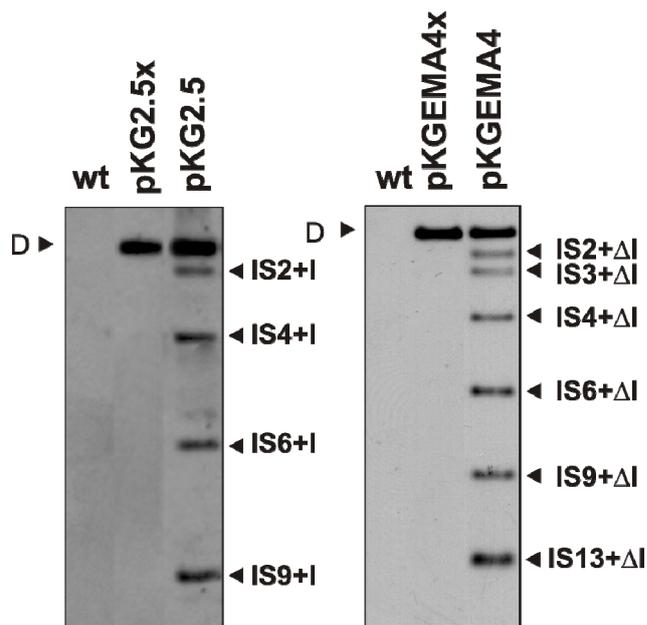


**Figura I.3.** Eficiencia de *homing* de RmInt1 y sus derivados ΔORF. A. Configuraciones genéticas de las construcciones donadoras. Pkm, promotor del gen de resistencia a kanamicina; B, *Bam*HI; Sp, *Spe*I; S, *Sac*I. En la parte derecha se muestra la predicción de la estructura secundaria del Dominio IV. Se muestran los codones de inicio y parada de la secuencia codificante de la IEP y mediante líneas se demarca la delección interna (ΔORF). Los subdominios DIVa y DIVb están indicados. B. Ensayos de *homing in vivo*. El DNA plasmídico de células RMO17 co-transformadas con los distintos plásmidos donadores

y receptores (pJB0.6LEAD o pJB0.6LAG) fueron analizados por hibridación DNA-DNA con una sonda específica de la secuencia de inserción *ISRm2011-2* que reconoce principalmente los plásmidos donadores con exones largos, los plásmidos receptores y los productos de *homing*. La proporción de invasión en cada ensayo de *homing* fue calculada como se describe en la sección de materiales y métodos y representados en histograma.

plásmidos expresaban el intrón desde el promotor constitutivo de resistencia a kanamicina. Como se representa en la figura 5A, las construcciones donadoras fueron pKG2, en la cual el intrón salvaje se encuentra flanqueado por exones -50/+146, y tres derivados  $\Delta$ ORF en los que se han deletado los nucleótidos comprendidos entre el 611 y el 1759 del intrón (Costa *et al.*, 2006a) y flanqueados por los mismos (-50/+146) o por exones más cortos (-20/+5) que expresan la IEP bien antes o después de la ribozima ( $\Delta$ ORF+IEP e IEP+ $\Delta$ ORF, respectivamente (Fig. I.3A). Cabe destacar que deleciones mayores del subdominio DIVb de *RmInt1* reducen o inhiben el *homing* (datos no mostrados). Los plásmidos receptores usados en los ensayos de *homing* fueron pJB0.6LEAD y pJB0.6LAG, en los cuales la diana de *RmInt1* está clonada en las dos orientaciones respecto a la dirección de la replicación del DNA (Martínez-Abarca *et al.*, 2004). Como se muestra en la figura 5B, el intrón salvaje (pKG2) mostró una eficiencia de *homing* sobre pJB0.6LEAD del 10% aproximadamente mientras que el uso de construcciones  $\Delta$ ORF+IEP flanqueadas por los mismos exones (-50/+146) dio lugar a una mayor eficiencia de *homing* (~ 23%). Curiosamente, las construcciones  $\Delta$ ORF+IEP e IEP+ $\Delta$ ORF flanqueadas por exones -20/+5 mostraban una eficiencia de *homing* incluso mayor (33% y 37% respectivamente). De acuerdo con la preferencia de *RmInt1* para invadir la hebra que sirve como molde para la síntesis de la hebra retardada en la horquilla de replicación (Martínez-Abarca *et al.*, 2004), el nivel de *retrohoming* en el plásmido receptor pJB0.6LAG se incrementaba considerablemente para todas las construcciones donadoras, aunque, nuevamente, la mayor eficiencia correspondió a construcciones  $\Delta$ ORF flanqueadas por exones cortos (~90%, Fig. I.3B). Además de un incremento en la eficiencia de *homing*, la combinación de intrón  $\Delta$ ORF y exones cortos dio lugar a niveles de maduración del intrón cercanos al 100% (datos no mostrados). Teniendo en cuenta todos estos datos podemos predecir que construcciones  $\Delta$ ORF y flanqueadas por exones cortos (-20/+5) serán la configuración más eficiente para la dispersión de *RmInt1 in vivo*.

### I.3.2. Dispersión de RmInt1 en el genoma de *S. meliloti*



**Figura 1.4.** *Retrohoming* de RmInt1 sobre el genoma sin intrones de RMO17. Los paneles muestran *Southern Blot* de DNA digerido con *SaII* de cepas de RMO17 salvaje, con pKG2.5 (donador del intrón salvaje, panel izquierdo), con pKGEMA4 (que expresa un derivado  $\Delta$ ORF, panel derecho), pKG2.5x o pKGEMA4x (ambos expresan intrones no móviles). Los distintos carriles corresponden con clones individuales y muestran todos los eventos de *homing*, usando como plásmido donador tanto pKG2.5 como pKGEMA4. La sonda utilizada es específica al extremo 5' de la secuencia de la ribozima del intrón. D hace referencia a la señal de hibridación correspondiente al DNA del plásmido donador. También están marcadas las inserciones genómicas del intrón salvaje (IS+I) o de los variantes  $\Delta$ ORF (IS+ $\Delta$ I). Las ISs fueron nombradas con números (1-13) de acuerdo con el patrón de hibridación proporcionado por la digestión del DNA total con *SaII* y usando una sonda de *ISRm2011-2*. El número uno fue otorgado al fragmento de mayor tamaño y el 13 al de menor.

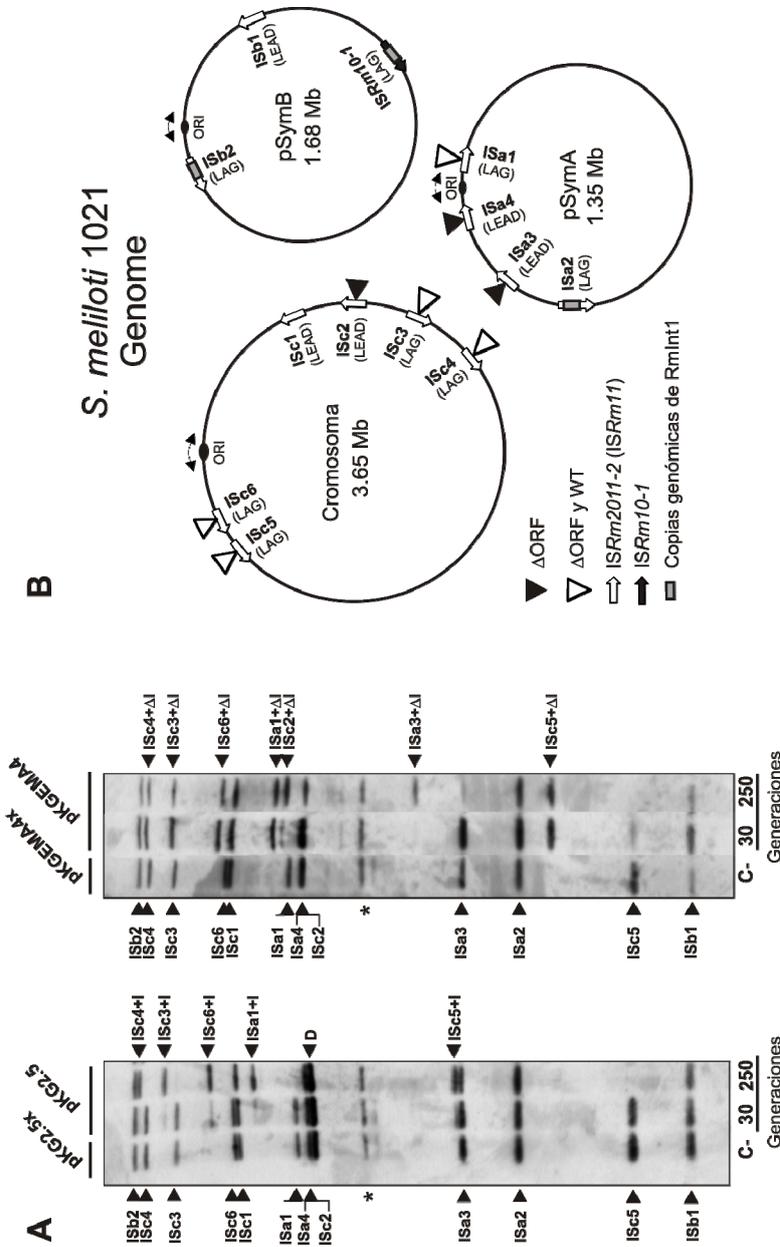
RMO17 es una cepa de *S. meliloti* que carece de RmInt1 aunque posee 13 copias de su sitio de inserción, *ISRm2011-2*. Para determinar cómo RmInt1 coloniza el genoma de *S. meliloti*, transferimos por conjugación a RMO17 el intrón salvaje (pKG2.5) y un derivado IEP+ $\Delta$ ORF flanqueado por exones cortos (pKGEMA4). Tras la adquisición de ambos plásmidos, pKG2.5 o pKGEMA4, el 100% de los transconjugantes mostraban eventos genómicos de invasión lo que pone de manifiesto la extraordinaria capacidad de dispersión de RmInt1. Esto eventos de *retrohoming* no tienen lugar cuando los plásmidos donadores de intrón codifican una IEP truncada en el dominio RT (pKG2.5x o pKGEMA4x; Fig. I.4), lo cual ratifica la necesidad de una IEP funcional para la movilidad de RmInt1. Se analizaron seis derivados de RMO17

independientes que contenían pKG2.5 y curiosamente, en todos esos clones la inserción del intrón salvaje tuvo lugar en las mismas cuatro copias genómicas de *ISRm2011-2* (IS2, IS4, IS6 e IS9) (ver el patrón de hibridación de unos de los clones en el panel izquierdo de la figura I.4). De acuerdo con su mayor eficiencia de *retrohoming*, el intrón  $\Delta$ ORF se integra en un gran número de sitios de *homing* (de 5 a 6). En este caso, se analizaron siete clones independientes que contenían el plásmido donador pKGEMA4. Seis de

esos siete derivados de RMO17 mostraron un patrón de hibridación en el que se ven invadidas las mismas cuatro ISs que las cepas que contienen pKG2.5 pero se observa un evento de invasión adicional en IS3. La invasión de IS13 se consideró menos frecuente ya que sólo se detectó en uno de los siete clones analizados (ver patrón de hibridación de estos transconjugantes en el panel derecho de la figura I.4). Estos resultados sugieren que algunos sitios de homing favorecen la invasión por *RmInt1* en el genoma de *S. meliloti*.

Para profundizar en la ruta de dispersión de *RmInt1* en el genoma de *S. meliloti*, los plásmidos pKG2.5 y pKGEMA4 fueron transferidos por conjugación a la cepa 1021, cuyo genoma ha sido secuenciado (Galibert *et al.*, 2001). El genoma tripartido de *S. meliloti* 1021 es de 6.7 megabases (Mb) y está constituido por un cromosoma de 3.65-Mb y por dos megaplásmidos simbióticos denominados pSymA (1.35-Mb) y pSymB (1.68-Mb). La cepa 1021 presenta trece posibles sitios diana para la inserción de *RmInt1* donde se incluyen doce copias de *ISRm2011-2* y una secuencia de inserción muy relacionada con la anterior, *ISRm10-1*. *RmInt1* se encuentra invadiendo dos copias de *ISRm2011-2* (una en pSymA y otra en pSymB), así como la única copia de *ISRm10-1* localizada en pSymB (Fig. I.5B). Estos tres sitios invadidos por *RmInt1* se encuentran en la hebra molde para la síntesis de la cadena retardada, lo cual coincide con la ruta de *retrohoming* preferida por *RmInt1* y otros intrones bacterianos del grupo II (Ichiyanagi *et al.*, 2002; Belhocine *et al.*, 2004; Martínez-Abarca *et al.*, 2004).

Las secuencias nucleotídicas de las restantes copias no invadidas de *ISRm2011-2* de la cepa 1021 contienen en todos los casos el sitio de inserción salvaje de *RmInt1*. Una vez adquirido pKG2.5 hasta cinco de los diez sitios genómicos de *homing* no invadidos de la cepa 1021 fueron gradualmente ocupados por *RmInt1*. Este hecho se ve reflejado en un paulatino descenso de la intensidad de banda de hibridación correspondiente a los sitios de *homing* y en la aparición paralela de nuevas bandas que contienen el intrón, que además incrementan su intensidad con el paso de generaciones (comparar los patrones de hibridación después de 30 y 250 generaciones en el panel izquierdo de la figura I.5A). Como era de esperar los derivados  $\Delta$ ORF del intrón se integraron en un mayor número de dianas de invasión (7 de 10 después de 250 generaciones; panel derecho en la figura I.5A). El Intrón salvaje expresado desde pKG2.5 se integró en cuatro copias cromosómicas de *ISRm2011-2* (ISc3, ISc4, ISc5 e ISc6) y una copia en pSymA (ISa1), en todos los casos las dianas están localizadas en la hebra molde para la síntesis de la cadena retardada, mientras que los derivados  $\Delta$ ORF se insertaron en dos dianas adicionales, una en pSymA (ISa3) y otra



**Figura 1.5.** *Retrohoming* de Rmlnt1 en el genoma de la cepa de referencia de *S. meliloti* 1021. A. *Southern blots* del DNA, digerido con *Bam*HI, de transconjuntantes de 1021 que contienen el plásmido pKG2.5 o el plásmido pKGEMA4 crecidos desde 30 hasta 250 generaciones. Las membranas fueron hibridadas con una sonda específica de la región 3' de *ISRm2011-2*. El primer carril de cada panel (C-) muestra un patrón de hibridación de *ISRm2011-2* de derivados de 1021 que expresan intrones no funcionales desde los plásmidos pKG2.5x o pKGEMA4x. También se indican las inserciones de intrón salvaje (IS+I) o de derivados ORF (IS+□). D indica la señal de hibridación correspondiente al DNA de los plásmidos pKG2.5 y pKG2.5x y el asterisco corresponde a una digestión parcial del fragmento que contiene a ISb1. B Distribución de Rmlnt1 y de las nuevas copias adquiridas (salvaje o ΔORF) dentro del genoma de 1021. El esquema muestra el cromosoma, pSymA y pSymb con indicaciones sobre su origen de replicación, localización de las dianas de Rmlnt1, las copias residentes del intrón y las nuevas inserciones tras la adquisición de los plásmidos pKG2.5 o pKGEMA4 como se deduce de A y de acuerdo con la anotación del genoma de *S. meliloti* en <http://bioinfo.genopole-toulouse.prd.fr/annotation/iANT/bacteria/rhime/>. Las dianas IS corresponden a la siguiente anotación de genes: SMc02205 (ISc1), SMc00848 (ISc2), SMc02411 (ISc3), SMc01382 (ISc4), SMc03084 (ISc5), SMc03305 (ISc6), SMa3000 (ISa1), SMa3014 (ISa2), SMa3015 (ISa3), SMa3018 (ISa4), SMb21616 (ISb1), SMb21703-5 (ISb2) e *ISRm10-1* (Smb21709-10).

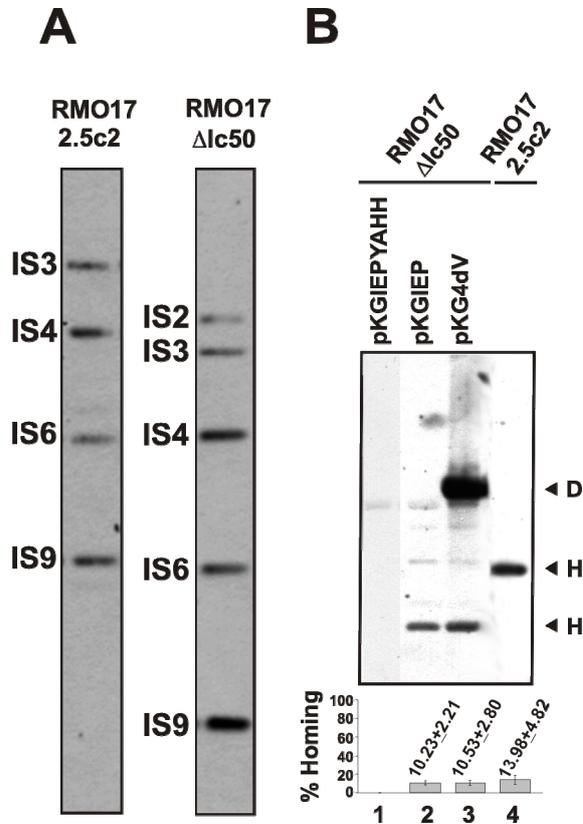
en el cromosoma (ISc2) (Fig. I.5B). Otras de las colonias analizadas presentaban una inserción adicional en ISa4 (dato no mostrado). Curiosamente, todos los eventos de *retrohoming* adicionales que tienen lugar en los derivados  $\Delta$ ORF ocurren en dianas localizadas en la hebra molde para la síntesis de la cadena conductora (Fig.I.5B). Todos estos resultados sugieren que la ruta de *retrohoming* seguida por RmInt1 para su dispersión en el genoma de *S. meliloti* está ligada a la replicación del cromosoma y los megaplásmidos simbióticos y muestra preferencia por la invasión de la hebra molde para la síntesis de la hebra retardada.

### **I.3.3. Capacidad de *retrohoming* de las copias de intrón insertadas en el genoma y complementación en *trans* del intrón $\Delta$ ORF con la secuencia codificante de la IEP**

Los resultados mostrados anteriormente sugieren que las dianas de invasión en el genoma de *S. meliloti* van a ser colonizadas gradual y secuencialmente por RmInt1, dependiendo de la expresión del intrón en la bacteria. Para comprobar si las copias de intrón recién adquiridas en el genoma retienen la capacidad de *retrohoming*, las cepas derivadas de RMO17 fueron curadas de los plásmidos donadores (pKG2.5 o pKGEMA4) para dar lugar a las cepas RMO17-2.5c2 y RMO17- $\Delta$ Ic50, respectivamente (Fig. I.6A) y fueron transformadas por conjugación con el plásmido receptor pJB0.6LAG. La eficiencia de *retrohoming* de los intrones salvajes adquiridos en RMO17-2.5c2 fue del 10-13% (Fig. I.6B, carril 4), mientras que no pudimos detectar eventos de inserción en la cepa RMO17- $\Delta$ Ic50 colonizada por intrones  $\Delta$ ORF (dato no mostrado). Estos resultados indican que las copias genómicas de intrón salvaje retienen la capacidad de movilidad y por tanto podrían contribuir a la posterior dispersión del intrón; además el *retrohoming* de las copias genómicas del intrón requiere la IEP, que no puede ser reemplazado por el fondo genético del hospedador.

La madurasa del intrón Ll.LtrB (LtrA) funciona más eficientemente cuando se expresa en *cis* tanto desde su localización normal dentro del intrón como desde una posición por debajo del exón 3'. La madurasa también funciona cuando se expresa en *trans* desde otro plásmido independiente, pero la maduración es inferior en un 10% a la maduración de la configuración en *cis* (Cui *et al.*, 2004). Para demostrar que la complementación en *trans* de la IEP de RmInt1 tiene lugar, la cepa RMO17- $\Delta$ Ic50 que contenía el plásmido receptor pJB0.6LAG fue transformada con los plásmidos pKGIEP y pKG4DV de forma independiente. Ambos plásmidos

expresan una IEP salvaje bajo el promotor constitutivo del gen de resistencia a kanamicina, bien sola (pKGIEP) o bien acompañada de una ribozima defectuosa en maduración que contiene una mutación dentro del DV catalítico (GUU→CGA; pKG4DV).



**Figura I.6.** Movilidad de las copias genómicas del intrón y *trans* complementación del derivado  $\Delta$ ORF con la IEP. **A** Patrón de hibridación de RmInt1 y su derivado  $\Delta$ ORF en RMO17 curada de los plásmidos donadores pKG2.5 (RMO17-2.5c2) o pKGEMA4 (RMO17- $\Delta$ lc50). El DNA total de ambas cepas fue digerido con *SaI* e hibridado con una sonda específica del extremo 5' de la secuencia de la ribozima del intrón. **B.** *Homing in vivo* de las copias de intrón adquiridas. El plásmido receptor pJB0.6LAG fue movilizado a las cepas RMO17-2.5c2 y RMO17- $\Delta$ lc50. Entonces se conjugaron con la cepa RMO17- $\Delta$ lc50, que contiene ya el receptor, plásmidos compatibles pKGIEP o PKG4dV que expresan una IEP funcional para complementar las inserciones  $\Delta$ ORF. La eficiencia de *homing* de RmInt1 y del derivado  $\Delta$ ORF se determinó como ya se ha descrito. Como control negativo de la complementación se usó la construcción pKGIEPYAHH. Las proporciones de invasión se calcularon como se describe en la sección de material y métodos y representadas en el histograma. D corresponde al plásmido donador y H al producto de *homing*.

Como control negativo de la complementación se usó pKGIEPYAHH que expresa una IEP no funcional con una mutación en el dominio RT (YADD→YAHH). Curiosamente, cuando se complementó en *trans* con una IEP funcional bien sola o acompañada de una ribozima defectuosa, las copias genómicas del intrón  $\Delta$ ORF mostraron una eficiencia de *homing* de ~10% sobre pJB0.6LAG, la cual es comparable a la que tiene lugar para el intrón salvaje (Fig. I.6B, carriles 2 y 3). Como esperábamos, la proteína mutada en el dominio RT no fue capaz de movilizar el intrón  $\Delta$ ORF (Fig. I.6B carril 1). Estos resultados podrían reflejar la estabilidad de la proteína libre y del RNA precursor del intrón no maduro en el hospedador *S. meliloti* y nos permite hipotetizar que copias de intrón inactivas podrían contribuir a la dispersión del intrón siempre y cuando sean complementadas en *trans* con una IEP activa.

## I.4 DISCUSIÓN

La dispersión de los intrones del grupo II en los genomas bacterianos ha sido estudiada fundamentalmente analizando la retrotransposición del intrón Ll.LtrB (Ichiyanagi *et al.*, 2002, 2003; Belhocine *et al.*, 2004; Conlan *et al.*, 2005; Coros *et al.*, 2005). También se da la diseminación por *retrohoming*, pero estos estudios se han limitado al análisis de un número reducido de eventos cromosómicos y a los eventos de invasión de plásmidos receptores (Belhocine *et al.*, 2004; Belhocine *et al.*, 2005) o al uso de intrones donadores modificados (targetron) que podrían estar influenciados por muchos factores inherentes a los mismos experimentos (Zhong *et al.*, 2003; Yao *et al.*, 2005). En este estudio, determinamos como un genoma bacteriano es colonizado por un intrón del grupo II tras su adquisición por transferencia conjugativa. Este análisis fue posible gracias a la existencia de numerosos sitios de *homing* de RmInt1 desocupados en el genoma de algunas cepas de *S. meliloti*, esto nos puede ayudar a comprender la dispersión de este intrón en poblaciones bacterianas naturales.

En primer lugar se ha mostrado que la maduración de RmInt1 *in vivo* y su eficiencia de *homing* se ve muy mejorada por la eliminación de la mayoría de la secuencia codificante de la proteína del lazo terminal del DIV y usando exones cortos (20 nt. en el exón 5' y 5 nt. en el exón 3'). Esta mejora en la maduración ya había sido observada para RmInt1 en reacciones de maduración *in vitro* (Costa *et al.*, 2006a, b), aunque la influencia de un exón 3' más corto aun no ha sido investigada. Además, las frecuencias de *homing* de derivados  $\Delta$ ORF de Ll.LtrB que expresan la IEP justo debajo del exón 3' aumentaron mucho, presumiblemente debido a una gran resistencia del intrón  $\Delta$ ORF a la rotura nucleotídica en el DIV (Guo *et al.*, 2000). Sin embargo no hay datos sobre las frecuencias y/o eficiencias de movimiento de construcciones que expresen la IEP por encima del exón 5'. Nosotros hemos determinado para RmInt1 que la IEP funciona de forma eficiente tanto por encima como por debajo del intrón  $\Delta$ ORF. Los derivados  $\Delta$ ORF usados en nuestros ensayos todavía mantienen un tallo del DIVb de ~30 pares de bases (Fig. I.3A). La delección del DIVb de Ll.LtrB reducía la maduración en ~60% (Cui *et al.*, 2004) y delecciones similares en la secuencia codificante del intrón de levaduras aI2 inhibía la maduración reversa promovida por la madurasa en ensayos *in vitro*, pero este efecto no se observó *in vivo*. Nosotros hemos visto que una delección mayor en el dominio DIVb de RmInt1 reduce la eficiencia de *homing* y que una delección completa inhibe el *homing* (datos no mostrados). Por lo tanto, el DIVb se requiere para

un *homing* eficiente de RmInt1 y presumiblemente para la movilidad de los intrones del grupo II pertenecientes a la subclase IIB.

Tan pronto como la bacteria adquiere el intrón salvaje o su variante IEP+ $\Delta$ ORF (con una mejora notable en la eficiencia de *homing*) las dianas de invasión presentes en el genoma de *S. meliloti* comienzan a ser ocupadas. Esta colonización tiene lugar con una extraordinaria frecuencia puesto que somos capaces de detectar eventos de *retrohoming* en el 100% de los transconjugantes analizados. Con anterioridad, nuestro grupo ha descrito la preferencia de RmInt1 por la ruta de *retrohoming* por la cual el RNA del intrón se integra en DNA de cadena simple de la horquilla de replicación, usando como cebador para la reverso-transcripción la cadena retardada que se está sintetizando (Martínez-Abarca *et al.*, 2004) de manera similar a la transposición ectópica de otros intrones del grupo II (Ichiyangi *et al.*, 2002). La colonización de los sitios de *homing* en el genoma de *S. meliloti* responde a esta preferencia en la ruta de invasión y las dianas localizadas en la hebra molde para la síntesis de la cadena retardada son las primeras en ser invadidas por el intrón, aunque con posterioridad también podrán ser ocupadas las dianas que se encuentren en el molde para la síntesis de la hebra conductora. Curiosamente, aunque sabemos que existe un sitio de invasión ectópica en *S. meliloti* denominado *oxi1* (Muñoz *et al.*, 2001) no hemos podido detectar en estos ensayos eventos de retrotransposición. Ya que este es un fenómeno muy poco frecuente, el *retrohoming* parece ser el mecanismo de movilidad elegido para la diseminación del intrón en el genoma de *S. meliloti*. Una vez que los sitios de *homing* estén saturados la retrotransposición para podría contribuir a la posterior dispersión del intrón.

Sabemos que Ll.LtrB también se retrotranspone a plásmidos receptores con un sesgo hacia el molde para la síntesis de la cadena retardada aun mayor que el que tiene lugar en los saltos ectópicos al cromosoma (Ichiyangi *et al.*, 2003). Se ha especulado que este sesgo podría estar relacionado con la diferente localización en la célula de los plásmidos y el cromosoma durante la replicación. Nuestros resultados no muestran tal sesgo en el *retrohoming* de RmInt1 en el genoma de *S. meliloti* y de hecho las primeras dianas en ser ocupadas son aquellas que están localizadas en el cromosoma. No obstante los dos megaplásmidos simbióticos, pSymA y pSymB, se consideran replicones de tipo cromosómico (Galibert *et al.*, 2001) y pudieran mostrar las mismas propiedades que el cromosoma principal estando localizados en un sitio similar durante la división celular e igualmente accesibles para el intrón. Además, nuestros resultados no muestran una preferencia para la integración del intrón en los sitios diana

cerca del origen de replicación como se ha observado en la retrotransposición de Ll.LtrB (Ichihyanagi *et al.*, 2003; Coros *et al.*, 2004) y el *retrohoming* de intrones modificados para la invasión de diferentes genes (targetron) en el genoma de *E. coli* (Zhong *et al.*, 2003).

También hemos mostrado que los intrones salvajes adquiridos por el genoma de *S. meliloti* retienen su capacidad de movimiento, pero no los variantes  $\Delta$ ORF, lo cual implica que las funciones llevadas a cabo por la IEP durante la movilidad del intrón no pueden ser sustituidas por ningún otro factor presente en el fondo genético del hospedador. La complementación en *trans* con LtrA ha permitido la movilización del derivado  $\Delta$ ORF de Ll.LtrB, aunque la capacidad de maduración es un 10% inferior a la que tiene lugar en la complementación en *cis* (Cui *et al.*, 2004). De manera similar, los derivados  $\Delta$ ORF de RmInt1 insertados en el genoma de *S. meliloti* podrían ser también movilizados suministrando una IEP funcional en *trans*, bien sola o en el contexto de una ribozima defectuosa, mostrando unos niveles de invasión similares al que presenta el intrón salvaje. Estos datos sugieren que podríamos contribuir a la dispersión eficiente de intrones mutados (truncados) aportando una IEP funcional.



## **CAPÍTULO II**



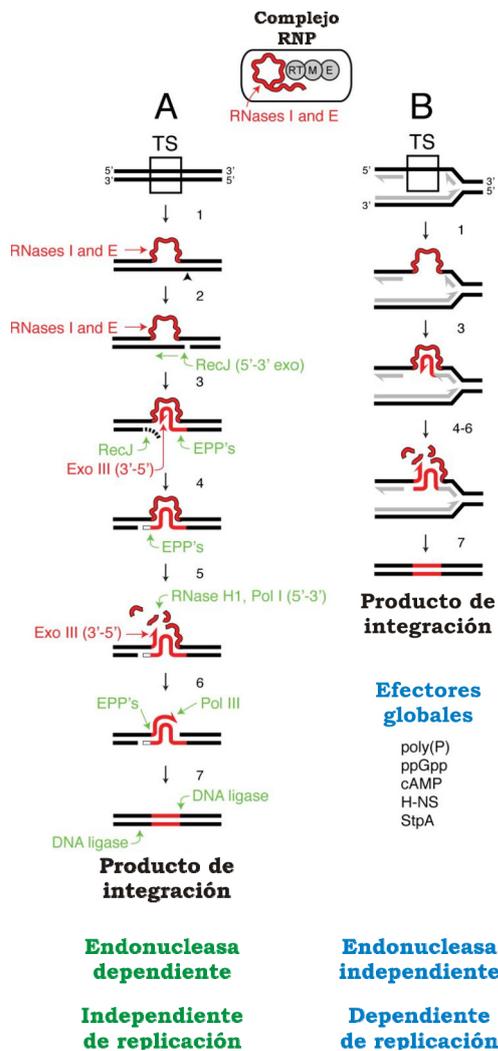
## RmInt1 Y *Sinorhizobium meliloti*: ¿Quid pro quo?

### II.1. INTRODUCCIÓN

Los intrones del grupo II tienen un especial interés, debido a que se piensa que son los antecesores de los intrones espliceosomales nucleares (Jacquier, 1990; Cavalier-Smith, 1991; Sharp, 1991), y se encuentran en los tres dominios de la naturaleza (Dai *et al.*, 2003; Dai y Zimmerly, 2003). Estos intrones son ribozimas que catalizan el *splicing* de sus exones (Mitchel y Ferat, 1995; Bonen y Vogel, 2001), y algunos son retroelementos móviles (Lambowitz *et al.*, 1999; Belfort *et al.*, 2002). Los intrones móviles son capaces de insertarse en su diana natural o en sitios ectópicos, mediante los procesos denominados *retrohoming* y retrotransposición respectivamente, mediante una proteína codificada por el intrón (IEP) que tiene actividad transcriptasa inversa (RT) (Fig. II.1).

El *retrohoming*, que tiene lugar con una frecuencia cercana al 100%, ha sido estudiado detalladamente en dos intrones mitocondriales de levadura (aI1 y aI2) y en tres intrones bacterianos (L1.LtrB de *L. lactis*, RmInt1 de *S. meliloti* y Eci5 de *E. coli*) (Belfort *et al.*, 2002; Lambowitz y Zimmerly, 2004; Martínez-Abarca *et al.*, 2004; Zhuang *et al.*, 2009a). La principal vía de *retrohoming* ocurre mediante un proceso en el cual el RNA del intrón escindido sufre *splicing* reverso sobre una de las cadenas del DNA diana, y entonces es transcrito desde RNA a DNA por la IEP, tomando como cebador de la reverso transcripción el DNA de la diana (Fig. II.1 A y B)). Este proceso es mediado por un complejo ribonucleoprotéico (RNP) que se forma durante el *splicing* de RNA y que está compuesto por la IEP y por el RNA del intrón escindido en forma *lariat* (lazo).

El *retrohoming* de los intrones aI1, aI2, L1.LtrB y Eci5 es dependiente de tres actividades de su IEP, además del RNA del intrón catalíticamente activo (Lambowitz y Zimmerly, 2010) (Fig. II.1A). Estas actividades son: RNA madurasa, para estabilizar la estructura catalíticamente activa del RNA del intrón durante el *splicing* directo e inverso del RNA; DNA endonucleasa, para cortar el DNA de la diana y generar un cebador para la transcripción reversa; y actividad RT, para sintetizar el cDNA copia del RNA del intrón. Después del *splicing* reverso del intrón sobre el DNA (Fig. II.1A, paso 1) y la rotura endonucleolítica de la segunda cadena, aguas abajo del sitio de



**Figura II.1.** Rutas de retromovilidad de los intrones del grupo II bacterianos. **A:** Ruta dependiente de endonucleasa vía dsDNA. Los pasos que sigue son: (1) splicing reverso sobre la diana de DNA, (2) rotura de la cadena de abajo por la actividad endonucleasa de la IEP, (3 y 4) síntesis del cDNA por la RT de la IEP, representado con la línea roja, (5) eliminación del RNA del intrón, (6) síntesis de la segunda cadena, y (7) ligación. Funciones celulares estimuladoras (verde) e inhibitoras (rojo) son representadas sobre el esquema, indicando con una flecha el posible sitio de actuación. **B:** Ruta independiente de endonucleasa vía ssDNA en la horquilla de replicación. El DNA que se está sintetizando se representa en gris. Los pasos numerados corresponden con A. Los efectores globales que influyen en el retromovimiento se indican en la tabla II.1.

inserción del intrón (Fig. II.1A, paso 2), la síntesis completa del cDNA se realiza mediante transcripción reversa usando como cebador el DNA de la diana (TPRT; Fig. II.1A, pasos 3 y 4). Los pasos finales del proceso pueden divergir: mientras que los intrones Ll.LtrB y EcI5 (Mills *et al.*, 1997; Cousineau *et al.*, 1998; Zhuang *et al.*, 2009a) son independientes de la función recombinasa RecA, los intrones de levadura pueden usar la recombinación homóloga (Eskes *et al.*, 2000).

Muchos intrones del grupo II carecen de dominio endonucleasa y, por lo tanto, no son capaces de llevar a cabo el corte en el DNA de la diana para insertarse. Estos intrones probablemente usen intermediarios replicativos como cebador en la síntesis del cDNA. El mejor estudiado de los intrones del grupo II que carecen de dominio endonucleasa es RmInt1 de *S. meliloti* (Martínez-Abarca *et al.*, 2000), el cual usa principalmente la cadena naciente retardada (*lagging*) en la horquilla de replicación como cebador para la síntesis de cDNA (Martínez-Abarca *et al.*, 2004) (Fig. II.1B).

La retrotransposición ha sido muy bien estudiada en el intrón Ll.LtrB de *L. lactis*, siendo un proceso al menos  $10^4$  veces menos eficiente que el *retrohoming* (Cousineau *et al.*, 2000). Además de tener unos requerimientos de secuencia más relajados que el *retrohoming*, la retrotransposición se produce de una forma preferente sobre DNA de cadena sencilla que de cadena doble (Fig. II.1A).

No requiere función endonucleasa, usando la horquilla de replicación como cebador en la síntesis de cDNA (Ichiyanagi *et al.*, 2002). Esta ruta es similar al *retrohoming* de RmInt1 (Martínez-Abarca *et al.*, 2004) y al *retrohoming* de mutantes del intrón Ll.LtrB en la actividad endonucleasa (Zhong y Lambowitz, 2003). En contraste a la retrotransposición en *L. lactis*, Ll.LtrB en *E. coli*, se retrotranspone de forma preferente en DNA de cadena doble, indicando que además del intrón, la célula hospedadora influye en la selección de la ruta de movilidad (Coros *et al.*, 2005; Beauregard *et al.*, 2006).

<b>Factor</b>	<b>Efecto<sup>a</sup></b>	<b>Función</b>	<b>Posible efecto sobre el intrón</b>	<b>Ref.</b>
cAMP	E*	Regulador global molecular	Promueve frecuencia movilidad	b
Exo III	I	3'-5' exonucleasa	Degrada el cDNA naciente	c
H-NS	E*	Componente del nucleoide, regulador transcripción	Promueve frecuencia movilidad-global	d
Ligasa	E	DNA ligasa	Sellado de muescas de DNA	c
MutD	E	3'-5' exonucleasa subunidad $\epsilon$ de la Pol III ( <i>dnaQ</i> )	Reparación en la síntesis de la segunda cadena de cDNA	c
Pol I	E	5'-3' exonucleasa; eliminación de los cebadores RNA de los fragmentos de Okazaki	Eliminación del molde RNA del intrón	c
Pol II	E	Polimerasa de reparación ( <i>polB</i> )	Reparación de uniones DNA-RNA	c
Pol III	E	Polimerasa de replicación	Síntesis de la segunda cadena de cDNA	c
Pol IV	E	Polimerasa de reparación ( <i>dinB</i> )	Reparación en la polimerización	c
Pol V	E	Polimerasa de reparación ( <i>umuDC</i> )	Reparación en la polimerización	c
poly(P)	E*	Regulador global molecular	Alteración de la localización de la IEP y de la integración del intrón	e
ppGpp	E*	Regulador global molecular	Promueve frecuencia movilidad	b
RecJ	E	5'-3' exonucleasa	Degradación 5'-3' del DNA	c
RNase E	I	Ribonucleasa; parte del degradosoma de RNA	Reduce la vida media del RNA del intrón	c
RNase H1	E	Ribonucleasa; elimina la cadena de RNA en híbridos RNA/DNA	Eliminación del molde RNA del intrón	c
RNase I	I	Ribonucleasa	Reduce la vida media del RNA del intrón	c
StpA	E*	Componente del nucleoide, chaperona de RNA	Promueve frecuencia movilidad-global	d

**Tabla II.1.** Factores del hospedador que afectan a la movilidad del intrón del grupo II Ll.LtrB en *E. coli*. **a:** E, estimula el retromovimiento; E\*, estimula el retromovimineto sólo al cromosoma; I, inhibe el movimiento; **b,** Coros *et al.*, 2009; **c,** Smith *et al.*, 2005; **d,** Beauregard *et al.*, 2006; **e,** Zhao *et al.*, 2008.

Se han llevado a cabo varios estudios sobre los efectos de determinados factores del hospedador sobre la movilidad del intrón Ll.LtrB en una colección de mutantes de *E. coli*. Se determinó que los intrones del grupo II se podrían comportar como parásitos, utilizando enzimas celulares

como polimerasas, nucleasas y DNA ligasas para completar el *retrohoming* y la integración del RNA del intrón (Smith *et al.*, 2005), y que otras enzimas disminuían los niveles de *retrohoming* sobre el cromosoma (Coros *et al.*, 2008). También se ha relacionado la proliferación de estos elementos con la supervivencia del hospedador (Beauregard *et al.*, 2006; 2008). Recientemente se ha descubierto que determinados reguladores globales del estado celular también influyen sobre la retrotransposición de Ll.LtrB en *E. coli* (Coros *et al.*, 2009) (Tabla II.1).

Para intentar dilucidar qué factores de *S. meliloti* pueden afectar a la movilidad de RmInt1 hemos usado dos estrategias. Una similar a la de los estudios ya realizados, usando una colección de mutantes en las cepas de *S. meliloti* 1021 y 2011 (de las que conocemos su secuencia completa); con la otra estrategia se ha estudiado cómo varía el transcriptoma de la cepa de *S. meliloti* RMO17 (que carece de copias de intrón) cuando en su genoma aparece el intrón, y de la cepa 1021 (con tres copias de intrón en su genoma) cuando aumenta el número de copias de RmInt1.

## II.2. MATERIAL Y MÉTODOS

### II.2.1. Cepas bacterianas y condiciones de crecimiento

Las cepas de *Sinorhizobium meliloti* usadas en este trabajo fueron: RMO17, cepa que carece de copias de RmInt1 (Villadas *et al.*, 1995), para determinar si el plásmido pGm4 es equiparable a pKGEMA4 como donador de intrón; 1021 (cepa cuyo genoma ha sido secuenciado; Galibert *et al.*, 2001) y 2011, que son cepas derivadas independientes del aislado silvestre *S. meliloti* SU47 resistentes a estreptomycin (Meade y Signer, 1977; Wais *et al.*, 2002) y en cuyo fondo genético se ha conseguido una colección de mutantes mediante inserción de plásmido o del minitransposón Tn5, respectivamente (Pobigaylo *et al.*, 2006). *Escherichia coli* DH5 $\alpha$  se usó para la clonación y mantenimiento de plásmidos. Las cepas de *S. meliloti* se crecieron a 28°C en medio completo (TY) o en medio mínimo definido (MM) (Robertsen *et al.*, 1981) y *E. coli* DH5 $\alpha$  a 37°C en medio Luria-Bertani (LB). Los medios fueron suplementados, según los requerimientos de las cepas bacterianas, con antibióticos a las siguientes concentraciones: kanamicina, 200  $\mu\text{g/ml}$  para rhizobios y 50  $\mu\text{g/ml}$  para *E. coli*; tetraciclina, 10  $\mu\text{g/ml}$ ; ampicilina, 200  $\mu\text{g/ml}$ ; estreptomycin, 600  $\mu\text{g/ml}$  (1021 y 2011); ácido nalidíxico 10  $\mu\text{g/ml}$  (1021 y 2011), gentamicina 50  $\mu\text{g/ml}$  (1021 y 2011).

Las cepas curadas con inserciones de intrón en el genoma se obtuvieron introduciendo el plásmido donador de intrón (pKG2,5, en el caso de precisar inserciones de intrón completo; o pKGEMA4, si sólo queremos que se movilice la ribozima), que confiere resistencia a kanamicina, por conjugación triparental desde *E.coli* DH5α a las diferentes cepas de *S. meliloti*, 1021 y RMO17. Para facilitar que la bacteria pierda el plásmido una vez que el intrón se ha insertado, se seleccionaron colonias aisladas, y se subcultivaron sin suplementar el medio con antibiótico. Se seleccionaron cinco clones sensibles a kanamicina (habían perdido el vector donador) de cada uno de los cruces para continuar con el análisis.

### *II.2.2 Hibridación DNA-DNA*

Para la determinación del patrón de inserciones en las distintas cepas se extrajo el DNA total procedente de 2 ml de cultivo bacteriano en fase estacionaria temprana usando el kit comercial *AquaPure genomic DNA isolation* (Bio-Rad). Se digieren 5 µl del DNA total resultante en un volumen final de 30 µl bien con *Bam*HI, bien con *Sa*II. Ninguna de las dos enzimas corta dentro de la secuencia de RmInt1 ni tampoco en la secuencia de inserción *ISRm2011-2*.

Se cargaron 10 µl de cada digestión en un gel de agarosa 0,8% en TAE de 30 cm de longitud. La electroforesis se mantiene a 50V durante 24 horas para conseguir una buena separación de los diferentes tamaños. La transferencia se hace por vacío en presencia de NaOH 1N según se describe en el material y métodos generales. Cada membrana se hibrida con dos sondas: aquella que identifica la diana *ISRm2011-2* (cebadores b1 y b2, ver material y métodos generales) y la que se corresponde con la secuencia de la ribozima (ε1-18R.0).

### *II.2.3. Hibridación genómica comparativa*

El DNA total usado para los ensayos de hibridación genómica comparativa se obtuvo usando kit comercial *AquaPure genomic DNA isolation* (Bio-Rad) pero partiendo de un mayor volumen de cultivo saturado, 6 ml. Se resuspendió en 100 µl y se digirieron 20 µl con 100 unidades de *Taq*I (NEB) a 65°C toda la noche. Se limpia la digestión usando el kit de limpieza de productos de PCR (Qiagen) y el producto eluido se seca completamente en speed-vac (se puede conservar a -20°C). Para usarlo, se resuspende en 50 µl de TE1x.

Para el marcaje de las muestras se procede de acuerdo al protocolo descrito en el material y métodos generales usando Bioprime Labeling kit (Invitrogen). Se toman 22  $\mu$ l del correspondiente DNA y se mezclan con 25  $\mu$ l de la mezcla de reacción conteniendo el fluoróforo. Tras la adición de 3  $\mu$ l de Exo-Klenow se incuban durante 2 h a 37°C. Se limpia el fluoróforo no incorporado mediante el uso de las columnas suministradas por la casa comercial. Se realizaron dos réplicas técnicas de cada comparación intercambiando los fluoróforos en el marcaje. Una vez pretratado el microarray, se lleva a cabo la hibridación en baño de agua a 65°C durante al menos 16 h. Tras los lavados, el escaneado se llevó a cabo con GenePix™ Personal 4100A Microarray Scanner. La normalización basada en regresión local y el tratamiento estadístico (test de Holm) de los datos se realizó con el programa EMMA 2.8.2 desarrollado por el Centro de Biotecnología de la Universidad de Bielefeld (Dondrup *et al.*, 2003). La ausencia o multiplicidad de los genes se consideró significativa de acuerdo a los siguientes criterios: el intervalo de confianza debe superar el 99% ( $p < 0,01$ ); la intensidad media por señal supera 8 ( $A \geq 8$ ) y la razón entre las intensidades de cada fluoróforo supera las 4 veces de diferencia ( $M \geq 2$  o  $\leq -2$ ); además deben corresponderse con al menos 3 señales de las 4 posibles en los chips.

#### II.2.4 Análisis mediante hibridación en microarrays.

Las muestras de RNA total se obtuvieron a partir de cultivos de las diferentes cepas curadas seleccionadas o bien a partir de las cepas control *S. meliloti* 1021 y RMO17 como se describe previamente (Rüberg *et al.*, 2003). Se parte de 10 ml de cultivo bacteriano a una  $DO_{600}$  de 0,6 unidades que se recogen en 7 tubos de 1,5 ml mediante una centrifugación de un minuto; rápidamente se lavan con 180  $\mu$ l de sarcosil 0,1%TE e inmediatamente se congelan en  $N_2$  líquido. Para la extracción se usó el RNeasy Mini Kit (Qiagen). Cualquier posible resto de DNA presente en la muestra fue eliminado por digestión con DNaseI (Qiagen) y una nueva purificación usando las columnas del kit. La integridad del RNA se comprobó mediante gel de agarosa desnaturalizante en MOPS/formaldehído al 1,4%, su concentración se determinó espectrofotométricamente con Nanodrop y la ausencia de contaminación con DNA mediante PCR usando oligos específicos para *nifH*.

El protocolo de marcaje de los cDNA se realizó de acuerdo al protocolo descrito en el material y métodos generales. Se sintetiza el cDNA a partir de 10  $\mu$ l de RNA total bacteriano extraído con kit. Se analiza el perfil de expresión de 3 réplicas biológicas, cada una de las cuales se analiza, a su

vez, por duplicado. En el caso de *S. meliloti* 1021, el marcaje de la cepa control siempre se llevó a cabo usando el fluoróforo Cy3, mientras que el cDNA procedente de la cepa con inserciones de intrón se acoplaba con Cy5. Sin embargo, para *S. meliloti* RMO17 los marcajes se intercambiaron en cada una de las réplicas técnicas. La hibridación tiene lugar en un baño de agua a 42°C durante al menos 16 h. La adquisición de las imágenes se llevó a cabo con GenePix™ Personal 4100A Microarray Scanner, así como la evaluación de las intensidades de hibridación. La normalización (*print-tip-lowess*) y el tratamiento estadístico (test de Holm) de los datos se realizó con el programa EMMA 2.8.2. Se considera que un gen presenta expresión diferencial si el indicador de confianza no supera 0,05 ( $p < 0,05$ ), la intensidad media se encuentra por encima de 7 ( $A \geq 7$ ) y la razón entre las intensidades se encuentra entre 1 y -1 ( $M \geq 1$  o  $\leq -1$ ), los que se traduce en 2 veces inducción o represión; se eliminaron del análisis todos aquellos genes en los que no se obtuvo señal en, al menos, 8 de las réplicas de las 12 presentes en los 6 microarrays.

#### II.2.5. RT-qPCR

Para la amplificación del pre-mRNA de RmInt1 se utilizaron la pareja de oligonucleótidos UP (5'-GAAGCACTGCCCGGAATTG-3') y P (5'-TGAAAGCCGATCCCGGAG) que generan un fragmento de 150 bp; los cebadores 3Q\_RIB1 (5'-GCGTAAAGCTGCGTGAATGAT) e ICI (5'-AGGATGACGAAACGGTCCTG) amplifican un producto de 240 bp que se corresponde con el extremo 5' de la ribozima de RmInt1; finalmente, la secuencia de la IEP se cuantificó mediante el uso de los oligonucleótidos 21.0 (5'-AGAAAAGACGTCAAATGCAA) y E17.0 (5'-AGTGAACAAGACCATCGTCG). La eficiencia de amplificación de los oligonucleótidos se determinó mediante el uso de diluciones seriadas de DNA plasmídico correspondiente a pKG2,5. El análisis de las curvas de apertura de las cadenas indica la especificidad del producto amplificado.

El cDNA se sintetiza a partir de 10 µg del RNA total que se utilizó en la hibridación en microarrays, usando hexámeros de secuencia aleatoria y AMV RT durante 3 h a 42°C. La PCR a tiempo real se lleva a cabo en iCycler System (Bio-Rad) utilizando como molde 1 µl del cDNA sintetizado, SYBR Green diluido 1:10000 y Platinum Taq DNA Polymerase (Invitrogen) (para ver el proceso detallado, consultar el apartado correspondiente en la sección de material y métodos generales). Cada muestra fue ensayada por duplicado. El

estudio se completa con controles sin molde o con RNA total no sometido a reverso transcripción (datos no mostrados).

Los niveles de expresión de ribozima e IEP se normalizan de acuerdo a la cantidad de intrón no escindido (pre-mRNA) presente en las muestras.

### II.2.6. Colección de mutantes de *S. meliloti* 1021 y 2011

Tomando como base los estudios de movilidad del intrón del grupo II realizados en *E. coli* (Smith *et al.*, 2005), se buscaron mediante identidad de secuencia aminoacídica las proteínas que probablemente participen también en la movilidad de RmInt1 en el fondo genético de *S. meliloti* 1021 o 2011 (mediante las herramientas BLASTp: <http://blast.ncbi.nlm.nih.gov/Blast.cgi>; Proteome Navigator: Bowers *et al.*, 2004; y la base de datos de *S. meliloti* <http://iant.toulouse.inra.fr/bacteria/annotation/cgi/rhime.cgi>). También se analizó la bibliografía existente sobre otros factores que pueden influir en la actividad de los intrones del grupo II (Sontheimer *et al.*, 1999; Jenkins y Barkan, 2001; Hertweck y Mueller, 2001; Gregan *et al.*, 2001; Noah y Lambowitz, 2003; Perron *et al.*, 2004; Bock-Taferner y Wank, 2004; Grohman *et al.*, 2007). Con toda la información se obtuvieron los mutantes de la colección existente en la Universidad de Bielefeld en Alemania (Pobigaylo *et al.*, 2006) y que se encuentra accesible desde la plataforma RhizoGATE (<http://www.rhizogate.de>) (Tabla II.2).

Gen mutado (Pos. en el genoma)	Nombre del mutante	Posición de la mutación	Función probable
<b>Ligasas</b>			
<b>SMA0424</b> (230330-229293)	2011mTn5STM.1.01.G09	229373- 229374	putativa DNA ligasa
<b>SMb21044</b> (665732- 663825)	2011mTn5STM.3.02.F10	664637- 664638	putativa ATP-dep. DNA ligasa
<b>SMb20912</b> (1320904- 1319813)	2011mTn5STM.3.10.E09	1319975-1319976	putativa ATP-dep. DNA ligasa
<b>SMb20686</b> (1489827- 1490708)	SmPI_1021.12.08C2 2011mTn5STM.2.11.C07 2011mTn5STM.4.11.G05	1490183-1490184 1490319-1490320 1490615-1490616	putativa ATP-dep. DNA ligasa
<b>SMc03835</b> (3501745- 3501161)	2011mTn5STM.1.07.E12 2011mTn5STM.2.07.C01	3501736-3501737 3501405-3501406	2'-5' RNA ligasa ( <b>ligT</b> )
<b>Helicasas</b>			
<b>SMc00228</b>	2011mTn5STM.3.10.D01	1791486-1791487	3'-5' helicasa ( <b>RecG</b> ), ATP-

(1791477-1793582)			dep.
<b>SMc01461</b> (2306850-2309423)	2011mTn5STM.2.03.F05	2307492-2307493	3'-5' helicasa de híbridos RNA/DNA <b>(UvrD1)</b>
	SmPI_1021.12.01H5	2308993-2308994	
<b>SMc01566</b> (2508771-2506861)	2011mTn5STM.3.02.G02	2507083-2507082	3'-5' DNA helicasa ( <b>RecQ</b> ), ATP-dep.
<b>SMc02586</b> (77448- 74989)	2011mTn5STM.2.03.B09	77299- 77300	ATP-dep. helicasa ( <b>helO</b> )
<b>SMc01090</b> (469665- 467776)	2011mTn5STM.1.06.H02	469071- 469072	ATP-dep. RNA helicasa ( <b>deaD</b> )
	2011mTn5STM.3.04.B07	468014- 468015	
<b>SMc02152</b> (560199- 556690)	2011mTn5STM.1.04.D01	558440- 558439	helicasa
<b>SMc02237</b> (603400- 600854)	2011mTn5STM.1.06.F08	602008- 602009	ATP-dep. helicasa
<b>Proteínas reparadoras</b>			
<b>SMc00556</b> (1227431-1225971)	2011mTn5STM.2.02.C03	1226930- 226931	Proteína reparadora de DNA ( <b>radA</b> )
	2011mTn5STM.4.04.E12	1226787-1226788	
	2011mTn5STM.1.14.F08	1226170-1226171	
<b>SMc01235</b> (1689506-1686588)	2011mTn5STM.4.03.A06	1689490-1689489	excinucleasa ABC subunidad A (reparación DNA ATP-dep.) <b>uvrA</b> <b>(OPERÓN CON SMc01236)</b>
	2011mTn5STM.1.09.G09	1687233-1687234	
<b>SMc01236</b> (1686495-1685692)	2011mTn5STM.2.12.B04	1686489-1686488	Proteína hipotética conservada
<b>SMc01877</b> (2331847-2330177)	2011mTn5STM.1.02.G07	2330305-2330306	Proteína de reparación del DNA ( <b>RecN</b> )
<b>Prot. Rel. con polimerasas.</b>			
<b>SMc03788</b> (3455860-3452510) <b>dnaE2</b>	2011mTn5STM.5.12.C02	3453150-3453149	Principal enzima replicativa ( <b>DNA Pol III</b> )
	2011mTn5STM.2.04.G08	3453492-3453491	
	2011mTn5STM.3.02.F01	3452935-3452934	
	2011mTn5STM.4.13.F08	3452877-3452878	
	2011mTn5STM.3.09.F02	3455685-3455684	
	2011mTn5STM.4.02.F11	3453248-3453249	
	2011mTn5STM.1.02.B04	3453643-3453644	
SmPI_1021.12.03H10	3453649-3453648		
<b>SMA0892</b> (497392-500655) <b>dnaE3</b>	2011mTn5STM.3.11.G05	498225-498224	
	2011mTn5STM.2.02.E09	499917-499916	
	2011mTn5STM.5.10.E04	498845-498846	
	2011mTn5STM.3.01.H03	498854-498853	
	2011mTn5STM.5.07.D02	497793-497794	
	2011mTn5STM.3.11.A08	500111-500110	

	2011mTn5STM.1.14.F07	498293-498292	
<b>SMc01373</b> (1409085-1407793)	2011mTn5STM.1.02.D07	1408298-1408299	DNA pol IV 1, extiende Horquilla parada ( <b>dinB1</b> )
	2011mTn5STM.4.10.F04	1408427-1408426	
<b>SMa2355</b> (1327042-1328151)	2011mTn5STM.4.05.D10	1327889-1327890	Nucleotidyltransferasa/DNA polymerasa
<b>SMb21448</b> (1378745-1379383)	2011mTn5STM.3.11.F08	1379151-1379150	putativa DNA polymerase
	2011mTn5STM.5.09.B01	1379368-1379367	
<b>SMc01419</b> (2267073-2267624)	2011mTn5STM.5.08.D04	2267397-2267396	RNA polymerasa sigma factor ( <b>rpoE1</b> )
	2011mTn5STM.3.07.G07	2267443-2267444	
<b>SMc01506</b> (2624014-2623463)	2011mTn5STM.1.04.C05	2623877-2623878	RNA polymerasa sigma-E factor (sigma-24) ( <b>rpoE2</b> ) (Operón con SMc01505 Anti-sigma factor: <b>NM</b> )
	2011mTn5STM.3.12.B10	2623780-2623779	
<b>SMc04051</b> (3055740-3055207)	2011mTn5STM.1.13.E03	3055308-3055309	RNA polymerasa sigma-E factor (sigma-24) ( <b>rpoE4</b> )
<b>SMc03873</b> (3539345-3538509)	2011mTn5STM.4.09.D11	3538958-3538959	RNA polymerasa sigma factor ( <b>rpoH2 sigC</b> )
<b>Nucleasas</b>			
<b>SMc01336</b> (1444408-1441634)	2011mTn5STM.4.07.G10	1442384-1442383	procesamiento y eliminación del RNA, <b>RNaseE</b> ( <i>rne</i> )
<b>SMc02327</b> (660513-660004)	2011mTn5STM.1.13.B02	660227-660228	<i>mhA2</i> , degradación del RNA de híbridos RNA/DNA ( <b>RNase H1</b> )
<b>SMc00867</b> (920230-919580)	SmPI_1021.12.04C9	919972-919971	<i>mhB</i> , degradación del RNA de híbridos RNA/DNA ( <b>RNase H2</b> )
<b>SMc02760</b> (38065-34496)	2011mTn5STM.4.10.G03	36777-36776	ssDNA exonucleasa ( <b>RecB</b> ), ATP-dep.
<b>SMb20689</b> (1492447-1493220)	2011mTn5STM.1.13.G07	1492508-1492509	3'-5'dsDNA exonucleasa ( <b>XthA4</b> )
	SmPI_1021.12.08C5	1492693-1492694	
<b>SMa0934</b> (515693-520312)	SmPI_1021.12.06E7	519062-519063	TraA1 proteína de transfer. conjugativa identidad con ssDNA exonucleasa ( <b>RecD</b> )
<b>SMc02843</b> (192909-193697)	SmPI_1021.12.01A5	193210-193211	endonucleasa III ( <b>nth</b> )
	2011mTn5STM.1.02.E08	193450-193451	
<b>SMc00378</b> (331801-333378)	2011mTn5STM.3.06.H09	332097-332096	exodeoxyribonucleasa ( <b>xseA</b> )
	2011mTn5STM.2.11.A01	332918-332917	
<b>SMc00606</b> (1277704-1276949)	2011mTn5STM.4.09.C03	1277434-1277435	ribonucleasa
<b>SMc01929</b>	2011mTn5STM.3.10.C03	1394989-1394990	Proteína hipotética

(1394904-1396571)	2011mTn5STM.3.13.E11	1395054-1395055	<b>(Ribonucleasa J)</b>
<b>SMc00622</b> (1297874-1296726)	2011mTn5STM.4.07.H06	1297771-1297770	ribonucleasa D ( <b>rnd1</b> )
<b>SMc04231</b> (2055780-2058734)	2011mTn5STM.2.08.A01 2011mTn5STM.4.03.D11	2055881-2055880 2057901-2057900	excinucleasa ABC subunidad B ( <b>uvrB</b> )
<b>Chaperonas</b>			
<b>SMc02857</b> (209764-211686)	2011mTn5STM.3.05.F08	211305- 211306	Proteína de choque térmico 70 (HSP70) chaperona ( <b>dnaK</b> )
<b>SMc00003</b> (985233- 986288)	SmPI_1021.12.01B9 2011mTn5STM.3.04.G09	985333- 985334 985818- 985819	chaperona
<b>SMc00913</b> (870772- 869138)	2011mTn5STM.1.08.G04 SmPI_1021.12.04B10 2011mTn5STM.2.08.H06	870554- 870555 870320- 870319 870019- 870020	60 KD chaperonina A ( <b>groEL1</b> )
<b>SMc01758</b> (1305463-1303838)	2011mTn5STM.2.07.C12 2011mTn5STM.3.03.B05 2011mTn5STM.2.05.F01 2011mTn5STM.1.03.G02	1305234-1305235 1304928-1304927 1304421-1304420 1304205-1304206	60 KD chaperonina B ( <b>groEL4</b> )
<b>SMc04456</b> (2527640-2527302)	2011mTn5STM.2.09.D09	2527480-2527479	chaperona ( <b>csaA</b> )
<b>SMA0124</b> (67826- 66195)	2011mTn5STM.1.04.A01 2011mTn5STM.1.05.H11	66337- 66336 66228- 66229	<b>GroEL3</b> chaperonina
<b>SMB21183</b> (932442- 930556)	2011mTn5STM.2.12.G12	931519- 931518	probable chaperonina, familia de las proteínas de choque térmico hsp90 ( <b>htpG</b> )
<b>Miscelanea</b>			
<b>SMc00321</b> (273485- 274417)	2011mTn5STM.4.10.C12 2011mTn5STM.2.08.G08 2011mTn5STM.4.11.B07	273584- 273583 273737- 273736 274206- 274205	tRNA pseudouridine sintasa B <b>truB</b>
<b>SMc02693</b> (2543265-2543981)	2011mTn5STM.2.07.H02	2543849-2543850	peptidyl-tRNA hidrolasa ( <b>pth</b> )
<b>SMc00713</b> (2867621-2867070)	2011mTn5STM.1.11.B04 SmPI_1021.12.05E3	2867465-2867466 2867447-2867446	Transporte de cationes ( <b>chaC</b> )
<b>SMc00399</b> (349712- 348726)	2011mTn5STM.2.03.D08 2011mTn5STM.5.03.D01	348992- 348993 348721- 348722	Proteína transmembrana de transporte de Mg. y Co. ( <b>corA1</b> )
<b>SMc00874</b> (911877- 912854)	2011mTn5STM.2.08.F04	912221- 912222	Proteína transmembrana de transporte de Mg. y Co. ( <b>corA2</b> )
<b>SMc01023</b> (1548277-1549044)	2011mTn5STM.4.08.B01	1548617-1548616	triosephosphate isomerase ( <b>tpiA1</b> )

<b>SMc03979</b> (2975447-2976454)	2011mTn5STM.2.11.H09	2975948-2975947	glyceraldehyde 3-phosphate dehydrogenase ( <b>gap</b> ) <b>OPERON gap, pgk, tkt2 y SMc03970</b>
<b>SMc03981</b> (2977053-2978252)	2011mTn5STM.1.09.G04 2011mTn5STM.3.08.F07	2977533-2977534 2977789 2977790	Probable phosphoglycerate kinase ( <b>pgk</b> )
<b>SMc03980</b> (2976463-2977053)	2011mTn5STM.4.09.B01	2976689-2976688	Proteína hipotética conservada
<b>SMc03978</b> (2973395-2975377)	2011mTn5STM.1.12.D08	2974676-2974675	Probable transketolase ( <b>tkt2</b> )
<b>SMc02163</b> (532370- 530748)	2011mTn5STM.1.01.H03 2011mTn5STM.4.12.B11	531452- 531453 531127- 531126	glucose-6-phosphate isomerase ( <b>pgi</b> )
<b>Smb21579</b> (1119502-1119945)	2011mTn5STM.4.11.A09 SmPI_1021.12.07H8	1119529-1119530 1119609-1119610	Regulador transcripcional putativo, familia de la proteína merR ( <b>hmrR2</b> )
<b>SMc00109</b> (1072051-1071584)	2011mTn5STM.1.05.G08 2011mTn5STM.3.07.E03	1071836-1071835 1071790-1071791	Probable regulador transcripcional
<b>SMc02659</b> (1164419-1166644)	2011mTn5STM.2.12.C04	1165740-1165741	Putative GTP pyrophosphokinase (ATP:GTP 3'-pyrophosphotransferase) ( <b>relA</b> )
<b>SMc02496</b> (3298640-3296436)	SmPI_1021.12.03F10	3297084-3297083	Probable primosomal protein N' (replication factor Y) <b>priA</b>
<b>SMc00339</b> (296513-294516)	2011mTn5STM.4.08.H11 2011mTn5STM.1.14.A01	296298-296299 295320-295319	Adenylate cyclase 1 protein ( <b>cyaA</b> )
<b>Smb21488</b> (1409489-1411498)	2011mTn5STM.4.03.B02 2011mTn5STM.4.06.A01	1410055-1410054 1411114-1411115	putative cytochrome o ubiquinol oxidase chain I protein ( <b>cyoB</b> )

**Tabla II.2.** Listado de los mutantes utilizados en este trabajo. En la primera columna se detalla el nombre de la región y la posición en el genoma de 1021, en la segunda el nombre del mutante. En la tercera columna la posición en la que se ha producido la mutación por inserción. En la última columna se hace referencia a la posible función desempeñada por el gen junto con el nombre asignado. La tabla se ha dividido por identidad con los genes del trabajo de Smith *et al.* y, a continuación, por funciones que podrían resultar interesantes en cuanto a su implicación en la actividad de los intrones del grupo II.

### II.2.7. Construcción del plásmido donador de intrón pGm4

Debido a que los mutantes en 1021 y 2011 son resistentes a kanamicina, no se puede usar el vector donador de intrón pKGEMA4 (capítulo I; Nisa-Martínez *et al.*, 2007) para los ensayos de movilidad, por lo que se desarrolló otro plásmido donador, pero que confiere resistencia a

gentamicina. Está basado en el plásmido pBBR1MCS-5 (Kovach *et al.*, 1994) en el cual se insertó el derivado  $\Delta$ ORF de RmInt1 en la diana *SacI* del sitio de clonación múltiple, en el mismo sentido que el gen promotor T3. A continuación se clonó aguas arriba un fragmento *Sall-XhoI* proveniente de la construcción pKGEMA4, que incluye el promotor del gen de resistencia a kanamicina y la secuencia codificante de la IEP, quedando todo en la misma disposición que en el pKGEMA4.

### II.2.8. Ensayos de movilidad

La movilidad del intrón expresado desde el plásmido donador pGm4, para compararlo con el donador pKGEMA4, se determinó en el fondo genético de *S. meliloti* RMO17 (carente de copias de intrón). Se llevó a cabo un ensayo de doble plásmido usando como plásmido aceptor la construcción pJB0,6LAG, y como control negativo pJB $\Delta$ 129. Para ello se introdujeron los plásmidos en RMO17 mediante transferencia conjugativa triparental, usando como cepa ayudante HB101:pRK2013. Los transconjugantes fueron seleccionados en MM suplementado con Ap/Tc/Km o Gm dependiendo del donador a seleccionar.

Para determinar la movilidad de RmInt1- $\Delta$ ORF en los mutantes de 1021 y 2011 seleccionados, se llevó a cabo el mismo ensayo de movilidad anteriormente explicado, pero los transconjugantes fueron seleccionados en medio TY suplementado con Sm/Ap/Gm.

Para cuantificar la eficiencia de movilidad se realizaron *Southern blots* con el DNA plasmídico extraído. En el caso de los mutantes, se realizó una mezcla del DNA plasmídico proveniente de entre 6 y 10 colonias diferentes elegidas al azar. La hibridación se realizó con una sonda que reconoce la diana (ISRm2011b1-b2). La cuantificación se realizó con la fórmula  $H/(H+A)$ , donde H es la banda correspondiente al *homing* y A es la banda correspondiente a la diana sin invadir.

## II.3. RESULTADOS

### II.3.1. Selección de las cepas para los ensayos con microarrays

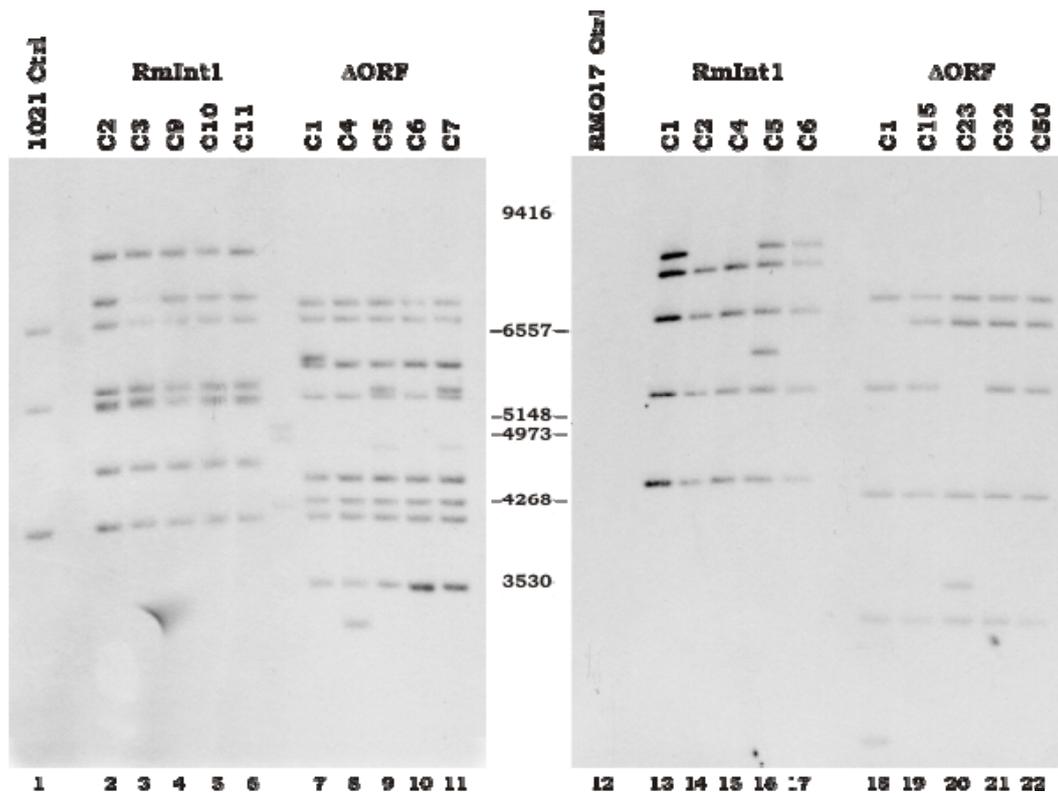
Con el fin de seleccionar cepas isogénicas en cuanto a presencia de intrón se refiere, se analizó el perfil de dispersión de RmInt1 en diferentes fondos genéticos de *S. meliloti*, 1021 –cepa cuyo genoma se encuentra secuenciado y que presenta 3 copias de RmInt1 en su genoma- y RMO17 –cepa que carece de copias de RmInt1 en su genoma (Figura II.2, carriles 1 y 12). De manera similar al procedimiento descrito en el capítulo I, se introduce un plásmido donador que expresa el intrón completo (ribozima + IEP) o solamente la ribozima en los diferentes fondos genéticos; transcurridas varias generaciones, las secuencias de inserción no invadidas, pasan a estarlo según el patrón descrito. Se analiza el perfil de inserciones en un conjunto de clones previamente curados del plásmido donador de intrón mediante *Southern blot* de digestiones BamHI o Sall de DNA total. Se obtuvieron distintos patrones de invasión en los que el número total de inserciones varía entre las 4 que se producen en algunos clones de *S. meliloti* RMO17 invadidos por intrón completo y las 9 en el caso de *S. meliloti* 1021, con  $\Delta$ ORF. Destacar que la mayoría de los clones analizados muestran 5 nuevas copias de intrón interrumpiendo las correspondientes secuencias de inserción (Figura II.2).

El análisis de los perfiles de invasión de *S. meliloti* 1021, nos permitió seleccionar dos cepas en las cuales las secuencias de inserción invadidas coinciden exactamente tanto para las invasiones con intrón completo como por la ribozima: 1021 2,5 C2 y 1021  $\Delta$ ORF C5 (carriles 2 y 9). En ambos casos las inserciones se mapearon en las secuencias de inserción identificadas como ISc4, ISc3, ISc6, ISa1 e ISc5, además de los eventos de invasión en ISa2, ISb2 e ISRm10-1 previamente identificados en *S. meliloti* 1021 (carril 1). Por otro lado, el patrón de hibridación parece indicar que las cepas RMO17 2,5 C1 y C6 y RMO17  $\Delta$ ORF C15, C32 y C50 presentan las mismas secuencias de inserción invadidas, concretamente las identificadas como IS2, 3, 4, 6 y 9. Para continuar con nuestro estudio seleccionamos las cepas RMO17 2,5 C1 y RMO17  $\Delta$ ORF C15 (carriles 13 y 19), pero desconocemos si las secuencias de inserción invadidas en *S. meliloti* RMO17 se corresponden con las identificadas en *S. meliloti* 1021.

Así, podemos decir que, atendiendo a cada cepa, los clones seleccionados con inserciones de intrón completo y  $\Delta$ ORF son isogénicos, pudiendo afirmar que no lo son entre las distintas cepas de acuerdo a la diferencia en el número de intrones insertados.

### II.3.2. Genes diferencialmente expresados en *S. meliloti* 1021 en presencia de intrón.

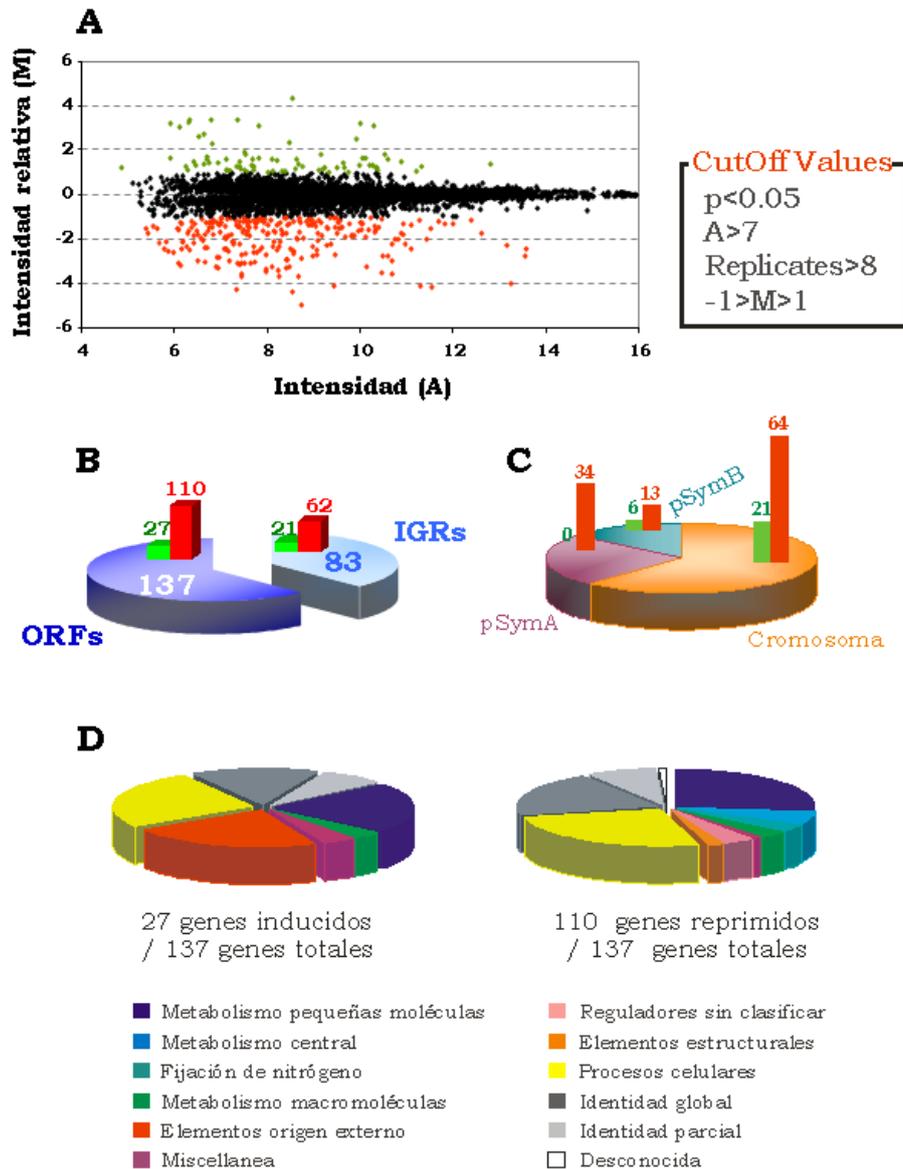
Para conseguir tener una idea global de los genes que se encuentran diferencialmente expresados en *S. meliloti* 1021 debido a la presencia de RmInt1 utilizamos la hibridación de microarrays. Disponemos de chips de DNA basados en la secuencia de *S. meliloti* 1021 en los cuales se encuentran



**Figura II.2** Distribución de las copias de intrón en *S. meliloti*. El DNA total procedente de las diferentes cepas curadas así como de los correspondientes controles fue digerido con Sall, separado en geles de agarosa, transferido a membrana e hibridado con sonda de la ribozima ( $\epsilon$ 1/18R.0). En el panel izquierdo, se muestra el patrón de hibridación con de las cepas curadas *S. meliloti* 1021 que han adquirido nuevas copias de intrón tanto completo (RmInt1) como en su versión modificada ( $\Delta$ ORF). La hibridación de la derecha se corresponde con idéntica descripción en *S. meliloti* RMO17. En ambos casos, la cepa que se utilizó como control se muestra en el primer carril.

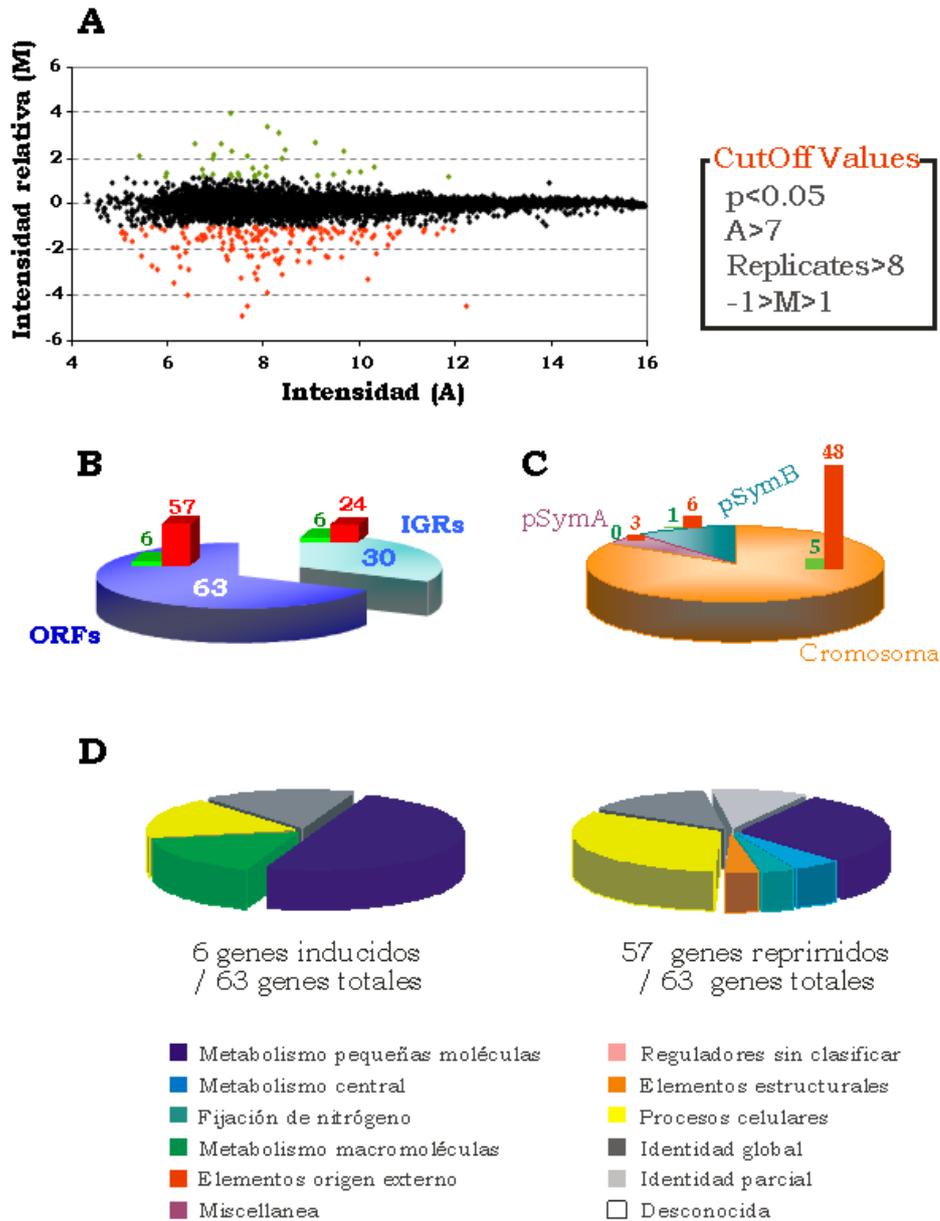
representadas tanto las regiones codificantes como las intergénicas, aunque en nuestro análisis nos vamos a centrar principalmente en las primeras. Nos proponemos distinguir entre aquellos genes que presentan regulación debido a la funcionalidad de RmInt1, frente a aquellos que pudieran verse afectados simplemente por la polaridad de la inserción. Para ello estudiamos comparativamente el efecto producido por la presencia de intrones completos totalmente funcionales, con cepas isogénicas en cuanto a dianas invadidas por intrones inactivos ( $\Delta$ ORF). Por un lado, la ribozima en ausencia de la IEP (intrón  $\Delta$ ORF) es incapaz de escindir, por lo que podría generar una alteración de la expresión de los genes circundantes; por otro lado, al tratarse de intrones truncados funcionalmente, aquellos genes que pudieran estar requiriéndose para los procesos de movilidad y escisión podrían no ver modificada su expresión. Partimos del supuesto de que la presencia de un mayor número de intrones pudiese generar diferencias de expresión de aquellos genes del hospedador que están implicados en procesos de maduración y movilidad de estos elementos genéticos.

El cDNA sintetizado a partir del RNA total de *S. meliloti* 1021 se marca con Cy3 mientras que el resultante de la transcripción reversa del RNA extraído de 1021 2,5 C2 y/o 1021  $\Delta$ ORF C5 se acopla con el fluoróforo Cy5. Los datos se analizaron tanto a nivel de hibridación individual con la posterior integración de los resultados, como agrupando las 6 réplicas de que disponemos en cada experimento a la hora de la normalización y tratamiento estadístico. Observamos que, al normalizar individualmente, el número de secuencias diferencialmente expresadas se reduce comparado con el análisis del grupo. Esto sería una ventaja si no fuese porque detectamos que la mayoría de nuestros genes o secuencias control no se encuentran representadas en el conjunto final; esto es, sólo observamos señal significativa en alguno de los oligonucleótidos que se corresponden con la ribozima o la IEP de RmInt1, mientras que en el caso del análisis del grupo podemos identificar todos los controles. Así, podemos concluir que la normalización conjunta de las hibridaciones individuales proporciona una información más completa de los genes diferencialmente expresados en nuestras comparaciones. Por tanto, los datos presentados a continuación se corresponden con el análisis de las réplicas agrupadas.



**Figura II. 3.** Representación gráfica de los datos obtenidos en la hibridación de microarrays para *S. meliloti* 1021 frente a 1021 2,5 C2. **A** Gráfica M/A donde se representan las señales estadísticamente significativas sin discriminar si corresponden con regiones intergénicas (IGRs) o secuencias codificantes (ORFs). En rojo se representan las secuencias que se encuentran reprimidas en presencia de 8 copias de intrón completo frente a las 3 presentes en la cepa control, mientras el verde se muestran las señales correspondientes a inducción. **B** Distribución de las señales de acuerdo a la naturaleza de la región donde se encuentran, ORFs o IGRs. **C** Localización de los 137 genes que se encuentran diferencialmente expresados, así como patrón de expresión. **D** Distribución de los genes que se encuentran inducidos (panel izquierdo) o reprimidos (derecho) de acuerdo a su funcionalidad según las bases de datos de *S. meliloti* y KEGG.

Como promedio, en las hibridaciones correspondientes a los cruces 1021 control con 1021 2,5 C2 se obtuvo señal en 11060 oligonucleótidos de los 14976 impresos en el chip, variando en intensidad entre 4,3 y 16. Sin embargo, sólo el 40,5% de esos valores resultó ser estadísticamente significativo ( $p < 0,05$ , figura II.3 A). El análisis de los datos resultó en 220 genes diferencialmente expresados, 137 se corresponden con oligonucleótidos que identifican ORFs anotadas en el genoma de *S. meliloti* 1021 mientras que los 83 restantes se identifican en regiones intergénicas (figura II.3 B). En su mayoría, estos genes se encuentran reprimidos y se localizan principalmente en el cromosoma, seguido de pSymA y el menor número se identificó en pSymB (figura II.3 C). Si analizamos la funcionalidad de los genes (figura II.3 D), encontramos que un elevado porcentaje se corresponde con proteínas de función desconocida o proteínas hipotéticas (27,7%), mientras el otro gran grupo de genes codifica proteínas implicadas en distintos procesos metabólicos (34,3%) y procesos celulares (26,3%). El gen que presenta unos niveles de represión más elevados (en torno a 16 veces con respecto a la cepa silvestre) se corresponde con una proteína hipotética que presenta identidad con *FixM*, que codifica una flavoproteína-oxidoreductasa implicada en regulación y fijación de nitrógeno en *S. meliloti* (Cosseau *et al.*, 2002) ; a su vez, este gen se localiza aguas abajo de otros reguladores transcripcionales que también se han identificado reprimidos en nuestros datos, *FixT/K*. Destacar que otro de los genes que presenta regulación negativa es *NifA*, que a su vez es uno de los reguladores más importantes del proceso de fijación de nitrógeno en *S. meliloti*. Englobadas en el grupo de proteínas implicadas en metabolismo se encuentran un importante número de permeasas, proteínas transmembrana y transportadores. La IEP de RmInt1 es uno de los pocos genes que se encuentran activados presentando a su vez uno de los mayores índices de expresión, resultado que corrobora que las nuevas copias de RmInt1 que se se han insertado en el genoma se expresan activamente. Seguidamente encontramos genes implicados en el catabolismo de glicina, así como un significativo número de transportadores. En el límite aplicado para la selección de genes, encontramos algunos genes relacionados con la síntesis de flagelos.



**Figura II. 4.** Representación gráfica de los datos obtenidos en la hibridación de microarrays para *S.meliloti* 1021 frente a 1021  $\Delta$ ORF C5. **A** Gráfica donde se representa la intensidad de cada punto frente a la relación de las intensidades entre ambas muestras correspondientes a las señales estadísticamente significativas incluyendo regiones intergénicas (IGRs) y secuencias codificantes (ORFs). En rojo se representan las secuencias que se encuentran reprimidas, mientras en verde se muestran las señales correspondientes a inducción. **B** Representación gráfica de las señales de inducción y represión correspondientes a todos los valores que se ajustan a los criterios de significación. **C** Distribución de los genes que se encuentran diferencialmente expresados, así como patrón de expresión. **D** Adscripción de los genes identificados a las categorías funcionales definidas en KEGG y adoptadas en las bases de datos de *S.meliloti* 1021.

Si atendemos a los resultados generados por la hibridación competitiva de 1021 control frente a 1021  $\Delta$ ORF C5, sólo observamos señal en el 66% de los oligonucleótidos presentes en el microarray determinando que sólo el 42% de esas señales eran estadísticamente significativas (intervalo de confianza del 95%, figura II.4 A). El número total de genes diferencialmente expresados fue menor que en el caso de las inserciones de intrón completo, 93 genes, pero igualmente correspondiendo principalmente con secuencias codificantes de proteínas (figura II.4 B). Siguiendo con la tendencia mostrada en el análisis con intrón completo, la mayoría de los genes resultantes se encontraban reprimidos y formando parte del cromosoma (figura II.4 C). Sólo 6 genes parecían encontrarse activados en este análisis, entre los que destacamos *gcvTHP*, genes que codifican el complejo multienzimático de degradación de glicina. En este caso no identificamos los oligonucleótidos que representan la IEP de RmInt1 puesto que las inserciones se corresponden solamente con la parte ribozímica de nuestro intrón, secuencias que han sido excluidas del presente análisis (aunque como veremos más adelante, somos capaces de identificar aquellos oligonucleótidos que hibridan en la ribozima). En cuanto a los genes reprimidos (figura II.4 D), se agrupan en las 3 categorías enumeradas en el caso de inserciones de intrón completo: proteínas hipotéticas con identidad global o parcial (24,5%), y procesos metabólicos (38%) y celulares (33%). Los genes que presentan unos niveles de expresión más reducidos respecto a la cepa control se corresponden con numerosos transportadores de membrana, permeasas y proteínas transmembrana, siendo SMc02634 el que presenta los mayores niveles de represión (32 veces inhibido respecto a la cepa silvestre). Como en el caso anterior, también algunos componentes reguladores como *FixK* se encuentran reprimidos.

#### *II.3.2.1. La presencia de intrones no implica diferencias en el genotipo.*

Existen diferentes herramientas para determinar cambios en el genotipo bacteriano, pero una de las tecnologías más utilizadas es la hibridación genómica comparativa o CGH. Se basa, en el marcaje con diferentes fluoróforos del DNA total extraído a partir de los fondos genéticos a comparar, procediendo después a la hibridación competitiva sobre el chip. Si las cepas son muy similares el color predominante será el amarillo; la observación de puntos rojos o verdes va a significar que la bacteria carece de esos genes o por el contrario que podría presentar más de una copia,

dependiendo del marcaje que se haya llevado a cabo. Tanto el principio como el análisis son similares a la hibridación en microarrays pero con este tipo de aproximación sólo tenemos información genética, no de expresión diferencial.

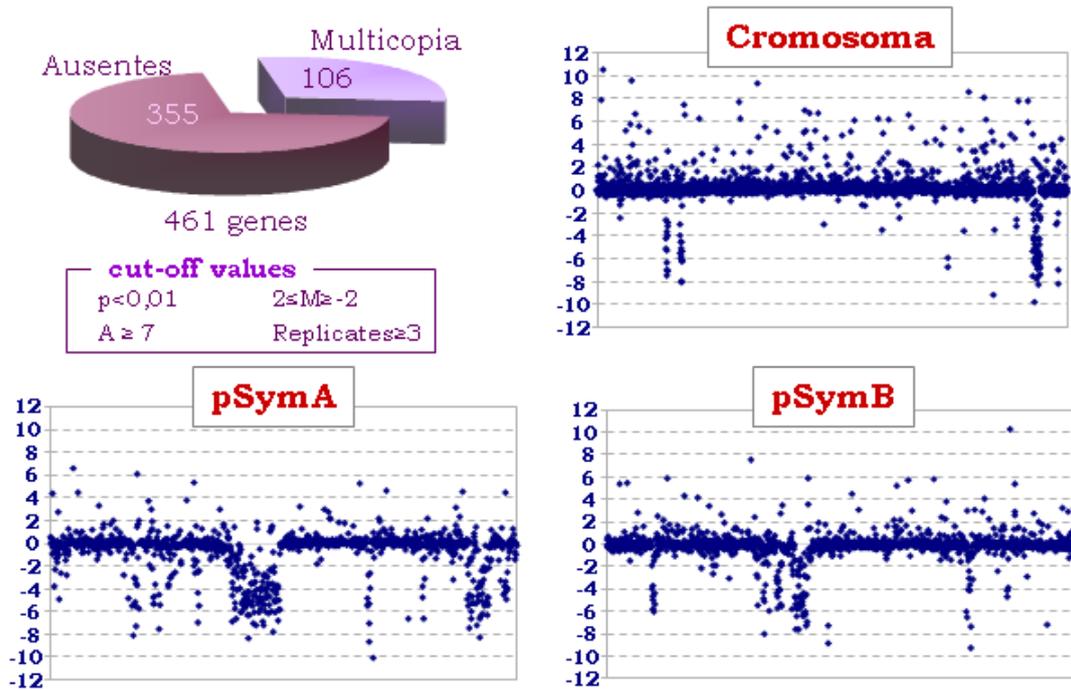
Analizamos los perfiles genéticos de la cepa *S. meliloti* 1021 control comparada con aquellos que se corresponden con las cepas con nuevas inserciones de intrón, tanto completo como en su versión  $\Delta$ ORF. Se observa un incremento en la intensidad de la señal en una serie de regiones intergénicas compartidas en los dos análisis que se corresponden con los oligonucleótidos que identifican la ribozima (estas secuencias están pendientes de ser anotadas correctamente en el genoma de *S. meliloti* 1021). Además, en el caso de la hibridación con la cepa que presenta las inserciones de intrón completo, pudimos identificar los oligonucleótidos que se corresponden con la IEP. De estos datos concluimos que, a diferencia de otros elementos móviles que, como las secuencias de inserción, introducen plasticidad en los genomas generando reordenaciones genéticas, la presencia de RmInt1 no provoca ninguna alteración genética en *S. meliloti* 1021, al menos a corto plazo. Así pues, las diferencias de expresión que hemos observado se producen como consecuencia, bien de pequeñas diferencias a nivel metabólico durante el proceso de cultivo y recogida celular, o bien a la propia expresión de RmInt1.

### **II.3.3. Perfil genético de *S. meliloti* RMO17 comparado con *S. meliloti* 1021.**

Aunque desconocemos los genes que podrían estar afectados por la presencia y movilidad de intrones y, lo que es más importante, el grado en que se estarían induciendo o reprimiendo, el hecho de que *S. meliloti* 1021 contenga 3 copias de RmInt1 insertadas en su genoma no facilita la interpretación de los resultados. Históricamente en nuestro laboratorio hemos trabajado con *S. meliloti* RMO17, una cepa aislada de nódulos de *Medicago orbicularis* que carece de copias de RmInt1 en su genoma, para llevar a cabo estudios de movilidad y maduración. Además, RmInt1 ha demostrado moverse más eficientemente en este fondo genético (Martínez-Abarca *et al.*, 2004). Por todo, nos proponemos llevar a cabo estudios transcriptómicos en cepas de *S. meliloti* RMO17 curadas de plásmido donador de intrón en las que algunas copias de IS*Rm2011-2* han sido invadidas por intrón completo o bien únicamente por ribozima. Sin embargo, nos encontramos con una limitación: desconocemos la organización del

genoma de *S. meliloti* RMO17 y sus niveles de identidad respecto a 1021. Así, nuestro primer objetivo consistió en validar el uso de los DNA microarrays diseñados para *S. meliloti* 1021 para su hibridación con material procedente de *S. meliloti* RMO17; una vez establecidas las diferencias genéticas, podremos valorar las posibles diferencias de expresión.

Cuando comparamos los genomas de *S. meliloti* 1021 y RMO17, ambas cepas silvestres (figura II.5), encontramos que 1241 oligonucleótidos presentan intensidad diferencial en la señal correspondiendo más de la mitad a secuencias intergénicas (780 puntos), entre las que identificamos la ribozima de RmInt1. Si nos restringimos al análisis de los oligonucleótidos que hibridan en secuencias codificantes, más del 5% parecen estar ausentes en *S. meliloti* RMO17 mientras que 106 podría decirse que se encuentran en multicopia. Entre los ausentes destacamos la IEP codificada por RmInt1, así como un considerable número de secuencias de inserción entre las que se encuentran *ISRm1*, 14, 17, 26, 28 y 30. El otro gran grupo representado en estas ausencias es el de transportadores, proteínas transmembrana, permeasas, etc.; a su vez, se podrían enumerar gran cantidad de proteínas hipotéticas. Dos proteínas podrían resultar interesantes para un estudio posterior, una ausente en *S. meliloti* RMO17 y la otra representada en multicopia. SMc03764 no se encuentra en *S. meliloti* RMO17; esta proteína, que se identifica con una DNA endonucleasa que repara desapareamientos, podría participar en la degradación del intrón una vez insertado, justificando las diferencias en eficiencia de movilidad entre las dos cepas. SMA0124 parece presentar un mayor número de copias en *S. meliloti* RMO17 si lo comparamos con 1021; el producto de este gen se identifica con una chaperona, GroEL3, que se encarga del correcto plegamiento de proteínas



**Figura II. 5.** Resultados de hibridación genómica comparativa entre *S. meliloti* 1021 y RMO17. En el panel superior izquierdo se representa el balance de genes que pierde RMO17 en comparación con 1021 así como posibles duplicaciones. También se incluye una tabla con los valores adoptados como criterio de significación en los estudios de CGH. Los restantes gráficos representan la disposición genómica de las secuencias adquiridas o perdidas por RMO17 distribuida por replicones: cromosoma y plásmidos simbióticos A y B. En el eje de ordenadas se representa el valor de M frente a la distribución ordenada de las secuencias en los diferentes replicones.

Como conclusión, podemos afirmar que el cromosoma parece ser el replicón menos variable, con sólo 3 regiones ausentes muy bien definidas, seguido de pSymB. Como se ha observado en el caso de otros aislados de *S. meliloti*, pSymA refleja los mayores cambios genéticos (Giuntini *et al.*, 2005; Stiens *et al.*, 2008).

### II.3.3.1. Genes diferencialmente expresados en *S. meliloti* RMO17 respecto a *S. meliloti* 1021.

Una vez filtrados los datos en función de los resultados obtenidos en CGH, nos encontramos con 104 genes diferencialmente expresados, 47 reprimidos y 57 inducidos. Entre los genes reprimidos, se identifica el operón de captación de fosfato *phoEDC* así como otras proteínas de unión a fosfato. Un elevado número de los genes activados se corresponden con sistemas transportadores de membrana, aunque también encontramos bastantes

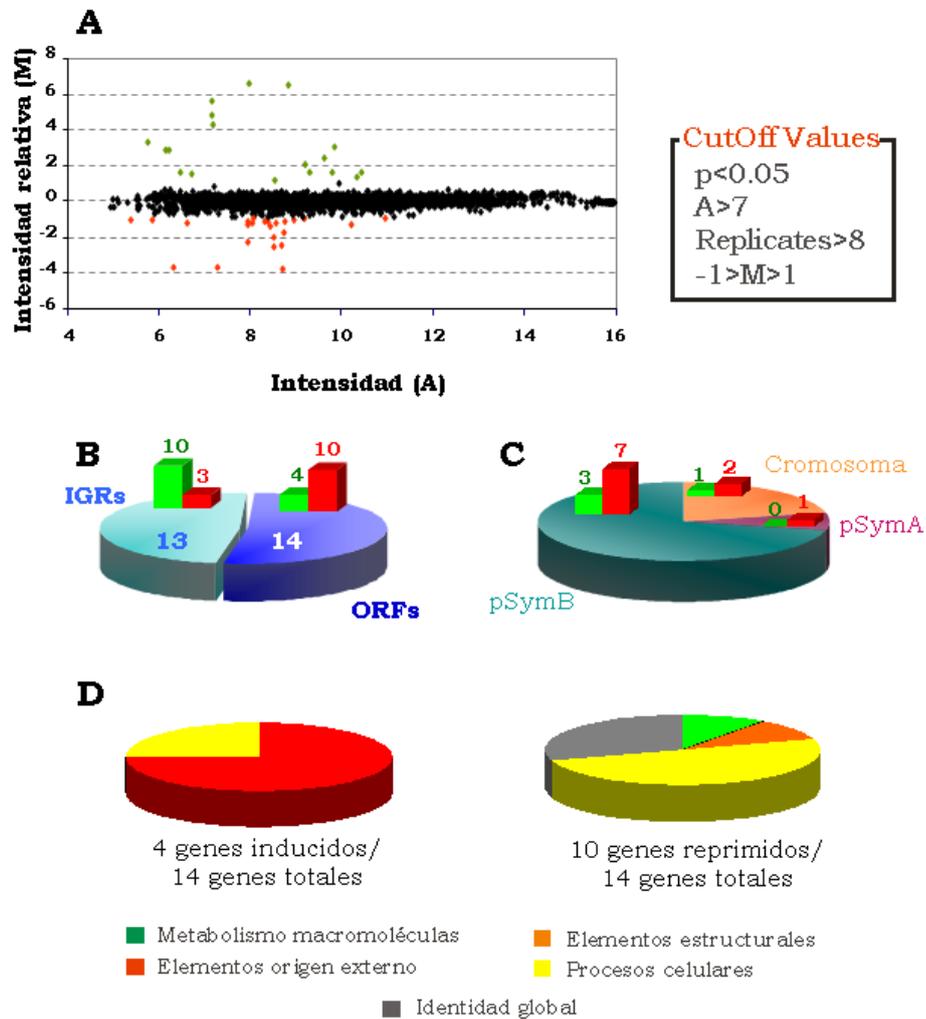
proteínas hipotéticas. Es importante destacar, que los oligonucleótidos correspondientes a la IEP o a la ribozima de los intrones que *S. meliloti* 1021 presenta en su genoma no mostraron señal diferencial. La comparación de los genomas (CGH) nos indica que no se produce hibridación de los oligonucleótidos correspondientes a ribozima e IEP de RmInt1 en *S. meliloti* RMO17, reafirmando el hecho de que esta cepa carece de copias de RmInt1 en su genoma. Sin embargo, cuando estudiamos los niveles de expresión de los transcritos derivados de RmInt1, no encontramos hibridación diferencial entre ambas cepas cuando lo esperable sería observar señal en *S. meliloti* 1021 pero no en RMO17. Este hecho se podría interpretar como falta de expresión de las 3 copias de intrón que *S. meliloti* 1021 tiene insertadas en su genoma.

#### **II.3.4. Genes diferencialmente expresados en *S. meliloti* RMO17 en presencia de intrón.**

Una vez determinados los genes que diferencian *S. meliloti* 1021 de RMO17, así como aquellos que presentan expresión diferencial entre ambas cepas, nos proponemos determinar la influencia de la presencia de RmInt1 en el genoma bacteriano. De manera similar al análisis que se ha descrito en *S. meliloti* 1021, se analiza el perfil transcriptómico de cepas de *S. meliloti* RMO17 con inserciones de intrón completo o sólo la ribozima frente a *S. meliloti* RMO17 control.

Obtuvimos señal en el 82-84,5% de los oligonucleótidos representados en el chip, pero solamente el 21-24% presentó un valor estadísticamente significativo (figura II.6 A). Se obtuvieron 27 genes diferencialmente expresados en presencia de intrón completo; el 50% de estas señales, se corresponde con secuencias intergénicas en su mayoría inducidas (figura II.6 B). Es importante reseñar que estas secuencias intergénicas inducidas se corresponden principalmente con la ribozima de RmInt1. 10 de las 14 ORFs identificadas se encontraban reprimidas mientras sólo 4 estaban inducidos. Como en el caso de las regiones intergénicas, 3 de esos genes inducidos se corresponden con la IEP de RmInt1 mientras el gen restante codifica un represor de la síntesis de flagelina. En cuanto a los genes reprimidos, identificamos un operón compuesto por SMB21036/37/38/39/40/41/42 que parece codificar un sistema transportador de membrana relacionado con procesamiento de oligopéptidos de la pared celular. A diferencia de lo que habíamos observado hasta el momento, la mayoría de los genes

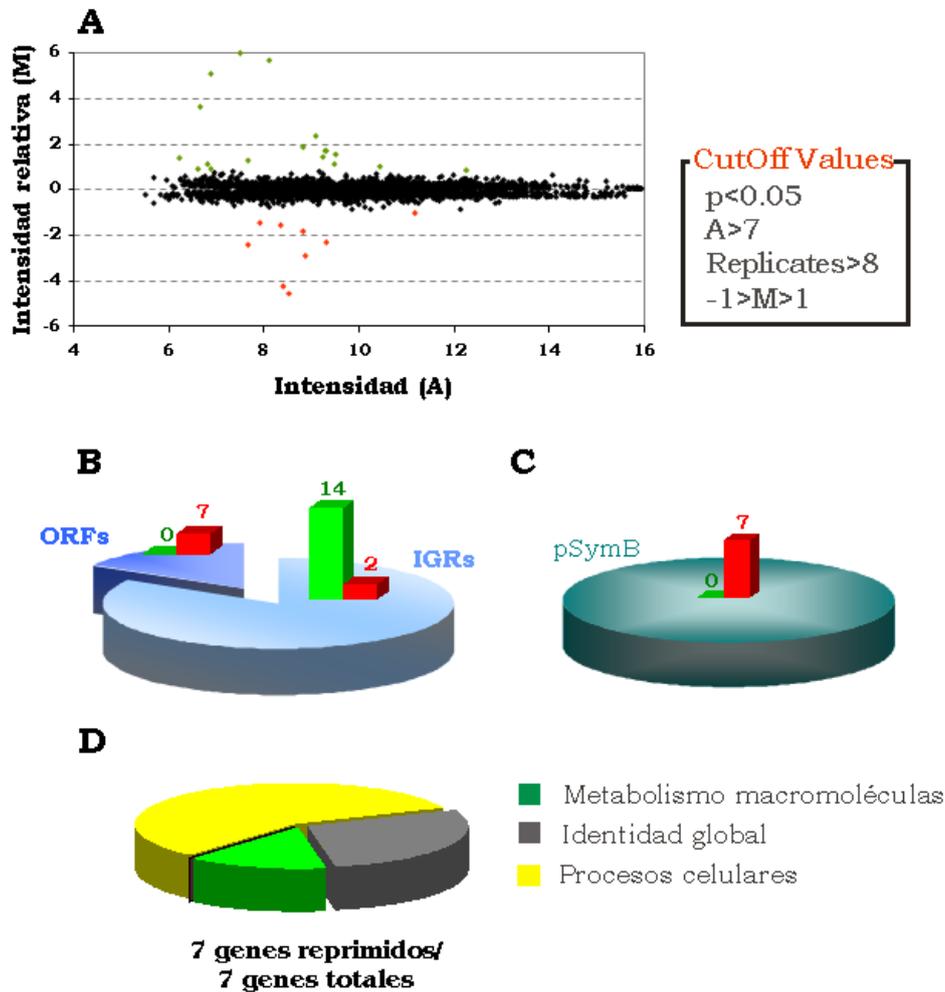
diferencialmente expresados se encuentran localizados en pSymB, 3 en el cromosoma y uno en pSymA (figura II.6 C).



**Figura II. 6.** Representación gráfica de los datos obtenidos en la hibridación de microarrays para *S. meliloti* RMO17 frente a RMO17 2,5 C1. A Diagrama de puntos que enfrenta la intensidad de la señal frente a la relación de las intensidades entre ambos fluoróforos de aquellos datos considerados estadísticamente significativos (incluye regiones intergénicas y secuencias codificantes). Las señales en rojo indican represión, mientras las representadas en verde se corresponden con inducción. B Distribución de las señales dependiendo si se encuentran en genes (ORFs) o en regiones intergénicas (IGRs). C Localización de los genes diferencialmente expresados en los distintos replicones de *S. meliloti*. D Representación gráfica de los genes identificados según las categorías funcionales definidas en KEGG y que se corresponden con la anotación de los genes en las bases de datos de *S. meliloti* 1021.

Cuando analizamos los datos derivados de la influencia de la presencia de la ribozima en el fondo genético estudiado (figura II.7), nos encontramos con que la mayoría de las señales correspondientes a expresión

diferencial se localizan en regiones intergénicas (16/23 genes). A su vez, estas señales reflejan mayoritariamente inducción y se corresponden con los oligonucleótidos que representan la ribozima (9 de las 14 señales de activación detectadas). Nuevamente volvemos a detectar la represión del operón SMb21036 a 42, siendo estos genes las únicas señales correspondientes a ORFs que observamos.



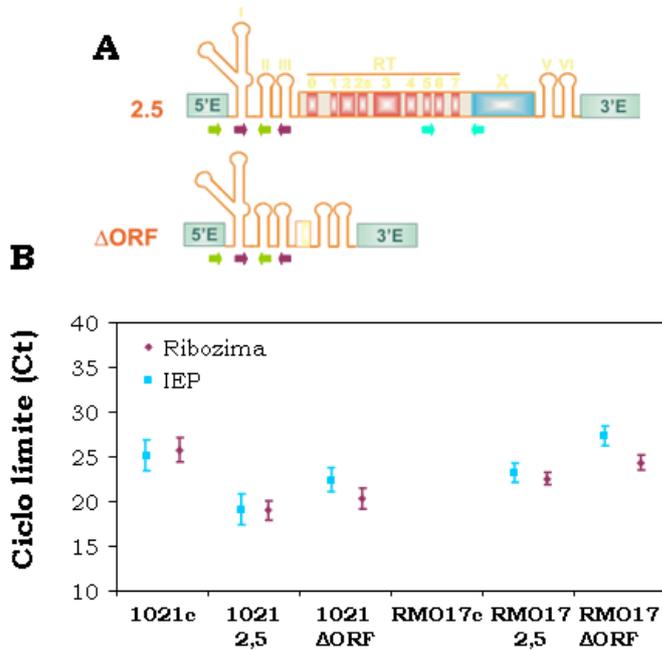
**Figura II. 7.** Representación gráfica de los datos obtenidos en la hibridación de microarrays para *S. meliloti* RMO17 frente a RMO17  $\Delta$ ORF C15. **A** Gráfica M/A correspondiente a las señales estadísticamente significativas incluyendo regiones intergénicas (IGRs) y secuencias codificantes (ORFs). Las secuencias destacadas en rojo se encuentran reprimidas, mientras las representadas en verde se corresponden a secuencias inducidas. **B** Diagrama donde se representa el número total de señales estadísticamente significativas según si se encuentran en ORFs o en IGRs. **C** Distribución de los genes por réplicas. **D** Categorías funcionales de los genes identificados según KEGG y las bases de datos de *S. meliloti* 1021.

### II.3.5. Verificación de los niveles de expresión de RmInt1 mediante RT-qPCR.

La técnica de hibridación en microarrays proporciona un elevado número de datos de naturaleza semicuantitativa. Así, se observa una diferencia en las señales de hibridación correspondientes a la ribozima y a la IEP dependiendo del oligonucleótido de que se trate. En el microarray se encuentran impresos 3 oligonucleótidos que se corresponden con la secuencia que codifica el dominio madurasa en la IEP. De los tres, el complementario a los 70 últimos nucleótidos que codifican la proteína suele dar la señal menos intensa; solapa 8 nucleótidos con el siguiente oligonucleótido que es, por el contrario, el que da una señal más intensa en todos los estudios llevados a cabo. Algo similar ocurre con los oligonucleótidos correspondientes a las regiones intergénicas donde se localizan las ribozimas de los diferentes intrones (que en realidad son iguales para las tres copias, por tanto en el array encontramos 9 oligonucleótidos que nos identifican la ribozima de RmInt1). Para conseguir determinar de una manera cuantitativa la expresión de los intrones en las distintas cepas utilizadas en este estudio, se utilizó RT-qPCR. Esta técnica nos va a permitir indirectamente validar los resultados obtenidos en la hibridación de microarrays.

Ribozima, IEP y pre-mRNA de RmInt1 fueron amplificados con distintas parejas de cebadores en todas las cepas descritas y utilizadas en este estudio (figura II.8 A). Como se esperaba, *S. meliloti* RMO17, que carece de copias de intrón en su genoma, genera amplificación inespecífica que se pone de manifiesto a partir del ciclo 40 tras la normalización, por lo que no aparece representada en la gráfica (figura II.8 B). En *S. meliloti* 1021 control (1021c en figura II.8 B), ambas parejas de cebadores comenzaron a amplificar alrededor del ciclo 25. Como cabría esperar en la cepa en que se inserta el intrón completo, ambas parejas de cebadores comienzan a amplificar 5 ciclos antes pero con la misma eficiencia; si sólo se inserta la ribozima se observa que la pareja de oligonucleótidos específica de esta región comienza antes la amplificación a diferencia de la que se corresponde con la IEP. Un patrón similar se observa en *S. meliloti* RMO17. Así pues la expresión de RmInt1 en una cepa con 3 copias de intrón se detecta en el ciclo 25, con 5 copias en el ciclo 23 y con 8 copias a partir del ciclo 19. Por tanto, podríamos decir que hay una variación en la cuantificación en función de las copias de RmInt1 presentes en el genoma: como promedio, por cada

intrón adicional insertado en el cromosoma, se reduce un ciclo el comienzo de la amplificación.



**Figura II. 8.** Evaluación de la expresión de RmInt1 en las cepas curadas. **A** Representación esquemática de las inserciones de RmInt1 como intrón completo (panel superior) o solamente la ribozima ( $\Delta$ ORF, panel inferior). Así mismo, se indica la localización de cada una de las parejas de cebadores utilizadas en la RT-qPCR: en verde, la pareja correspondiente al precursor (UP y P); en morado, los oligonucleótidos 3 e ICI que amplifican la ribozima; finalmente, en azul, los cebadores que hibridan en la IEP (21.0 y E17.0). **B** Los niveles de expresión de ribozima e IEP fueron evaluados en función de la cantidad de precursor (gen referencia) y relativizados a RMO17 control, cepa que carece de intrones tipo RmInt1 en su genoma. Cuanto menor es el ciclo límite, mayor es la cantidad de moléculas presente en las muestras.

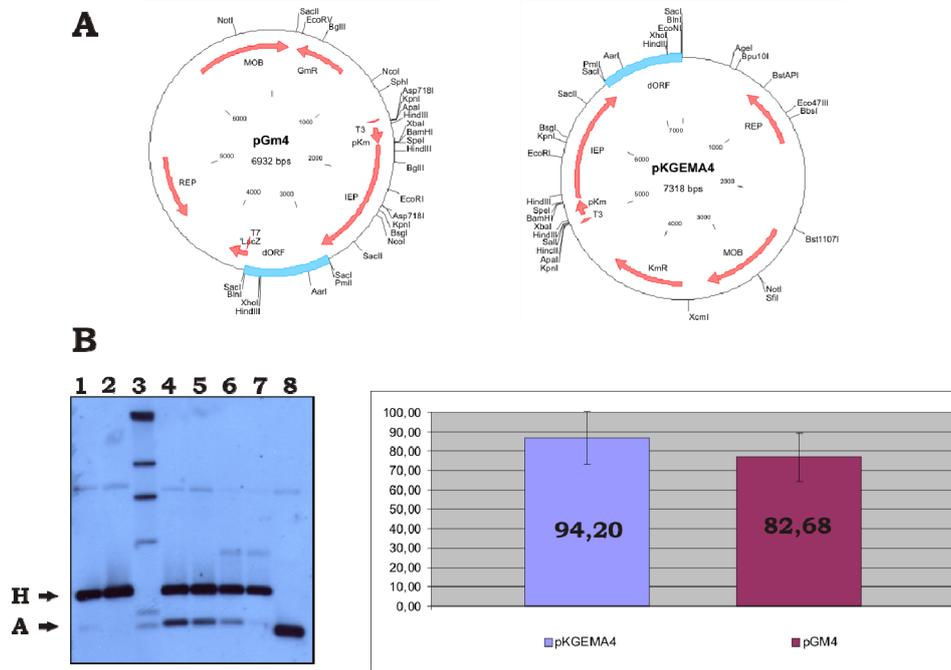
Los datos representados resultaron del análisis del RNA extraído a partir de 3 réplicas biológicas.

De los datos de DNA microarrays habíamos concluido que no se detectaba expresión de las copias genómicas presentes en *S. meliloti* 1021. Sin embargo, la cuantificación mediante RT-qPCR señala expresión diferencial entre esta cepa y *S. meliloti* RMO17, sin intrones en su genoma. Este dato nos indica que los intrones presentes en *S. meliloti* 1021 se están expresando pero quizá la hibridación en microarrays no es lo suficientemente sensible para detectarlo. No debemos olvidar que el PCR magnifica los resultados puesto que las diferencias se incrementan de manera exponencial conforme aumenta el número de ciclos. Nuestros resultados son una nueva evidencia de que la hibridación en microarrays no es una técnica cuantitativa.

### II.3.6. Movilidad de pGm4

Se analizó la movilidad del intrón contenido en el plásmido donador pGm4 en *S. meliloti* RMO17 mediante ensayo de doble plásmido y posterior *Southern blot*. Se determinó que no había diferencias significativas en cuanto a la eficiencia de invasión con el vector pKGEMA4, como se observa en la

figura II.9, por lo que se usó en los ensayos de movilidad llevados a cabo en los mutantes de 1021 y 2011.



**Figura II.9. A:** Esquema representando el plásmido pCm4 en comparación con pKGEMA4. **B:** *Southern blot* del DNA plasmídico obtenido del ensayo de movilidad de pCm4; Carriles: 1 y 2 pKGEMA4 x pJB0,6LAG; 3 Marcador de peso molecular II; 4, 5, 6 y 7 pCm4 x pJB0,6LAG; 8 pCm4 x pJBΔ129 (control negativo); H: banda de *homing*; A: banda del aceptor sin invadir. Histograma representando los valores del porcentaje de invasión obtenidos de la cuantificación del *Southern blot*.

### II.3.7. Ensayos de movilidad en mutantes de *S. meliloti* 1021 y 2011

Los ensayos de movilidad se llevaron a cabo mediante un ensayo de doble plásmido. Estos ensayos se realizaron en las cepas silvestres de *S. meliloti* 1021 y 2011 y en los mutantes obtenidos de la colección existente en la Universidad de Bielefeld. Los resultados de *homing* obtenidos en los mutantes, fueron relativizados a los obtenidos en las cepas silvestres para evitar los posibles errores metodológicos (se consideró el valor de *homing* en el silvestre como 100%). De algunos de los mutantes enumerados en la tabla

II.2 no se pudieron obtener resultados por no conseguir movilizar a su interior los plásmidos necesarios para el ensayo de movilidad. Los resultados de movilidad obtenidos aparecen en las siguientes tablas.

En la mayoría de los mutantes analizados, el nivel de invasión no se ha visto afectado o lo ha sido en un bajo porcentaje. En otros casos, dependiendo del mutante analizado dentro del mismo gen, se han visto tanto aumento como disminución pronunciados del *homing*.

### II.3.7.1. Ligasas

Los mutantes defectivos en algunas de las funciones DNA ligasa sólo han mostrado variación al alza de un 20-30% en el caso de los locus SMb21044 y SMb20686 (mutante 3.10.E09) al igual que el mutante en la LigT (2'-5' RNA ligasa). En el caso de los demás mutantes, no se han obtenido diferencias significativas (Tabla II.3).

**Tabla II.3.** *Homing* en mutantes defectivos en distintas locus con función ligasa atribuida.

Gen mutado	Nombre del mutante	Función probable	% de inv. rel. al WT
<b>SMA0424</b>	2011mTn5STM.1.01.G09	putativa DNA ligasa	103
<b>SMb21044</b>	2011mTn5STM.3.02.F10	putativa ATP-dep.DNA ligasa	129
<b>SMb20912</b>	2011mTn5STM.3.10.E09	putativa ATP-dep.DNA ligasa	105
<b>SMb20686</b>	SmPI_1021.12.08C2	putativa ATP-dep.DNA ligasa	121
	2011mTn5STM.2.11.C07		100
<b>SMc03835</b>	2011mTn5STM.1.07.E12	2'-5' RNA ligasa ( <b>ligT</b> )	110
	2011mTn5STM.2.07.C01		128

### II.3.7.2. Helicasas

De los genes con función helicasa (DNA, RNA o híbridos RNA-DNA) mutados, se ha visto que en dos de ellos no hay efecto (SMc02237 y SMc01090-1.06.H02); en otros se observa un aumento de la eficiencia de *homing* de entre un 20-30% (SMc00228, SMc01461-01H5, SMc01090-3.04.B07); entre los mutantes en los que se observa una reducción de la invasión existen cuatro niveles, leve (SMc02586, 85%), moderada (SMc02152, 65%), acusada (SMc01566, 26%) y drástica (SMc01461-2.03.F05, 0%) (Tabla II.4).

**Tabla II.4.** *Homing* en mutantes defectivos en distintas locus con función helicasa atribuida.

Gen mutado	Nombre del mutante	Función probable	% de inv. rel. al WT
<b>SMc00228</b>	2011mTn5STM.3.10.D01	3'-5' helicasa ( <b>RecG</b> ), ATP-dep.	120
<b>SMc01461</b>	2011mTn5STM.2.03.F05 SmPI_1021.12.01H5	3'-5' helicasa de híbridos RNA/DNA ( <b>UvrD1</b> )	0 131
<b>SMc01566</b>	2011mTn5STM.3.02.G02	3'-5' DNA helicasa ( <b>RecQ</b> ), ATP-dep.	26
<b>SMc02586</b>	2011mTn5STM.2.03.B09	ATP-dep. helicasa ( <b>helO</b> )	85
<b>SMc01090</b>	2011mTn5STM.1.06.H02 2011mTn5STM.3.04.B07	ATP-dep. RNA helicasa ( <b>deaD</b> )	93 125
<b>SMc02152</b>	2011mTn5STM.1.04.D01	helicasa	65
<b>SMc02237</b>	2011mTn5STM.1.06.F08	ATP-dep. helicasa	109

### II.3.7.3. Proteínas reparadoras del DNA

Cuando analizamos genes mutados con función reparadora de DNA también observamos resultados heterogéneos, desde aquellos con una reducción importante en el porcentaje de invasión (SMc01877, 59%) hasta aquellos en los que el efecto es al contrario, aumentando el *homing* en un 50% (SMc00556-2.02.C03, 150% y -4.04.E12, 148%), pasando por mutantes en los que la variación es muy leve o casi inapreciable (Tabla II.5).

**Tabla II.5.** *Homing* en mutantes defectivos en distintas locus con función de reparación de DNA.

Gen mutado	Nombre del mutante	Función probable	% de inv. rel. al WT
<b>SMc00556</b>	2011mTn5STM.2.02.C03 2011mTn5STM.4.04.E12 2011mTn5STM.1.14.F08	Proteína reparadora de DNA ( <b>radA</b> )	150 148 109
<b>SMc01235</b>	2011mTn5STM.4.03.A06 2011mTn5STM.1.09.G09	excinucleasa ABC subunidad A (reparación DNA ATP-dep.) <b>uvrA</b> ( <b>OPERÓN CON SMc01236</b> )	114 81
<b>SMc01236</b>	2011mTn5STM.2.12.B04	Proteína hipotética conservada	123
<b>SMc01877</b>	2011mTn5STM.1.02.G07	Proteína de reparación del DNA ( <b>RecN</b> )	59

### II.3.7.4. DNAs Polimerasas y RNAs polimerasas

Sólo se han podido analizar mutantes en las DNA Polimerasas III y IV, ya que en las demás (Pol I, II y V) no existían mutantes disponibles. Existe una gran heterogeneidad entre diferentes mutantes del locus SMc03788

(DNA PolIII), obteniendo tanto inhibición total de la invasión como un aumento de ésta hasta en un 75%. En el caso de la PolIV, los dos mutantes analizados han demostrado que una mutación en este gen produce una inhibición del *homing*. En cuanto a la RNA polimerasa, al analizar mutantes en los factores sigma, se ha visto que en el mutante SMc01419-5.08.D04 (*rpoE1*) la eficiencia de invasión aumenta hasta en un 80%, mientras que en otros factores sigma, la variación es mucho menor, tanto positiva como negativamente (Tabla II.6).

**Tabla II.6.** *Homing* en cepas mutantes en las DNA y RNA polimerasas.

Gen mutado	Nombre del mutante	Función probable	% de inv. rel. al WT
<b>SMc03788</b> <i>dnaE2</i>	2011mTn5STM.5.12.C02	Principal enzima replicativa ( <b>DNA Pol III</b> )	106
	2011mTn5STM.2.04.G08		91
	2011mTn5STM.3.02.F01		95
	2011mTn5STM.4.13.F08		0
	2011mTn5STM.3.09.F02		111
	2011mTn5STM.4.02.F11		2
	2011mTn5STM.1.02.B04		32
	SmPI_1021.12.03H10		173
<b>SMA0892</b> <i>dnaE3</i>	2011mTn5STM.3.11.G05		175
	2011mTn5STM.2.02.E09		49
	2011mTn5STM.5.10.E04		78
	2011mTn5STM.3.01.H03		0
	2011mTn5STM.5.07.D02		171
	2011mTn5STM.3.11.A08		141
	2011mTn5STM.1.14.F07		0
<b>SMc01373</b>	2011mTn5STM.1.02.D07	DNA pol IV 1, extiende horquilla parada ( <i>dinB1</i> )	0
	2011mTn5STM.4.10.F04		0
<b>SMA2355</b>	2011mTn5STM.4.05.D10	Nucleotidyltransferasa/DNA polimerasa	29
<b>SMB21448</b>	2011mTn5STM.3.11.F08	putativa DNA polymerase	95
<b>SMc01419</b>	2011mTn5STM.5.08.D04	RNA polimerasa sigma factor ( <b>rpoE1</b> )	179
	2011mTn5STM.3.07.G07		120
<b>SMc01506</b>	2011mTn5STM.1.04.C05	RNA polimerasa sigma-E factor (sigma-24) ( <b>rpoE2</b> )	97
	2011mTn5STM.3.12.B10		107
<b>SMc04051</b>	2011mTn5STM.1.13.E03	RNA polimerasa sigma-E factor (sigma-24) ( <b>rpoE4</b> )	79
<b>SMc03873</b>	2011mTn5STM.4.09.D11	RNA polimerasa sigma factor ( <b>rpoH2 sigC</b> )	98

## II.3.7.5. Nucleasas

De los mutantes analizados, existen tres de ellos en los que el nivel de invasión se ve reducido drásticamente, a un 3% en el caso de SMc02760 (nucleasa de DNA de cadena simple) y hasta desaparecer en el caso de SMb20689 (exonucleasa de DNA de doble cadena) y en el mutante en SMa0934 (nucleasa de DNA de cadena simple). En otros mutantes, la reducción de *homing* es mucho menor (entre un 20% y un 40%) como en el caso de los genes SMc02327 (RNAsa H1), SMc00378 (*xseA*) y SMc00606.

Existen genes que al ser mutados se aprecia un aumento en el nivel de invasión (30%-50%), es el caso de los genes SMc01336 (RNAsa E), SMc00867 (RNAsa H2) y SMc00622 (*rnd1*).

Los niveles de variación en los demás mutantes en nucleasas analizados varían sólo en un 10% (positiva o negativamente), por lo que no serán considerados (Tabla II.7).

**Tabla II.7.** Niveles de invasión en mutantes de genes que codifican nucleasas.

Gen mutado	Nombre del mutante	Función probable	% de inv. rel. al WT
<b>SMc01336</b>	2011mTn5STM.4.07.G10	procesamiento y eliminación del RNA, <b>RNaseE</b> ( <i>rne</i> )	143
<b>SMc02327</b>	2011mTn5STM.1.13.B02	<i>mhA2</i> , degradación del RNA de híbridos RNA/DNA ( <b>RNase H1</b> )	67
<b>SMc00867</b>	SmPI_1021.12.04C9	<i>mhB</i> , degradación del RNA de híbridos RNA/DNA ( <b>RNase H2</b> )	152
<b>SMc02760</b>	2011mTn5STM.4.10.G03	ssDNA exonucleasa ATP-dep. <b>RecB</b> )	3
<b>SMb20689</b>	2011mTn5STM.1.13.G07 SmPI_1021.12.08C5	3'-5'dsDNA exonucleasa ( <b>XthA4</b> )	0 0
<b>SMa0934</b>	SmPI_1021.12.06E7	TraA1 proteína de transfer. conjugativa identidad con ssDNA exonucleasa ( <b>RecD</b> )	0
<b>SMc02843</b>	SmPI_1021.12.01A5 2011mTn5STM.1.02.E08	endonucleasa III ( <b>nth</b> )	97 112
<b>SMc00378</b>	2011mTn5STM.3.06.H09	exodeoxyribonucleasa ( <b>xseA</b> )	61
<b>SMc00606</b>	2011mTn5STM.4.09.C03	ribonucleasa	82
<b>SMc01929</b>	2011mTn5STM.3.10.C03 2011mTn5STM.3.13.E11	Proteína hipotética ( <b>Ribonucleasa J</b> )	96 102
<b>SMc00622</b>	2011mTn5STM.4.07.H06	ribonucleasa D ( <b>rnd1</b> )	133
<b>SMc04231</b>	2011mTn5STM.2.08.A01 2011mTn5STM.4.03.D11	excinucleasa ABC subunidad B ( <b>uvrB</b> )	116 92

## II.3.7.6. Chaperonas

El efecto de los mutantes en algunas de las chaperonas existentes sobre el nivel de invasión de RmInt1 no es excesivamente importante. La mayor variación se ha observado en el mutante del gen SMc02857 (*dnaK*), en el cual el *homing* aumenta en un 50%. También se ve aumentada en un 20%-30% en el caso de mutantes en los genes SMc04456 (*csaA*) y SMb21183 (*htpG*). El nivel mayor de reducción de la invasión en un mutante se ha observado en el caso del gen SMc00003 (entre un 20% y un 30% dependiendo del mutante que se analice) tratándose de una chaperona no caracterizada, al igual de lo que ocurre con los mutantes del gen SMc01758 (*GroEL4*) en los que los niveles o no se alteran o lo hacen en niveles entre un 10% y un 20%. Los valores obtenidos de los demás mutantes no fueron relevantes (Tabla II.8).

**Tabla II.8.** Efecto de mutaciones en chaperonas sobre la movilidad de RmInt1.

Gen mutado	Nombre del mutante	Función probable	% de inv. rel. al WT
<b>SMc02857</b>	2011mTn5STM.3.05.F08	Proteína de choque térmico 70 (HSP70) chaperona ( <b>dnaK</b> )	149
<b>SMc00003</b>	SmPI_1021.12.01B9	chaperona	69
	2011mTn5STM.3.04.G09		81
<b>SMc00913</b>	2011mTn5STM.1.08.G04	60 KD chaperonina A ( <b>groEL1</b> )	111
	SmPI_1021.12.04B10		95
	2011mTn5STM.2.08.H06		112
<b>SMc01758</b>	2011mTn5STM.3.03.B05	60 KD chaperonina B ( <b>groEL4</b> )	82
	2011mTn5STM.2.05.F01		100
	2011mTn5STM.1.03.G02		89
<b>SMc04456</b>	2011mTn5STM.2.09.D09	chaperona ( <b>csaA</b> )	123
<b>SMA0124</b>	2011mTn5STM.1.04.A01	<b>GroEL3</b> chaperonina	106
	2011mTn5STM.1.05.H11		94
<b>SMb21183</b>	2011mTn5STM.2.12.G12	probable chaperonina, familia de las proteínas de choque térmico hsp90( <b>htpG</b> )	129

## II.3.7.7. Mutantes en otras funciones

En este apartado se ha estudiado genes con otras funciones, y que según la bibliografía intervienen en la actividad de otros intrones del grupo II, por lo que quizá también influyan sobre RmInt1. Entre las funciones estudiadas hay proteínas relacionadas con los RNAs de transferencia, transportadores iónicos de membrana, enzimas del metabolismo general y algunos posibles factores de transcripción entre otros. En general, no se ha

visto influencia de las mutaciones en estos genes sobre la movilidad de RmInt1, a excepción de algunos mutantes. En los dos mutantes del gen *corA1* (SMc00399: proteína transmembrana transportadora de iones magnesio y cobalto) el nivel de invasión se ve aumentado en un 50%. También aumenta, aunque en menor nivel (35%) el mutante en la gliceraldehido-3-fosfato deshidrogenasa (*gap*: SMc03979), mientras que los mutantes en el resto de genes que forman el operón el comportamiento es diferente, disminuyendo la movilidad en un 20%-25% en dos de ellos, SMc03981 y SMc03978 (una probable fosfoglicerato kinasa y una probable transcetolasa). La mayor disminución en el nivel de invasión se produce en el mutante en la proteína transportadora de cationes *chaC* SMc00713 (1.11.B04), en más de un 40%, seguido del mutante en una peptidil tRNA hidrolasa (SMc02693) en el que se reduce en casi un 35%.

**Tabla II.9.** Homing en mutantes defectivos en funciones generales.

Gen mutado	Nombre del mutante	Función probable	% de inv. rel. al WT
<b>SMc00321</b>	2011mTn5STM.4.10.C12	tRNA pseudouridine sintasa B <b>truB</b>	104
	2011mTn5STM.2.08.G08		127
	2011mTn5STM.4.11.B07		113
<b>SMc02693</b>	2011mTn5STM.2.07.H02	peptidyl-tRNA hidrolasa ( <b>pth</b> )	67
<b>SMc00713</b>	2011mTn5STM.1.11.B04	Transporte de cationes ( <b>chaC</b> )	59
	SmPI_1021.12.05E3		93
<b>SMc00399</b>	2011mTn5STM.2.03.D08	Proteína transmembrana de transporte de Mg. y Co. ( <b>corA1</b> )	145
	2011mTn5STM.5.03.D01		158
<b>SMc00874</b>	2011mTn5STM.2.08.F04	Proteína transmembrana de transporte de Mg. y Co. ( <b>corA2</b> )	105
<b>SMc01023</b>	2011mTn5STM.4.08.B01	triosephosphate isomerase ( <b>tpiA1</b> )	98
<b>SMc03979</b>	2011mTn5STM.2.11.H09	glyceraldehyde 3-phosphate dehydrogenase ( <b>gap</b> ) <b>OPERON gap, pgk, tkt2 y SMc03970</b>	134
<b>SMc03981</b>	2011mTn5STM.1.09.G04	Probable phosphoglycerate kinase ( <b>pgk</b> )	76
	2011mTn5STM.3.08.F07		117
<b>SMc03978</b>	2011mTn5STM.1.12.D08	Probable transketolase ( <b>tkt2</b> )	79
<b>SMc02163</b>	2011mTn5STM.4.12.B11	glucose-6-phosphate isomerase ( <b>pgi</b> )	103
<b>SMb21579</b>	2011mTn5STM.4.11.A09	Regulador transcripcional putativo, familia de la proteína merR ( <b>hmrR2</b> )	82
	SmPI_1021.12.07H8		59
<b>SMc00109</b>	2011mTn5STM.1.05.G08	Probable regulador transcripcional	125
	2011mTn5STM.3.07.E03		121
<b>SMc02659</b>	2011mTn5STM.2.12.C04	Putative GTP pyrophosphokinase (ATP:GTP 3'-pyrophosphotransferase) ( <b>relA</b> )	87

<b>SMc02496</b>	SmPI_1021.12.03F10	Probable primosomal protein N' (replication factor Y) <b>priA</b>	121
<b>SMc00339</b>	2011mTn5STM.4.08.H11 2011mTn5STM.1.14.A01	Adenylate cyclase 1 protein ( <b>cyaA</b> )	95 71
<b>SMb21488</b>	2011mTn5STM.4.03.B02 2011mTn5STM.4.06.A01	putative cytochrome o ubiquinol oxidase chain I protein ( <b>cyoB</b> )	89 123

## II.4. DISCUSIÓN

### II.4.1. Análisis genómico y transcriptómico por microarrays.

Se ha analizado el perfil genómico y transcriptómico de cepas de *S.meliloti* 1021 y RMO17 que presentan diferente número de secuencias de inserción interrumpidas por RmInt1 en su versión completa, ribozima junto con IEP, como en su versión modificada (IEP delecionada del dominio IV de la ribozima). En general, observamos una mayor incidencia de genes reprimidos que se localizan en el cromosoma en el caso de 1021 y en pSymB, en RMO17. En los chips se presentan una serie de oligonucleótidos que se corresponden con la ribozima (comprendidos dentro de las regiones intergénicas) y tres, concretamente, que hibridan en la secuencia que codifica la IEP (representando las tres copias de intrón que *S.meliloti* 1021 contiene en su genoma). Nuestros resultados muestran una perfecta correlación entre el patrón de hibridación y el tipo de intrón insertado: cuando se produce la integración de intrón completo observamos hibridación en ambos grupos de oligonucleótidos; sin embargo, si sólo se inserta la ribozima, encontramos señal únicamente en los primeros. Finalmente, hemos comprobado los niveles de expresión de RmInt1 en las diferentes cepas, llegando a la conclusión de que la expresión depende del número de copias insertadas. En cuanto al perfil genómico, hemos comprobado que la única modificación que genera la presencia de intrones en *S.meliloti* 1021 es un incremento en las señales de hibridación en aquellos oligonucleótidos que se corresponden con RmInt1, ribozima y/o IEP dependiendo de la naturaleza de intrón insertado. Si atendemos a las diferencias entre *S.meliloti* 1021 y RMO17, el ~7,5% de las secuencias que codifican proteínas presenta un patrón diferencial de hibridación entre ambas cepas, lo que hace factible el uso de los DNA microarrays diseñados en *S.meliloti* 1021 para los estudios en RMO17.

Uno de los primeros aspectos que revelan nuestros estudios es que parece importante el diseño técnico de los experimentos: el número de genes que aparecen diferencialmente expresados en *S.meliloti* 1021 es considerablemente superior al que se pone de manifiesto en el caso de RMO17. En este sentido podemos aducir dos argumentos no excluyentes: por un lado, podemos pensar que las diferencias genómicas que hemos puesto de manifiesto entre las dos cepas podrían estar restando genes en nuestros análisis de expresión; por otro lado, podría estar influenciado por el hecho de que se hayan intercambiado los marcajes de las muestras referencia (cepa silvestre) y problema (cepas con inserciones) en el caso de *S.meliloti* RMO17 pero no en 1021. Se han descrito diferencias en cuanto a apantallamiento de señales entre Cy3 y Cy5; mientras Cy3 es un fluoróforo bastante estable y que proporciona señales constantes en el tiempo, Cy5 se describe como un fluoróforo muy sensible a la luz, mostrando diferencias de intensidad, por ejemplo, en dos escaneados consecutivos (información proporcionada por la casa comercial, GE healthcare). Por razones que desconocemos, el escáner utilizado para el revelado de las imágenes de microarrays nos obligaba a incrementar diferencialmente el balance de excitación del láser a 635 nm, canal de excitación del Cy5, frente a 532 nm, que se corresponde con la intensidad de excitación del fluoróforo verde. Este hecho puede provocar un cierto sesgo en los datos que se corrige cuando se intercambian los marcajes y se analizan los datos conjuntamente (la normalización y el posterior tratamiento estadístico eliminarían esos posibles sesgos, dando lugar a resultados más fiables y ajustados a la realidad). Así, pensamos que el intercambio de marcajes en las réplicas técnicas en el caso de *S.meliloti* RMO17 es la razón más importante por la que se reduce el número de genes diferencialmente expresados respecto a 1021.

Consecuencia de todo lo anteriormente expuesto podría ser el hecho de que al comparar los resultados obtenidos en *S.meliloti* 1021 y RMO17, los únicos genes de los que obtenemos un patrón común son los correspondientes a la IEP de RmInt1 (figura II.10). Este mismo análisis es

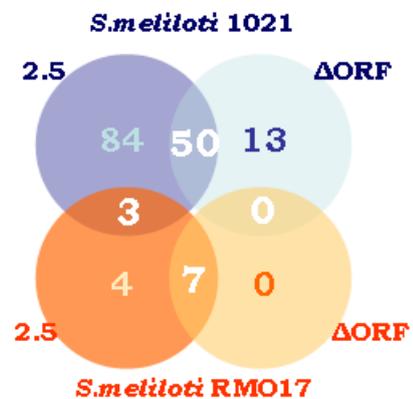


Figura II. 10 Diagrama de intersecciones de las secuencias que corresponden a genes.

válido para el caso de las regiones intergénicas correspondientes con la ribozima (dato no mostrado).

En cuanto a los genes diferencialmente expresados compartidos en el análisis de *S.meliloti* 1021 con inserciones de intrón completo y  $\Delta$ ORF (figura II.10), encontramos una gran variedad de transportadores de membrana. La relación entre la represión de estos genes y la expresión de intrones en el genoma bacteriano, no parece evidente. En cualquier caso, lo esperable sería la inducción de algunos transportadores de iones divalentes como  $Mg^{2+}$ , cuya importancia para el correcto plegamiento y funcionalidad de intrones del grupo II ha sido demostrada sobradamente. Sin embargo, lejos de esta asunción, lo que observamos es represión. De manera similar, nuestros resultados también ponen de manifiesto la represión que tiene lugar sobre reguladores del proceso de fijación de nitrógeno, concretamente sobre los genes *nifA* y *fixK/T*. Así mismo, observamos la inducción de genes implicados en el metabolismo de glicina (*gcvTHP*) y en la formación de flagelos. Es difícil establecer una relación entre la funcionalidad de los genes detectados y la presencia de intrones aunque podemos afirmar que, en ningún, caso estos genes se encuentran cercanos a las inserciones localizadas en el genoma bacteriano, por lo que descartamos que la inserción de RmInt1 provoque un efecto regulador posicional.

Si atendemos a los 7 genes reprimidos compartidos en las diferentes hibridaciones en *S.meliloti* RMO17 (figura II.10), se corresponden con un operón que codifica el conjunto de componentes de un transportador ABC de péptidos, amina y aminoácidos. El análisis genómico apunta a que estaría presente en *S.meliloti* RMO17 puesto que, en la mayoría de los casos, no se observa hibridación diferencial en los datos de CGH. Sin embargo, es importante destacar que cuando estudiamos la expresión en *S.meliloti* 1021 frente a RMO17 sin copias de intrón, estos genes parecen encontrarse inducidos (ver tabla II.10). Por el contrario, si analizamos la cepa en presencia de intrones pasan a estar reprimidos. En este caso, la explicación más sencilla podría ser la existencia de algún efecto insercional del intrón. En *S.meliloti* 1021, se ha descrito que el intrón identificado en la secuencia de inserción *ISRm10-1* se localiza ~13 Kbp aguas abajo de este operón. Esto no significa que en *S.meliloti* RMO17 su localización sea la misma; de hecho, una represión debida a algún tipo de polaridad en la inserción debería generarse si el intrón se encontrase aguas arriba, por ejemplo en la región promotora. En este sentido, es interesante destacar que en la región 5'UTR de este operón se ha predicho, mediante el uso de herramientas

bioinformáticas, un RNA antisentido que podría regular la expresión del operón. Generalmente, este tipo de reguladores son secuencias altamente estructuradas que interfieren en la traducción de los genes mediante el apareamiento de bases con la región donde se encuentra la secuencia Shine-Dalgarno. Desconocemos si podría existir algún tipo de interacción entre la región 5'UTR del operón, el supuesto RNA antisentido y el intrón; esta interacción tendría lugar específicamente en RMO17, probablemente como consecuencia de algún factor presente en esta cepa que no se encuentra en 1021.

**Tabla II.10** Datos de expresión diferencial del operón SMb21036/37/38/39/40/41/42. Sólo se muestran aquellos que son estadísticamente significativos.

<b>Reporter Name</b>	<b>1021xRMO17</b>	<b>RMO17xWT</b>	<b>RMO17xΔORF</b>
SMb21036	3,794839729	-3,683675271	-4,224499534
SMb21037	1,411857742	-2,571815938	-2,290064408
SMb21038		-2,011505641	-1,822960697
SMb21039	1,121555351	-2,412707667	-2,894115207
SMb21040		-1,242383061	-1,539440025
SMb21041		-1,162046178	-1,428404276
SMb21042		-2,249645525	-2,426527924

#### **II.4.2. Análisis de la movilidad de RmInt1 en mutantes de *S. meliloti***

Por otro lado, se han analizado más de 100 mutantes en ORFs con función conocida o predicha por identidad de secuencia aminoacídica. En la mayor parte de los casos, no se han producido grandes efectos sobre la movilidad de RmInt1. Consideraremos que la mutación a tenido un efecto sobre la movilidad cuando la variación haya supuesto un aumento o disminución igual o mayor al 50%, en relación con la movilidad de RmInt1 en las cepas silvestres. Efectos menores pueden indicar varias cosas: *i*) que la mutación no tiene influencia sobre el gen; *ii*) que realmente el gen no participa en la movilidad de RmInt1; *iii*) que existen otros factores que pueden realizar la misma función.

Genes que favorecen el *homing* (la movilidad en los mutantes correspondientes se ha visto reducida) son: SMc01461-2.03.F05 (UvrD1: helicasa 3'-5' de híbridos RNA/DNA); SMc03788-4.13.F08, -4.02.F11 y -1.02.B04 (dnaE2: principal enzima replicativa DNA Pol III); SMA0892-

2.02.E09, -3.01.H03 y -1.14.F07 (dnaE3: principal enzima replicativa DNA Pol III); SMc01373-1.02.D07 y -4.10.F04 (dinB1: probable DNA polimerasa IV 1); Sma2355 (dinB2: DNA polimerasa IV); SMc02760 (RecB: ssDNA exonucleasa ATP-dependiente); Smb20689 (XthA4: 3'-5' dsDNA exonucleasa y RNasa H); Sma0934 (TraA1: proteína de transferencia conjugativa que presenta una cierta identidad con RecD: ssDNA exonucleasa).

Los genes analizados que reprimen el *homing* de RmInt1 (cuando se analiza el mutante, el porcentaje de invasión aumenta) son: SMc00556-2.02.C03 y -4.04.E12 (radA: proteína reparadora de DNA); SMc03788-03H10 (DnaE2); Sma0892-3.11.G05 y -5.07.D02 (DnaE3); SMc01419-5.08.D04 (RpoE1: factor sigma de la RNA polimerasa); SMc01336 (RNasaE procesamiento y eliminación del RNA); SMc00867 (RNasa H2: degradación del RNA de híbridos RNA/DNA; degradación de los fragmentos de Okazaki); SMc02857 (DnaK: chaperona); SMc00399 (corA1; proteína transmembrana de transporte de iones magnesio y cobalto).

Nuestros resultados nos hacen pensar que, aunque puedan existir otros factores del hospedador que intervengan en el proceso de movilidad de RmInt1 y que debido a las múltiples funciones proteicas o efectos indirectos la interpretación de los resultados pueda ser complicada, cuando integramos estos resultados en el proceso global de invasión de una diana se hace patente la necesidad de determinadas funciones celulares. En la figura II.11 se muestra una representación esquemática de la posible participación en el proceso de *retrohoming* de las proteínas estudiadas.

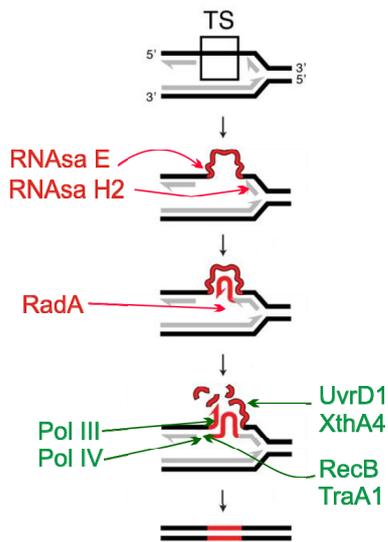
#### II.4.2.1 Efectos sobre las etapas iniciales del *homing*

El proceso de invasión se puede ver impedido ya desde la transcripción, por el factor sigma de la RNA polimerasa (rpoE1), cuya función es reconocer las regiones promotoras para iniciar la transcripción de DNA a RNA. Se cree que la expresión de RmInt1 se realiza a partir de promotores situados aguas arriba, como por ejemplo el propio promotor de la transposasa de la secuencia de inserción *ISRm2011-2*, que también es un elemento genético móvil. Si rpoE1 estuviese regulando de alguna manera la transcripción de la transposasa de la IS (para evitar reordenamientos génicos), podría afectar colateralmente a la expresión de las copias de RmInt1 que se hayan insertado en ella, por lo que una mutación en este factor haría menos estrictos los requerimientos de reconocimiento para el inicio de la transcripción, permitiendo la expresión de RmInt1.

Otro factor que puede jugar un papel importante es la RNasa E degradando el RNA del intrón antes del *splicing* reverso o de la síntesis del cDNA, impidiendo por tanto los pasos posteriores en la invasión.

La RNasa H2 juega un papel muy importante, ya que degrada los fragmentos de Okazaki, por lo tanto podría actuar eliminando los cebadores usados por la RT de RmInt1 para la síntesis del cDNA, impidiendo que se complete el *retrohoming*, aunque por ahora desconocemos si son los fragmentos de Okazaki o los cebadores de RNA los usados por la RT de RmInt1.

Las proteínas con actividad desoxirribonucleasa parecen favorecer el proceso de invasión, como es el caso de RecB, TraA1. Éstas podrían participar, mediante su actividad 5'-3' exonucleasa, recortando o desapareando el exón 5' para permitir que el cDNA que se ha sintetizado pueda aparear con la cadena de arriba.



**Figura II.11.** Esquema representativo de la posible participación en el *retrohoming* de RmInt1 en *S. meliloti* de los genes estudiados. En verde aparecen las proteínas facilitadoras, mientras que en rojo se muestran las proteínas que perjudican la movilidad.

#### II.4.2.2. Degradación del RNA del intrón tras la transcripción reversa

Después de la transcripción reversa, el RNA del intrón debe ser eliminado para permitir la síntesis de la segunda cadena, usando el cDNA como molde. Los intrones del grupo II, a diferencia de los retrovirus y de los retrotransposones LTR, no codifican sus propias RNAsas H (Malik y Eickbush, 2001). De nuestros resultados podemos deducir que UvrD1 es necesaria en este paso, ya que podría facilitar la degradación del RNA gracias a su actividad helicasa de híbridos RNA-DNA, mientras que Xth4, gracias a su actividad RNasa H (Mol *et al.*, 1995), podría eliminar el RNA del intrón.

#### II.4.2.3 Papel de las DNA polimerasas

Se han estudiado de forma independiente mutantes en dos genes que codifican la cadena alfa de la Polimerasa III, *dnaE2* y *dnaE3*. Dependiendo del mutante analizado, se han obtenido resultados antagónicos, aunque hay que resaltar que en determinados

mutantes los niveles de invasión se hacían indetectables. Los resultados de inhibición del *homing* son explicables debido a que Pol III participaría en la síntesis de la segunda cadena de DNA, usando como molde el cDNA sintetizado por la RT de RmInt1, al igual que ocurre en el caso del intrón de *L. lactis* Ll.LtrB (Smith *et al.*, 2005). Los resultados que indican que Pol III podría inhibir el *retrohoming* podrían deberse a efectos colaterales producidos por la inserción del minitransposón Tn5 o del plásmido recombinante usados para la generación de los mutantes.

El caso de la polimerasa Pol IV es más claro debido a su función reparadora (Kuban *et al.*, 2004), ejerciendo un efecto positivo sobre el *retrohoming* de RmInt1. El producto de los genes *dinB1* y *dinB2* actuaría reparando las uniones entre el DNA que ha sido invadido y el DNA que se ha sintetizado a partir del intrón que se ha insertado, concluyendo así el proceso de invasión.

#### II.4.2.4. Otras proteínas

RadA es una proteína que ayuda a desbloquear horquillas de replicación. Se cree que es una proteasa ATP-dependiente involucrada en la reparación del DNA y en la degradación de proteínas, péptidos y glicopéptidos (Beam *et al.*, 2001). Esta proteína podría actuar mientras se está produciendo el *retrohoming* en el seno de la horquilla de replicación, el cual puede estar causando impedimentos estéricos debido al tamaño de las RNPs formadas por la IEP y el RNA del intrón. RadA ayudaría a eliminar las RNPs de la horquilla de replicación dificultando el proceso de invasión.

Otra de las proteínas que interfieren en el *homing* de RmInt1 es DnaK. Es una chaperona que participa junto con DnaJ y GrpE (Schröder *et al.*, 1993) en el plegamiento de proteínas, por lo que, una posible explicación a que la movilidad de RmInt1 se vea reducida en su presencia es que interactúe con la IEP, interfiriendo en su plegamiento y por lo tanto en su actividad.

Por último, cuando se muta CorA (proteína de transporte de magnesio y cobalto; Wang *et al.*, 2006) se observa que el *homing* aumenta, por lo que el cambio en la homeostasis interna de la célula estaría favoreciendo la movilidad de RmInt1. Esto indicaría que, de alguna manera, la concentración de magnesio, cuando CorA es activa, no es la más favorable para la actividad de la ribozima.

De todos estos resultados podemos concluir que Rmint1 necesita funciones celulares y de la maquinaria enzimática del hospedador para mantenerse en el genoma que ha invadido. También hemos comprobado que otras funciones celulares no favorecen su dispersión, manteniéndolo de alguna forma “controlado” y confinado.

Para comprobar y corroborar estos resultados, será preciso en el futuro realizar ensayos de complementación en *trans* y ver si el fenotipo silvestre se restituye.



## **CAPÍTULO III**



## INCORPORACIÓN DE MARCADORES DE RESISTENCIA EN LA SECUENCIA DE RmInt1 PARA SU USO EN MUTAGÉNESIS INSERCIONAL.

### III.1 INTRODUCCIÓN

El proceso de *retrohoming* de los intrones del grupo II está mediado por el complejo RNP formado por la IEP y el RNA escindido del intrón. Se sabe que tanto la IEP como el RNA participan en el reconocimiento del sitio diana en el DNA (Guo *et al.*, 1997; Matsuura *et al.*, 2001; Singh y Lambowitz, 2001; Jiménez-Zurdo *et al.*, 2003; Noah *et al.*, 2006). Debido a que el apareamiento de bases entre el RNA del intrón y el sitio diana es el requisito más importante en la especificidad de diana, los intrones del grupo II pueden ser redirigidos hacia sitios de interés modificando las secuencias de reconocimiento de los exones situadas en el DI del RNA del intrón (Jones III *et al.*, 2005; Zhuang *et al.*, 2009a; García-Rodríguez *et al.*, 2011).

El Intrón del grupo II Ll.LtrB de *L. lactis* ha sido ampliamente usado para su desarrollo en mutagénesis insercional (Targetron) en un amplio número de especies bacterianas Gram-negativas (*E. coli*, *Shigella flexneri*, *Salmonella enterica* serovar Typhimurium, *Francisella tularensis*, *Azospirillum braziliensis*, *Pseudomonas aeruginosa* y *Agrobacterium tumefaciens*; Karberg *et al.*, 2001; Yao y Lambowitz, 2007; Malhotra y Srivastava, 2008) y Gram-positivas (*L. lactis*, *Clostridium perfringens*, *Spaphylococcus aureus* y *Paenibacillus alvei*; Chen *et al.*, 2005; Yao, 2006; Zarschler *et al.*, 2009). Otro elemento genético móvil usado recientemente para la inserción génica es un nuevo intrón de la clase IIB1/CL1, EcI5, que se localiza en un plásmido de virulencia de *E. coli*. Por último, el intrón del grupo II de *S. meliloti* perteneciente a la clase IIB3/D RmInt1, también ha sido redirigido con éxito recientemente para llevar a cabo mutaciones dirigidas en sitios específicos del genoma de *E. coli* (García-Rodríguez *et al.*, 2011). Hay que destacar que, mientras que las IEPs de Ll.LtrB y de EcI5 mantienen los dominios de unión a DNA (D) y endonucleasa (En) RmInt1 carece de ellos, sin embargo, tanto su frecuencia como eficiencia de *retrohoming* es muy elevada (como se ha demostrado en el capítulo I de resultados de esta tesis).

Para la interrupción de genes en bacteria, el intrón es introducido dentro de un plásmido donador que expresa un intrón  $\Delta$ ORF con exones cortos flanqueantes. El intrón  $\Delta$ ORF tiene la mayoría del DIV deletado, y la proteína se expresa desde una localización aguas abajo del exón 3' en el caso de las construcciones con Ll.LtrB o aguas arriba en las construcciones

con RmInt1 (Guo *et al.*, 2000; Karberg *et al.*, 2001; García-Rodríguez *et al.*, 2011). Además, el intrón puede ser rediseñado para que sea capaz de insertarse en distintas orientaciones, redirigiéndolo hacia la cadena sentido o antisentido de la diana, haciendo posible obtener inserciones condicionadas o no condicionadas dependiendo de la expresión de la IEP y del splicing posterior (Frazier *et al.*, 2003; Yao *et al.*, 2006).

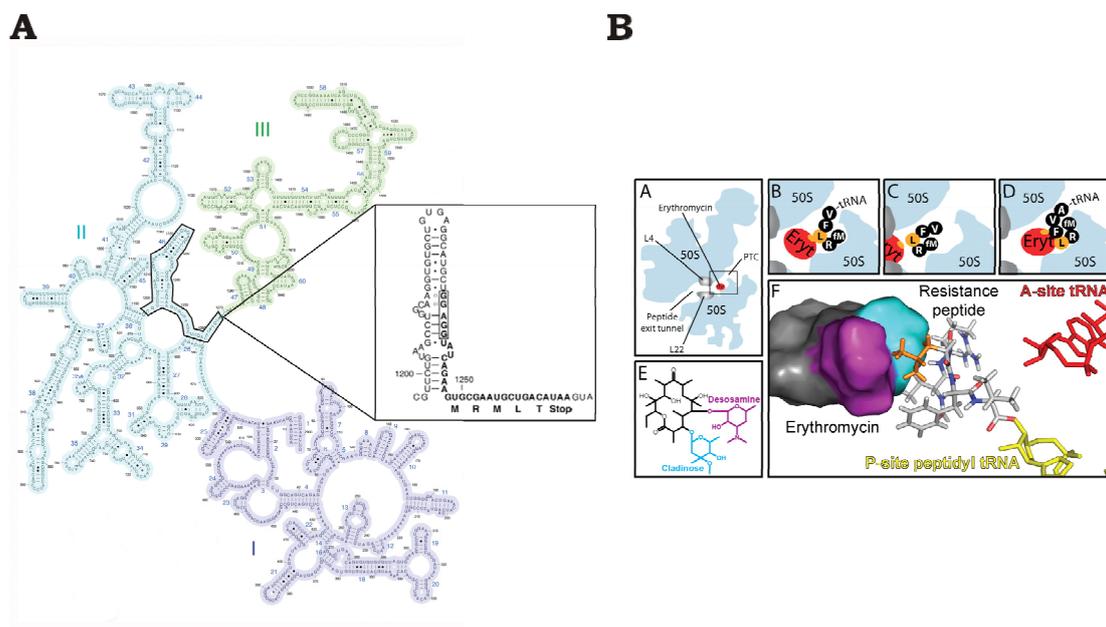
Gracias a la eliminación de casi toda la secuencia codificante de la IEP dentro del DIV pueden ser insertados marcadores de selección en su lugar para facilitar la identificación de los eventos de integración deseados, los cuales pueden ser detectados con una eficiencia de casi el 100% después de la selección (Zhong *et al.*, 2003).

<b>Vector (Referencia)</b>	<b>R<sup>a</sup></b>	<b>Inserción en dIV (en nt)</b>	<b>Frecuencia Mov. (%)<sup>b</sup></b>	<b>Invasión Intrón (%)<sup>c</sup></b>
pACD3 (wt) <sup>d</sup> (Guo <i>et al.</i> , 2000)	-	-	94±2%	ND
pACD3-RAM (Zhong <i>et al.</i> , 2003)	Tm	313	18.8 %	ND
Ll.LtrB donator (wt) <sup>d</sup> (Plante y Cousineau, 2006)	-	-	--	34.4
Ll.LtrB donator pHw+ (“)	-	64	--	13.2 (38.4)
Ll.LtrB donator pHw-(“)	-	64	--	12.0 (34.9)
Ll.LtrB donator pHs+ (“)	-	79	--	24.0 (69.8)
Ll.LtrB donator pHs-(“)	-	79	--	1.6 (4.7)
Ll.LtrB donator holin/lysin+ (“)	-	1180	--	0.3 (0.9)
Ll.LtrB donator pHw/holin/lysin-(“)	-	1244	--	ND
Ll.LtrB donator NaCl-(“)	-	1251	--	ND
Ll.LtrB donator pHs/holin/lysine-(“)	-	1269	--	0.4 (1.2)
Ll.LtrB donator pHs/ <i>acmA</i> + (“)	-	1469	--	ND
RmInt1 donator pKGEMA4(wt) <sup>d</sup> (Nisa-Martínez <i>et al.</i> , 2007)	-	-	100	85.4±3.5
RmInt1 donator pKGEMA4 T7 (García-Rodríguez <i>et al.</i> , 2011)	-	31	100	63.3±16.1 (74.1)

**Tabla III.1** Efecto del tamaño de inserto en el dominio IV en la eficiencia de invasión de intrones del grupo II. <sup>a</sup> Fenotipo de resistencia: (-) no resistente (Tm) resistente a Trimetroprima. <sup>b</sup> Frecuencia de células en las que al menos ha habido un evento de salto. <sup>c</sup> Porcentaje de invasión de la diana. Entre paréntesis el valor respecto al vector de referencia sin inserto. <sup>d</sup> Vectores de referencia en cada ensayo de movilidad. ND: sin determinar.

Los intrones del grupo II desarrollados para la mutagénesis contienen secuencias heterólogas insertadas en el dominio IV del intrón y pueden, por tanto, ser vehículos potenciales para la introducción de un determinado transgén. Sin embargo, estudios recientes demuestran que la movilidad de estos elementos se ve muy reducida (hasta en 4 órdenes de magnitud; Tabla

III.1) por inserciones de alrededor de 1 Kb (Plante y Cousineau, 2006; Heap *et al.*, 2007). Como elementos de mutagénesis dirigida, la incorporación en los intrones de marcadores de resistencia que contribuyan a facilitar la selección de los mutantes, sin una merma en su capacidad de movimiento, constituye uno de los aspectos pendientes por desarrollar en esta tecnología. Hasta ahora la unidad más pequeña incorpora un gen de resistencia al antibiótico Trimetroprima, modificado hasta una longitud de 313 nt, reduciéndose la movilidad del intrón a un 20% sobre el silvestre (Tabla III.1; Zhong *et al.*, 2003). Por tanto, la búsqueda de unidades de resistencia de un tamaño menor, constituye un paso crítico y supone una importante limitación actual en el desarrollo de estas herramientas de mutagénesis.



**Figura III.1. A:** Estructura secundaria de la región 5' del rRNA 23S de *E. coli*. Se ha ampliado la región en la que se encuentra la secuencia codificante del pentapéptido que confiere resistencia a Eritromicina. En negrita se encuentra destacada dicha secuencia enmarcando el Shine-Dalgarno en un rectángulo. Bajo la secuencia se especifica los aminoácidos que formarían el pentapéptido por su abreviatura de una letra, y el codón de parada como Stop. **B:** Modelo estructural del mecanismo de acción del péptido. En el panel A se muestra una representación de la subunidad grande del ribosoma seccionada a lo largo del túnel de salida del péptido naciente en el que se puede observar una molécula de eritromicina unida (en rojo). El rectángulo negro indica la sección que se muestra en los paneles B-D, en los que se muestra la hipótesis del mecanismo de acción del pentapéptido. En el panel B se representa al pentapeptidil-tRNA interactuando con la eritromicina. El panel C muestra al péptido de resistencia que ha desplazado a la eritromicina durante la terminación de la síntesis de proteínas. En el panel D aparece un hexapeptidil-tRNA que ha perdido el contacto con la eritromicina y permite al peptidil-tRNA liberarse. El panel E representa la estructura de la eritromicina. En el panel F se representa la interacción entre la leucina conservada (naranja) del pentapéptido de resistencia y la eritromicina. (Lovmar *et al.*, 2005).

En este sentido, a finales de los 90 se determinó un nuevo mecanismo de resistencia al antibiótico Eritromicina, mediada por un pentapéptido

codificado en el rRNA de la subunidad 23S del ribosoma de *E. coli* (Tenson *et al.*, 1996; Dam *et al.*, 1996), el cual resulta de especial interés por el pequeño tamaño de la secuencia codificante. Ésta se encuentra entre los nucleótidos 1248 y 1265 en la unión de los dominios II y III del rRNA 23S de *E. coli* y su Shine-Dalgarno (GGAGG) ocho nucleótidos antes (Fig. III.1A). Determinadas mutaciones, que desestabilizan la estructura secundaria en la que se encuentra inmersa la secuencia S-D, hacen posible el acceso a éste, permitiendo que se transcriba y traduzca el pentapéptido. Se sabe que el pentapéptido actúa en *cis*, es decir, que debe ser sintetizado por los ribosomas para conferirles la resistencia al macrólido, ya que se ha demostrado que la adición del péptido al medio no confiere resistencia a la eritromicina. Por comparación de secuencias se ha observado que, a excepción de en los géneros *Rhodobacter* y *Flexibacter*, la secuencia está muy conservada en distintas bacterias, incluyendo *E. coli* y *S. meliloti*, las cuales sólo se diferencian en el codón de parada (UAA en *E. coli* mientras que UGA en *S. meliloti*). También se ha comprobado que el tamaño del péptido es fundamental para su actividad, estando el óptimo entre cuatro y seis aminoácidos, ya que cualquier mutación que elimine el codón de parada también elimina la resistencia a Eritromicina. Existe una prevalencia de aminoácidos hidrofóbicos en las posiciones tercera y quinta, mientras que en las posiciones segunda y cuarta hay una mayor diversidad, pero se ha determinado que la secuencia que confiere una mayor resistencia es la formada por los aminoácidos MSLKV (Met. Ser. Leu. Lys. Val.; Tenson *et al.*, 1997; 2001). El mecanismo de resistencia propuesto se basa en que el pentapéptido interactúa con la eritromicina expulsándola de su sitio de unión en el ribosoma, permitiendo así que los sucesivos ARN mensajeros puedan completar su traducción (Fig. III.1B).

De los resultados obtenidos en este capítulo se ha desarrollado una patente, en la actualidad en solicitada a la Oficina Española de Patentes y Marcas (Ref. 201130782) con el título: “Construcción de una unidad de resistencia a antibióticos de 60-200 nucleótidos en tamaño.”; inventores: Toro, N., Martínez-Abarca F. y Nisa-Martínez R.”

## III.2 MATERIAL Y MÉTODOS

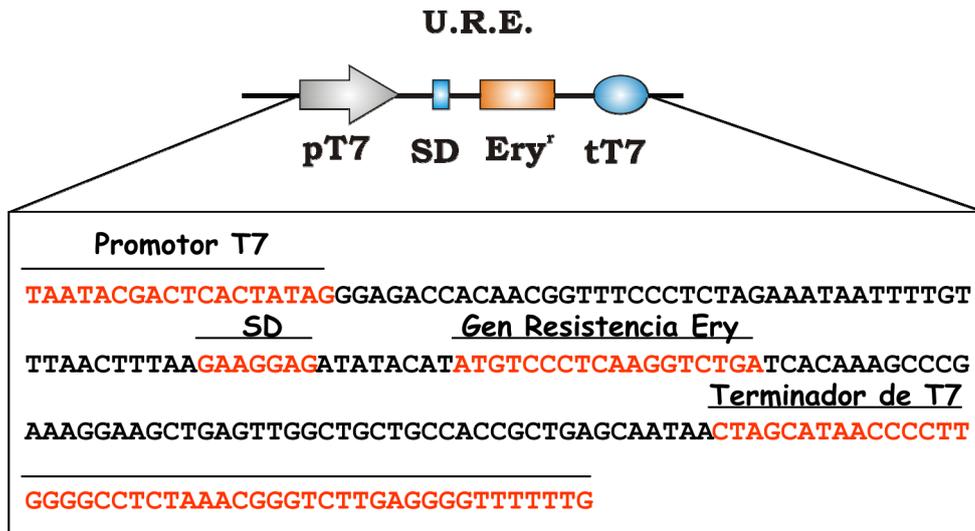
### III.2.1. Cepas bacterianas y condiciones de crecimiento

La cepa de *Sinorhizobium meliloti* usada en este trabajo fue RMO17, cepa que carece de copias de *RmInt1* (Villadas *et al.*, 1995). *Escherichia coli* DH5 $\alpha$  se usó para la clonación y mantenimiento de plásmidos; B121(DE3) y DH5 $\alpha$  fueron usadas para algunos de los ensayos de resistencia y movilidad. Las cepas de *S. meliloti* se crecieron en medio completo (TY) o en medio mínimo definido (MM) (Robertsen *et al.*, 1981) a 28°C y *E. coli* DH5 $\alpha$  en medio Luria-Bertani (LB) a 37°C para mantener las construcciones a 28°C, cuando el medio era suplementado con eritromicina para minimizar inactivación por temperatura. Los medios fueron suplementados, según los requerimientos de las cepas bacterianas, con antibióticos a las siguientes concentraciones: kanamicina, 200  $\mu\text{g/ml}$  para rhizobios y 50  $\mu\text{g/ml}$  para *E. coli*; tetraciclina, 10  $\mu\text{g/ml}$ , ampicilina, 200  $\mu\text{g/ml}$  y eritromicina a diferentes concentraciones. También se suplementó con Isopropil- $\beta$ -D-1-tiogalactopiranosido (IPTG) a diferentes concentraciones en los ensayos con inducción.

### III.2.2. Construcción de la Unidad de Resistencia a Eritromicina (U.R.E.)

Para introducir el oligonucleótido de 18 nt que codifica el pentapéptido que confiere resistencia a eritromicina bajo el promotor de la RNA polimerasa del Fago T7, se diseñó una estrategia mediante PCR divergente utilizando como molde el plásmido comercial pET3a (New England Biolabs) usando los cebadores “4” (5'-GCGTGATCAGACCTTGAGGGACATATGTATATCTCCTTCTTAAAGTT-3') y “5” (5' GCGTGATCACAAAGCCC-GAAAGGAAGCTG 3').

El producto resultante fue purificado y tras ser digerido con la enzima de restricción *Bcl*I, religado en condiciones favorecedoras de la autoligación. Se transformó *E. coli* DH5 $\alpha$  para mantener la construcción denominada pETURE.

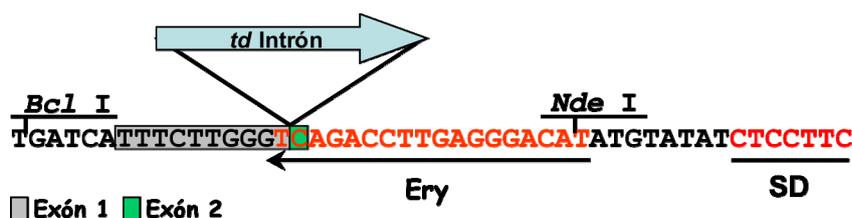


**Figura III. 2.** Esquema de la unidad de resistencia a eritromicina (U.R.E.) constituido por el promotor de la enzima RNA polimerasa del fago T7, una secuencia espaciadora, a continuación la secuencia de unión del ribosoma (SD), la secuencia que codifica al pentapéptido, otra secuencia espaciadora, y por último la secuencia terminadora de la transcripción el mismo fago.

A continuación, sobre esta mismo plásmido se realizó una nueva PCR con los cebadores Salery (5'-GCGTCGACTTAATACGACTCACTATAGG-3') y Eryxho2 (5'-CCGGAGCTCCAAAAACCCCTCAAGAC-3') para obtener la denominada Unidad de Resistencia a Eritromicina (U.R.E.) constituido por el promotor de T7, un espaciador, la secuencia Shine-Dalgarno, otro pequeño espaciador, el pentapéptido, espaciador y el terminador de la polimerasa de T7, abarcando en total 198 nucleótidos como secuencia efectiva (Fig. III.2A). La correcta amplificación fue comprobada mediante el clonaje del amplicón en el vector pGEMT Easy (Promega) y posterior secuenciación. Una vez verificada la secuencia se procedió a incorporar esta unidad de resistencia en el sitio de restricción *XhoI* existente en el dominio IV del  $\Delta$ ORF de pKGEMA4 ya que se le incorporaron dianas a tal efecto (*SalI* en 5' y *XhoI* en 3'). Una vez comprobada la correcta inserción mediante el perfil de restricción y habiendo obtenido las construcciones pKG4UREs (la unidad esta clonada en el sentido de la secuencia del intrón RmInt1) y pKG4UREas (la unidad esta clonada en el sentido contrario a la secuencia del intrón RmInt1; Fig. III.2), se llevó a cabo los estudios de movilidad (*homing*) de las construcciones modificadas *in vivo*.

### III.2.3. Construcción de un marcador de resistencia activable mediante *retrohoming*

Estudios previos realizados demostraron que el movimiento de *RmInt1* se producía, al igual que otros intrones del grupo II, mediante intermediarios de RNA, esto es mediante *retrohoming* (Martínez-Abarca *et al* 2004). Este hecho venía demostrado de construcciones doble-intrón (*twintron*; Cousineau *et al* 1998, Martínez-Abarca *et al* 2004); construcciones en las que el intrón del grupo I *tdI* se insertaba dentro del dominio IV de la ribozima. Las formas finalmente insertadas en la diana perdían el intrón del grupo I, tras sufrir *splicing* via RNA. Este hecho permite desarrollar vectores en los cuales la resistencia derivada se active sólo cuando el *splicing* del intrón *tdI* haya tenido lugar (Zhong *et al* 2003). Con el objetivo de diseñar una estrategia parecida sobre el marcador de resistencia a la eritromicina se realizó una PCR con los cebadores *tdiEx1* (5'-CCGAATTCTGATCATTTCTTGGGTTAATTGAGGCC-3') y *tdiEx2* (5'-CGGAATTCATATGTCCTCAAGGTCTGCATTATGTTTCAGATAAG-3') usando como molde la construcción pKGAD1 que contenía el intrón *tdI* activo (Martínez-Abarca *et al.*, 2004).



**Figura III. 3.** Detalle de cómo el intrón *tdI* interrumpe el codón de parada del pentapéptido y de cómo se restituiría este al sufrir este intrón *splicing* y quedar unidos los exones, lo que permitiría su traducción.

El amplificado de 445 pb estaba compuesto por la secuencia *tdΔ1-3* (Cousineau *et al.*, 1998) que conserva las secuencias exónicas imprescindibles para su automaduración. Se encuentra flanqueada por las dianas de restricción *BclI* en 5' y *NdeI* en 3', que nos permite clonarlo en el extremo 3' de la secuencia U.R.E. interrumpiendo el codón de parada de la traducción del pentapéptido (Fig. III.3). Debido a que péptidos de mayor tamaño a 5 no confieren resistencia a eritromicina, ésta solo se dará cuando la construcción no contenga el intrón *tdI* y se recomponga la secuencia original que confiere resistencia (Fig III.3). El fragmento *SalI-XhoI* compuesto por la U.R.E. interrumpida por el intrón *tdI* (al cual llamaremos UT) se clonó en el sitio *XhoI* de pKGEMA4 antes descrito, y se usó la construcción en la

que el *tdI* se encuentra en la misma orientación que el  $\Delta$ ORF, obteniendo el plásmido denominado pKG4UT (Fig III.6A).

Se analizó la movilidad de este *twintron* tanto por hibridación de ADN plásmídico (*Southern blot*) para estimar la posible reducción en la eficiencia de *homing* respecto a la construcción silvestre, como por selección directa de clones resistentes. Para ello, mezclas de plásmidos obtenidas en *S. meliloti* se usaron para transformar *E. coli* DH5 $\alpha$  y posterior siembra en placas de LB suplementadas con los antibióticos Ap. y Tc. (a los que confiere resistencia el plásmido aceptor pJB0.6LAG) para estimar los transformantes totales y (Ap/Tc y Ery) para evaluar los eventos de invasión.

#### III.2.4. Ensayos de *homing in vivo*.

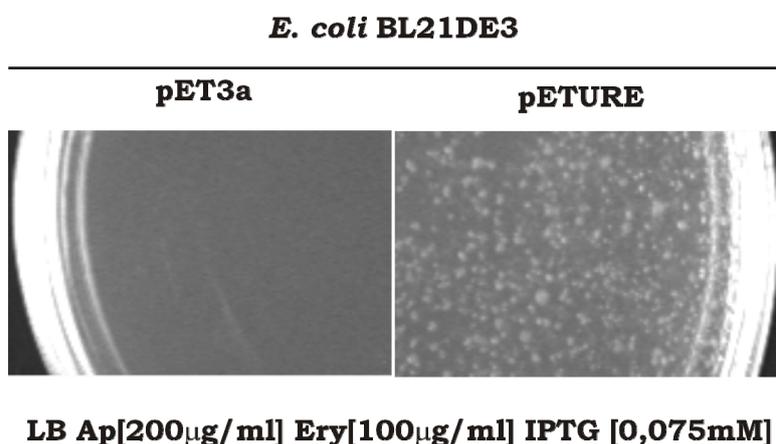
Los ensayos de movilidad *in vivo* fueron llevados a cabo mediante la introducción de las construcciones donadoras de intrón en cepas de *S. meliloti* RMO17 (que carece de copias de RmInt1) que contenían los plásmidos aceptores pJB0.6LAG y pJB $\Delta$ 129 (control negativo de la invasión, ya que carece del sitio de inserción). Los plásmidos donadores fueron transferidos mediante conjugación triparental, mediada por el plásmido pRK2013 contenido en la cepa *E. coli* HB101. La mezcla de las tres cepas se realizó sobre TY, y tras un día creciendo a 28°C se extendió una proporción en placas de MM. Cuando aparecieron colonias individuales se analizó la presencia de *E. coli* mediante réplicas en endoagar. Las unidades formadoras de colonia libres de *E. coli* fueron crecidas en medio TY líquido para la posterior extracción de su DNA plasmídico.

Los *Southern blots* se llevaron a cabo utilizando el DNA plasmídico obtenido mediante lisis alcalina y digerido con la enzima de restricción *SaII* (nos permite discriminar por tamaño entre los fragmentos provenientes de eventos de invasión de los que no lo son). Se hibridó con una sonda de la diana de RmInt1, lo que nos permite calcular la proporción de diana invadida. La eficiencia de *homing* se calculó mediante la fórmula  $(H / [H+A]) \times 100$  donde H hace referencia al producto de *homing* y A al plásmido receptor sin invadir. Los datos son la media de al menos cuatro determinaciones independientes con su correspondiente desviación estándar (Jiménez-Zurdo *et al.*, 2003; Muñoz-Adelantado *et al.*, 2003; Martínez-Abarca *et al.*, 2004). También se analizó la proporción de invasión mediante la transformación, con el mismo DNA plasmídico, de *E. coli* DH5 $\alpha$  y su posterior selección con los antibióticos Ap/Tc o con Ap/Tc/Ery mediante la ecuación  $[(Ap/Tc/Ery) \div (Ap/Tc)] \times 100$ .

### III.3. RESULTADOS

#### III.3.1. El casete U.R.E. confiere resistencia a eritromicina

Una vez construido el casete U.R.E. (Fig. III.2) en el plásmido pET3a, dando lugar a la construcción pETURE, se analizó si realmente las células bacterianas que contenían este plásmido serían resistentes a Eritromicina o no. Para ello se realizó una curva de concentraciones crecientes de Eritromicina, en placas de cultivo con medio LB suplementado con los antibióticos Ampicilina y Eritromicina, e IPTG 0.075M para inducir su expresión, comparando células BL21 (DE3) conteniendo el vector pET3a con aquellas que contenían su derivado pETURE que expresaba el pentapéptido. Se determinó que a una concentración de 100 $\mu$ /ml de Eritromicina el crecimiento de las células que contenían pET3a era inhibido casi en su totalidad, mientras que las que portaban a pETURE todavía eran capaces de crecer (Fig. III.4).

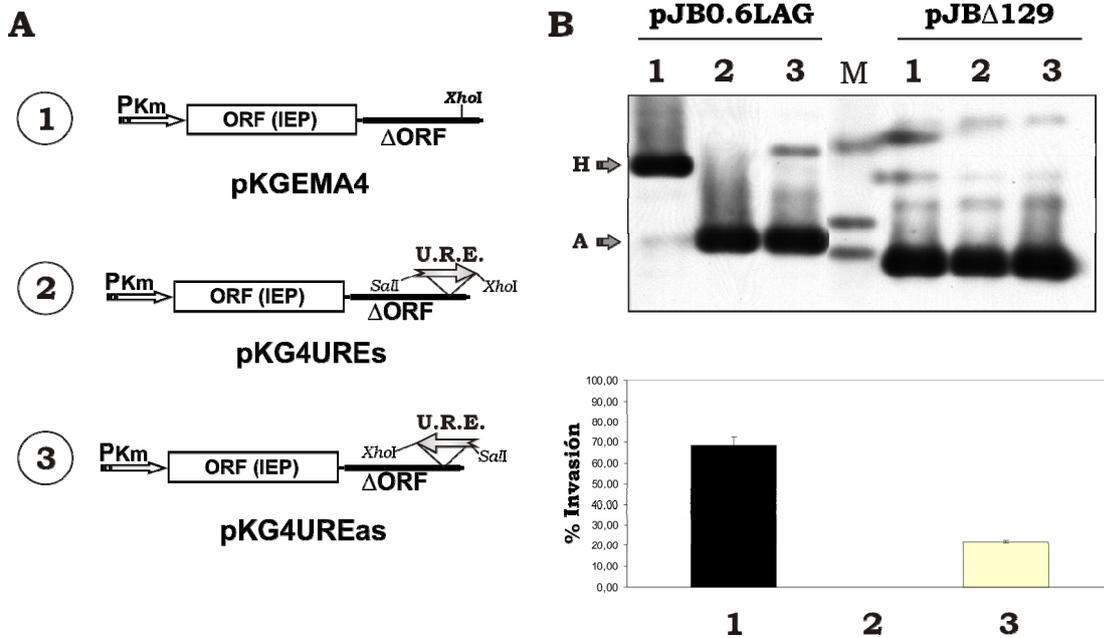


**Figura III.4.** Secciones correspondientes a dos placas de crecimiento que ilustran la resistencia conferida por el pentapéptido a las células bacterianas que lo portan, a una concentración de eritromicina de 100 $\mu$ g/ml, bajo una inducción con 0,075mM de IPTG.

#### III.3.2. Construcción de un marcador de resistencia activable por retrohoming y su incorporación a vectores donadores de intrón.

Una vez obtenidas las construcciones pKG4UREs y pKG4UREas (Fig III.5A) se llevó a cabo el análisis de su movilidad hacia la diana silvestre contenida en el plásmido pJB0.6LAG (sobre el cual se ha comprobado que la eficiencia de *homing* es mayor) junto con el control negativo pJB $\Delta$ 129 (que

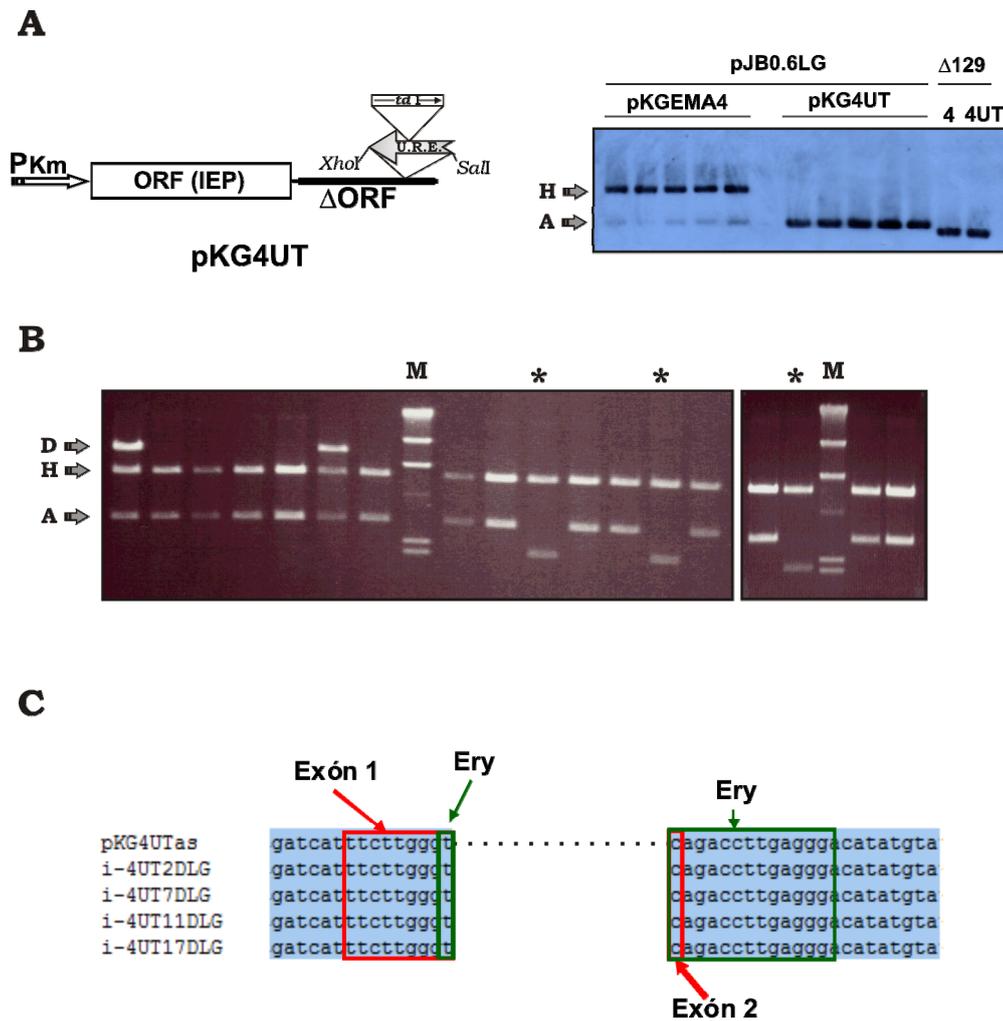
carece del sitio de inserción) observando diferentes resultados entre ambas construcciones. Se vio que, en el caso del donador pKG4UREs (con el casete URE insertado en el mismo sentido que el  $\Delta$ ORF) el fenómeno de invasión no era detectado mediante hibridación con una sonda de la diana, al contrario de lo que ocurre con la otra orientación. En el caso de pKG4UREas, sí se consigue detectar los eventos de salto, aunque su eficiencia se ve reducida en un 80% sobre el donador silvestre pKGEMA4 (Fig III.5B).



**Figura III.5. A:** Esquema de las construcciones donadoras de intrón pKGEMA4, PKG4UREs y pKG4UREas (1, 2 y 3 respectivamente) en las que se indica el sitio XhoI existente en el  $\Delta$ ORF donde se ha insertado, en ambas orientaciones, el casete U.R.E. **B:** Southern blot e histograma correspondiente a las cuantificaciones del porcentaje de invasión sobre la diana silvestre que se encuentra en el plásmido pJB0.6LAG. Se indican como 1, 2 y 3 los carriles correspondientes a cada una de las construcciones donadoras de intrón, discriminando entre diana silvestre (pJB0.6LAG) y control negativo (pJBΔ129). En el histograma sólo se han representado los porcentajes correspondientes a la cuantificación del salto sobre la diana silvestre de cada uno de los donadores, siendo estos:  $68,46 \pm 4,16$  para el pKGEMA4; No detectado para el pKG4UREs y  $21,87 \pm 0,6$  en el caso de pKGEMA4UREas. H (*homing*): banda correspondiente a la invasión; A (aceptor): banda correspondiente a la diana aun sin invadir.

Cuando se toma una mezcla de ADN plasmídico proveniente de lisis alcalina de cultivos celulares en los que se ha producido eventos de salto y, una vez digerido el plásmido donador con enzimas de restricción que no actúen sobre el plásmido receptor, al transformar *E. coli*, podemos observar que el plásmido aceptor que ha sufrido invasión del  $\Delta$ ORF con el inserto UREas, también es capaz de conferir resistencia a eritromicina a la célula que lo ha adquirido, favoreciendo su selección. Pero, para evitar tener que

extraer DNA plasmídico y la posterior transformación, lo ideal sería poder seleccionar los eventos de invasión de una forma directa, es decir, que la resistencia a eritromicina sólo se haga patente una vez que se ha producido el *retrohoming*. Para ello se ha usado una estrategia cuya efectividad está contrastada (Zhong et al., 2003) y que consiste en la inserción en el dominio IV del  $\Delta$ ORF, en orientación inversa, de un marcador de resistencia con su propio promotor (U.R.E.), pero que está interrumpido por el intrón del grupo I *td* (que es capaz de sufrir auto-splicing de forma muy eficiente) en el mismo sentido que el  $\Delta$ ORF. Durante el *retrohoming* del intrón, que ocurre vía intermediario de RNA, el intrón *td* sufre *splicing*, activando el marcador de selección, que puede ser seleccionado después de la integración en el DNA.



**Figura III.6. A:** En la parte izquierda se muestra en detalle parte de la construcción pKG4UT, formada por el promotor del gen de resistencia a kanamicina a partir del cual se transcribe tanto la IEP como el

twintron formado por la secuencia □ORF de RmInt1 que incorpora el casete U.R.E. interrumpido por el intrón del grupo I td. En la parte derecha se observa el southern blot correspondiente al ensayo de invasión de plásmido a plásmido del que se puede deducir que el porcentaje de invasión es menor del 8-10% (límite de detección de esta técnica para este tipo de ensayo). H: banda correspondiente al *homing*; A: banda correspondiente al aceptor sin invadir. **B:** electroforesis en gel de agarosa en la que se analizaron mediante restricción con Sall si las células resistentes a eritromicina contenían el plásmido aceptor invadido o por el contrario correspondían a resistentes espontáneos. De los analizados (18) sólo 3 resultaron ser negativos (marcados con \*). M: marcador de peso molecular II de Roche. **C:** Comparativa entre la secuencia correspondiente al donador (pKG4UT td+) y las secuencias provenientes de cuatro colonias en las que el plásmido aceptor ha sido invadido y, por lo tanto, se han unido los exones del intrón tdI (marcados en rojo) recuperándose el codón de parada del pentapéptido (en verde), obteniéndose resistencia a eritromicina.

Una vez obtenida la construcción pKG4UT (*tdI* +) (Fig. III.6A) se llevó a cabo un ensayo de movilidad de plásmido a plásmido en *S. meliloti* RMO17 para poner de manifiesto el *homing* mediante *Southern blot* y así poder evaluarlo, pero no fue detectada ninguna banda correspondiente al evento de movilidad. Para analizar si se había producido o no *homing*, se transformó DH5 $\alpha$  con DNA plasmídico proveniente de las colonias usadas en el *southern blot*. La selección con eritromicina puso de manifiesto colonias resistentes que, tras un análisis por restricción, se comprobó que los plásmidos que contenían eran en su mayoría vectores aceptores que habían sufrido invasión (83%) junto con una baja proporción de falsos positivos. En cuanto al porcentaje total de *homing*, se ve claramente reducido habiendo sido estimado con un valor de un 0,01%, es decir, sufrirían invasión 10 dianas de cada 1000 disponibles. Fueron secuenciadas 4 de esas formas invadidas, y se comprobó que el intrón del grupo I se había escindido de forma limpia y, por lo tanto, que la secuencia que confiere resistencia a eritromicina se había restituido adecuadamente (Fig. III.6).

#### III.4. DISCUSIÓN

Tanto los intrones del grupo II Ll.LtrB y Eci5 (que contienen dominios de unión a DNA y endonucleasa) como RmInt1 (que carece de estos), han sido usados para la interrupción de genes mediante la modificación de las secuencias de reconocimiento del RNA del intrón (Guo et al., 2000; Mohr et al., 2000; Zhuang et al., 2009; García-Rodríguez et al., 2011). Pero sólo en el caso de Ll.LtrB se le ha incorporado un sistema de selección de eventos de movilidad mediante resistencia a antibióticos (Zhong et al., 2003). En este capítulo hemos desarrollado un sistema de selección directa de los eventos de movilidad mediante un marcador de resistencia activable por *retrohoming*.

Esta estrategia se basa en otra usada por primera vez para estudiar la movilidad del transposón de levaduras Ty1 y que, posteriormente, fue adaptado para el estudio de la transposición de los intrones del grupo II a sitios ectópicos (Ichiyangi *et al.*, 2002), y para nuestro mismo propósito por el grupo del doctor Lambowitz (Zhong *et al.*, 2003).

La construcción usada se basa en el plásmido donador del intrón pKGEMA4, descrito en el capítulo I, el cual contiene un derivado  $\Delta$ ORF del intrón de *Sinorhizobium meliloti* RmInt1, flanqueado por exones cortos -20/+5, clonado debajo del ORF de la IEP, y expresado a partir del promotor del gen de resistencia a kanamicina. En este intrón modificado, se ha insertado tanto en orientación directa como inversa un marcador seleccionable dentro del dominio IV del intrón del grupo II. Este marcador está compuesto por el promotor de la T7 RNA polimerasa, una secuencia espaciadora hasta el sitio de unión al ribosoma, una secuencia codificante de un pentapéptido que confiere resistencia a eritromicina, otra secuencia espaciadora y por último, la secuencia terminadora de la transcripción de la misma T7 RNA polimerasa, pudiendo reducir en nuestro caso el tamaño del marcador de resistencia que se introduce en el dIV del RNA del intrón a 198 nt, frente a los 313 nt del gen de resistencia a trimetroprima usado por Zhong y colaboradores, lo que nos ha permitido obtener una eficiencia de salto sobre su diana natural (en ensayos de *homing* plásmido a plásmido) de un 21% aproximadamente, y con una frecuencia de salto de un 100%. Se ha observado mediante ensayos de movilidad que, mientras que cuando el inserto está orientado en el mismo sentido que el RNA del intrón no se detectan eventos de inserción mediante *southern blot*, estos eventos sí son detectados para la construcción en la que la U.R.E. se encuentra en la orientación inversa. Cuando se interrumpe con el intrón del grupo I *tdI* en orientación directa, permite seleccionar los eventos de movilidad que provienen de la vía *retrohoming*, evitando las posibles inserciones que se pueden producir por recombinación debida a la homología de secuencias, pero la eficiencia de invasión se reduce hasta un 0,01%, corroborando que cuando se vuelve a introducir en el dIV información genética extraña, del que previamente se ha delecionado la secuencia codificante de la IEP, la eficiencia de movilidad se ve reducida a niveles inferiores a la que se obtiene cuando el intrón no ha sido modificado.

El promotor desde el que se expresa el gen de resistencia a eritromicina es el promotor T7, inducible por IPTG, lo cual permite una mejor selección del marcador en el fondo genético adecuado, es decir, en aquellos organismos que contengan la RNA polimerasa de T7, inducible por IPTG.

Otra ventaja de esta estrategia es que el marcador de resistencia es inactivo antes del *retrohoming*, por lo cual no sería necesaria la eliminación del plásmido donador antes de seleccionar las colonias resistentes a eritromicina, lo cual es una ventaja frente a otras herramientas que usan también marcadores de resistencia.

Una vez comprobado que esta construcción es capaz de invadir dianas libres de intrón y estos eventos pueden ser seleccionados mediante la resistencia a eritromicina proporcionada por el marcador que incorpora el intrón, se podrán modificar las secuencias de reconocimiento para redirigir el intrón hacia secuencias de interés y poder así detectar las mutaciones insercionales mediante la activación del gen de resistencia a eritromicina, evitando tener que poner de manifiesto mediante PCR los mutantes obtenidos (García-Rodríguez *et al.*, 2011), lo cual resulta costoso en tiempo y dinero.

## **CAPÍTULO IV**



## **ACTIVIDAD DE RmInt1 EN HOSPEDADORES HETERÓLOGOS BACTERIANOS Y EUCARIOTAS (PLANTAS).**

### IV.1 INTRODUCCIÓN

RmInt1 es un intrón del grupo II perteneciente a la subclase IIB3/D (Toro., 2003). Fue encontrado en la bacteria simbiótica fijadora de nitrógeno *S. meliloti*, la cual establece simbiosis con la raíces de plantas leguminosas del género *Medicago* (Jones *et al.*, 2007). Se ha demostrado que este intrón tiene una alta eficiencia de movilidad dentro de su hospedador natural (Nisa-Martínez *et al.*, 2007). RmInt1 está ampliamente distribuido, siendo detectado en el 90% de las cepas estudiadas (Biondi *et al.*, 2011) siendo muy abundante en algunas de ellas, que pueden contener hasta 11 copias. Ésto nos indica que RmInt1 se encuentra perfectamente adaptado a su hospedador natural *S. meliloti*, pero, debido a su interés como herramienta de mutagénesis insercional, es muy importante comprobar su comportamiento en otros hospedadores de interés. En este capítulo se ha comprobado la capacidad de movimiento de RmInt1 en diferentes bacterias, tales como *E. coli* y *P. putida*, en cuyos genomas también han sido encontrados intrones del grupo II (Zhuang *et al.*, 2009a; Yeo *et al.*, 2001). Si RmInt1 (descubierto en alpha-proteobacterias) retiene la capacidad de retromovilidad en otras bacterias filogenéticamente alejadas, el rango de aplicabilidad de la tecnología basada en este elemento móvil como herramienta de mutagénesis insercional se ve ampliado. Es muy probable que, debido a las características específicas tanto del intrón como de los hospedadores, existan restricciones al movimiento de RmInt1 en estos fondos genéticos.

También hemos realizado una primera aproximación al estudio de éste elemento genético móvil en células eucariotas, específicamente células de la planta *A. thaliana* como organismo vegetal modelo, en cuyos orgánulos se han encontrado copias de intrones del grupo II.

Un aspecto interesante a tener en cuenta en la expresión de RmInt1 en otros hospedadores diferentes a *S. meliloti* es el estudio de la localización subcelular de la proteína codificada por él (IEP), ya que nos puede ofrecer pistas de cómo y donde se expresa esta proteína, y de con qué elementos del hospedador puede estar interaccionando. En el caso de las células vegetales es importante conocer su localización debido a la existencia de compartimentación celular, que podría convertirse en un importante inconveniente en el caso de querer usar estas herramientas genéticas como

elementos mutágenos, ya que es imprescindible que las RNPs se encuentren en contacto con el material genético para poder llevar a cabo la inserción en el genoma, cosa que no ocurrirá si están excluidas del núcleo.

## IV.2 MATERIAL Y MÉTODOS

### IV.2.1. Cepas bacterianas y condiciones de crecimiento

Las cepas de *Escherichia coli* usadas fueron DH5 $\alpha$  para propagar todas las construcciones excepto el vector pK7WGF2, ya que contiene el operón *ccdB*, por lo que se usó la cepa DB3.1; y HMS174 (DE3) (Novagen) para los ensayos de movilidad. *E. coli* creció en medio LB a 37°C, excepto en los ensayos de movilidad que creció a 28°C.

La cepa de *Sinorhizobium meliloti* usada para comprobar la movilidad de determinadas construcciones fue RMO17, que carece de copias de RmInt1 (Villadas *et al.*, 1995). *S. meliloti* se creció en medio completo (TY) o en medio mínimo definido (MM) (Robertsen *et al.*, 1981) a 28°C.

Para el análisis de la movilidad de RmInt1 en *Pseudomonas putida* se usó la cepa KT2440 (Bagdasarian *et al.*, 1981) y fue crecida en medio LB a 28°C. Los medios fueron suplementados con antibióticos, según los requerimientos de las cepas bacterianas, a las siguientes concentraciones: kanamicina, 200  $\mu\text{g/ml}$  para *S. meliloti* y 50  $\mu\text{g/ml}$  para *E. coli* y *P. putida*; tetraciclina, 20  $\mu\text{g/ml}$  para *P. putida* y 10  $\mu\text{g/ml}$  para *S. meliloti* y *E. coli*; ampicilina, 200  $\mu\text{g/ml}$  para *S. meliloti* y *E. coli*; estreptomycin, 100  $\mu\text{g/ml}$  para *E. coli*; y cloranfenicol, 35  $\mu\text{g/ml}$ .

### IV.2.2. Construcción de plásmidos

El plásmido donador de intrón pCm4 es un derivado del vector pBBR1MCS (Kovach *et al.*, 1994) que confiere resistencia al antibiótico cloranfenicol, en el cual se introdujo mediante restricción con las enzimas *ApaI-SacI* un inserto proveniente del plásmido pKGEMA4 (Nisa-martínez *et al.*, 2007) que contiene el promotor del gen de resistencia a kanamicina, y bajo este, la secuencia codificante de la IEP. Posteriormente fue introducido el derivado  $\Delta\text{ORF}$  de RmInt1 mediante extremos *SacI*. De esta forma se obtuvo un plásmido donador de intrón con las mismas características que pKGEMA4 pero con diferente resistencia a antibiótico. Su movilidad fue ensayada en *S. meliloti* RMO17, comprobando que no existían diferencias con pKGEMA4 (datos no mostrados).

Las fusiones traduccionales IEP-GFP para la localización subcelular en protoplastos de *A. thaliana* se usó la tecnología Gateway® (Invitrogen). Primero fue necesario generar mediante recombinación los plásmidos “Entry” a partir del vector pDONR™221 y de fragmentos de PCR específicos para cada construcción. A partir de estos se generaron los vectores de expresión definitivos sobre el vector pK7WGF2 (Karimi et al., 2002). Los cebadores necesarios para las distintas construcciones se encuentran especificados en la tabla IV.1. Como pareja de los cebadores N-IEPsattB2 y N-  $\Delta$ C29attB2 se usó el oligonucleótido N-IEPattB1. Para amplificar los diferentes mutantes de por sustitución la IEP se usó la pareja de cebadores N-IEPattB1 / N-IEPattB2. El mutante en el NLS putativo se obtuvo usando el kit de Promega *Altered Sites*® II Mutagenesis System.

FUSIÓN	CEBADOR	SECUENCIA
GFP:IEP	N-IEPattB1	5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTC ATGACTTCGGAAAGTACGACA-3'
	N-IEPattB2	5'-GGGGACCACTTTGTACAAGAAAGCTGG GTCTCAGGTAAACGTGTTTCGTTCC-3'
IEP:GFP	C-IEPattB1	5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACC ATGACTTCGGAAAGTACGACA-3'
	C-IEPattB2	5'-GGGGACCACTTTGTACAAGAAAGCTGG GTCGGTAAACGTGTTTCGTTCCGAA-3'
GFP:Mat	N-MattB1	5'-GGGGACAAGTTTGTACAAAAAAGCAGGC TTCTACTGCAAGGATCAACGGCG-3'
	N-MattB2	5'-GGGGACCACTTTGTACAAGAAAGCTGG GTCTCAGGTAAACGTGTTTCGTTCCGA-3'
GFP-RT	N-RTattB1	5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTC CTTGCAGCAAACCTCTACAA-3'
	N-RTattB2	5'-GGGGACCACTTTGTACAAGAAAGCTGG GTCTCAGAAGAACTCGTCCCCTGTG-3'
GFP:IEPs	N-IEPsattB2	5'-GGGGACCACTTTGTACAAGAAAGCTGG GTCTCAGGTAAACGTGTTTCGTTCC-3'
GFP: $\Delta$ C29	N- $\Delta$ C29attB2	5'-GGGGACCACTTTGTACAAGAAAGCTGG GTCTCATGTCTTATGGGACTGAAAGCG-3'

**Tabla IV.1.** Cebadores usados para la construcción de las fusiones traduccionales que se han usado en la localización subcelular de la IEP en protoplastos de *A. thaliana*.

Usando como molde las construcciones pK7-NIEP, pK7-CIEP y pK7-NLSm se obtuvieron los plásmidos derivados de pKGEMA4: pKG4nGFP, pKG4cGFP y pKG4NLSm respectivamente, para su posterior análisis en movilidad y de localización intracelular en bacteria en el caso de las dos fusiones con GFP.

#### **IV.2.3. Ensayos de homing in vivo**

Los ensayos *in vivo* de movilidad en cepas de *S. meliloti* RMO17 fueron llevados a cabo mediante la introducción (por transferencia conjugativa triparental) de las construcciones donadoras de intrón en bacterias que ya contenían los plásmidos aceptores pJB0.6LAG y pJB $\Delta$ 129 (control negativo de la invasión que carece del sitio de inserción). Los ensayos de movilidad en *E. coli* y *P. putida* se llevaron a cabo mediante la introducción simultánea del plásmido donador y del aceptor por electroporación. Los subcultivos para el análisis de las distintas generaciones se realizaron mediante una dilución 1/10000 del cultivo crecido durante la noche, del cual se tomaba una muestra para la extracción del DNA plasmídico. El cálculo de las generaciones se realizó mediante la fórmula:  $n = \log \text{dilución} / \log 2$  donde  $n$  es el número de generaciones.

Los *Southern blots* se llevaron a cabo utilizando el DNA plasmídico obtenido mediante lisis alcalina y digerido con la enzima de restricción *SaII* (nos permite discriminar por tamaño entre los fragmentos provenientes de eventos de invasión de los que no lo son). Se hibridó con una sonda que reconoce la diana de RmInt1. Los valores de invasión se obtuvieron tras cuantificar mediante el programa informático Quantity One® (Biorad) las señales correspondientes tanto a la diana invadida como sin invadir.

#### **IV.2.4. Localización de las fusiones IEP-GFP**

La localización en bacteria se llevó a cabo depositando sobre un portaobjetos de vidrio 5 $\mu$ l de un cultivo saturado y examinando la fluorescencia mediante un microscopio Leica DMI6000 B y usando un objetivo de inmersión x100/1.40 con un filtro para GFP.

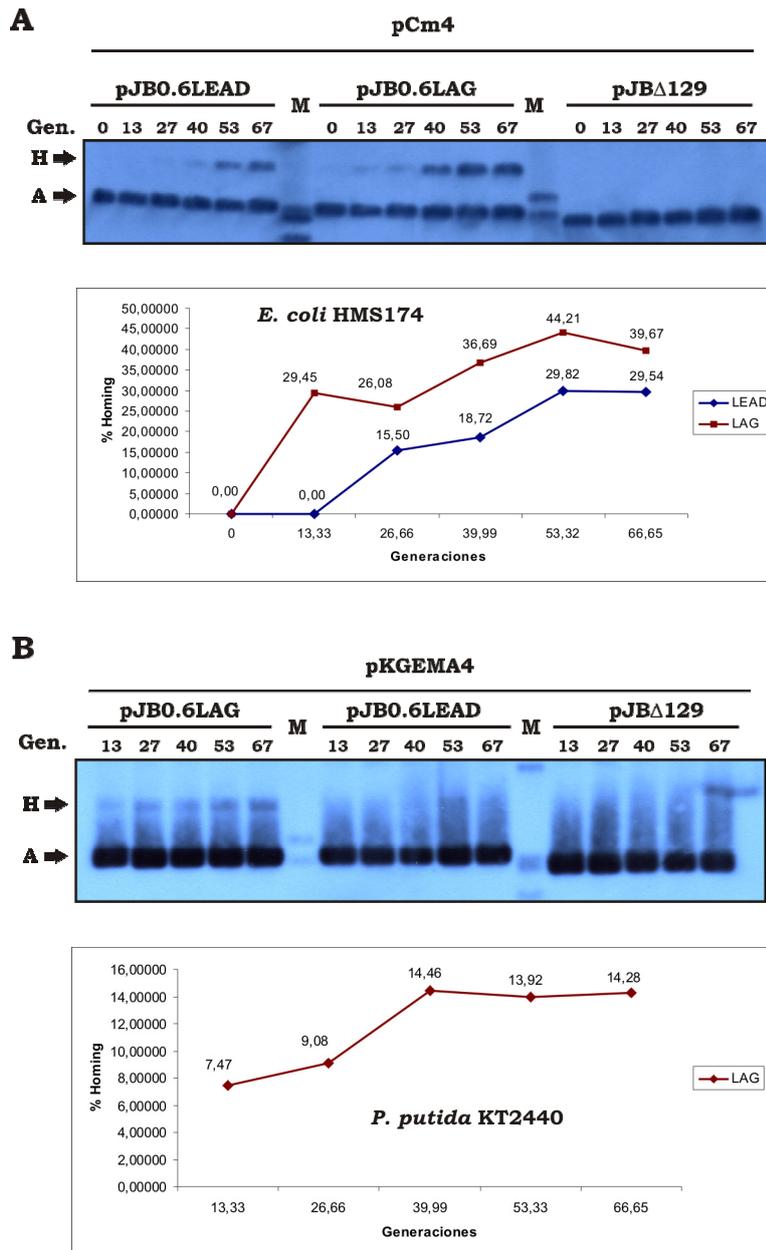
La localización en los protoplastos, obtenidos a partir de un cultivo celular mantenido de *A. thaliana* ColT87, se realizó mediante expresión transitoria por transformación con polietilenglicol (PEG; Mathur y Koncz, 1998). Tras 18 horas de expresión se observaron también en el microscopio Leica DMI6000 B pero con un objetivo de inmersión x63/1.40. Los protoplastos co-transformados con la construcción SRP34-RFP fueron observadas con el microscopio láser confocal Leica TCS SP (Leica Microsystems).

### IV.3. RESULTADOS

#### **IV.3.1. RmInt1 es móvil en *E. coli* y *P. putida*.**

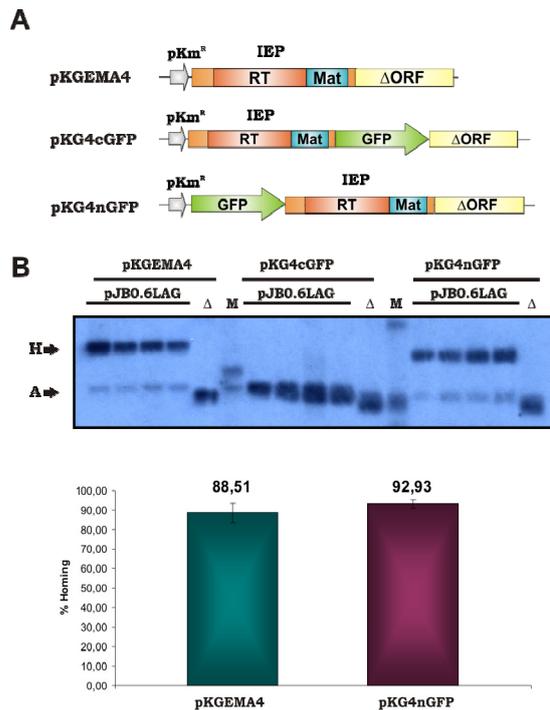
En nuestro grupo ha sido demostrado que el *retrohoming* de RmInt1 en *S. meliloti* está sujeto a una preferencia en cuanto a la orientación en la que se encuentre la diana libre respecto al sentido de la replicación, llegando a ser invadidas todas aquellas que se encuentren en el mismo sentido que la horquilla de replicación (denominadas LAG; Martínez-abarca *et al.*, 2004). Para comprobar si ocurría lo mismo tanto en *E. coli* como en *P. putida*, se realizaron sendos experimentos de *homing* utilizando como donador del intrón la construcción pCm4 en el caso de *E. coli* HMS174 y pKGEMA4 en el caso de *P. putida* KT2440. Como plásmidos aceptores fueron usados pJB0.6 LEAD, pJB0.6LAG y pJBΔ129 (control negativo). Una vez introducidos ambos plásmido (donador y aceptor) mediante electroporación, y tras realizar subcultivos de estas colonias durante casi 100 generaciones en el caso de *E. coli* y 70 en el caso de *P. putida*, se analizaron los eventos de invasión mediante *Southern blot*. Se detectó *homing* tanto en *E. coli* como en *P. putida*, pero en proporciones de invasión diferentes. En el caso de *E. coli* (Fig. IV.1.A) fuimos capaces de detectar y cuantificar el porcentaje de eventos de invasión en LEAD (hasta un 38,31%) y en LAG (hasta un 46,49%). En ambos casos se observa un incremento del porcentaje de invasión durante las primeras generaciones, alcanzándose una especie de plató en las últimas. También se observa una diferencia aproximada de un 10% de invasión durante todo el experimento entre los dos aceptores.

En el caso de *P. putida* KT2440 (Fig. IV.1.A), al igual que en *E. coli*, se detectó invasión en ambos tipos de plásmidos susceptibles de ser invadidos, pero en el caso del pJB0.6LEAD no se pudo cuantificar la invasión debido a que la sensibilidad de la técnica no fue suficiente. Se realizaron subcultivos hasta casi alcanzar las 70 generaciones, llegando a obtener un máximo de porcentaje de invasión del 14,46%. Al igual que en *E. coli*, se observó un incremento de la invasión en las primeras generaciones, alcanzándose un equilibrio entre los aceptores invadidos y los sin invadir durante las últimas generaciones.



**Figura IV.1.** Movilidad de RmInt1 en huéspedes heterólogos. Los valores de % *homing* se obtuvieron a partir de la fórmula  $[H/(H+A)] \times 100$ , donde H corresponde a la banda de la diana invadida y A es la diana sin invadir. En las gráficas se representan los valores obtenidos a partir de los *Southern blots* que se encuentran sobre ellas, mostrando en el eje Y los porcentajes de invasión y en el eje X las generaciones transcurridas (Gen.). **A:** Movilidad en *E. coli* HMS174: se ha podido detectar y cuantificar mediante *Southern blot* los eventos de invasión tanto en LEAD como en LAG. **B:** Movilidad en *P. putida* KT2440: aunque habiendo detectado invasión en ambas orientaciones, sólo se pudo cuantificar los valores correspondientes a la construcción pJB0.6LAG.

### IV.3.2. Construcción de fusiones IEP-GFP para su uso en bacteria



**Figura IV.2.** Análisis de la movilidad de las fusiones IEP-GFP. **A:** Construcción silvestre pKGEMA4 y fusiones de la IEP con GFP en C y N terminal (pKG4cGFP y pKG4nGFP respectivamente). Son expresadas desde el promotor del gen de resistencia a la kanamicina  $pkm^r$ . **B:** Southern blot de las construcciones en el que se observa que la fusión en C terminal no es activa, mientras que la N terminal sí lo es. En el histograma se representan los % de homing obtenidos de la cuantificación de este Southern blot. No existen diferencias significativas entre la construcción silvestre (pKGEMA4) y la construcción con la fusión en el N terminal de la IEP (pKG4nGFP).

Para estudiar la localización intracelular en bacteria de la IEP de RmInt1, se construyeron fusiones traduccionales de la GFP, tanto con el extremo C como N terminal de la IEP, en el contexto del plásmido donador pKGEMA4, obteniendo las construcciones pKG4cGFP y pKG4nGFP respectivamente (Fig. IV.2.A). Este vector contiene un derivado de RmInt1 tipo  $\Delta$ ORF con un tamaño de 0,75-kb, flanqueado por exones cortos -20/+5, y con la IEP expresada desde aguas arriba del exón 5'. Como control de la localización de la GFP sin fusionar se usó la construcción pKGFP, que carece tanto de la IEP como de la secuencia  $\Delta$ ORF del intrón.

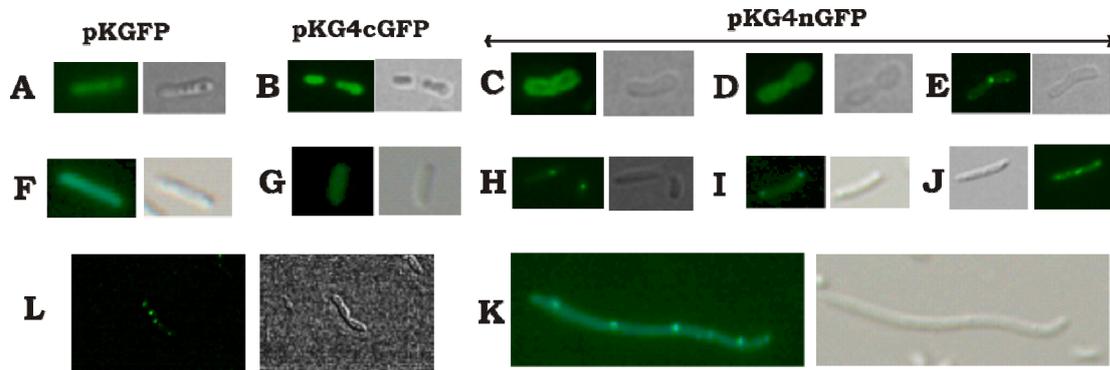
Para determinar si las construcciones eran activas, se llevó a cabo ensayos de movilidad de plásmido a plásmido en *S. meliloti* RMO17. Se puso de manifiesto que la fusión en el extremo C terminal de la IEP no es activa, ya que no se pudo detectar invasión, mientras que la fusión en el N terminal de la IEP permite la movilidad del derivado

RmInt1- $\Delta$ ORF al nivel silvestre (Fig. IV.2.B).

### IV.3.3. Localización intrabacteriana de las fusiones IEP-GFP

Se observó que cuando los cultivos bacterianos eran crecidos a 30°C la intensidad de fluorescencia de las fusiones era mayor que cuando eran crecidos a 37°C, por lo que en el caso de *E. coli* los experimentos fueron

llevados a cabo a la temperatura menor. Este hecho puede estar relacionado con una mayor estabilidad de la IEP a 30°C debido al nicho ecológico que ocupa su hospedador natural, *S. meliloti*.



**Figura IV.3.** Localización intracelular bacteriana de las fusiones IEP-GFP mediante microscopía de fluorescencia y campo claro tanto en *S. meliloti* (paneles A-E) como en *E. coli* (paneles F-L). El panel L muestra la localización en células en división de la fusión GFP:IEP en *E. coli* usando la técnica de deconvolución, (como alternativa al microscopio confocal) que nos permite poner de manifiesto determinados detalles.

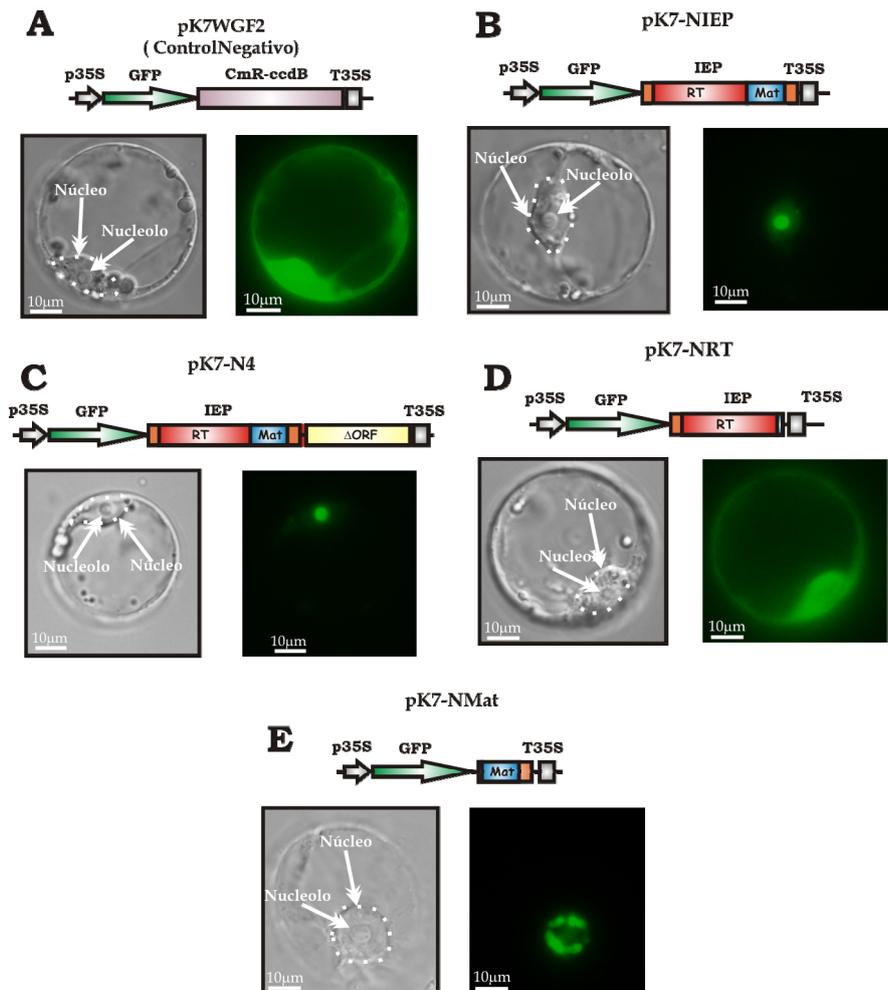
Se examinó la localización de las fusiones IEP-GFP en la cepa de *S. meliloti* RMO17 mediante microscopía de fluorescencia. Se observó que tanto las células que expresaban la GFP sola como fusionada al C terminal de la IEP mostraban una distribución homogénea de la fluorescencia en toda la célula (Fig. IV.3 A y B). En el caso de la fusión N- terminal GFP:IEP se observaron tres patrones diferentes: uno mayoritario (79% de las células) en el que la fluorescencia es semejante al control negativo (Fig. IV.3D); un patrón secundario en el que la fluorescencia se concentraba de forma preferente adyacente a la pared celular (19%) (Fig. IV.3C); y, en mucha menor proporción, se observaron casos de localización de la fluorescencia en un sólo punto central (2%) (Fig. IV.3E).

Cuando se analiza la localización en *E. coli* DH5 $\alpha$  se observa que, al igual que en el caso anterior, la construcción control (Fig. IV.3F) y la fusión C terminal (Fig. IV.3G) muestran el mismo patrón de difusión de la fluorescencia por toda la célula. Cuando la GFP es fusionada en el extremo N terminal de la IEP aparecen células con un sólo foco de localización de la fluorescencia (73%) (Fig. IV.3H), pero a diferencia de lo que ocurre en *S. meliloti*, existen células en las que se puede apreciar dos (Fig. IV.3I) o más (Fig. IV.3J y K) focos de fluorescencia (29%), pudiendo pertenecer a células en división como se observa tras el tratamiento de la imagen por deconvolución de una de las células (Fig. IV.3L). No se aprecian bacterias con una distribución de la fluorescencia adyacente a la pared celular.

#### IV.3.4 Localización subcelular en protoplastos de *A. thaliana*

Estos experimentos fueron llevados a cabo en el laboratorio perteneciente al C.N.R.S. (Centre National de la Recherche Scientifique) liderado por el Doctor Martín Crespi.

Debido a los resultados obtenidos anteriormente, los experimentos de localización se llevaron a cabo con fusiones traduccionales de la GFP en el extremo N terminal de la IEP, tanto silvestre como modificada por determinadas mutaciones y deleciones.



**Figura IV.4.** Los paneles muestran imágenes de microscopía de fluorescencia y en campo claro. Sobre cada uno de ellos se dispone un esquema de la construcción que se ha sido usada para la expresión transitoria de las fusiones en protoplastos de *A. thaliana*. En las imágenes en campo claro se indican las estructuras subcelulares correspondientes tanto al núcleo (delimitado por una línea de puntos) como al

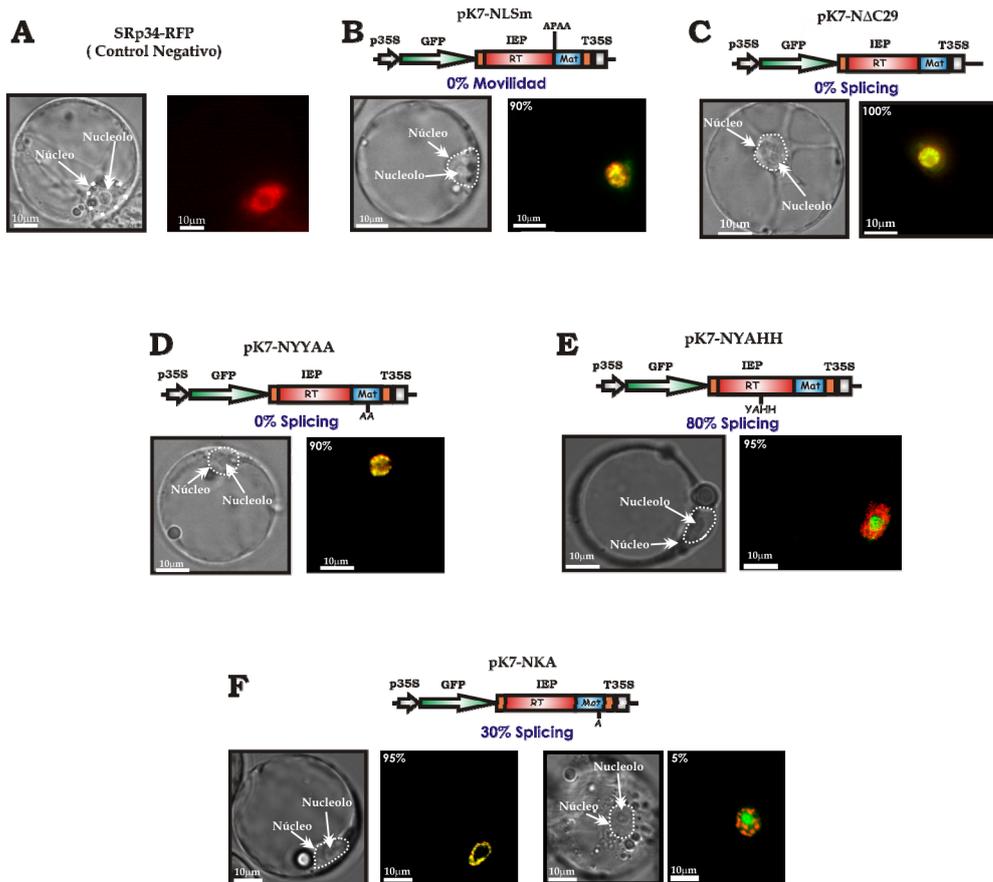
nucleolo (estructura en forma de botón que se encuentra dentro del núcleo). En el panel A se muestra el patrón de fluorescencia de la construcción control pK7WGF2.

Como control se usó el vector pK7WGF2 (Fig. IV.4A), donde posteriormente se construyeron todas las fusiones, que muestra un patrón de la fluorescencia difuso por todo el citoplasma y el núcleo del protoplasto. Cuando fusionamos la IEP (pK7-NIEP) vemos como la GFP pasa a tener una localización nucleolar (Fig. IV.4B), hecho que no cambia si desde el mismo plásmido se expresa el RNA del intrón (pK7-N4; Fig. IV.4C). Decidimos dividir la proteína y fusionar por separado a la GFP la secuencia de DNA correspondiente a sus dos dominios funcionales. En el caso de la fusión GFP:RT (pK7-NRT; Fig. IVD) la localización muestra un patrón igual que el del control, mientras que si lo que fusionamos es el dominio madurasa (pK7-NMat) se ve una localización nuclear formada por un patrón punteado, pero se pierde la localización nucleolar (Fig. IV.4E). Mediante el programa de predicción de localización subcelular PSORT se realizó una búsqueda de posibles motivos que pudiesen actuar como una señal de localización nuclear (NLS). Se obtuvo como resultado que en la posición aminoacídica 291 de la IEP (en el inicio del dominio madurasa) existe una secuencia de cuatro aminoácidos (RPRR) que podrían actuar como una NLS monopartita. Tras mutar las argininas (R) por alaninas, se observó que, si bien se perdía la localización nucleolar, la fluorescencia permanecía restringida en el núcleo (Fig. IV.5B), por lo que el NLS putativo no era tal. Ensayos de *homíng* mostraron que en presencia de esta IEP mutada el intrón no era móvil (datos no mostrados).

Se analizaron determinados mutantes de la IEP para determinar cómo podían afectar a la localización de esta proteína en protoplastos de *A. thaliana*. Para determinar donde localizan se realizaron experimentos de co-localización con la proteína de mamíferos SRP34 (con un patrón de localización similar al de alguna de nuestras construcciones) fusionada con RFP y que localiza en los *speckles*<sup>1</sup> (Fig. IV.5A; Lorkovic *et al.*, 2004). El mutante  $\Delta$ C29 (pK7-N $\Delta$ C29; Fig. IV.5C) tiene delecionados los últimos 29 aminoácidos del extremo C-terminal de la IEP provocando la inhibición del *splicing* (Molina-Sánchez *et al.*, 2010), y como se puede apreciar en la figura co-localiza con la proteína SPR34 en los *speckles*. El mutante YYAA de la madurasa (pK7-NYYAA; Fig. IV.5D), que tampoco tiene *splicing* (Molina-Sánchez *et al.*, 2010), también localiza en los *speckles*, al igual que el mutante en el NLS putativo (pK7-NLSm). Cuando analizamos el mutante en el sitio activo del dominio RT (pK7-NYAAH; que conserva un 80% de *splicing*)

---

<sup>1</sup> Estructuras subnucleares ricas en factores de *splicing* y que se encuentran en el nucleoplasma.



**Figura IV.5.** Co-localización entre la fusión SRP34-RFP (que localiza en los *speckles*) y fusiones de la GFP con determinados mutantes de la IEP. En el panel A se muestra la localización de la proteína SRP34 usada como control. Sobre cada imagen se encuentra representado un esquema de la construcción con GFP usada en la co-localización.

observamos que, en el 95% de los protoplastos analizados sigue manteniendo la localización nucleolar (Fig. IV.5E). El mutante en el residuo conservado K381 (pK7-NKA) de la madurasa (que tiene un 30% de *splicing*) muestra una pérdida de la localización nucleolar en el 95% de los protoplastos analizados (Fig. IV.5F) mientras que en el 5% restante sí se apreció una localización nucleolar.

#### IV.4. DISCUSIÓN

En este capítulo se ha estudiado la movilidad del derivado  $\Delta$ ORF de RmInt1 en hospedadores diferentes al natural (*S. meliloti*), como son *E. coli* HMS174 (DE3) Y *P. putida* KT2440, con el objeto de determinar su uso como

herramienta genética de mutación insercional en distintos organismos bacterianos. Mediante los ensayos de movilidad *in vivo* hemos podido detectar y cuantificar mediante *Southern blot* los eventos de *homing* que se han producido en estos fondos genéticos, aunque con resultados diferentes. Los experimentos llevados a cabo en *E. coli* han demostrado que RmInt1 en su versión  $\Delta$ ORF tiene capacidad de movimiento en esta bacteria, siendo capaz de invadir dianas en las dos orientaciones respecto a la replicación (LAG y LEAD). El porcentaje de invasión de las dianas se ve incrementado con el aumento del número de generaciones, tanto en LAG como en LEAD, pero ocurre que en torno a la generación 54 parece establecerse un equilibrio entre las formas invadidas y las sin invadir sin llegar a alcanzar (durante las generaciones estudiadas) la plena invasión. Esto contrasta con la dinámica de invasión en dianas LAG que se da en *S. meliloti*, en la cual puede llegar a alcanzarse una invasión del 100% de las dianas en LAG en muy pocas generaciones, mientras que en el caso de las dianas en LEAD sí se alcanza este equilibrio, obteniéndose unos niveles de invasión similares en ambos hospedadores. En el caso de *P. putida* aunque si bien hemos podido detectar invasión en dianas en LEAD, el porcentaje de *homing* no ha podido ser cuantificado. Sí hemos podido cuantificar el nivel de invasión en LAG, obteniéndose hasta casi un 15%, observando la misma tendencia que en *E. coli*, pero con unos niveles reducidos a menos de la mitad de lo obtenido en *E. coli*. Estos datos nos indican que, aunque la capacidad de movimiento de los intrones del grupo II depende mínimamente del hospedador, es evidente que deben existir determinados factores o condiciones que hacen que la capacidad de movimiento de RmInt1 sea mejor en su hospedador natural que en otras bacterias. Los datos certifican que, gracias a su capacidad móvil, RmInt1 podría ser usado en hospedadores bacterianos heterólogos como herramienta de mutagénesis insercional.

Debido a que los intrones del grupo II cuya IEP carece de dominio endonucleasa dependen del uso de las cadenas nacientes en la horquilla de replicación como cebadores de la transcripción reversa del RNA del intrón insertado, y la replicación del material genético está ligado a la división celular, se investigó mediante microscopía de fluorescencia si podría existir una localización intracelular preferencial de la IEP (que forma parte de las RNPs) al igual que ocurre en el caso del intrón de *L. lactis* Ll.LtrB. Para ello se realizaron fusiones traduccionales IEP-GFP. Al determinar si estas fusiones seguían permitiendo la movilidad de RmInt1 se observó mediante *Southern blot* que la fusión en el extremo N terminal sí seguía siendo activa, incluso a niveles silvestres, mientras que la fusión en el extremo C terminal

inhibía el movimiento del intrón, lo que hace pensar que esta fusión podría causar un plegamiento incorrecto de la IEP u otro tipo de impedimentos que hagan que esta construcción no sea móvil. Cuando se analiza en *S. meliloti* RMO17 la localización de la fusión IEP:GFP se observa que muestra el mismo patrón difuso por toda la célula que el control de la GFP sin fusionar. La construcción pKG4nGFP, aun ofreciendo en algunas de las células el mismo patrón que el control, también muestra otros dos patrones diferentes: un patrón mayoritario en el que la fluorescencia se acumula bordeando la pared celular, lo que nos podría hacer pensar que interacciona con algún factor sin determinar que hace que se acumule en esta localización; el otro patrón es la aparición de un solo foco de localización de fluorescencia. Cuando usamos esta última construcción para la localización en *E. coli*, aparecen células con un mayor número de focos de fluorescencia, situándose en los polos de las células de forma similar a lo que ocurre en estudios realizados con el intrón de *L. lactis* Ll.LtrB (Zhao y Lambowitz., 2005), en los cuales se observa que su IEP (LtrA) también localiza en estas regiones, concluyéndose que el que la IEP localice en estos polos se encuentra relacionado con la migración del nucleóide hacia los extremos de la célula durante el proceso de división. Otro factor que puede contribuir a la localización polar de la IEP es que, al depender RmInt1 de la existencia de regiones del DNA en monohebra para poder insertarse debido a que carece de actividad endonucleasa, en los polos de las células en división es donde se va a localizar la región Ori del cromosoma (Khang y Shapiro, 2003), siendo más probable encontrar este tipo de hebras de DNA. Por todo esto se propone que, debido a que la división celular en *S. meliloti* es más lenta (se divide cada dos horas aproximadamente), es más improbable encontrar localización en focos de la fluorescencia, mientras que *E. coli*, al tener una mayor dinámica de división favorece la aparición de focos polares de fluorescencia.

Para poder utilizar RmInt1 en *A. thaliana*, debido a la compartimentación subcelular existente en la célula eucariota, es importante determinar dónde se localiza la IEP. Hemos determinado que la fusión GFP:IEP se localiza en el nucleolo de la célula por sí sola o desde construcciones que también expresan el RNA del intrón. Esta localización cambia cuando lo que se fusiona a la GFP son los dos dominios principales de los que consta la IEP: el dominio RT por sí solo no influye en la localización específica de la fluorescencia, difundiéndose por toda la célula excepto por la vacuola, mientras que el dominio madurasa contribuye a una localización nuclear de la fluorescencia, pero no nucleolar. Podemos concluir

que el RNA del intrón no afecta a la localización nucleolar de la IEP. También podemos decir que la fusión GFP:IEP localiza en el núcleo gracias a alguna señal presente en el dominio madurasa y que no se corresponde con la Señal de Localización Nuclear (NLS\*) predicha bioinformáticamente que además no actúa como tal ya que cuando es mutada la IEP permanece en el núcleo. Otra conclusión que obtenemos de estos resultados es que la presencia del dominio RT es necesaria para la localización nucleolar de la IEP.

Cuando la localización se realiza usando diferentes mutantes de la IEP (por deleciones o sustituciones) vemos que las construcciones que conservan actividad (pK7-KA y pK7-NYAAH) muestran localización nucleolar de la IEP en mayor o menor grado, mientras que aquellas en las que la movilidad o el *splicing* se ven comprometidos (pK7-NLSm, pK- $\Delta$ C29, pK7-YYAA) siguen localizados en el núcleo, concretamente en los llamados *speckles*, pero no en el nucleolo.

Aunque la localización de la IEP en el nucleolo pudiera ser un mecanismo de defensa y secuestro por parte de la célula vegetal, cuando se analiza en la literatura la clasificación de las proteínas presentes en el nucleolo de *A. thaliana* (Brown et al., 2005) observamos que el mayor porcentaje pertenece a proteínas ribosomales ya que es aquí donde se procesan y ensamblan los ribosomas. Menos representadas se encuentran otras clases proteicas, pero todas ellas relacionadas con el procesamiento de RNA. También se ha visto que otras transcriptasa reversas (como la de la telomerasa humana) también se localizan en el nucleolo de las células humanas (Lin et al., 2008) y que esta localización se encuentra relacionada con su función (Yang et al., 2002). Estos datos, junto con nuestros resultados, nos llevan a lanzar la hipótesis de que, en el caso de la IEP de RmInt1, su funcionalidad y capacidad de unión a RNA están íntimamente relacionadas con la localización subcelular en protoplastos de *A. thaliana*.

## **CONCLUSIONES**



1. Las modificaciones introducidas en RmInt1 tales como, deletar la mayor parte de la secuencia que codifica la IEP del dominio IV para ser expresada en *cis* aguas arriba o abajo de la secuencia codificante del RNA del intrón (intrón  $\Delta$ ORF) y recortar la longitud de los exones, hacen a este intrón más eficiente en su movilidad.
2. RmInt1 tiene una gran capacidad de dispersión, fundamentalmente hacia a su diana natural. La ruta de dispersión en el genoma se encuentra ligada a la replicación y se inserta de forma preferente en la hebra que sirve de molde para la hebra retardada. Las nuevas copias de intrón adquiridas retienen la capacidad de movimiento, contribuyendo a su posterior dispersión. Cuando las copias genómicas de intrón  $\Delta$ ORF son complementadas con una IEP funcional, aún presentan capacidad de sufrir *retrohoming*.
3. De los datos obtenidos mediante la técnica de microarrays concluimos que la presencia de nuevas copias de RmInt1 en el genoma de *S. meliloti* tiene efectos muy reducidos sobre su transcriptoma, pudiendo verse debidos tanto a la posición de la inserción en el genoma como a la expresión de éste elemento genético móvil.
4. RmInt1 necesita de determinadas funciones celulares generales (polimerasas, helicasas, exonucleasas) para llevar a cabo su movilidad y completar el proceso de *retrohoming*. Por el contrario, otras funciones (como determinadas RNAsas) interfieren de forma negativa en la movilidad de RmInt1, haciendo más difícil su dispersión.
5. Se ha desarrollado un sistema de selección de eventos de *retrohoming* usando como marcador la resistencia al antibiótico eritromicina y aprovechando la capacidad de RmInt1- $\Delta$ ORF de transportar en su dominio IV información genética externa. Esta herramienta podrá ser utilizada para hacer más eficiente el método de selección de mutantes obtenidos mediante la modificación de las secuencias de reconocimiento de la diana de RmInt1.
6. RmInt1 es móvil en hospedadores diferentes a *S. meliloti* (*E. coli* y *P. putida*). Su capacidad de movimiento se ve reducida, indicando que puede existir una adaptación al fondo genético de *S. meliloti*. La capacidad de dispersión de RmInt1 en hospedadores heterólogos amplía la posibilidad de utilización de RmInt1 como herramienta de mutagénesis insercional.

7. Cuando fusionamos al extremo N-terminal de la IEP al gen reportero de la proteína fluorescente verde (GFP) podemos observar la localización subcelular de esta proteína. En *S. meliloti* RMO17 no se distribuye igual en todas las células, siendo la localización preferente alrededor de la pared celular. En *E. coli* DH5 $\alpha$  el patrón de distribución es diferente, siendo más abundantes las células que presentan una distribución en focos de fluorescencia.

8. El análisis de la localización de las fusiones traduccionales GFP:IEP en protoplastos de *A. thaliana* nos ha permitido determinar que la IEP se localiza en el nucleolo de estas células. Los estudios realizados con IEPs delecionadas nos permite decir que la localización nuclear de esta proteína depende de la secuencia del dominio madurasa, mientras que el dominio RT es determinante para la localización nucleolar de la IEP. Estos resultados junto con los obtenidos con mutantes funcionales de la IEP nos permiten aventurar la siguiente hipótesis: la localización subcelular de la IEP de RmInt1 podría estar relacionada con su capacidad de unirse a RNA y con su funcionalidad.

# **BIBLIOGRAFÍA**





## BIBLIOGRAFÍA

**Abelson, J., Trotta, C.R., y Li, H.** (1998). tRNA splicing. *J. Biol. Chem.* **273**; 12685–12688.

**Abramovitz DL, Friedman RA and Pyle AM.** 1996. Catalytic role of 2'-hydroxyl groups within a group II intron active site. *Science* **271**:1410–1413.

**Aizawa, Y., Xiang, Q., Lambowitz, A.M., y Pyle, A.M.** (2003). The pathway of DNA recognition and RNA integration by a group II intron retrotransposon. *Mol. Cell* **11**; 795–805.

**Asakura, Y. y Barkan, A.** (2006) *Arabidopsis* orthologs of maize chloroplast splicing factors promote splicing of orthologous and species-specific group II introns. *Plant Phys.*, **142**: 1656-1663.

**Altschul, S.F., Gish, W., Miller, W., Myers, E.W. y Lipman, D.J.** (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**(3); 403-410.

**Asakura, Y. y Barkan, A.** (2007) A CRM domain protein functions dually in group I and group II intron splicing in land plant chloroplasts. *Plant Cell*, **19**: 3864-75.

**Asakura, Y., Bayraktar, O.A. y Barkan, A.** (2008) Two CRM protein subfamilies cooperate in the splicing of group IIB introns in chloroplasts. *RNA*, **14**:2319-32.

**Augustin, S., Mueller, M.W., y Schweyen, R.J.** (1990). Reverse self-splicing of group II intron RNAs *in vitro*. *Nature* **343**; 383–386.

**Bae, K., Kwon, Y. D., Shin, H., Hwang, M., Ryu, E., Park, K., et al.** (2003). Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nature Biotechnology*, **21**(3), 275-280.

**Ban, N., Nissen, P., Hansen, J., Moore, P.B., Steitz, T.A.** (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**; 905–920.

**Bagdasarian, M., Lurz, R., Ruckert, B., Franklin, F. C., Bagdasarian, M. M., Frey, J., y Timmis K. N.** (1981). Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene* **16**:237-47.

**Barkan, A.** (2004). Intron splicing in plant organelles. En *Molecular Biology and Biotechnology of Plant Organelles*, H. Daniell and C. Chase, eds (Dordrecht, The Netherlands: Kluwer Academic Publishers), pp. 281–308.

**Barkan, A., Klipcan, L., Ostersetzer, O., Kawamura, T., Asakura, Y. y Watkins, K.** (2007) The CRM domain: an RNA binding module derived from an ancient ribosome-associated protein. *RNA*, 13:55-64.

**Barkan A.** (2009). Genome-wide analysis of RNA-protein interactions in plants. *Methods Mol. Biol.* **553**: 13–37.

**Barrientos-Durán, A.M.** (2008) Bases moleculares para la aplicación biotecnológica del intrón del grupo II RmInt1 de *Sinorhizobium meliloti*. Tesis doctoral. Universidad de Granada.

**Barrientos-Duran A, Chillon I, Martinez-Abarca F, Toro N.** (2011) Exon sequence requirements for excision in vivo of the bacterial group II intron RmInt1. *BMC Mol Biol.* **12(1)**:24.

**Beringer, J.E.** (1974) R factor transfer in *Rhizobium leguminosarum*. *J. Gen. Microbiol.* **84**: 188-198.

**Beam, C. E., Saveson, C. J. y Lovett S. T.** (2002) Role for *radA/sms* in Recombination Intermediate Processing in *Escherichia coli*. *Journal of Barteriology* **184**: 6836–6844

**Beauregard A, Chalamcharla VR, Piazza CL, Belfort M, Coros CJ.** (2006). Bipolar localization of the group II intron Ll.LtrB is maintained in *Escherichia coli* deficient in nucleoid condensation, chromosome partitioning and DNA replication. *Mol. Microbiol.* **62**; 709–22.

**Beauregard, A., Curcio, M.J., and Belfort, M.** (2008). The take and give between retrotransposable elements and their hosts. *Annu. Rev. Genet.* **42**; 587–617.



---

**Belfort, M., Derbyshire, V., Parker, M.M., Cousineau, B. y Lambowitz, A.M.** (2002) Mobile Introns: Pathways and Proteins. *Mobile DNA II*. pp 761-781. ASM Press, Washington, D.C.

**Belhocine, K., Plante, I. y Cousineau, B.** (2004) Conjugation mediates transfer of the Ll.ltrB group II intron between different bacterial species. *Mol. Microbiol.*, **51**; 1459-1469.

**Belhocine, K., Yam, K.K., y Cousineau, B.** (2005) Conjugative transfer of the *Lactococcus lactis* chromosomal sex factor promotes dissemination of the Ll.LtrB group II intron. *J Bacteriol* **187**: 930-939.

**Benton, B. M., J. P. Zhang, S. Bond, C. Pope, T. Christian, L. Lee, K. M. Winterberg, M. B. Schmid y J. M. Buysse** (2004) Large-scale identification of genes required for full virulence of *Staphylococcus aureus*. *J Bacteriol*, 186, 8478- 8489.

**Biondi, E. G., Toro, N., Bazzicalupo, M. y Martínez-Abarca, F.** (2011) Spread of the group II intron RmInt1 and its insertion sequence target sites in the plant endosymbiont *Sinorhizobium meliloti* **1(1)**: 1-6.

**Birnboim, H.C. y Doly, J.** (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* **7**: 1513-1524.

**Blocker, F.J.H., Mohr, G., Conlan, L.H., Qi, L., Belfort, M. y Lambowitz, A.M.** (2005) Domain structure and three-dimensional model of a group II intron-encoded reverse transcriptase. *RNA*, **11**; 14-28.

**Bock-Taferner, P., y Wank, H.** (2004) GAPDH enhances group II intron splicing in vitro. *Biol. Chem.* **385**; 615-621.

**Bonen L.** (1993) *Trans*-splicing of pre-mRNA in plants, animals and protists. *FASEB J*, **7**: 40-46.

**Bonen, L. y Vogel, J.** (2001). The ins and outs of group II introns. *Trends Genet.* **17**; 322-331.

**Bonen, L.** (2008) *Cis-* and *trans*-splicing of group II introns in plant mitochondria. *Mitochondrion*, **8(1)**; 26-34.

**Boudvillain, M. y Pyle, A.M.** (1998). Defining functional groups, core structural features and inter-domain tertiary contacts essential for group II intron self-splicing: a NAIM analysis. *EMBO J.* **17**; 7091–7104.

**Boudvillain, M., de Lencastre, A., y Pyle, A.M.** (2000). A tertiary interaction that links active-site domains to the 5' splice site of a group II intron. *Nature* **406**; 315–318.

**Boulanger, S.C., Belcher, S.M., Schmidt, U., Dib-Hajj, S.D., Schmidt, T., y Perlman, P.S.** (1995). Studies of point mutants define three essential paired nucleotides in the domain 5 substructure of a group II intron. *Mol. Cell. Biol.* **15**; 4479–4488.

**Bowater, R. y Doherty, A.J.** (2006). Making ends meet: Repairing breaks in bacterial DNA by non-homologous end-joining. *PLoS Genet* **2**: e8.

**Bowers, P.M., Pellegrini, M., Thompson, M.J., Fierro, J. Yeates, T.O. y Eisenberg, D.** (2004) Prolinks: a database of protein functional linkages derived from coevolution. *Genome Biol.* **5**; R35.

**Brown, J.W.S., Shaw, P.J., Shaw, P. and Marshall, D.F.** (2005). The Arabidopsis nucleolar protein database. *Nucleic Acids Research* **33**: 633-636.

**Bunse, A., Nickelsen, J. y Kück, U.** (2001) Intron-specific RNA binding proteins in the chloroplast of the green alga *Chlamydomonas reinhardtii*. *Biochem Biophys Acta* **1519**: 46-54

**Burge, C.B., Tuschl, T., y Sharp, P.A.** (1999). Splicing of precursors to mRNAs by the spliceosomes. **In:** *The RNA World, 2nd edition*. pp. 525–560, Cold Spring Harbor Laboratory Press, New York.

**Burger, G., Yan, Y., Javadi, P. y Lang, B.F.** (2009). Group I-intron transsplicing and mRNA editing in the mitochondria of placozoan animals. *Trends Genet.* **25**: 381–386.



---

**Burt, A.** (2003) Site-specific selfish genes as tools for the control and genetic engineering of natural populations. *Proc. R. Soc. Lond. B*, **270**; 921-928.

**Carignani, G., Groudinsky, O., Frezza, D., Schiavon, E., Bergantino, E., y Slonimski, P.P.** (1983). An mRNA maturase is encoded by the first intron of the mitochondrial gene for the subunit I of cytochrome oxidase in *S. cerevisiae*. *Cell* **35**; 733-742.

**Casadesús, J. y Olivares, J.** (1979) Rough and fine linkage mapping of the *Rhizobium meliloti* chromosome. *Mol. Gen. Genet.* **174**; 203-209.

**Casse, F., C. Boucher, J. S. Julliot, M. Michell, and J. Dénarié.** (1979). Identification and characterization of large plasmids in *Rhizobium meliloti* using agarose gel electrophoresis. *J. Bacteriol.* **113**: 229-242.

**Cathomen, T., y Joung, J. K.** (2008). Zinc-finger nucleases: the next generation emerges. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, **16(7)**, 1200-1207.

**Cavalier-Smith T** (1991) Intron phylogeny: a new hypothesis. *Trends Genet* **7**; 145-148.

**Cabanes, D., Boistard, P. y Batut, J.** (2000) Identification of *Sinorhizobium meliloti* genes regulated during symbiosis. *J. Bacteriol.*, **182** (13); 3632-3637.

**Cech, T.K.** (1986) The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell*, **44**; 207-210.

**Chalamcharla, V.R., Curcio, M.J. y Belfort, M** (2010) Nuclear expression of a group II intron is consistent with spliceosomal intron ancestry. *Genes & Dev.*, **24**: 827 - 836.

**Chandrasegaran, S., y Smith, J.** (1999). Chimeric restriction enzymes: what is next? *Biological Chemistry*, **380(7-8)**, 841-848.

**Chanfreau, G. y Jacquier, A.** (1994). Catalytic site components common to both splicing steps of a group II intron. *Science*, **266**; 1383-1387.

**Chanfreau, G. y Jacquier, A.** (1996). An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO J.* **15**; 3466–3476.

**Chen, Y., McClane, B. A., Fisher, D.J., Rood, J. I., y Gupta, P.** (2005) Construction of an alpha toxin gene knockout mutant of *Clostridium perfringens* type A by use of a mobile group II intron. *Appl Environ Microbiol* **71**; 7542–7547.

**Chillón I, Martínez-Abarca F, Toro N.** (2011) Splicing of the *Sinorhizobium meliloti* RmInt1 group II intron provides evidence of retroelement behavior. *Nucleic Acids Res.* **39(3)**:1095-104

**Chin, K. y Pyle, A.M.** (1995). Branch-point attack in group II introns is a highly reversible transesterification, providing a potential proofreading mechanism for 5'-splice site selection. *RNA* **1**; 391–406.

**Choquet Y., Goldschmidt-Clermont M., Girard-Bascou J., Kuck U., Bennoun P. y Rochaix J. D.** (1988) Mutant phenotypes support a trans-splicing mechanism for the expression of the tripartite *psaA* gene in the *C. reinhardtii* chloroplast. *Cell*, **52**: 903-13.

**Christopher, D.A. y Hallick, R.B. (1989).** *Euglena gracilis* chloroplast ribosomal protein operon: a new chloroplast gene for ribosomal protein L5 and description of a novel organelle intron category designated group III. *Nucleic Acids Res.* **17**; 7591–7608.

**Chu, V.T., Liu, Q., Podar, M., Perlman, P.S., y Pyle, A.M.** (1998). More than one way to splice an RNA: branching without a bulge and splicing without branching in group II introns. *RNA* **4**; 1186–1202.

**Chu, V.T., Adamidi, C., Liu, Q., Perlman, P.S., y Pyle, A.M.** (2001). Control of branch-site choice by a group II intron. *EMBO J.* **20**; 6866–6876.

**Conlan L.H., Stanger M.J., Ichiyanagi K., y Belfort M.** (2005) Localization, mobility and fidelity of retrotransposed group II introns in rRNA genes. *Nucleic Acids Res* **33**: 5262-5270.



---

**Copertino, D.W. y Hallick, R.B.** (1993). Group II and group III introns of twintrons: potential relationship with nuclear pre-mRNA introns. *Trends Biochem. Sci.* **18**; 467–471.

**Cornu, T. I., Thibodeau-Beganny, S., Guhl, E., Alwin, S., Eichinger, M., Joung, J. K., et al.** (2008). DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, **16(2)**, 352-358.

**Coros CJ, Landthaler M, Piazza CL, Beauregard A, Esposito D, et al.** (2005). Retrotransposition strategies of the *Lactococcus lactis* Ll.LtrB group II intron are dictated by host identity and cellular environment. *Mol. Microbiol.* **56**; 509–24.

**Coros, C.J., Piazza, C.L., Chalamcharla, V.R., y Belfort, M.** (2008). A mutant screen reveals RNase E as a silencer of group II intron retromobility in *Escherichia coli*. *RNA* **14**; 2634–2644.

**Coros, C.J., Piazza, C.L., Chalamcharla, V.R., Smith D., y Belfort, M.** (2009). Global regulators orchestrate group II intron retromobility. *Mol. Cell* **34**; 250-256.

**Cosseau, C., Garnerone, A. M. y Batut, J.** (2002) The FixM Flavoprotein Modulates Inhibition by AICAR or 5'AMP of Respiratory and Nitrogen Fixation Gene Expression in *Sinorhizobium meliloti*. *MPMI* **15(6)** 598–607.

**Costa, M. y Michel, F.** (1995). Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.* **14**; 1276–1285.

**Costa, M., Fontaine, J.M., Loiseaux-de Goer, S., y Michel, F.** (1997a). A group II self-splicing intron from the brown alga *Pylaiella littoralis* is active at unusually low magnesium concentrations and forms populations of molecules with a uniform conformation. *J. Mol. Biol.* **274**; 353–364.

**Costa, M., Deme, E., Jacquier, A., y Michel, F.** (1997b). Multiple tertiary interactions involving domain II of group II self-splicing introns. *J. Mol. Biol.* **267**; 520–536.

**Costa, M., Michel, F., y Westhof, E.** (2000). A three-dimensional perspective on exon binding by a group II self-splicing intron. *EMBO J.* **19**; 5007–5018.

**Costa, M., Michel, F., Molina-Sanchez, M.D., Martinez-Abarca, F., y Toro, N.** (2006a). An alternative intron-exon pairing scheme implied by unexpected *in vitro* activities of group II intron RmInt1 from *Sinorhizobium meliloti*. *Biochimie* **88**; 711–717.

**Costa, M., Michel, F., y Toro, N.** (2006b). Potential for alternative intron-exon pairings in group II intron RmInt1 from *Sinorhizobium meliloti* and its relatives. *RNA* **12**; 338–341.

**Cousineau B, Smith D, Lawrence-Cavanagh S, Mueller JE, Yang J, et al.** (1998). Retrohoming of a bacterial group II intron: mobility via complete reverse splicing, independent of homologous DNA recombination. *Cell* **94**; 451–62

**Cousineau B, Lawrence S, Smith D, Belfort M.** (2000). Retrotransposition of a bacterial group II intron. *Nature* **404**; 1018–21. Erratum. 2001 *Nature* **414**; 84.

**Cousineau, B., Lawrence, S., Smith, D., y Belfort, M.** (2001). Retrotransposition of a bacterial group II intron (Erratum). *Nature*, **414**; 84.

**Crick, F.H.C.** (1968). The origin of the genetic code, *J. Mol. Biol.* **38**; 367–379.

**Cui, X., Matsuura, M., Wang, Q., Ma, H. y Lambowitz, A.M.** (2004) A group II intron-encoded maturase functions preferentially in *cis* and requires both the reverse transcriptase and X domains to promote RNA splicing. *J Mol Biol*, **340**; 211–231.

**Cui, X. y Davis, G.** (2007) Mobile group II intron targeting: applications in prokaryotes and perspectives in eukaryotes. *Frontiers in Bioscience*, **12**; 4972-4985.

**Curcio, M.J. y Belfort, M.** (1996). Retrohoming: cDNA mediated mobility of group II introns requires a catalytic RNA. *Cell* **84**; 9–12.



---

**Dai, L. y Zimmerly, S.** (2002). Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior. *Nucleic Acids Res.* **30**; 1091–1102.

**Dai L, Toor N, Olson R, Keeping A, Zimmerly S.** (2003). Database for mobile group II introns. *Nucleic Acids Res.* **31**; 424–26

**Dai, L. y Zimmerly, S.** (2003). ORF-less and reversetranscriptase-encoding group II introns in archaeobacteria, with a pattern of homing into related group II intron ORFs. *RNA* **9**; 14–19.

**Dai, L., Chai D Gu, S.Q., Gabel, J., Noskov, S.Y., Blocker, F.J., Lambowitz, A.M. y Zimmerly, S.** (2008). A three-dimensional model of a group II intron RNA and its interaction with the intron-encoded reverse transcriptase. *Molecular Cell* **30**: 472–485.

**Dam, M., Douthwaite, S., Tenson, T. y Mankin, A. S.** (1996). Mutations in domain II of 23S rRNA facilitate translation of a 23S rRNA-encoded pentapeptide conferring erythromycin resistance. *J. Mol. Biol.* **259**, 1-6.

**Daniels, D., Michels, W.J., y Pyle, A.M.** (1996). Two competing pathways for self-splicing by group II introns; a quantitative analysis of *in-vitro* reaction rates and products. *J. Mol. Biol.* **256**; 31–49.

**Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A. y Scazzochio, C.** (1982). Making ends meet: a model for RNA splicing in fungal mitochondria. *Nature* **300**; 719–724.

**Del Campo, M., Tijerina, P., Bhaskaran, H., Mohr, S., Yang, Q., Jankowsky, E., Russell, R. y Lambowitz, A.M.** (2007). Do DEAD-box proteins promote group II intron splicing without unwinding RNA? *Mol Cell* **28**: 159–166.

**Del Campo, M., Mohr, S., Jiang, Y., Jia, H., Jankowsky, E. y Lambowitz AM.** (2009). Unwinding by local strand separation is critical for the function of DEAD-box proteins as RNA chaperones. *J Mol Biol* **389**: 674–693.

**DeLencastre, A., Hamill, S. y Pyle, A.M.** (2005). A single active-site region

for a group II intron. *Nat Struct Mol Biol* **12**: 626–627.

**Dellaporta, S.L., Xu, A., Sagasser, S., Jakob, W., Moreno, M.A., Buss, L.W. y Schierwater, B.** (2006). Mitochondrial genome of *Trichoplax adhaerens* supports placozoa as the basal lower metazoan phylum. *Proc Natl Acad Sci USA* **103**: 8751–8756.

**De Longevialle., A. F., Hendrickson, L., Taylor, N.L., Delannoy, E., Lurin, C., et al.,** (2008) The pentatricopeptide repeat gene *OTP51* with two LAGLIDADG motifs is required for the *cis*-splicing of plastid *ycf3* intron 2 in *Arabidopsis thaliana*, *The Plant Journal* **56**: 157-168

**Deppenmeir, U., Johann, A., Hartsch, T., Merkl, R., Schimits, R.A. et al.** (2002) The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between Bacteria and Archaea. *J. Mol. Micro. Biotechnol.*, **4**; 453-461.

**Dib-Hajj, S.D., Boulanger, S.C., Hebbar, S.K., Peebles, C.L., Franzen, J.S., y Perlman, P.S.** (1993). Domain 5 interacts with domain 6 and influences the second transesterification reaction of group II intron selfsplicing. *Nucleic Acids Res.* **21**; 1797–1804.

**Dickson, L., Huang, H.R., Liu, L., Matsuura, M., Lambowitz, A.M., y Perlman, P.S.** (2001). Retrotransposition of a yeast group II intron occurs by reverse splicing directly into ectopic DNA sites. *Proc. Natl. Acad. Sci. USA* **98**; 13207–13212.

**Ditta, G., Stanfield, S., Corbin, D. y Helinski, D.R.** (1980) Broad host range DNA cloning system for Gram-negative bacteria. Construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. USA* **77**: 7347-7351.

**Doetsch, N.A., Thompson, M.D., y Hallick, R.B.** (1998). A maturase-encoding group III twintron is conserved in deeply rooted euglenoid species: are group III introns the chicken or the egg? *Mol. Biol. Evol.* **15**; 76–86.

**Doetsch, N.A., Thompson, M.D., Favreau, M.R., y Hallick, R.B.** (2001). Comparison of *psbK* operon organization and group III intron content in chloroplast genomes of 12 Euglenoid species. *Mol. Gen. Genet.* **264**; 682–690.



---

**Dondrup, M., Goesmann, A., Bartels, D., Kalinowski, J., Krause, L., Linke, B., Rupp, O., Sczyrba, A., Pühler, A. y Meyer F** (2003) EMMA: a platform for consistent storage and efficient analysis of microarray data. *J Biotechnol*, **106(2-3)**:135-146.

**Eckstein, F. y Lilley, D.M.J.** (1996) *Nucleic Acids and Molecular Biology. Catalytic RNA*, **Vol. 10**. Springer, Berlin.

**Eickbush, T. H.** (1994) Origins and evolutionary relationships of retroelements. *The Evolutionary Biology of Viruses* (Morse SS, ed), pp. 121–157. Raven Press, New York.

**Eisen, J. S., y Smith, J. C.** (2008). Controlling morpholino experiments: don't stop making antisense. *Development (Cambridge, England)*, **135(10)**, 1735-1743.

**Ems, S.C., Morden, C.W., Dixon, C.K., Wolfe, K.H., dePamphilis, C.W. y Palmer, J.D.** (1995). Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant *Epifagus virginiana*. *Plant Mol Biol* **29**: 721–733.

**Eskes, R., Yang, J., Lambowitz, A.M., y Perlman, P.S.** (1997). Mobility of yeast mitochondrial group II introns: engineering a new site specificity and retrohoming via full reverse splicing. *Cell* **88**; 865–874.

**Eskes R, Liu L, Ma H, Chao MY, Dickson L, et al.** 2000. Multiple homing pathways used by yeast mitochondrial group II introns. *Mol. Cell Biol.* **20**; 8432–46

**Fedorova, O., Mitros, T., y Pyle, A.M.** (2003). Domains 2 and 3 interact to form critical elements of the group II intron active site. *J. Mol. Biol.* **330**; 197–209.

**Fedorova, O. y Pyle, A.M.** (2005). Linking the group II intron catalytic domains: tertiary contacts and structural features of domain 3. *EMBO J.*, **24**; 3906–3916.

**Fedorova, O. y Zingler, N.** (2007) Group II introns: structure, folding and splicing mechanism. *Biol. Chem.*, **388**; 665–678

**Ferat, J. L. y Michel, F.** (1993) Group II self-splicing introns in bacteria. *Nature*, **364**; 358–361.

**Ferat, J. L., Le Gouar, M. y Michel, F.** (1994) Multiple group II selfsplicing introns in mobile DNA from *Escherichia coli*. *C R Acad Sci III* **317**; 141–148.

**Ferat, J. L., Le Gouar, M. y Michel, F.** (2003) A group II intron has invaded the genus *Azotobacter* and inserted within the termination codon of the essential *groEL* gene. *Mol Microbiol.*, **49**; 1407–1423.

**Fernández-López, M., Muñoz-Adelantado, E., Gillis, M., Willems, A. y Toro, N.** (2005) Dispersal and evolution of the *Sinorhizobium meliloti* group II RmInt1 intron in bacteria that interact with plants. *Mol Biol Evol* **22**; 1518–1528

**Fontaine, J.M., Goux, D., Kloareg, B., Loiseaux-de Goër, S.** (1997). The reverse-transcriptase-like proteins encoded by group II introns in the mitochondrial genome of the brown alga *Pylaiella littoralis* belong to two different lineages which apparently coevolved with the group II ribozyme lineages. *J Mol Evol* **44**: 33–42.

**Frazier, C.L., San Filippo, J., Lambowitz, A.M., y Mills, D.A.** (2003). Genetic manipulation of *Lactococcus lactis* by using targeted group II introns: generation of stable insertions without selection. *Appl. Environ. Microbiol.* **69**; 1121–1128.

**Galibert, F., Finan, T.M., Long, S.R., Pühler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M.J., Becker, A., Boistard, P. et al.** (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293**; 668-672.

**García-Rodríguez, F. M., Barrientos-Durán, A., Díaz-Prado, V., Fernández-López, M. y N. Toro** (2011). Use of RmInt1, a Group IIB Intron Lacking the Intron-Encoded Protein Endonuclease Domain, in Gene Targeting *App Env Microbiol* **77**: 854–861.

**Giuntini, E., Mengoni, A., De Filippo, C., Cavalieri, D., Aubin-Horth, N., Landry, C.R., et al..** (2005). Large-scale genetic variation of the symbiosis-



---

required megaplasmid pSymA revealed by comparative genomic analysis of *Sinorhizobium meliloti* natural strains. *BMC Genomics* **6(1)**: 158-169.

**Galagan, J.E., Nusbaum, C., Roy, A., Endrizzi, M.G., Macdonald, P., et al.** (2002) The Genome of *Methanosarcina acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res.* **12**; 532-542.

**Glanz, S., Bunse, A., Wimbert, A., Balczun, C. y Kück, U.** (2006) A nucleosome assembly protein-like polypeptide binds to chloroplast group II intron RNA in *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* **34(18)**: 5337-5351.

**Glanz, S., y Kück, U.** (2009). *Trans-splicing of organelle introns—a detour to continuous RNAs.* *Bioessays* **31**: 921–934.

**Goldschmidt-Clermont M., Choquet Y., Girard-Bascou J., Michel F., Schirmer-Rahire M. y Rochaix J. D.** (1991) A small chloroplast RNA may be required for trans-splicing in *Chlamydomonas reinhardtii*. *Cell*, **65**: 135-143.

**Gorbalenya, A.E.** (1994) Self-splicing group I and group II introns encode homologous (putative) DNA endonucleases of a new family. *Protein Sci.* **3**: 1117-1120.

**Gordon, P.M. y Piccirilli, J.A.** (2001). Metal ion coordination by the AGC triad in domain 5 contributes to group II intron catalysis. *Nat. Struct. Biol.* **8**; 893–898.

**Gordon, P.M., Fong, R. y Piccirilli, J.A.** (2007). A second divalent metal ion in the group II intron reaction center. *Chem. & Biol.*, **14**; 607-612.

**Granlund, M., Michel, F. y Norgren, M.** (2001). Mutually exclusive distribution of IS1548 and GBSi1, an active group II intron identified in human isolates of group B streptococci. *J Bacteriol* **183**:2560–2569.

- Gregan, J., Kolisek, M., y Schweyen, R.J.** (2001) Mitochondrial Mg<sup>2+</sup> homeostasis es critical for group II intron splicing in vivo. *Genes and Dev.* **15**; 2229-2237.
- Griffin, E.A., Jr., Qin, Z., Michels, W.J., Jr., y Pyle, A.M.** (1995). Group II intron ribozymes that cleave DNA and RNA linkages with similar efficiency, and lack contacts with substrate 2'-hydroxyl groups. *Chem. Biol.* **2**; 761-770.
- Grohman, J.K., Del Campo, M., Bhaskaran, H., Tijerina, P., Lambowitz, A.M., y Russell, R.** (2007). Probing the mechanisms of DEAD.Box proteins as general RNA chaperones: the C-terminal Domain of CYT-19 mediates general recognition of RNA. *Biochemistry*, **46**; 3013-3022.
- Gu, S-Q., Cui, X., Mou, S., Mohr, S., Yao, J., Lambowitz, A.M.** (2010). Genetic identification of potential RNA-binding regions in a group II intron-encoded reverse transcriptase. *RNA* **16**: 732-747.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. y Altman, S.** (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**; 849-857.
- Guo, H., Zimmerly, S., Perlman, P.S., y Lambowitz, A.M.** (1997). Group II intron endonucleases use both RNA and protein subunits for recognition of specific sequences in double-stranded DNA. *EMBO J.* **16**; 6835-6848.
- Guo, H., Karberg, M., Long, M., Jones, J. P. 3rd, Sullenger, B. y Lambowitz, A. M.** (2000). Group II introns designed to insert into therapeutically relevant DNA target sites in human cells. *Science* **289**; 452-57
- Hallick, R. B., Hong, L., Drager, R. G., Favreau, M. R., Monfort, A., Orsat, B., Spielmann, A., y Stutz, E.** (1993). Complete sequence of *Euglena gracilis* chloroplast DNA *Nucleic Acids Res.*, **21**; 3537 - 3544.
- Hamill, S. y Pyle, A.M.** (2006). The receptor for branch-site docking within a group II intron active site. *Mol. Cell* **23**: 831-840.



- Heap, J.T., Pennington, O.J., Cartman, S.T., Carter, G.P., Minton, N.P.** (2007). The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *J.Microbiol.Methods* **70**: 452–64.
- Hebbar, S.K., Belcher, S.M., y Perlman, P.S.** (1992). A maturase-encoding Group IIA intron of yeast mitochondria selfsplices *in vitro*. *Nucleic Acids Res.* **20**; 1747–1754.
- Hertweck, M., y Mueller, M.W.** (2001) Mapping divalent metal ion binding sites in a group II intron by Mn<sup>2+</sup> - y Zn<sup>2+</sup> -induced site specific RNA cleavage. *Eur. J. Biochem.* **268**; 4610-4620.
- Huang, H.R., Chao, M.Y., Armstrong, B., Wank, Y, Lambowitz, A.M. y Perlman, P.S.** (2003) The DIVa maturase binding site in the yeast group II intron aI2 is essential for intron homing but not for *in vivo* splicing. *Mol. Cell. Biol.*, **23**; 8809-8819.
- Huang, H.R., Rowe, C.E., Mohr, S., Jiang, Y., Lambowitz, A.M. y Perlman PS.** (2005). The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box proteinwith RNA chaperone function. *Proc Natl Acad Sci* **102**: 163–168.
- Huang, T., Shaikh, T. R., Gupta, K., Contreras-Martin, L. M., Grassucci, R. A., Van Duyne, G. D., Frank J., y Belfort, M.** (2010) The group II intron ribonucleoprotein precursor is a large, loosely packed structure, *Nuc Acid Res*,
- Ichiyanagi K, Beauregard A, Lawrence S, Smith D, Cousineau B, Belfort M.** (2002). Multiple pathways for the Ll.LtrB group II intron include reverse splicing into DNA targets. *Mol. Microbiol.* **46**; 1259–71
- Ichiyanagi K, Beauregard A y Belfort M** (2003) A bacterial group II intron favors retrotransposition into plasmid targets. *Proc. Natl Acad Sci USA* **100**; 15742–15747.
- Ish-Horowicz, D. y Burke, J.F.** (1981) Rapid and efficient cosmid vector cloning. *Nucleic Acids Res.* **9**: 2989-2998.

**Iversen, P. L., y Newbry, S.** (2005). Manipulation of zebrafish embryogenesis by phosphorodiamidate morpholino oligomers indicates minimal non-specific teratogenesis. *Current Opinion in Molecular Therapeutics*, **7(2)**, 104-108.

**Jacquier A.** (1990). Self-splicing Group II and nuclear premRNA introns: How similar are they? *Trends Biochem. Sci.* **15**; 351-54.

**Jacquier, A. y Michel, F.** (1987). Multiple exon-binding sites in class II self-splicing introns. *Cell* **50**; 17-29.

**Jacquier, A. y Michel, F.** (1990). Base-pairing interactions involving the 5' and 3'-terminal nucleotides of group II self-splicing introns. *J. Mol. Biol.* **213**; 437-447.

**Jacquier, A. y Jacquesson-Breuleux, N.** (1991). Splice site selection and role of the lariat in a group II intron. *J. Mol. Biol.* **219**; 415-428

**Jacobs, J., Glanz, S., Bunse-Grassmann, A., Kruse, O. y Kück, U.** (2010) RNA trans-splicing: identification of components of a putative chloroplast spliceosome. *Eur J Cell Biol.* **89(12)**: 932-939

**Jarrell, K.A., Peebles, C.L., Dietrich, R.C., Romiti, S.L., y Perlman, P.S.** (1988). Group II intron self-splicing: alternative reaction conditions yield novel products. *J. Biol. Chem.* **263**; 3432-3439.

**Jarrell, K.A.** (1993). Inverse splicing of a group II intron. *Proc Natl Acad Sci* **90**: 8624-8627.

**Jenkins, B.D., y Barkan, A.** (2001). Recruitment of a peptidyl-tRNA hydrolase as a facilitator of group II intron splicing in chloroplast. *The EMBO Journal*. **20**; 872-879.

**Jestin, J.-L., Deme, E., y Jacquier, A.** (1997). Identification of structural elements critical for inter-domain interactions in a group II self-splicing intron. *EMBO J.* **16**; 2945-2954.

**Jiménez-Zurdo, J. I., García-Rodríguez, F. M., Barrientos-Durán, A., Toro, N.** (2003). DNA target site requirements for homing in vivo of a



bacterial group II intron encoding a protein lacking the DNA endonuclease domain. *J. Mol. Biol.* **326**; 413–23.

**Jones, K.M., Kobayashi, H, Davies, B.W., Taga, M.E. y Walker, G.C.** (2007) How rhizobial symbionts invade plants: the Sinorhizobium-Meficago model. *Nat Rev Microbiol* **5**: 619-633.

**Jones III, J.P., Kerlin, M.N., Coon; R.G., Perutka, J., Lambowitz, A.M. y Sullenger, B.A.** (2005). Retargeting mobile group II introns to repair mutant genes. *Mol. Therapy*, **11 (5)**; 687-694.

**Kahng, L.S. y Shapiro, L.** (2003). Polar localization of replication origins in the multipartite genomes of *Agrobacterium tumefaciens* and *Sinorhizobium meliloti*. *J. Bacteriol.* **185**: 3384-3391.

**Karimi, M., Inzé, D., Depicker, A.** (2002) Gateway vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci.* **7(5)**: 193-195.

**Karberg, M., Guo, H., Zhong, J., Coon, R., Perutka, J., y Lambowitz, A.M.** (2001). Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. *Nat. Biotechnol.* **19**; 1162–1167.

**Karberg, M.S.** (2005). Group II intron mobility and its applications in biotechnology and gene therapy. PhD. Dissertation, University of Texas at Austin.

**Keating, K.S., Toor, N. y Pyle, A.M.** (2008). The GANC tetraloop: a novel motif in the group IIC intron structure. *J Mol Biol* **383**:475–481.

**Keating, K.S., Toor, N., Perlman, P.S. y Pyle, A.M.** (2010). A structural analysis of the group II intron active site and implications for the plicesome. *RNA* **16**:1–9.

**Kennell, J.C., Moran, J.V., Perlman, P.S., Butow, R.A., y Lambowitz, A.M.** (1993). Reverse transcriptase activity associated with maturase-encoding group II introns in yeast mitochondria. *Cell* **73**; 133–146.

**Keren, I., Bezawork-Geleta, A., Koltun, M., Maayan, I., Belausov, E., Levy, M., Mett, A., Gidoni, D., Shaya, F. y Ostersetzer-Biran, O.** (2009). AtnMat2, a nuclear-encoded maturase required for splicing of group-II introns in *Arabidopsis* mitochondria. *RNA* **15**: 2299–2311.

**Köhler, D., Schmidt-Gattung, S. y Binder, S.** (2010). The DEAD-box protein PMH2 is required for efficient group II intron splicing in mitochondria of *Arabidopsis thaliana*. *Plant Mol Biol* **72**: 459–467.

**Klein, J.R. y Dunny, G.M.** (2002). Bacterial group II introns and their association with mobile genetic elements. *Frontiers in Bioscience*, **7**; 1843–1856.

**Koch, J.L., Boulanger, S.C., Dib-Hajj, S.D., Hebbar, S.K., y Perlman, P.S.** (1992). Group II introns deleted for multiple substructures retain self-splicing activity. *Mol. Cell. Biol.* **12**; 1950–1958.

**Kohlstaedt L.A., Wang J., Friedman J.M., Rice P.A., Steitz T.A.** (1992) Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science*, **256**: 1783–1790.

**Knoop, V., Brennicke, A.** (1993). Group II introns in plant mitochondria - Trans-splicing, RNA editing, evolution, and promiscuity. In: *Plant Mitochondria*. Brennicke, A., Kück, U. (eds.) Weinheim: VCH Chemie, pp. 221-232.

**Knoop, V., Altwasser, M. y Brennicke, A.** (1997). A tripartite group II intron in mitochondria of an angiosperm plant. *Mol Gen Genet*, **255(3)**; 269-76.

**Koonin, E. V.** (2006). The origins of introns and their role in eukaryogenesis: a compromise solution to introns-early vs. introns-late debate? *Biol Direct* **1**; 39.

**Konforti, B.B., Liu, Q., y Pyle, A.M.** (1998a). A map of the binding site for catalytic domain 5 in the core of a group II intron ribozyme. *EMBO J.* **17**; 7105–7117.



---

**Konforti, B.B., Abramovitz, D.L., Duarte, C.M., Karpeisky, A., Beigelman, L., y Pyle, A.M.** (1998b). Ribozyme catalysis from the major groove of group II intron domain 5. *Mol. Cell* **1**; 433–441.

**Kono, M., Satoh, H., Okabe, Y., Abe, Y., Nakayama, K. y Okada, M.** (1991). Nucleotide sequence of the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from the green alga *Bryopsis maxima*. *Plant Mol Biol*, **17(3)**; 505-508.

**Kovach, M.E., Phillips, R.W., Elzer, P.H., Roop, R.P. 2<sup>nd</sup> y Peterson, K.M.** (1994). pBBR1MCS: a broad-host-range cloning vector. *Biotechniques*. **16**; 800-802.

**Kroeger, T.S., Watkins, K.P., Friso, G., van Wijk, K.J. y Barkan, A.** (2009). A plant-specific RNA-binding domain revealed through analysis of chloroplast group II intron splicing. *Proc Natl Acad Sci* **106**: 4537–4542.

**Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., y Cech, T.R.** (1982). Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **31**: 147-157.

**Kuban, W., Jonczyk, P., Gawel, D., Malanowska, K., Schaaper, R. M. y Fijalkowska I. J.** (2004) Role of *Escherichia coli* DNA Polymerase IV in *In Vivo* Replication Fidelity. *Journal of Bacteriology*, **186**: 4802-4807.

**Kück, U., Choquet, Y., Schneider, M., Dron, M. y Bennoun, P.** (1987). Structural and transcription analysis of two homologous genes for the P700 chlorophyll a-apoproteins in *Chlamydomonas reinhardtii*: evidence for in vivo *trans*-splicing. *EMBO J*, **6(8)**; 2185-2195.

**Kück, U.** (1989). The intron of a plastid gene from a green alga contains an open reading frame for a reverse transcriptase-like enzyme. *Mol Gen Genet*, **218(2)**; 257-65

**Lambowitz AM.** (1989). Infectious introns. *Cell* **56**; 323–326.

**Lambowitz, A.M. y Belfort, M.** (1993). Introns as mobile genetic elements. *Annu. Rev. Biochem.* **62**; 587–622.

**Lambowitz, A.M., Caprara, M.G., Zimmerly, S., Perlman, P.S.** (1999). Group I and group II ribozymes as RNPs: clues to the past and guides to the future. En *The RNA World* 2nd ed., pp. 451–485. Cold Spring Harbor Lab. Press, New York.

**Lambowitz, A.M. y Zimmerly, S.** (2004). Mobile group II introns. *Ann Rev Genet* **38**; 1–35.

**Lambowitz, A.M. y Zimmerly, S.** (2010). Group II introns: Mobile Ribozymes that invades DNA. En: *The RNA World*, 4th Edition (R.F. Gesteland, T.R. Cech, and J.F. Atkins, Editors), Cold Spring Harbor Perspect Biol 1:a003616.

**Lazowska, J., Meunier, B., y Macadre, C.** (1994). Homing of a group II intron in yeast mitochondrial DNA is accompanied by unidirectional co-conversion of upstreamlocated markers. *EMBO J.* **13**; 4963–4972

**Lehmann, K. y Schmidt, U.** (2003). Group II introns: structural and catalytic versatility of large natural ribozymes. *Crit. Rev. Biochem. Mol. Biol.* **38**; 249–303.

**Léon, G. y Roy, P.H.** (2009). Group IIC intron mobility into attC sites involves a bulged DNA stem-loop motif. *RNA* **15**: 1543–1553.

**Lewin, A.S. y Hauswirth, W.W.** (2001). Ribozyme gene therapy: applications for molecular medicine. *TRENDS Mol. Med.*, **7 (5)**; 221-228.

**Lewin, R.** (1986). RNA catalysis gives fresh perspective on the origin of life. *Science* **231**; 545-546.

**Liberati, N. T., J. M. Urbach, S. Miyata, D. G. Lee, E. Drenkard, G. Wu, J. Villanueva, T. Wei y F. M. Ausubel** (2006) An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proc Natl Acad Sci U S A* **103(8)**:2833-2838

**Lin J, Jin R, Zhang B, Chen H, Bai YX, Yang PX, et al.** (2008) Nucleolar localization of TERT is unrelated to telomerase function in human cells. *J Cell Sci.* **121**: 2169–2176.



---

**Lin, X., Kaul, S., Rounsley, S, Shea, T.P., Benito, M.I. et al.** (1999). Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature*, **402**; 761-768.

**Linder P, Lasko PF, Ashburner M, Leroy P, Nielsen PJ, Nishi K, Schnier J, Slonimski PP** (1989). "Birth of the D-E-A-D box". *Nature* **337** (6203): 121–122.

**Li-Pook-Than, J. y Bonen, L.** (2006). Multiple physical forms of excised group II intron RNAs in wheat mitochondria. *Nucleic Acids Res.* **34**; 2782–2790.

**Long, M.B., Jones III, J.P., Sullenger, B.A. y Byun, J.** (2003). Ribozyme-mediated revision of RNA and DNA. *J. Clin. Invest.*, **112 (3)**; 312-318.

**Lorkovic, Z.J., J. Hilscher, Y A. Barta** (2004) Use of fluorescent protein tags to study nuclear organisation of the spliceosomal machinery in transiently transformed living plant cells. *Mol. Biol. Cell* **15**: 3233-3243.

**Lovmar, M., K. Nilsson, V. Vimberg, T. Tenson, M. Nervall, and M. Ehrenberg.** (2006). The molecular mechanism of peptide-mediated erythromycin resistance. *J. Biol. Chem.* **281**:6742-6750.

**Lykke-Andersen, J., Aagaard, C., Semionenkov, M., y Garrett, R.A.** (1997). Archaeal introns: splicing, intercellular mobility and evolution. *Trends Biochem. Sci.* **22**; 326–331.

**Luan, D.D., Korman, M.H., Jakubczak, J.L y Eickbush, T.H.** (1993). Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72**; 595–605.

**Madhani, H.D. y Guthrie, C.** (1992). A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell*, **71**; 803-817.

**Malhotra, M. y Srivastava, S.** (2008). An *ipdC* gene knock-out of *Azospirillum brasilense* strain SM and its implications on indole-3-acetic acid

biosynthesis and plant growth promotion. *Antonie Van Leeuwenhoek*. **93**: 425-433.

**Malik, H.S., Burke, W.D., Eickbush, T.H.** (1999). The age and evolution of non-LTR retrotransposable elements. *Mol Biol Evol* **16**: 793–805.

**Manhart, J.R. y Von der Haar, R.A.** (1991). Intron revealed by nucleotide sequence of large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Codium fragile* (Chlorophyta): Phylogenetic analysis. *J. Phycol.* **27**; 613-617.

**Marienfeld, J., Unseld, M. y Brennicke, A.** (1999). The mitochondrial genome of *Arabidopsis* is composed of both native and immigrant information. *Trends Plant Sci.*, **4**; 495-502.

**Marqués, S., Ramos, J.L. y Timmis, K.N.** (1993) Analysis of the mRNA structure of the *Pseudomonas putida* TOL meta fission pathway operon around the transcription initiation point, the xylTE and the xylFJ regions. *Biochem. Biophys. Acta* **1.216**: 227-236.

**Martin, W. y Koonin, E.V.** (2006). Introns and the origin of nucleus-cytosol compartmentalization. *Nature* **440**: 41–45.

**Martínez-Abarca, F., Zekri, S. y Toro, N.** (1998). Characterization and splicing *in vivo* of a *Sinorhizobium meliloti* group II intron associated with particular insertion sequences of the IS630- Tc1/IS3 retroposon superfamily. *Mol Microbiol* **28**; 1295–1306.

**Martínez-Abarca, F., García-Rodríguez, F.M., Muñoz, E., y Toro, N.** (1999). Biochemical demonstration of reverse transcriptase activity and splicing efficiency of a bacterial group II intron from *Sinorhizobium meliloti*. *Nucleic Acids Symp. Ser.* **41**; 117–119.

**Martínez-Abarca, F. y Toro, N.** (2000a). Group II introns in the bacterial world. *Mol. Microbiol.* **38**; 917–926.

**Martínez-Abarca, F. y Toro, N.** (2000b). RecA-independent ectopic transposition *in vivo* of a bacterial group II intron. *Nucleic Acids Res.* **28**; 4397-4402.



---

**Martínez-Abarca, F., García-Rodríguez, F.M. y Toro, N.** (2000). Homing of a bacterial group II intron with an intron-encoded protein lacking a recognizable endonuclease domain. *Mol Microbiol* **35**; 1405–1412.

**Martínez-Abarca, F., Barrientos-Durán, A., Fernández-López, M. y Toro, N.** (2004). The RmInt1 group II intron has two different retrohoming pathways for mobility using predominantly the nascent lagging strand at DNA replication forks for priming. *Nucleic Acids Res* **32**; 2880–2888.

**Mastroianni M., K. Watanabe, T.B. White, F. Zhuang, J. Vernon, M. Matsuura, J. Wallingford y A.M. Lambowitz.** (2008) Group II intron-based gene targeting reactions in eukaryotes. *PLoS ONE*. **3(9)**: e3121.

**Mathur, J. y Koncz, C.** En *Methods in Molecular Biology, Arabidopsis Protocols* Ed.: Martínez-Zapater J . y Salinas J. :Humana Press Inc., Totowa, NJ **Vol. 82**;, 267-276.

**Matsuura, M., Saldanha, R., Ma, H., Wank, H., Yang, J., Mohr, G., Cavanagh, S., Dunny, G.M., Belfort, M., y Lambowitz, A.M.** (1997). A bacterial group II intron encoding reverse transcriptase, maturase, and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron. *Genes Dev.* **11**; 2910–2924.

**Matsuura, M., Noah, J.W., y Lambowitz, A.M.** (2001). Mechanism of maturase-promoted group II intron splicing. *EMBO J.* **20**; 7259–7270.

**Meade, H. M. y Signer, E. R.** (1977). Genetic mapping of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. U.S.A.* **74**; 2076-2078.

**Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., y Tuschl, T.** (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Molecular Cell*, **15(2)**, 185-197.

**Meng, Q., Wang, Y., y Liu, X.Q.** (2005). An intron-encoded protein assists RNA splicing of multiple similar introns of different bacterial genes. *J Biol Chem* **280**: 35085–35088.

**Merendino, L., Perron, K., Rahire, M., Howald, I., Rochaix, J.D. y Goldschmidt-Clermont M.** (2006). A novel multifunctional factor involved in transsplicing of chloroplast introns in *Chlamydomonas*. *Nucleic Acids Res* **34**: 262–274.

**Meunier, B., Tian, G.-L., Macadre, C., Slonimski, P.P. y Lazowska, J.** (1990). Group II introns transpose in yeast mitochondria. In *Structure Function and Biogenesis of Energy Transfer Systems*, ed. E Quagliariello, S Papa, F Palmieri, C Saccone, pp. 169–74. Amsterdam: Elsevier

**Michel, F., Jacquier, A., y Dujon, B.** (1982). Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. *Biochimie* **64**; 867–861.

**Michel, F. y Lang, B.F.** (1985). Mitochondrial class II introns encode proteins related to the reverse transcriptases of retroviruses. *Nature* **316**; 641–43

**Michel, F., Umesono, K., y Ozeki, H.** (1989). Comparative and functional anatomy of group II catalytic introns — a review. *Gene* **82**; 5–30.

**Michel, F. y Ferat, J.L.** (1995). Structure and activities of group II introns. *Annu Rev Biochem* **64**; 435–461.

**Michel, F., Costa, M., Doucet, A.J. y Ferat, J.-L.** (2007). Specialized lineages of bacterial group II introns. *Biochimie*, **89**; 542-553.

**Michel, F., Costa, M. y Westhof, E.** (2009). The ribozyme core of group II introns: a structure in want of partners. *Trends Biochem Sci* **34**:189–199.

**Mikheeva, S., Murray, H.L., Zhou, H., Turczyk, B.M., y Jarrell, K.A.** (2000). Deletion of a conserved dinucleotide inhibits the second step of group II intron splicing. *RNA* **6**; 1509–1515.

**Mills, D.A., McKay, L.L., y Dunny, G.M.** (1996). Splicing of a group II intron involved in the conjugative transfer of pRS01 in Lactococci. *J. Bacteriol.* **178**; 3531–3538.



---

**Mills, D.A., Manias, D.A., McKay, L.L., y Dunny, G.M.** (1997). Homing of a group II intron from *Lactococcus lactis* subsp. *lactis* ML3. *J. Bacteriol.* **179**; 6107-6111.

**Mills, D.A.** (2001). Mutagenesis in the post genomics era: tools for generating insertional mutations in the lactic acid bacteria. *Curr. Opin. Biotech.* **12**; 503-509.

**Mittal, V.** (2004). Improving the efficiency of RNA interference in mammals. *Nature Reviews. Genetics*, **5(5)**, 355-365.

**Mohr G., Perlman P. S. y Lambowitz A. M.** (1993) Evolutionary relationships among group II intron-encoded proteins and identification of a conserved domain that may be related to maturase function. *Nucleic Acids Res*, **21**: 4991-4997.

**Mohr G., Smith D., Belfort M. y Lambowitz A. M.** (2000) Rules for DNA target-site recognition by a lactococcal group II intron enable retargeting of the intron to specific DNA sequences. *Genes Dev*, **14**: 559-573.

**Mohr G. y Lambowitz, A. M.** (2003). Putative proteins related to group II intron reverse transcriptase/maturases are encoded by nuclear genes in higher plants. *Nucleic Acids Res*, **31**: 647-652.

**Mohr S., Matsuura M., Perlman P. S., y Lambowitz A. M.** (2006) A DEAD-box protein alone promotes group II intron splicing and reverse splicing by acting as an RNA chaperone. *Proc Natl Acad Sci USA*, **103**: 3569-3574.

**Mol, C.D., Kuo, C.F., Thayer, M.M., Cunningham, R.P., y Tainer, J.A.** (1995) Structure and function of the multifunctional DNA-repair enzyme exonuclease III. *Nature*, **374(6520)**; 381-386.

**Molina-Sánchez M. D., Martínez-Abarca F. y Toro N.** (2006) Excision of the *Sinorhizobium meliloti* group II intron RmInt1 as circles in vivo. *J Biol Chem*, **281**: 28737-28744.

**Molina-Sánchez, M.D.** (2008). Intrones del grupo II en *Shinorhizobium meliloti*: contribución de la proteína y la ribozima codificadas por RmInt1 a su escisión y movilidad. *Tesis doctoral Universidad de Granada*.

**Molina-Sánchez, M.D., Martínez-Abarca, F. y Toro, N.** (2010). Structural features in the C-terminal region of the *Sinorhizobium meliloti* RmInt1 group II intrón-encoded proteína contribuye to its maturase and intrón DNA-insertion function. *FEBS J.*, **277**: 244-255

**Moore, M.J., Query, C.C. y Sharp, P.A.** (1993). Splicing of precursors to mRNA by the spliceosome. *The RNA World*, pp. 303-357, Cold Spring Harbor Laboratory Press, New York.

**Moran, J.V., Mecklenburg, K.L., Sass, P., Belcher, S.M., Mahnke, D., Lewin, A., y Perlman, P.** (1994). Splicing defective mutants of the *COXI* gene of yeast mitochondrial DNA: initial definition of the maturase domain of the group II intron aI2. *Nucleic Acids Res.* **22**; 2057–2064.

**Moran, J.V., Zimmerly, S., Eskes, R., Kennell, J.C., Lambowitz, A.M., Butow, R.A., y Perlman, P.S.** (1995).. Mobile group II introns of yeast mitochondrial DNA are novel site-specific retroelements. *Mol. Cell. Biol.* **15**; 2828–2838.

**Mörl, M. y Schmelzer, C.** (1990a). Integration of group II intron bI1 into a foreign RNA by reversal of the self-splicing reaction *in vitro*. *Cell* **60**; 629–636.

**Mörl, M. y Schmelzer, C.** (1990b). Group II intron RNA-catalyzed recombination of RNA *in vitro*. *Nucleic Acids Res.* **18**; 6545–6551.

**Mörl, M., Niemer, I., y Schmelzer, C.** (1992). New reactions catalyzed by a group II intron ribozyme with RNA and DNA substrates. *Cell* **70**; 803–810.

**Meunier, B., Tian, G.-L., Macadre, C., Slonimski, P.P. y Lazowska, J.** (1990). Group II introns transpose in yeast mitochondria. En *Structure Function and Biogenesis of Energy Transfer Systems*, pp. 169–74. Elsevier Amsterdam.



**Mueller, M.W., Hetzer, M., y Schweyen, R.J.** (1993a). Group II intron RNA catalysis of progressive nucleotide insertion: a model for RNA editing. *Science* **261**; 1035–1038.

**Mueller, M.W., Allmaier, M., Eskes, R., y Schweyen, R.J.** (1993b). Transposition of group II intron a11 in yeast and invasion of mitochondrial genes at new locations. *Nature* **366**; 174–176.

**Muñoz, E., Villadas, P.J., y Toro, N.** (2001). Ectopic transposition of a group II intron in natural bacterial populations. *Mol. Microbiol.* **41**; 645–652.

**Muñoz-Adelantado, E., San Filippo, J., Martínez-Abarca, F., García-Rodríguez, F.M., Lambowitz, A.M. y Toro, N.** (2003). Mobility of the *Sinorhizobium meliloti* group II intron RmInt1 occurs by reverse splicing into DNA, but requires an unknown reverse transcriptase priming mechanism. *J Mol Biol* **327**; 931–943.

**Muñoz Adelantado, E.** (2003). El Intrón del Grupo II de *Sinorhizobium meliloti*, RmInt1: Estudios Ecológicos y Caracterización Molecular de las Actividades Bioquímicas Implicadas en su Movilidad. Tesis Doctoral. Universidad de Granada.

**Muth, G.W., Ortoleva-Donnelly, L., Strobel, S.A.** (2000) A single adenosine with a neutral pKa in the ribosomal peptidyl transferase center. *Science* **289**; 947–950.

**Murray, H.L., Mikheeva, S., Coljee, V.W., Turczyk, B.M., Donahue, W.F., Bar-Shalom, A. y Jarrel, K.A** (2001). Excision of group II introns as circles. *Mol Cell* **8**; 201–211.

**Nakamura, H., Oda, Y., Iwai, S., Inoue, H., Ohtsuka, E., Kanaya, S. et al.** (1991) How does RNase H recognize a DNA-RNA hybrid? *Proc. Natl. Acad. Sci. USA*, **88**; 11535-11539.

**Nakagawa, N. y Sakurai, N.** (2006). A mutation in *At-nMat1a*, which encodes a nuclear gene having high similarity to group II intron maturase, causes impaired splicing of mitochondrial NAD4 transcript and altered carbon metabolism in *Arabidopsis thaliana*. *Plant Cell Physiol* **47**: 772–783.

**Nasevicius, A., y Ekker, S. C.** (2000). Effective targeted gene 'knockdown' in zebrafish. *Nature Genetics*, **26(2)**, 216-220.

**Nielsen, H., Westhof, E. y Johansen, S.** (2005). An mRNA is capped by a 2'ϕ, 5'ϕ lariat catalyzed by a group I-like ribozyme. *Science* **309**:1584–1587.

**Nisa-Martínez, R., Jiménez-Zurdo, J.I., Martínez-Abarca, F., Muñoz-Adelantado, E. y Toro, N.** (2007). Dispersion of the RmInt1 group II intron in the *Sinorhizobium meliloti* genome upon acquisition by conjugative transfer. *Nucleic Acids Res* **35**; 214–222.

**Nissen, P., Hansen, J., Ban, N., Moore, P.B., Steitz, T.A.** (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**; 920–930.

**Noah, J.W., y Lambowitz, A.M.** (2003) Effects of maturase binding and Mg<sup>2+</sup> concentration on group II intron RNA folding investigated by UV cross-linking. *Biochemistry*, **42**; 12466-12480.

**Noah, J.W., Park, S., Whitt, J.T., Perutka, J., Frey, W. y Lambowitz, A.M.** (2006). Atomic force microscopy reveals DNA bending during group II intron ribonucleoprotein particle integration into double-stranded DNA. *Biochemistry* **45**; 12424–12435.

**Noller, H.F.** (2006) Evolution of ribosomes and translation from an RNA World. *The RNA World (3<sup>rd</sup> ed.)*, pp 287-307. Cold Spring Harbor Laboratory Press, New York.

**Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., et al.** (1992). Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. *J. Mol. Biol.* **223 (1)**; 1-7.

**Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., et al.** (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature*, **322 (6079)**; 572-74.

**Orgel, L.E.** (1968). Evolution of the genetic apparatus, *J. Mol. Biol.* **38**; 381-393.



- Ostheimer, G.J., Williams-Carrier, R., Belcher, S., Osborne, E., Gierke, J. y Barkan, A.** (2003). Group II intron splicing factors derived by diversification of an ancient RNA-binding domain. *EMBO J.* **22**: 3919–29
- Osiewacz H. D. y Esser K.** (1984) The mitochondrial plasmid of *Podospira anserina*: a mobile intron of a mitochondrial gene. *Curr Genet*, **8**: 299–305.
- Ostersetzer O, Cooke AM, Watkins KP and Barkan A.** (2005). CRS1, a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains. *Plant Cell* **17**:241–255.
- Paddison, P. J., Caudy, A. A., y Hannon, G. J.** (2002). Stable suppression of gene expression by RNAi in mammalian cells. *Proc Natl Acad Sci of the U.S.A* **99**(3), 1443-1448.
- Pace, N.R. y Marsh T.L.** (1985). RNA catalysis and the origin of life. *Origins Life Evol. Biosph.* **16**; 97-116.
- Padgett,R.A., Konarska,M.M., Grabowski,P.J., Hardy,S.F. y Sharp,P.A.** (1984) Lariat RNAs as intermediates and products in the splicing of messenger RNA precursors. *Science*, **225**; 89-903.
- Palmer, J.D. y Logsdon, J.M. Jr.** (1991) The recent origins of introns. *Curr. Opin. Gen. Develop.*, **1**; 470-477.
- Pearson, W. R. y Lipman, D. J.** (1988) Improved Tools for Biological Sequence Comparison. *Proc. Natl. Acad. Sci. USA*, **85**; 2444 - 2448.
- Peebles, C.L., Perlman, P.S., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A., y Cheng, H.L.** (1986). A self-splicing RNA excises an intron lariat. *Cell*, **44**; 213–223.
- Peebles, C.L., Benatan, E.J., Jarrell, K.A., y Perlman, P.S.** (1987). Group II intron self-splicing: development of alternative reaction conditions and identification of a predicted intermediate. *Cold Spring Harb. Symp. Quant. Biol.* **52**; 223–232.

- Peebles, C.L., Zhang, M., Perlman, P.S., y Franzen, J.S.** (1995). Catalytically critical nucleotide in domain 5 of a group II intron. *Proc. Natl. Acad. Sci. USA* **92**; 4422–4426.
- Perlman, P.S. y Butow, R.A.** (1989). Mobile introns and intron-encoded proteins. *Science* **246**:1106–1109.
- Perlman, P. y Podar, M.** (1996). Reactions catalyzed by group II introns in vitro. *Methods Enzymol* **264**:66–86.
- Perron, K., Goldschmidt-Clermont, M. y Rochaix, J.D.** (1999). A factor related to pseudouridine synthases is required for chloroplast group II intron trans-splicing in *Chlamydomonas reinhardtii*. *EMBO J* **18**: 6481–6490.
- Perron, K., Goldschmidt-Clermont, M., y Rochaix, J.D.** (2004) A multiple complex involved in chloroplast group II intron splicing. *RNA*, **10**; 704-711.
- Perutka, J., Wang, W., Goerlitz, D., y Lambowitz, A.M.** (2004). Use of computer-designed group II introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. *J. Mol. Biol.* **336**; 421–39
- Phizicky, E.M. y Hopper, A.K.** (2010) tRNA biology charges to the front, *Genes & Dev.*; **24**: 1832 – 1860.
- Plante, I. y Cousineau, B.** (2006) Restriction for gene insertion within the *Lactococcus lactis* Ll.LtrB group II intron. *RNA*, **12**; 1980–1992.
- Podar, M., Dib-Hajj, S., y Perlman, P.S.** (1995a). A UV-induced Mg<sup>2+</sup>-dependent cross-link traps an active form of domain 3 of a self-splicing group II intron. *RNA*, **1**; 828–840.
- Podar, M., Perlman, P.S., y Padgett, R.A.** (1995b) Stereochemical selectivity of group II intron splicing, reverse splicing, and hydrolysis reactions. *Mol. Cell. Biol.* **15**; 4466–4478.
- Podar, M., Zhuo, J., Zhang, M., Franzen, J.S., Perlman, P.S., y Peebles, C.L.** (1998a). Domain 5 binds near a highly conserved dinucleotide in the joiner linking domains 2 and 3 of a group II intron. *RNA*, **4**; 151–166.



---

**Podar, M., Chu, V.T., Pyle, A.M., y Perlman, P.S.** (1998b). Group II intron splicing *in vivo* by first-step hydrolysis. *Nature*, **391**; 915–918.

**Podar, M., Perlman, P.S., y Padgett, R.A.** (1998c). The two steps of group II intron self-splicing are mechanistically distinguishable. *RNA*, **4**; 890-900.

**Pobigaylo, N., Wetter, D., Szymczak, S., Schiller, U., Kurtz, S., Meyer, F., Nattkemper, T.W., y Becker, A.** (2006) Construction of a large sequence signature-tagged miniTn5 transposon library and its application to mutagenesis of *Sinorhizobium meliloti*. *Appl Environ Microbiol* **72**; 4329-4337.

**Porteus, M. H., Cathomen, T., Weitzman, M. D., y Baltimore, D.** (2003). Efficient gene targeting mediated by adeno-associated virus and DNA double-strand breaks. *Molecular and Cellular Biology*, **23(10)**, 3558-3565.

**Porteus, M. H., y Carroll, D.** (2005). Gene targeting using zinc finger nucleases. *Nature Biotechnology*, **23(8)**, 967-973.

**Pruett-Miller, S. M., Connelly, J. P., Maeder, M. L., Joung, J. K., & Porteus, M. H.** (2008). Comparison of zinc finger nucleases for use in gene targeting in mammalian cells. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, **16(4)**, 707-717.

**Puerta-Fernández, E., Romero-López, C., Barroso-delJesus, A. y Berzal-Herranz, A.** (2003) Ribozymes: recent advances in the development of RNA tools. *FEMS Microbiol. Rev.* **27**; 75-97.

**Pyle, A. M.** (2000) New tricks from an itinerant intron. *Nature Struct Biol* **7**: 352-354.

**Pyle, A.M.** (2005). Capping by branching: a new ribozyme makes tiny lariats. *Science* **309**:1530–1531.

**Pyle, A.M. y Lambowitz, A.M.** (2006) Group II introns: ribozymes that splice RNA and invade DNA. *RNAWorld*. 3rd edn (Getsland RF, Ceh TR & Atkins JF, eds), pp. 469–505. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

**Pyle, A.M., Fedorova, O., y Waldsich, C.** (2007). Folding of group II introns: a model system for large, multidomain RNAs? *Trends Biochem. Sci.* **32**; 138–145.

**Pyle, A.M.** (2008). Group II introns: catalysts for splicing, genomic change and evolution. In Lilley DMJ and Eckstein F (Eds) *Ribozymes and RNA Catalysis* New York: Springer-Verlag. p. 201–223.

**Pyle, A.M.** (2010) The tertiary structure of group II introns: implications for biological function and evolution. *Crit. Rev. Biochem. Mol. Biol.* **45**, 215-232.

**Qin, P.Z. y Pyle, A.M.** (1998). The architectural organization and mechanistic function of group II intron structural elements. *Curr. Opin. Struct. Biol.* **8**; 301–308.

**Qiu, Y.L. y Palmer, J.D.** (2004) Many independent origins of *trans*-splicing of a plant mitochondrial group II intron. *J Mol Evol*, **59(1)**; 80-9.

**Rambo, R.P. y Doudna, J.A.** (2004) Assembly of an active group II intron-maturase complex by protein dimerization. *Biochemistry*, **43**; 6486-6497.

**Ramirez, C. L., Foley, J. E., Wright, D. A., Müller-Lerch, F., Rahman, S. H., Cornu, T. I., et al.** (2008). Unexpected failure rates for modular assembly of engineered zinc fingers. *Nature Methods*, **5(5)**, 374-375.

**Reed, R.** (2000). Mechanisms of fidelity in pre-mRNA splicing. *Curr. Opin. Cell Biol.* **12**; 340–345.

**Rest, J.S. y Mindell, D.P.** (2003) Retroids in archaea: phylogeny and lateral origins. *Mol Biol Evol* **20**; 1134–1142.

**Robart A. R. y Zimmerly S.** (2005) Group II intron retroelements: function and diversity. *Cytogenet Genome Res*, **110**: 589-597.

**Robart, A.R., Seo, W. y Zimmerly, S.** (2007). Insertion of group II intron retroelements after intrinsic transcriptional terminators. *Proc Natl Acad Sci* **104**: 6620–6625.



---

**Robertsen, B.K., Aiman, P., Darvill, A.G., McNeil, M. y Albersheim, P.** (1981) The structure of acidic extracellular polysaccharides secreted by *Rhizobium leguminosarum* y *Rhizobium trifolii*. *Plant Physiol.* **67**: 389-400.

**Roitzsch, M., y Pyle, A.M.** (2009). The linear form of a group II intron catalyzes efficient autocatalytic reverse splicing, establishing a potential for mobility. *RNA* **15**: 473-482.

**Rüberg, S., Tian, Z-X., Krol, E., Linke, B., Meyer, F., Wang, Y., Puhler, A., Weidner, S. y Becker, A.** (2003) Construction and validation of a *Sinorhizobium meliloti* whole genome DNA microarray: genome-wide profiling of osmoadaptive gene expression. *J Biotechnol*, **106(2-3)**:255-268.

**Ruskin, B., Krainer, A.R., Maniatis, T. y Green, M.R.** (1984) Excision of an intact intron as a novel lariat structure during pre-mRNA splicing *in vitro*. *Cell*, **38**; 317-331.

**Russell, D. W., e Hirata, R. K.** (1998). Human gene targeting by viral vectors. *Nature Genetics*, **18(4)**, 325-330.

**Salama, N. R., B. Shepherd y S. Falkow** (2004) Global transposon mutagenesis and essential gene analysis of *Helicobacter pylori*. *J Bacteriol*, **186**: 7926-35

**Saldanha, R., Chen, B., Wank, H., Matsuura, M., Edwards, J., y Lambowitz, A.M.** (1999). RNA and protein catalysis in group II intron splicing and mobility reactions using purified components. *Biochemistry*, **38**; 9069-9083.

**Sambrook, J., Fritsch, E.F. y Maniatis, T.** (1989) Molecular cloning. A laboratory Manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY.

**Samulski, R. J., Chang, L. S., y Shenk, T.** (1989). Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. *Journal of Virology*, **63(9)**, 3822-3828.

**Sandy, P., Ventura, A., y Jacks, T.** (2005). Mammalian RNAi: a practical guide. *BioTechniques*, **39(2)**, 215-224.

**San Filippo, J. y Lambowitz, A.M.** (2002). Characterization of the C-terminal DNA-binding/DNA endonuclease region of a group II intron-encoded protein. *J. Mol. Biol.* **324**; 933–951.

**Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E. y Tabata, S.** (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. *DNA Res.*, **6**; 283-290.

**Schäfer, B., Gan, L. y Perlman, P.S.** (2003) Reverse transcriptase an reverse splicing activities encoded by the mobile group II intron COB11 of fission yeast mitochondrial DNA. *J. Mol. Biol.*, **329**; 191-206.

**Schmelzer, C., Schmidt, C., y Schweyen, R.J.** (1982). Identification of splicing signals in introns of yeast mitochondrial split genes: mutational alterations in intron bI1 and secondary structures in related introns. *Nucleic Acids Res.* **10**; 6797–6808.

**Schmelzer, C. y Schweyen, R.J.** (1986). Self-splicing of group II introns *in vitro*: mapping of the branch point and mutational inhibition of lariat formation. *Cell*, **46**; 557–565.

**Schmidt, U., Riederer, B., Morl, M., Schmelzer, C., y Stahl, U.** (1990). Self-splicing of the mobile group II intron of the filamentous fungus *Podospora anserina* (COI I1) *in vitro*. *EMBO J.* **9**; 2289–2298.

**Schmidt, U., Podar, M., Stahl, U., y Perlman, P.S.** (1996). Mutations of the two-nucleotide bulge of D5 of a group II intron block splicing *in vitro* and *in vivo*: phenotypes and suppressor mutations. *RNA*, **2**; 1161–1172.

**Schröder, H., Langer, T., Hartl, F.U., y Bukau, B.** (1993) DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. *EMBO J.* **12**: 4137-4144.

**Sellem, C.H., Lecellier, G., y Belcour, L.** (1993) Transposition of a group II intron. *Nature*, **366**; 176-178.



**Sèraphin, B., Simon, M., Boulet, A. y Faye, G.** (1989). Mitochondrial splicing requires a protein from a novel helicase family. *Nature* **337**: 84–87.

**Sharp, P.A.** (1985). On the origin of RNA splicing and introns. *Cell* **42**; 397–400.

**Sharp, P.A.** (1991) “Five easy pieces”. *Science*, **254**; 663.

**Shearman, C., Godon, J.J., y Gasson, M.** (1996). Splicing of a group II intron in a functional transfer gene of *Lactococcus lactis*. *Mol. Microbiol.* **21**; 45–53.

**Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., et al.** (1986). The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J*, **5(9)**; 2043–2049.

**Shub, D.A., Goodrich-Blair, H. y Eddy, S.R.** (1994) Amino acid sequence motif of group I endonucleases is conserved in open reading frames of group II introns. *Trends Biochem Sci* **19**: 402–404.

**Shukla, G.C. y Padgett, R.A.** (2002) A catalytically active group II intron domain 5 can function in the U12-dependent spliceosome. *Mol Cell*, **9(5)**; 1145–50.

**Simon, D.M., Clark, N.A., McNeil, B.A., Johnson, I., Pantuso, D., Dai, L., Chai, D. y Zimmerly, S.** (2008). Group II introns in eubacteria and archaea: ORF-less introns and new varieties. *RNA* **14**:1704–1713.

**Simon, D.M., Kelchner, S.A., Zimmerly, S.** (2009). A broad-scale phylogenetic analysis of group II intron RNAs and intron-encoded reverse transcriptases. *Mol Biol Evol* **26**: 2795–2808.

**Singh, N.N. y Lambowitz, A.M.** (2001). Interaction of a group II intron ribonucleoprotein endonuclease with its DNA target site investigated by DNA footprinting and modification interference. *J. Mol. Biol.* **309**; 361–386.

**Singh, R.N., Saldanha, R.J., D’Souza, L.M., y Lambowitz, A.M.** (2002). Binding of a group II intron-encoded reverse transcriptase/maturase to its

high affinity intron RNA binding site involves sequence-specific recognition and autoregulates translation. *J. Mol. Biol.* **318**; 287–303.

**Skelly, P.J., Hardy, C.M., y Clark-Walker, G.D.** (1991). A mobile group II intron of a naturally occurring rearranged mitochondrial genome in *Kluyveromyces lactis*. *Curr. Genet.* **20**; 115–120.

**Solem, A., Zingler, N., Pyle, A.M. y Li-Pook-Than J.** (2009). Group II introns and their protein collaborators. En Walter NG (Ed.) *Non-Protein Coding RNAs* Berlin: Springer-Verlag. p. 167–182.

**Smith D, Zhong J, Matsuura M, Lambowitz AM y Belfort M** (2005) Recruitment of host functions suggests a repair pathway for late steps in group II intron retrohoming. *Genes Dev.* **19**; 2477–2487.

**Sontheimer, E.J., Gordon, P.M., y Piccinilli, J.A.** (1999) Metal ion catalysis during group II intron self-splicing: parallels with the spliceosome. *Genes and Dev.* **13**; 1729-1741.

**Stabell, F.B., Tourasse, N.J., Ravnum, S., y Kolstø, A.B.** (2007). Group II intron in *Bacillus cereus* has an unusual 39 extension and splices 56 nucleotides downstream of the predicted site. *Nucleic Acids Res.*, **35**; 1612 - 1623.

**Stabell, F.B., Tourasse, N.J. y Kolsto, A.B.** (2009). A conserved 3' extension in unusual group II introns is important for efficient second-step splicing. *Nucleic Acids Res* 37:3202–3214.

**Steiner, M., Karunatilaka, K.S., Sigel, R.K. y Rueda, D.** (2008). Single molecule studies of group II intron ribozymes. *Proc Natl Acad Sci USA* **105**:13853–13858.

**Stern, D.B., Goldschmidt-Clermont, M. y Hanson, M.R.** (2010). Chloroplast RNA metabolism. *Ann Rev Plant Biol* **61**: 125-155.

**Stiens M, Becker A, Bekel T, Gödde V, Goesmann A, Niehaus K, et al.** (2008). Comparative genomic hybridisation and ultrafast pyrosequencing revealed remarkable differences between the *Sinorhizobium meliloti* genomes



---

of the model strain Rm1021 and the field isolate SM11. *J. Biotechnol.* **136(1-2)**: 31-37.

**Studier, F.W.** (1991) Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J. Mol Biol.* **219**: 37-44.

**Su, L.J., Waldsich, C., y Pyle, A.M.** (2005). An obligate intermediate along the slow folding pathway of a group II intron ribozyme. *Nucleic Acids Res.* **33**: 6674-6687.

**Symons, R.H.** (1992) Small catalytic RNAs. *Annu. Rev. Biochem.* **61**: 641-671.

**Tenson, T., DeBlasio, A., y Mankin, A.** (1996). A functional peptide encoded in the *Escherichia coli* 23S rRNA. *Proc. Natl. Acad. Sci USA.* **93**, 5641-5646.

**Tenson, T Xiong, L., Kloss, P. y Mankin, A. S.** (1997). Erythromycin resistance peptides selected from random peptide libraries. *J. Biol. Chem.* **272**, 17425-17430.

**Tenson, T. y Mankin, A. S.** (2001). Short peptides conferring resistance to macrolide antibiotics. *Peptides.* **22**, 1661-1668.

**Theimer, C.A., Blois, C.A. y Feigon, J.** (2005). Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol Cell* **17**:671-682.

**Till, B., Schmitz-Linneweber, C., Williams-Carrier, R. y Barkan, A.** (2001). CRS1 is a novel group II intron splicing factor that was derived from a domain of ancient origin. *RNA* **7**: 1227-1238.

**Toor, N., Hausner, G. y Zimmerly, S.** (2001) Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA* **7**; 1142-1152.

**Toor, N. y Zimmerly, S.** (2002) Identification of a family of group II introns encoding LAGLIDADG ORFs typical of group I introns. *RNA*, **8**; 1373-1377.

- Toor, N., Robart, A.R., Christianson, J. y Zimmerly, S.** (2006) Self-splicing of a group IIC intron: 50 exon recognition and alternative 50 splicing events implicate the stem-loop motif of a transcriptional terminator. *Nucleic Acids Res* **34**; 6561–6573.
- Toor, N., Keating, K.S., Taylor, S.D. y Pyle, A.M.** (2008a) Crystal structure of a self-spliced group II intron. *Science*, **320**; 77-82.
- Toor, N., Rajashankar, K., Keating, K.S. y Pyle, A.M.** (2008b). Structural basis for exon recognition by a group II intron. *Nat Struct Mol Biol* **15**:1221–1222.
- Toor, N., Keating, K.S. y Pyle, A.M.** (2009). Structural insights into RNA splicing. *Curr Opin Struct Biol* **19**:260-266
- Toor, N., Keating, A., Fedorova, O., Rajashankar, K., Wang, J. y Pyle, A.M.** (2010). The tertiary architecture of a group II intron. *RNA* **16**:57–69.
- Toro, N., Molina-Sánchez, M.D. y Fernández-López, M.** (2002) Identification and characterization of bacterial class E group II introns. *Gene* **299**: 245-250.
- Toro, N., Martínez-Abarca, F., Fernandez-Lopez, M. y Muñoz-Adelantado, E.** (2003) Diversity of group II introns in the genome of *Sinorhizobium meliloti* strain 1021: splicing and mobility of RmInt1. *Mol. Genet. Genomics*, **268**(5); 628-36.
- Toro, N.** (2003) Bacteria and archaea group II introns: additional mobile elements in the environment. *Environ Microbiol* **5**; 143–151.
- Toro, N., Jiménez-Zurdo, J.I. y García-Rodríguez, F.M.** (2007). Bacterial group II introns: not just splicing. *FEMS Microbiol. Rev.* **31**; 342–358.
- Tourasse, N.J., Stabell, F.B., Reiter, L., y Kolstø, A.B.** (2005). Unusual group II introns in bacteria of the *Bacillus cereus* group. *J. Bacteriol.* **187**; 5437–5451.



- Valadkhan, S., Mohammadi, A., Jaladat, Y. y Geisler, S.** (2009). Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc Natl Acad Sci* **106**: 11901–11906.
- Vallès, Y., Halanych, K.M. y Boore, J.L.** (2008) Group II introns break new boundaries: presence in a bilaterian's genome. *PLoS ONE*, **3(1)**; e1488.
- Van der Veen, R., Arnberg, A.C., van der Horst, G., Bonen, L., Tabak, H.F., y Grivell, L.A.** (1986). Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing *in vitro*. *Cell*, **44**; 225–234.
- Van der Veen, R., Kwakman, J.H., y Grivell, L.A.** (1987). Mutations at the lariat acceptor site allow self-splicing of a group II intron without lariat formation. *EMBO J.* **6**; 3827–3831.
- van Gent, D. C., Hoeijmakers, J. H., y Kanaar, R.** (2001). Chromosomal stability and the DNA double-stranded break connection. *Nature Reviews. Genetics*, **2(3)**, 196-206.
- Vernon, J.L.** (2008) Toward group II intron-based genome targeting in eukaryotic cells. *Dissertation* The University of Texas at Austin.
- Villadas, P.J., Velázquez, E., Martínez-Molina, E. y Toro, N.** (1995) Identification of nodule-dominant *Rhizobium meliloti* strains carrying pRmeGR4b-type plasmid within indigenous soil populations by PCR using primers derived from specific DNA sequences. *FEMS Microbiol. Ecol.* **17**;161-168.
- Vogel, J., Börner, T. y Hess, W.R.** (1999). Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plant mutants. *Nucleic Acids Res* **27**: 3866–3874.
- Vogel, J. y Börner, T.** (2002). Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. *EMBO J.* **21**; 3794–3803.
- Wais, R.J., Wells, D.H. y Long, S.R.** (2002). Analysis of differences between *Sinorhizobium meliloti* 1021 and 2011 strains using the host calcium spiking response. *M.P.M.I.* **15**; 1245-1252

**Waldsich, C. y Pyle, A.M.** (2007). A folding control element for tertiary collapse of a group II intron ribozyme. *Nat. Struct. Mol. Biol.* **14**; 37–44.

**Wang, B., y Zhou, J.** (2003). Specific genetic modifications of domestic animals by gene targeting and animal cloning. *Reproductive Biology and Endocrinology: RB&E*, **1**, 103.

**Wang, H. y Lambowitz, A.M.** (1993) The Mauriceville plasmid reverse transcriptase can initiate cDNA synthesis de novo and may be related to reverse transcriptase and DNA polymerase progenitor. *Cell*, **75**; 1071–1081.

**Wang, S., Chen, Y., Sun, Z., Zhou, Q. y Sui, S.** (2006) *Escherichia coli* CorA Periplasmic Domain Functions as a Homotetramer to Bind Substrate. *Journal of Biological Chemistry*, **281**: 26813-26820.

**Wank, H., San Filippo, J., Singh, R.N., Matsuura, M. y Lambowitz, A.M.** (1999) A reverse transcriptase/maturase promotes splicing by binding at its own coding segment in a group II intron RNA. *Mol. Cell*, **4**; 239-250.

**Watanabe, K. y Lambowitz, A.M.** (2004). High-affinity binding site for a group II intron-encoded reverse transcriptase/maturase within a stem-loop structure in the intron RNA. *RNA*, **10**; 1433–1443.

**Weghuber, J., Dietrich, F., Froschauer, E.M., Svidová, S. y Schweyen, J.R.** (2006) Mutational analysis of functional domains in Mrs2p, the mitochondrial Mg<sup>2+</sup> channel protein of *Saccharomyces cerevisiae*. *FEBS Journal*, **273**: 1198-1209.

**Woese C.** (1967). *The genetic code*, pp. 179-195. Harper & Row, New York.

**Xiang, Q., Qin, P.Z., Michels, W.J., Freeland, K., y Pyle, A.M.** (1998). Sequence specificity of a group II intron ribozyme: multiple mechanisms for promoting unusually high discrimination against mismatched targets. *Biochemistry* **37**; 3839–3849.

**Xiong, Y. y Eickbush, T.H.** (1990). Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* **9**; 3353–3362.



- Yan, Z., Sun, X., y Engelhardt, J. F.** (2009). Progress and prospects: techniques for sitedirected mutagenesis in animal models. *Gene Therapy*, **16(5)**, 581-588.
- Yang, J., Zimmerly, S., Perlman, P.S., y Lambowitz, A.M.** (1996). Efficient integration of an intron RNA into double-stranded DNA by reverse splicing. *Nature*, **381**; 332-335.
- Yang, J., Mohr, G., Perlman, P.S., y Lambowitz, A.M.** (1998). Group II intron mobility in yeast mitochondria: target DNA-primed reverse transcription activity of aI1 and reverse splicing into DNA transposition sites *in vitro*. *J. Mol. Biol.* **282**; 505-523.
- Yang, Y., Chen, Y., Zhang, C., Huang, H. y Weissman, S.M.** (2002) Nucleolar localization of hTERT protein is associated with telomerase function. *Exp. Cell Res.* **277**:201-209.
- Yao, J., Zhong, J. y Lambowitz, A.M.** (2005). Gene targeting using randomly inserted group II introns (targetrons) recovered from an *Escherichia coli* gene disruption library. *Nucleic Acids Res*, **33**; 3351-62.
- Yao, J., Zhong, J., Fang, Y., Geisinger, E., Novick, R.P. y Lambowitz, A.M.** (2006). Use of targetrons to disrupt essential and non-essential genes in *Staphaphylococcus aureus* reveals temperature-sensitivity of group II intron splicing. *RNA*, **12**; 1271-81.
- Yao, J. y Lambowitz, A.M.** (2007). Gene targeting in Gram-negative bacteria by use of a mobile group II intron ("Targetron") expressed from a broad-host-range vector. *Appl Environ Microbiol*, **73**; 2735-43.
- Zarschler, K., B. Janesch, S. Zayni, C. Scha"ffer, and P. Messner.** (2009).Construction of a gene knockout system for application in *Paenibacillus alvei* CCM 2051T exemplified with the initiation enzyme WsfP of S-layer glycan biosynthesis. *Appl. Environ. Microbiol.* **75**: 3077-3085.
- Zhao J, Niu W, Marcotte E, Lambowitz A.** (2008). Group II intron protein localization and insertion sites are affected by polyphosphate. *PLOS Biol.* **6**; e150.

**Zhong J, Lambowitz AM.** (2003). Group II intron mobility using nascent strands at DNA replication forks to prime reverse transcription. *EMBO J.* **22**; 4555–65

**Zhong J., Karberg M., y Lambowitz A.M.** (2003) Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. *Nucleic Acids Res* **31**: 1656-1664.

**Zhuang, F., Karberg, M., Perutka, J. y Lambowitz, A.M.** (2009a) EcI5, a group IIB intron with high retrohoming frequency: DNA target site recognition and use in gene targeting. *RNA* **15**: 432-449.

**Zhuang, F., Mastroianni, M., White, T.B. y Lambowitz AM.** (2009b). Linear group II intron RNAs can retrohome in eukaryotes and may use nonhomologous end-joining for cDNA ligation. *Proc Natl Acad Sci* **106**: 18189–18194.

**Zimmerly S., Guo H., Perlman P. S. y Lambowitz A. M.** (1995a) Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell*, **82**: 545-554.

**Zimmerly S., Guo H., Eskes R., Yang J., Perlman P. S. y Lambowitz A. M.** (1995b) A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell*, **83**: 529-538.

**Zimmerly S., Hausner G. y Wu X. C.** (2001) Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res*, **29**: 1238-1250.

**Zoschke, R., Nakamura, M., Liere, K., Sugiura, M., Börner, T. y Linneweber-Schmitz, C.** (2010). An organellar maturase associated with multiple group II introns. *Proc Natl Acad Sci* **107**: 3245–3250

**Zuker, M.** (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31** (13); 3406-3415.

**ABREVIATURAS**

**Y**

**ANGLICISMOS**



**μg**: Microgramo  
**μl**: Microlitro  
**μm**: Micrómetro o micra  
**3'-ss**: sitio de maduración 3' o unión del extremo 3' del intrón con el exón 2, del inglés 3' Splice Site  
**5'-ss**: sitio de maduración 5' o unión del exón 1 con el extremo 5' del intrón, del inglés 5' Splice Site  
**A<sub>260nm</sub>**: Absorbancia a 600 nm de longitud de onda  
**A**: Adenina  
**Ap<sup>R</sup>**: Resistencia a ampicilina  
**Ap**: Ampicilina  
**APS**: Persulfato amónico  
**bp**: pares de bases  
**Branching/Branch**: ramificación, bifurcación  
**C**: Citosina  
**cdNA**: ADN complementario; ADN sintetizado a partir de ARN por una RT  
**Ci**: Curio  
**CIA**: mezcla de cloroformo:alcohol isoamílico en proporción 24:1  
**cm**: centímetro  
**CPM**: cuentas por minuto  
**cRT-PCR**: RT-PCR precedida de ligación con T4 ARN ligasa  
**CSPD**: sustrato que al hidrolizarse emite luz  
**D.O.<sub>260nm</sub>**: Densidad óptica a 260 nm  
**dcDNA**: ADN de cadena doble  
**ddNTP**: Di-deoxinucleósido trifosfato (ddATP, ddCTP, ddGTP, ddTTP)  
**dI- dVI**: dominios de la ribozima, dominio I-dominio VI  
**DMS**: Dimetil sulfato  
**DNA**: Ácido Desoxirribonucleico, ADN.  
**dNTP**: Deoxinucleósido trifosfato (dATP, dCTP, dGTP, dTTP, dUTP)  
**DSBR**: reparación de roturas en la doble cadena, del inglés Double Strand Break Repair  
**DTT**: 1,4 Ditiotreitol  
**EBS1, EBS2, EBS3**: Secuencias de unión al exón, del inglés Exon Binding Sites  
**EDTA**: Ácido etilendiaminotetraacético  
**et al.**: y colaboradores  
**FRT**: sitios de recombinación para la enzima Flp, del inglés Flipase Recognition Target  
**G**: Guanina

**GFP:** Proteína verde fluorescente, del inglés Green Fluorescent Protein

**HIV-1:** virus de la inmunodeficiencia adquirida tipo 1, del inglés Human Immunodeficiency Virus

**Homining:** Proceso de inserción del intrón del grupo II en una copia del alelo de su elemento hospedador original

**HPLC:** cromatografía líquida de alta afinidad, del inglés High-Performance Liquid Chromatography

**IBS1, IBS2, IBS3:** Secuencias de unión al intrón, del inglés Intron Binding Sites

**IEP:** Proteína codificada por el intrón, del inglés Intron Encoded Protein

**IS:** Secuencia de inserción, del inglés, Insertion Sequence

**K:** puede corresponderse con los nucleótidos G o U

**Kb:** Kilobases

**Kit:** Equipo

**Km:** Kanamicina

**Km<sup>R</sup>:** Resistencia a kanamicina

**LAG:** durante la replicación, la hebra que se sintetiza de manera discontinua o retardada, del inglés lagging. En nuestro caso hace referencia a la orientación del sitio diana del intrón respecto a la dirección de la replicación; así, el intrón se insertaría en la hebra que sirve como molde para la síntesis de la hebra retardada.

**Lariat:** forma escindida del intrón consistente en una característica forma en lazo

**LB:** medio Luria-Bertani

**LEAD:** durante la replicación, la hebra que se sintetiza de manera continua, del inglés leading. La diana para la inserción del intrón se encuentra en la hebra que sirve como molde para la síntesis de la hebra continua.

**LTR:** terminaciones largas repetidas del inglés (Long Terminal Repeats)

**M:** Molar

**Mb:** Megabase

**MDa:** Megadaltons

**min:** Minuto

**ml:** Mililitro

**mm:** Milímetro

**mM:** Milimolar

**mRNA:** ARN mensajero

**ng:** Nanogramo

**nm:** Nanometro

**nt:** nucleótido

**NTP o rNTP:** Ribonucleósido trifosfato (ATP, CTP, GTP, TTP, UTP)

**ORF:** Marco abierto de lectura, del inglés Open Reading Frame

**p/v:** peso/volumen

**pb:** Pares de bases

**PCR:** Reacción en cadena de la polimerasa, del inglés Polymerase Chain Reaction

**PEG:** Polietilenglicol

**pmol:** picomol

**PPE:** extensión de cebador envenenada, del inglés Primers-Poisoned Extension

**pT7:** promotor de la ARN polimerasa del fago T7

**r.p.m.:** Revoluciones por minuto

**R:** purina (A o G)

**RAM:** marcador de selección activado por retrotransposición, del inglés Retrotransposition Activated Marker

**RFP:** Proteína roja fluorescente, del inglés Red Fluorescent Protein

**Retrohoming:** Proceso de inserción del intrón del grupo II en una copia del alelo de su elemento hospedador original mediada por un intermediario RNA

**RMN:** Resonancia Magnética Nuclear

**RNA:** Ácido Ribonucleico

**RNP, RNPs:** Partículas de RiboNucleoProteína o complejos ribonucleoproteicos

**RPA:** ensayo de protección frente a digestión con RNAasas, del inglés Ribonuclease Protection Assay

**rRNA:** ARN ribosómico

**RT:** Reverso Transcriptasa y también, dependiendo del contexto, reverso transcripción.

**rTn:** retrotransposón

**RT-PCR:** combinación de ambas técnicas: síntesis de cDNA seguido de amplificación

**S.D.:** desviación estándar

**scDNA:** ADN de cadena simple o sencilla

**SDS:** Sodio dodecil sulfato

**SER:** reacción de reapertura de los exones, del inglés Spliced Exons Reopening

**snoRNA:** ARN nucleolar pequeño

**snRNA:** ARN nuclear pequeño

**Splicing:** Maduración del intrón del grupo II mediante un proceso por corte y empalme

**SSC:** Solución Salina Concentrada  
**Tc<sup>R</sup>:** Resistencia a tetraciclina  
**Tc:** Tetraciclina  
**TEMED:** N,N,N,N.tetrametiletenodiamina  
**Tetra<sup>loop</sup>:** región desapareada compuesta por 4 nt  
**TPRT:** Transcripción Reversa a partir de Cebador, del inglés Target DNA-  
Primed Reverse Transcription  
**tRNA:** ARN de transferencia  
**TY:** medio Triptona-Yeast  
**TYA:** medio Triptona-Yeast modificado  
**U:** Uracilo  
**UV:** ultravioleta  
**UV-crosslinking:** entrecruzamientos moleculares de proximidad inducidos  
por luz ultravioleta  
**v/v:** volumen/volumen  
**V:** Voltio  
**V<sub>F</sub>:** volumen final  
**W-C:** referente a apareamientos canónicos Watson-Crick  
**WT:** silvestre, del inglés Wild-Type  
**X-Gluc:** 5-bromo-4-cloro-3-indolil-β-D-glucuronida  
**Y:** pirimidina (C, T o U)  
**ΦCIA:** mezcla de fenol:cloroformo:alcohol isoamílico en proporción 25:24:1

# **AGRADECIMIENTOS**



## **AGRADECIMIENTOS**

Los que me conocéis sabéis que soy persona de pocas palabras, y este capítulo es el que más me va a costar escribir. Pero quiero expresar mi gratitud a todas aquellos que, de una forma u otra, me han ayudado a llegar hasta aquí.

En primer lugar, debo agradecer a mi director de tesis Nicolás Toro García el que me diese la oportunidad de formar parte de este grupo y de llevar a cabo este proyecto. Gracias por tu apoyo y confianza durante estos años.

Gracias José Ignacio (J.I. a partir de ahora). Gracias por tu paciencia conmigo en los inicios de mi vida científica, por saber enseñar (de la U.D.S. no diré nada). Gracias a Paco, por esas intensas y fructíferas discusiones científicas que tanto ayudan a desarrollar la mente científica. Gracias también a Manolo, por todos los esfuerzos que ha realizado por y para el grupo, a veces poco reconocidos, y por las charlas más allá de la ciencia. También a ti Pablo, gracias por pelearte con los pedidos, y por hacer más distendido el ambiente a veces un poco cargado de un laboratorio, no te preocupes, que el C.B. seguro que sube el año que viene. Gracias a Steffi, aprendí mucho de ella, lástima que se marchase tan pronto.

Gracias Fernando, tu apoyo y amistad desde que llegaste han sido muy importantes. Gracias por tu serenidad y por tener siempre un hueco para hablar.

Chema, llevas poco en el grupo, pero tu ayuda durante ese tiempo ha sido muy importante.

Muchas gracias a mis compañeros en precario (Merche, Gloria, Omar, Pepe, J.A., Antonio, Hurry, Isa, Laura, Alicia y Alex) por estar siempre dispuestos a ayudar y a compartir las experiencias, que son muchas horas metidos en el labo.

A Sensi, Tita, Vane y Tere, por vuestro apoyo al grupo, sin el cual todo sería más difícil y lento.

Muchas gracias a los compañeros de otros grupos, muchas gracias por vuestra ayuda en los momentos de necesidad. Sois muchos, por lo que para que nadie se me enfade por que se me olvide nombrarlo, os lo agradezco a todos, vosotros sabéis quienes sois.

Muchas gracias a todo el personal de la Estación Experimental del Zaidín, sobre todo a Mantenimiento (Los Antonios, Los Fernandos, Samu, etc.) por vuestra ayuda y disponibilidad cuando la he necesitado (que no ha

sido pocas veces). No me puedo olvidar de los tres mosqueteros (Jimi, Luis y José Luis), siempre arrancándome una sonrisa. Muchas gracias también a César, Javi y Pablo.

No me puedo olvidar de mis estancias en el grupo del Dr. Martin Crespi. Muchas gracias, sobre todo a Martín, Florian y Philippe, por recibirme con los brazos abiertos y por ayudarme en todo. Gracias a José Ángel y Amada por ocuparse de mí y por los buenos ratos pasados en la *France*.

Quiero agradecer a mi Padre (que no ha podido disfrutar de este día) y a mi Madre (que ha sabido "*tirar pa lante*"), por todos los esfuerzos y sacrificios que han hecho para que yo pueda estar aquí hoy, por la educación y el amor que me han dado. Muchas gracias a mis hermanos por su apoyo y sus consejos, nunca suficientes. A mi Tio Rafael, espero llegar a tu edad con ese espíritu y esa alegría.

Llega tu turno Lola, de no ser por ti, no habría comenzado esta andadura. Gracias por descubrirme este mundo, por estar a mi lado siempre, en los buenos y en los malos momentos, ayudándome y animándome, aún ahora en la distancia te noto a mi lado. Muchas gracias por estar ahí, eres lo mejor que me ha pasado. GRACIAS.

En definitiva, muchas gracias a todos, y perdón si me he olvidado de alguien. Aquí me tenéis para lo que queráis.