



UNIVERSIDAD DE JAÉN
FACULTAD DE CIENCIAS EXPERIMENTALES

Grado en Biología
Trabajo Fin de Grado



**Caracterización de una familia
de ADN repetitivo en el
genoma de *Heliotaurus
ruficollis* (Coleoptera,
Tenebrionidae)**

Jesús Vela Herrador

Jaén, 2013



UNIVERSIDAD DE JAÉN
FACULTAD DE CIENCIAS EXPERIMENTALES

Grado en Biología
Trabajo Fin de Grado



**Caracterización de una
familia de ADN repetitivo en
el genoma de *Heliotaurus
ruficollis* (Coleoptera,
Tenebrionidae)**

Fdo: Jesús Vela Herrador

INDICE

1.	ABSTRACT	1
2.	INTRODUCCION	3
2.1.	Generalidades de Tenebriónidos (Coleoptera, Tenebrionidae)	3
2.2.	ADN repetitivo de los genomas eucariotas	3
2.2.1.	ADN altamente repetitivo	4
2.2.2.	ADN moderadamente repetitivo	5
2.2.2.1.	ADN repetitivo disperso	5
2.2.2.2.	ADN repetitivo en tándem	5
2.3.	Características de los principales tipos de ADN repetitivo: ADN satélite y elementos transponibles	6
2.3.1.	ADN satélite	6
2.3.2.	Elementos transponibles	7
2.3.2.1.	Elementos de la Clase I	8
2.3.2.2.	Elementos de la Clase II	8
2.4.	ADN repetitivo en Tenebriónidos	9
3.	HIPÓTESIS Y OBJETIVOS	14
4.	MATERIAL Y MÉTODOS	
4.1	Material	16
4.2	Métodos	16
4.2.1.	Extracción de ADN genómico	16
4.2.2.	Aislamiento del ADN repetitivos mediante digestión del ADN genómico con enzimas de restricción	17

4.2.3.	Purificación de ADN a partir de agarosa	18
4.2.4.	Preparación del vector y reacciones de ligado de las bandas de ADN genómico	19
4.2.5.	Transformación	19
4.2.6.	Extracción de ADN plasmídico	19
4.2.7.	Marcaje de sondas con digoxigenina mediante "random priming"	20
4.2.8.	Comprobación del grado de marcaje de las sondas marcadas con digoxigenina	20
4.2.9.	Selección de plásmidos recombinantes mediante dot-blot	21
4.2.10.	Amplificación de ADN por PCR	22
4.2.11.	Reacciones de ligado de los productos de PCR	23
4.2.12.	Selección de plásmidos recombinantes mediante lisis rápida	23
4.2.13.	Conservación de los clones	23
4.2.14.	Preparación de ADN plasmídico para secuenciación	24
4.2.15.	Southern-blot del ADN genómico digerido con enzimas de restricción	24
4.2.16.	Análisis de las secuencias	25
4.2.17.	Composición y preparación de los reactivos utilizados	25
5.	RESULTADOS Y DISCUSIÓN	27
6.	BIBLIOGRAFÍA	44

1. ABSTRACT

Tenebrionids (Coleoptera, Tenebrionidae) are a good model for the study of repetitive DNA, because higher amounts of this special DNA are present in their genomes. The information about other repetitive DNAs is very limited unless *Tribolium castaneum*, which genome has been sequenced. Eight subfamilies are included in the Tenebrionidae family. All studies about repetitive DNA have been carried out only in three subfamilies: Tenebrioninae, Palorinae y Pimeliinae. The aim of this study is to begin the analysis of the repetitive DNA in others tenebrionid subfamilies, concretely in *Heliotaurus ruficollis* (Alleculinae subfamily). In this species a new repetitive DNA family has been characterized, named Hruf-Alu. Its features suggest that it is a LTR-retrotransposon like. The sequence analysis showed that it as defective an element which only conserves the reverse transcriptase domain.

2. INTRODUCCIÓN

2.1. Generalidades de Tenebriónidos (Coleoptera, Tenebrionidae)

Heliotaurus ruficollis Fabricius, objeto de estudio de este trabajo, es un insecto coleóptero que pertenece a la familia Tenebrionidae. Esta familia es bastante extensa con unos 1700 géneros y al menos 18000 especies. La mayor parte de los tenebriónidos están bien adaptados a hábitats xerófilos, colonizando algunos de ellos los ambientes desérticos. Son saprófagos, alimentándose de casi cualquier tipo de materia orgánica (restos vegetales, madera muerta, restos de algas o incluso excrementos). No obstante, existen también especies que son fungívoras, xilófagas e incluso necrófagas. Algunos miembros de la familia *Alleculidae*, como *H. ruficollis* son florícolas y se alimentan de flores y polen. Se le considera una especie polinizadora si bien el hecho de que se alimente de los órganos florales puede hacer que en gran número tengan un efecto perjudicial (Herrera 1988).

Algunas especies de esta familia son plagas de productos alimenticios como sustancias amiláceas (especies pertenecientes a los géneros *Tenebrio* y *Tribolium*) o café (especies del género *Gonocephalum*) (Booth et al. 1990, Oromí 1982). Cabe destacar el ejemplo de *Tenebrio molitor* cuyas larvas son también conocidas como "gusano de la harina" por ser plaga de este alimento.

Se pueden distinguir ocho subfamilias de tenebriónidos: *Alleculinae*, *Coelometopinae*, *Diaperinae*, *Lagriinae*, *Palorinae*, *Phrenapatinae*, *Pimeliinae* y *Tenebrioninae*. Dentro de la subfamilia *Alleculidae* se engloban diversos géneros, entre ellos *Heliotaurus* que es el que nos ocupa.

2.2. ADN repetitivo de los genomas eucariotas

Cuando el ADN de una especie se somete al estudio de su cinética de reasociación, la porción de este que primero se reasocia es el ADN repetitivo y la que más tarda corresponde al ADN de copia única (Britten & Davidson 1971). En los organismos eucariotas, el porcentaje de ADN no codificante es muy variable, pudiendo constituir entre el 30 y el 99% del genoma (Elgar & Vavouri 2008,

Cavalier-Smith 1985). Dentro de este ADN no codificante se encuentran las secuencias repetitivas cuyo dinamismo genera reordenamientos en el genoma y por tanto una divergencia evolutiva que tiene utilidad, por ejemplo, para identificar especies o para la realización de estudios filogenéticos. El ADN repetitivo está compuesto principalmente por repeticiones dispersas y en tándem (Slamovits & Rossi 2002). Pueden diferenciarse dos tipos básicos de ADN repetitivo: ADN altamente repetitivo y ADN moderadamente repetitivo (**Figura 1**).

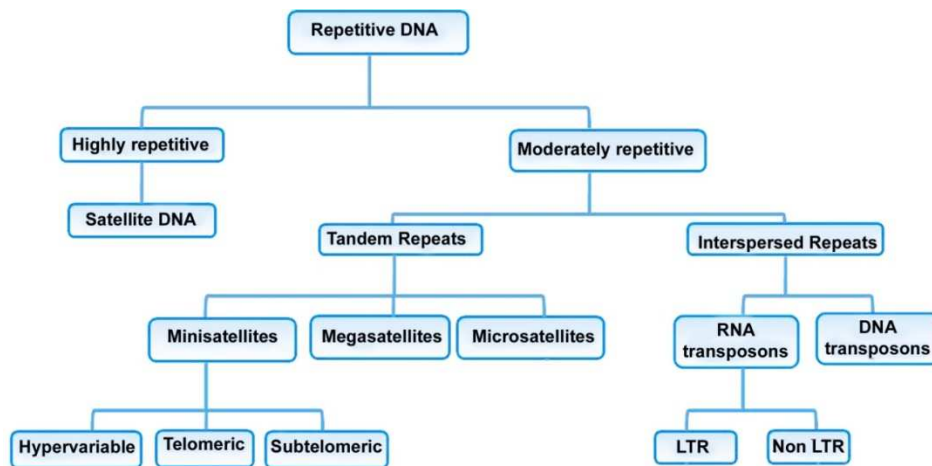


Figura 1. Clasificación de los diferentes tipos de ADN repetitivo. Tomada de Pathak & Ali (2012).

2.2.1. ADN altamente repetitivo

Se consideran ADN altamente repetitivo las secuencias cortas de 5 a 200 pb que se repiten miles o millones de veces en el genoma, llegando a ocupar entre el 10 y el 15% de éste, normalmente en forma de repeticiones en tándem. Este tipo de repetitivo, suele encontrarse en las regiones heterocromáticas, de centrómeros o telómeros.

Dentro del ADN altamente repetitivo destaca el ADN satélite (ADNsat), que puede considerarse el más abundante, al llegar a constituir entre el 1 y el 70% del genoma de un organismo (Charlesworth et al. 1994), aunque este porcentaje puede variar ampliamente entre especies, incluso pertenecientes al mismo género (Palomeque & Lorite, 2008). Cuando el ADN genómico se centrifuga en gradiente

de densidad con CsCl se forma una pequeña banda constituida por este tipo de ADN repetitivo, separada de la banda principal por su densidad diferente (Kit 1961).

2.2.2. ADN moderadamente repetitivo

Como ADN moderadamente repetitivo se incluyen dos tipos de secuencias, unas de 150 a 300 pb y algunas más largas (~ 5kb) que ocupan el 40% y el 1-2% del genoma respectivamente y localizadas en las zonas eucromáticas. Estas pueden estar organizadas en tándem o dispersas en el genoma (Slamovits & Rossi 2002).

2.2.2.1. ADN repetitivo disperso

Este tipo de ADN lo conforman secuencias que se encuentran dispersas a lo largo del genoma, habiéndose originado esta dispersión mediante transposición (Miller & Capy 2004, Brown 2002). Dependiendo del mecanismo de transposición que actúa sobre las secuencias, éstas se pueden clasificar en elementos de la clase I (retrotransposones) y elementos de la clase II (transposones de ADN). Los transposones de ARN o retrotransposones, son dependientes de retrotranscripción para su actividad y se distinguen dos clases dependiendo de su estructura: elementos LTR y elementos no-LTR, incluyéndose dentro de estos últimos los elementos LINE y SINE. Los transposones de ADN (también denominados elementos transponibles clase II) no necesitan un intermediario de ARN y se transponen directamente. En eucariotas, los transposones de ADN son menos comunes que los retrotransposones.

2.2.2.2. ADN repetitivo en tándem

Consisten en la repetición de unidades desde dos a varios miles. Pueden clasificarse atendiendo a sus características en diversos tipos:

El ADN megasatélite, se distingue por presentar repeticiones en tándem en las que la unidad de repetición aparece aproximadamente de 50 a 400 veces, generando bloques que pueden suponer varios cientos de kb. Algunos

megasatélites están compuestos por regiones codificantes, como por ejemplo los genes del ADN ribosómico.

Los minisatélites o VNTRs (*Variable Number of Tandem Repeats*) son copias en tándem de repeticiones de 6 a 100 nucleótidos de longitud (Tautz 1993). Las repeticiones teloméricas y subteloméricas son consideradas minisatélites. Los microsatélites o STRs (*Short Tandem Repeats*, Edwards et al. 1991) presentan una unidad repetitiva que va de 2 a 7 pb (Bruford & Wayne 1993) y se encuentran dispersos tanto en regiones no codificantes como codificantes de los genomas eucariotas y procariotas (Gur-Arie et al. 2000; Toth et al. 2000).

2.3. Características de los principales tipos de ADN repetitivo: ADN satélite y elementos transponibles

2.3.1. ADN satélite

Está formado por unidades de tamaño variable repetidas en tándem. El tamaño del monómero puede variar desde 30 pb hasta 4.746 pb, como el que constituye uno de los satélites humanos (Saitoh et al. 2000). En general, suelen presentar una longitud media comprendida entre 100 y 500 pb.

El ADNsat es un constituyente principal de la heterocromatina centromérica (Ugarkovic & Plohl 2002, Galián & Vogler 2003), telomérica (Pons et al. 1993, Hartley & Davidson 1994) o incluso de las regiones intersticiales de los cromosomas (Reed & Phillips 1995). La extensión del genoma es variable y puede llegar a ocupar hasta más de 100 Mb (Miklos & Gill, 1982), pudiendo coexistir varias familias de ADNsat en un mismo genoma (Mravinac & Plohl 2007).

El ADNsat puede estar organizado en una estructura simple formada por un único tipo de repetición o por HORs (*Higher-Order Repeats*, repeticiones de orden mayor), esto es, la unidad de repetición estaría constituida por distintas subunidades. Se da este caso en el ADNsat de *Chrysolina carnifex* que muestra un alto grado de complejidad, presentando tres tipos de unidades de repetición,

constituidas por seis unidades diferentes, probablemente con un origen común (Palomeque et al. 2005).

El ADNsat generalmente es rico en AT y suelen encontrarse agrupaciones de residuos de A-T ≥ 3 , lo que se ha relacionado con la curvatura y esta, con la condensación de la cromatina (Benfante et al. 1990). Aunque existan excepciones de ADNst ricos en GC, como el de *Drosophila hydei* (Burgtorf & Bünemann 1993). También hay que considerar que la secuencia de determinados ADN satélites, especialmente de insectos, contienen fragmentos palindrómicos que pueden dar lugar a estructuras cruciformes y en horquilla. Estas estructuras, especialmente en los casos en los que no existe curvatura, pueden tener un papel importante como señal para el posicionamiento de las histonas (Barceló et al., 1998). Han sido detectadas en numerosos organismos, por ejemplo, en coleópteros del género *Tribolium* (Mravinac et al. 2005).

2.3.2. Elementos transponibles

Los elementos transponibles son una fuente de cambio para los genomas (Eickbush 1992, Charleswotr et al. 1994). Son secuencias de ADN capaces de moverse de un sitio a otro dentro del genoma mediante la transposición. Pueden llegar a constituir una elevada proporción del genoma, por ejemplo, en el genoma humano, hasta el 45% ha derivado de este tipo de elementos (Venter et al. 2001).

La transposición se puede llevar a cabo de dos modos diferentes. En la *transposición replicativa* se genera una copia del transposón que se inserta en el ADN receptor, por tanto, el resultado es la generación de una copia. En la *transposición no replicativa* el transposón es cortado del ADN donante y transpuesto al ADN receptor. Así, en este caso, una única copia ha cambiado su posición en el genoma.

Los elementos transponibles pueden clasificarse en dos tipos principales dependiendo de su modo de transposición en elementos de clase I o de clase II (Finnegan 1992).

2.3.2.1. Elementos de la Clase I

Su mecanismo de transposición consiste en su transcripción a ARN que, gracias a la reversotranscriptasa que el propio transposón codifica, pasa a ADN, que puede ser insertado en el ADN receptor gracias a una transposasa (que también codifica el mismo transposón). Por ello son conocidos también como transposones de ARN o retrotransposones.

Los retrotransposones LTR (*Long Terminals Repeats*, repeticiones largas terminales) disponen, como su nombre indica, de repeticiones largas terminales que son necesarias para la transposición. Son capaces de codificar una poliproteína con actividad reversotranscriptasa, integrasa, proteasa y RNasa H gracias al gen *pol*. También es característico de este tipo el gen *gag* (codifica una proteína de la cápsida). Pueden clasificarse en dos familias: *Ty1/copia* y *Ty2/gypsy* diferentes en sus ORFs (*Open Reading Frame*, fase de lectura abierta). Dentro de este tipo se incluyen retrovirus de genoma ARN y retrovirus ERV, estos últimos, capaces de integrarse en el cromosoma para heredarse en el genoma hospedador. Aunque algunos ERV aún son activos, la mayoría son vestigiales y no tienen capacidad de formar nuevos virus (Patience et al. 1997).

Por otra parte, en los retrotransposones sin LTR, en su lugar presentan secuencias repetidas simples. Se distinguen en este grupo los elementos nucleares interdispersos largos (LINEs: *Long Interspersed Nuclear Elements*) y elementos nucleares interdispersos cortos (SINEs: *Short Interspersed Nuclear Elements*). Dentro de los LINE destaca L-1 de mamíferos que en el caso del genoma humano constituye el 16,9% del genoma (Lander et al. 2001). Entre los SINE, las secuencias Alu pueden constituir más del 10% del genoma humano (Lander et al. 2001).

2.3.2.2. Elementos de la Clase II

Están presentes en baja proporción respecto a otros elementos transponibles, por ejemplo, en humanos, solo se debe a la presencia de estos elementos el 3% del genoma y sin evidencia de actividad en los últimos 50 millones de años. Se diferencian de los elementos de la clase I en no necesitar ARN intermediario para su

transposición. Están constituidos por un gen que codifica una transposasa flanqueado por dos ITRs (*Inverted Terminal Repeats*, repeticiones terminales invertidas) que reconoce la transposasa para efectuar el corte y escisión.

Pueden agruparse en familias y subfamilias en base a la similitud de las secuencias, ITRs y del TSD (*target site duplication*, duplicación del sitio diana) generado en el proceso de inserción. Por ejemplo, *hAT* y *Tc1/mariner* son familias de transposones bastante comunes en eucariotas (Plasterk et al. 2002).

2.4. ADN repetitivo en Tenebriónidos

Los Tenebriónidos (Coleptera, Tenebrionidae) son un grupo de insectos que han sido usados como modelo para el estudio del ADNsat, aunque hasta ahora solo se han estudiado especies pertenecientes a seis géneros diferentes y de estos, los estudios se han centrado sobre todo en cuatro: *Palorus*, *Pimelia*, *Tenebrio* y *Tribolium*. Los estudios llevados a cabo han permitido establecer modelos de evolución del ADNsat y han ayudado a reconstruir el ciclo de vida de este tipo especial de ADN (Mestrovic et al. 1998, Mravinac et al. 2002, Mravinac & Plohl 2007), han permitido caracterizar estructuras que pueden estar bajo presión selectiva (Ugarkovik et al. 1996, Plohl et al. 1990, 2004, Mravinac et al. 2004) y entender los patrones de organización en el genoma (Mravinac & Plohl 2010, Plohl et al. 1992).

La utilización de la familia Tenebrionidae viene dada posiblemente por la alta proporción de este tipo de repetitivo en el genoma de algunas especies pertenecientes a ella. Por ejemplo, el 50% del genoma de *Tenebrio molitor* (Petitpierre et al. 1988) o el 43.7% en *Pimelia criba* (Pons et al. 1997) está constituido por ADNsat. En el genoma de los tenebriónidos es característica la presencia de bloques de heterocromatina pericentromérica, en los que aparece principalmente uno o dos tipos de ADN satélite extremadamente abundantes distribuidos en todos los cromosomas del complemento (revisado en Palomeque & Lorite 2008). Esto ocurre, por ejemplo, en *Palorus subdepressus*, donde todos los cromosomas del complemento presentan la misma familia de ADNst (Plohl et al. 1998); o en especies como *Misolampus goudoti* y *Tenebrio obscurus* en las que su

genoma contiene diferentes familias (Pons et al. 1993, Plohl & Ugarkovic 1994). Respecto a la distribución en todos los cromosomas, existen excepciones en las que el ADNsat es específico de cromosoma o no se encuentra en todos los del complemento. En *Pimelia sparsa*, por ejemplo, el satélite se localiza en todos los cromosomas excepto en X e Y (Barceló et al. 1997,1998).

El tamaño de las unidades repetitivas de los ADNsat en tenebriónidos oscila generalmente entre 100 y 700 pb, pero hay que tener en cuenta diversas excepciones, de las que pueden servir de ejemplo, la familia *PstI* de 1169 pb descrita en *Misolampus goudoti* (Pons et al. 1993, Juan et al. 1993, Pons 2004) y la familia TBREV de 1061 pb de *Tribolium brevicornis*, si bien este último parece haberse formado a partir de subunidades menores que están orientadas inversamente en el nuevo monómero (Mravinac et al. 2005).

Como se ha expuesto anteriormente, el estudio del ADNsat de la familia Tenebrionidae ha permitido establecer modelos de evolución de este tipo de ADN repetitivo. Cabe destacar que una de las hipótesis sobre la dinámica evolutiva del ADN satélite fuera demostrada experimentalmente gracias al estudio de tenebriónidos del género *Palorus*. La “hipótesis de la biblioteca”, descrita originalmente por Fry & Salser (1977), indica que hay diferentes familias de ADNsat en una especie antecesora común a diferentes especies. Posteriormente, al divergir las especies, cada familia de ADNsat se amplifica diferencialmente en cada una de ellas. Se constituye así en cada especie, uno o varios satélites mayoritarios que coexisten con restos de otras familias pudiendo ser estas mayoritarias en las otras especies provenientes del mismo antecesor. El genoma de cuatro especies de *Palorus* está compuesto en un 30% aproximadamente por una familia de ADNsat diferente en cada uno de ellos, pero también contienen en mucha menor proporción satélites de las otras especies (Mestrovic et al. 1998). También se ha comprobado en otros géneros de tenebriónidos como *Pimelia*. En este género se han descrito siete familias de ADNsat, las cuales parecen haberse originado a partir de una secuencia ancestral común mediante complejos procesos de mutación y recombinación (Pons et al. 2004). Estos ADNs satélite mayoritarios aparecen en una o varias especies del género. Bruvo-Madaric et al. (2007) demostraron que la

mayor parte de las familias están presentes en la mayor parte de las especies del género, si bien en pequeñas proporciones.

El contenido en AT en el ADNsat es especialmente alto, en torno al 60-70% pudiendo incluso llegar a superar el 79% como ocurre en la familia TDEST 140 pb de *Tribolium destructor* (Mravinac et al. 2004). Este hecho adquiere importancia debido a que estos nucleótidos tienden a formar agrupaciones periódicas de AT mayores o iguales a 3, lo cual se ha relacionado con la curvatura del ADN y ésta, a su vez, con la condensación de la cromatina y con la unión de proteínas específicas (Ugarkovic et al. 1992, Juan et al. 1993). En numerosas especies de *Pimelia* se ha demostrado la importancia de las agrupaciones de adenina para la curvatura del ADN (Barceló et al. 1997, Pons et al. 2004).

Además de la curvatura del ADN, la composición del ADNsat puede influir en la organización de estructuras secundarias que también pueden estar implicadas en la condensación de la cromatina o tener funciones relacionadas con el centrómero (Bigot et al. 1990, Jonstrup et al. 2008). Estas estructuras pueden tener su origen en la presencia de repeticiones invertidas de las que cabe destacar las encontradas en el género *Tribolium*. En este género, las unidades repetitivas suelen estar compuestas por subunidades de orientación inversa con capacidad de formar uniones intracatenarias que dan lugar a estructuras cruciformes (Mravinac & Plohl 2007, 2010, Mravinac et al. 2005, Ugarkovic et al. 1996).

Como se ha indicado los estudios sobre ADNsat en tenebriónidos ha sido muy intenso, posiblemente debido a que constituyen un buen modelo debido a la gran cantidad de ADN repetitivo que presentan. Sin embargo, los estudios sobre elementos transponibles ha sido muy escaso y fragmentado, todos ellos realizados sobre especies de los géneros *Tenebrio* y *Tribolium*.

Beeman et al. (1996) encontraron que una mutación del gen *Abdominal* en la especie *Tribolium castaneum* estaba asociada con la presencia de una inserción de 12 kb. La caracterización de esta secuencia permitió determinar que se trataba de un retrotransposón, denominado *Woot*, que pertenece a la familia *Ty3/Gypsy* dentro del grupo de los LTR retrotransposones. El número de copias de ese elemento en el

genoma de la especie era de 25-35, presentaba unas LTR de 3.6 kb rodeando una región codificante de 5 kb que incluía dos ORFs correspondientes a los genes *gag* y *pol*, presentes en otros retrotransposones y retrovirus. El gen *gag* codifica una poliproteína que da lugar a proteínas de la nucleocápside. De la poliproteína originada por el gen *pol* se generan una proteasa, una reversotranscriptasa, RNasa H y una integrasa. Mediante PCR se demostró que este elemento también estaba presente en otras especies del género, en concreto en *T. freemani* y *T. brevicornis*.

Posteriormente un segundo tipo de retrotransposón fue aislado también en la especie *Tribolium castaneum*, llamado *TcDirs1* y que pertenece al grupo de elementos DIRS (Goodwin et al. 2004). Los elementos DIRS son retrotransposones que parecen estar relacionados con los retrotransposones LTR pero difieren de ellos en que no codifican una integrasa y en su lugar tienen una tirosina recombinasa, encargada de insertar el intermediario de ADN en el genoma huésped (Duncan et al. 2002).

La información sobre transposones de la clase II o DNA transposones en tenebriónidos es aún más escasa. De hecho no se ha analizado ningún elemento completo de este grupo. La excepción es el estudio de un elemento MITE (*Miniature Inverted-repeat Transposable Element*) en la especie *Tenebrio molitor*. Los elementos MITE se han originado a partir de elementos transponibles de la clase II por pérdida de la región codificante de la transposasa, perdiendo su capacidad de transposición que dependería de una transposasa producida por el elemento autónomo del que es originario. El elemento MITE aislado en *T. molitor*, denominado DEC, presenta las características típicas de estos elementos, como es su pequeño tamaño, riqueza en AT, presencia de repeticiones terminales invertidas y la ausencia de ORF (Braquart et al. 1999).

Pons (2004) ha sugerido que un transposón de la clase II en el origen y distribución de la familia *PstI* de ADNsat en el tenebriónido *Misolampus goudoti*, basándose en la existencia de repeticiones invertidas y motivos con similitud con este tipo de transposones. La intervención de la transposición explicaría la presencia de este ADNsat en todas las regiones heterocromáticas, incluidas las correspondientes al cromosoma Y, a diferencia de lo que ocurre en otros satélites

aislados en esta misma especie, o en otras especies de coleópteros de la misma familia, donde la distribución de un determinado tipo de ADN no es tan amplia, restringiéndose a la región pericentromérica de los autosomas.

Sin embargo, lo que mayor cantidad de datos ha aportado sobre la presencia de elementos transponibles en tenebriónidos ha sido la secuenciación del genoma de la especie *Tribolium castaneum* (Richards et al. 2008) y los trabajos que a partir de este secuenciación se han publicado. Esta especie presenta varias familias tanto de transposones de clase I como de clase II, incluyendo tanto LTR como no-LTR retrotransposones. Estas secuencias constituyen aproximadamente el 6% del genoma (Wang et al. 2008a) (**Tabla 1**). Tal como ocurre en otros organismos los transposones de ADN de la superfamilia *Tc1-mariner* son frecuentes, habiéndose identificado 30 subfamilias. La mayor parte de estos elementos transponibles son inactivos. Estudios posteriores han permitido determinar que algunos transposones de clase II son activos, tal como ocurre con *TcBuster* de la familia *hAT* (Arensburger et al. 2011) o *TcPLE* del grupo *piggyBac* (Wang et al. 2008b). Del grupo de retrotransposones, es activo el elemento *Woot*, tal como se había determinado con anterioridad (Beeman et al. 1996).

Class	TE library* (kb)	Number of families	Percentage of genome	TE length range (bp)	Average length (bp)	Copy number (range)	Average copy number	GC content range (%)	Average GC content (%)
Non-LTR	238.1	69	2.0	786-6,820	3,363	1-2,556	161	27.15-57.94	38.14
LTR	290.2	48	1.7	3,292-11,097	6,019	1-1,634	202	30.61-53.21	39.31
DNA transposons	78.6	45	2.2	456-4,878	1,746	1-8,949	420	30.90-46.08	37.22

Tabla 1. Resumen de los transposones de ADN, LTR y no-LTR retrotransposones encontrados en el genoma de *Tribolium castaneum*. Tomada de Wang et al. (2008a).

3. HIPÓTESIS Y OBJETIVOS

Los Tenebriónidos han sido modelo de estudio de ADN repetitivo, especialmente del ADN satélite. Posiblemente la causa de esto haya sido el elevado porcentaje que este tipo de ADN representa en el genoma de algunas especies, en las que puede alcanzar el 50% del ADN genómico total, como ocurre con el escarabajo de la harina *Tenebrio molitor* (Petitpierre et al. 1988). Otro hecho importante que se ha determinado en tenebriónidos es que a pesar de que cada especie presenta una familia de ADNsat mayoritario, presenta restos de otras familias, que en otras especies constituyen su ADNsat mayoritario. Todas ellas derivarían de una familia ancestral (Pons et al. 2004, Bruvo-Madaric et al. 2007).

Esta familia es bastante extensa con unos 1700 géneros y al menos 18000 especies. En Europa esta familia está representada por 1478 especies y subespecies presentes en 63 países (Fattorini & Ulrich 2012). A pesar del elevado número de especies, los estudios sobre ADNsat, así como el de otros tipos de ADNs repetitivos se han llevado a cabo especies pertenecientes a seis géneros, incluidos en tres de las ocho subfamilias en las que se divide la familia Tenebrionidae; *Tenebrioninae*, *Palorinae* y *Pimeliinae*.

Los estudios llevados hasta el momento muestran que en líneas generales, al menos para el ADNsat el comportamiento es similar en las subfamilias, con grandes proporciones de ADNsat, con un elevado contenido en adenina-timina, presencia de numerosas repeticiones invertidas y trectos de adenina o timina iguales o mayores de tres residuos (Plohl 2010). Sobre otros tipos de ADN repetido los datos en la familia son escasos, con la excepción de la especie *Tribolium castaneum*, cuyo genoma ha sido secuenciado y analizado (Richards et al. 2008, Wang et al. 2008a).

Nada se sabe sobre el ADN repetitivo en ninguna de las otras cinco subfamilias de Tenebrionidae ni si presentan características similares a los presentes en las subfamilias analizadas. Por este motivo se propuso analizar especies pertenecientes a otras subfamilias, lo que mejoraría la visión global de la evolución de los ADNs repetitivos dentro de la familia Tenebrionidae. Esta nueva

línea de investigación se inicia en este trabajo, donde se comienza el estudio en la especie *Heliotaurus ruficollis*, especie perteneciente a la subfamilia *Alleculinae*.

La caracterización del ADN repetitivo en una especie es un estudio complejo, especialmente en especies, como los tenebriónidos, donde pueden existir múltiples familias de este tipo de ADN. La naturaleza repetitiva de estos ADNs hace que, ni aún con la secuenciación del genoma, se puede tener la certeza de haber caracterizado todos estos ADNs (Wang et al. 2008a). Por tanto, y tal como se ha comentado, el objetivo de presente trabajo es iniciar el estudio del ADN repetitivo en *Heliotaurus ruficollis* mediante la caracterización de alguna de las familias de ADN repetitivo presentes en su genoma.

4. MATERIAL Y MÉTODOS

4.1. Material

En este trabajo se han utilizado individuos de *Heliotaurus ruficollis* (Coleoptera, Tenebrionidae) (Figura 2) que han sido recolectados en el término municipal de Los Villares (Jaén). Principalmente en vegetación silvestre de lindes de parcelas de olivar en plantas del género *Cistus* y sobre flores de diversas compuestas.



Figura 2. Ejemplar de *Heliotaurus ruficollis*.

4.2. Métodos

4.2.1. Extracción de ADN genómico

El ADN genómico se ha extraído a partir de machos y hembras adultos mediante la técnica de Heinze et al. (1994) con algunas modificaciones. En cada extracción se utilizaron entre 5 y 7 individuos. El material es macerado para su homogeneización usando nitrógeno líquido. Al producto resultante se le añade 600 µl de tampón de extracción (Tris-ClH pH 7.5mM, ClNa 60 mM, EDTA 100 mM), 25 µl de una solución de proteinasa K (10 mg/ml), y 900 µl de solución de lisis (Tris-ClH 0.2 mM pH 9, EDTA 30 mM, SDS 2%). La mezcla se incuba a 56°C durante 6 horas. Posteriormente se añaden 5 µl de ARNasa (10 mg/ml) y se incuba a 37°C durante 3 horas. A continuación se realiza una fenolización: al homogeneizado se le añade un volumen de fenol, se mezcla por agitación manual durante algunos minutos, y se centrifuga durante 10 minutos a 2.500 g. La fase acuosa se pasa a un tubo nuevo y

se añade fenol:cloroformo:alcohol isoamílico (24:24:1) en proporción 1:1. Se mezcla por agitación manual, se vuelve a centrifugar durante 10 minutos a 2.500 g. A la fase acuosa se le añade un volumen de cloroformo:alcohol isoamílico (24:1), se vuelve a mezclar y centrifugar con las mismas condiciones. La fase acuosa se precipita añadiendo 1/20 del volumen de NaCl 5M y un volumen de isopropanol. Se mantiene al menos 30 minutos a -20°C para centrifugar después durante 5 minutos a 2.500 g. Se lava el precipitado con etanol al 70%. Tras dejar secar, se resuspende el precipitado en agua ultrapura estéril.

La cuantificación y la determinación de la pureza del ADN extraído se realizan por espectrofotometría, midiendo la absorbancia a 260 y 280nm. Considerando que una absorbancia igual a 1 midiendo a 260 nm corresponde con una concentración de ADN de 50 ng/μl. El cociente entre las absorbancias a 260 nm y 280 nm (A_{260}/A_{280}) nos da un valor de pureza, ya que indica si existe contaminación con proteínas; se considera que una disolución de ADN está limpia de proteínas, cuando este cociente tiene un valor entre 1,8 y 2 (Wilfinger et al. 1997).

4.2.2. Aislamiento del ADN repetitivos mediante digestión del ADN genómico con enzimas de restricción

Que el ADNsat tenga una organización molecular de monómeros repetidos en tándem permite que si una enzima de restricción corta en un único punto del monómero, se generarán diferentes fragmentos que derivan de la unidad de repetición (monómeros, dímeros, trímeros, etc). Al realizar una electroforesis en gel de agarosa, el alto número de fragmentos generados de igual tamaño hace que pueda observarse como una escalera formada por estos fragmentos (Figura 3A).

Otros ADNs repetitivos no organizados en tándem también pueden aislarse mediante digestión con enzimas de restricción. Si la secuencia repetida es cortada por un mismo enzima de restricción en al menos dos ocasiones se generan fragmentos de igual tamaño que podrán visualizarse en geles de agarosa siempre y cuando el número de fragmentos generados sea suficientemente alto (Figura 3B). En este caso no aparecen escaleras.

La digestión del ADN genómico de la especie en estudio se realizó usando una batería de enzimas de restricción. Dichas digestiones se llevaron a cabo con 4 μg de ADN en el tampón adecuado y siguiendo las instrucciones del proveedor, utilizándose 4 U de enzima por μg de ADN en un volumen final de 50 μl . Los ADNs digeridos fueron separados en geles de agarosa al 2%. Los geles fueron teñidos con bromuro de etidio y observados en un transiluminador. Las imágenes fueron recogidas gracias a un fotodocumentador de geles.

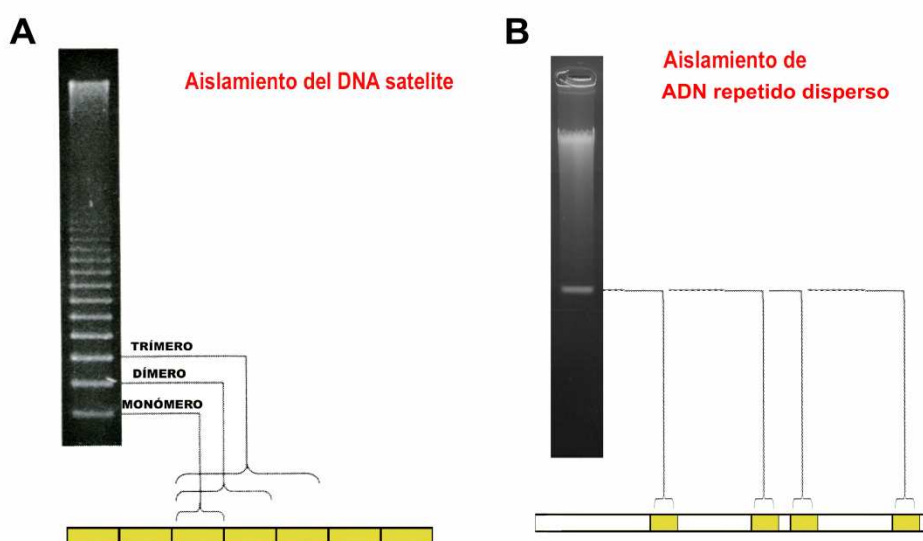


Figura 3. (A) Resultado de la digestión de un ADN repetitivo organizado en tándem con un enzima de restricción que corta la unidad de repetición en un solo sitio. Como consecuencia de esta organización se generan bandas que forman una escalera que difieren en el tamaño de una unidad de repetición. (B) Resultado de la digestión de un ADN repetitivo dispersos con un enzima de restricción que corta dos veces en la secuencia repetida. Como consecuencia de esta organización se genera una sola banda correspondiente a las copias de igual tamaño del fragmento generado por la digestión.

4.2.3. Purificación de ADN a partir de agarosa

Las bandas cortadas de los geles de agarosa se purificaron con el kit comercial *NucleoSpin Extract* de Macherey-Nagel, siguiendo las instrucciones del proveedor.

4.2.4. Preparación del vector y reacciones de ligado de las bandas de ADN genómico

El fragmento de ADN genómico de *H. ruficollis* obtenido tras la digestión con *A*/ul fue extraído de las bandas se clonó en el vector pUC19. Para realizar las clonaciones, se digiere el plásmido pUC19 con la enzima *Sma*I en su tampón correspondiente. Esta enzima genera extremos romos al igual que el enzima *A*/ul. Tras tres horas de incubación a 37°C, se realiza la electroforesis del ADN digerido en un gel de agarosa al 1%. El plásmido digerido se extrae de la agarosa y se desfosforila con 1 µl de fosfatasa alcalina (10 U/µl) (Roche) a 37°C durante varias horas. Posteriormente, el plásmido se migra en un gel de agarosa al 1% para eliminar la fosfatasa alcalina, y se purifica de la agarosa. Las reacciones de ligado se llevaron a cabo en un volumen de 15 µl, compuesto por 1,5 µl del tampón de la ligasa (10x), 1 µl de ligasa (3 U/µl) de Roche, 12 µl del ADN de la banda purificada que pretendemos clonar y 0,5 µl del vector pUC19.

4.2.5. Transformación

Tanto para los clones generados con pUC19 como para los generados con pGEMT, las transformaciones se llevaron a cabo directamente sobre bacterias *E. coli* DH5α competentes comerciales (Zymo). Para la transformación se siguieron las instrucciones del proveedor. 50 µl de bacterias previamente conservadas a -80°C se descongelan en hielo, se añade 4.5 µl de ligado y se incuba 5 minutos en hielo. La siembra se realizó sobre placas de cultivo LB (Luria-Bertani) con ampicilina (100 mg/ml), X-gal (2 mg/ml) e IPTG (4 mM). Las bacterias con inserto, y por lo tanto de color blanco, son sembradas en una nueva placa con ampicilina para tener suficiente material y con este llevar a cabo la selección de plásmidos recombinantes.

4.2.6. Extracción de ADN plasmídico

Con ayuda de un asa de siembra estéril se coge parte del cultivo en placa de cada clon y se hacen cultivos de 15 ml en medio LB líquido, al que se le ha añadido ampicilina. Este cultivo se deja crecer durante toda la noche y al día siguiente se

procede a la extracción del ADN plasmídico. Para ello, el cultivo se centrifuga a 5000 rpm durante 20 minutos y se desecha el sobrenadante. El pellet se resuspende en 100 µl de tampón de lisis I (glucosa 50 mM, EDTA pH=7.5-8 10 mM, Tris-ClH pH=8 25 mM). Después, se añaden 200 µl de tampón de lisis II (NaOH 0.2 N, SDS 1%), se agita el tubo y se añaden 150 µl de tampón de lisis III (acetato potásico 3 M, ácido fórmico 1.8 M). La mezcla resultante se deja durante 30 minutos en hielo y posteriormente se centrifuga a 14.000 rpm durante 15 minutos. Se recupera el sobrenadante y se vierte en un tubo nuevo. Los plásmidos se precipitan añadiendo un volumen de isopropanol frío al 100%. Después de dejarlos al menos 30 minutos a -20°C, se centrifuga a 14.000 rpm durante 15 minutos y se elimina el sobrenadante. El precipitado se lava con etanol al 70% y se resuspende en 20 µl de agua ultrapura estéril.

4.2.7. Marcaje de sondas con digoxigenina mediante "random priming"

Este marcaje se inicia con la desnaturalización del ADN molde. 1 µg de dicho ADN se calienta en un volumen de 15 µl de agua durante 10 minutos a 100° C. A continuación, se pasa a hielo, y se le añade 2 µl de la mezcla de nucleótidos con digoxigenina (dGTP 1mM, dATP 1mM, dCTP 1mM, dUTP 0.35 mM, DIG-dUTP 0.65 mM), 2 µl de la mezcla de hexanucleótidos y 1 µl de Klenow (10 U/µl) (Kit de Roche). Esta mezcla se incuba a 37°C durante 16 horas. Se para la reacción añadiendo 2 µl de EDTA 0.2 M pH 8. El ADN marcado se precipita y se resuspende en agua ultrapura estéril.

4.2.8. Comprobación del grado de marcaje de las sondas marcadas con digoxigenina

Para comprobar el grado de marcaje de cada una de las sondas utilizadas se realizan diluciones seriadas de la mismas (1/100, 1/200, 1/400, 1/800 ...). Con un ADN control marcado (Roche) se realizan las mismas diluciones. Estas diluciones se depositan en una membrana Hybond-N⁺ (Amersham), usando un aparato de dot-blot. Tras dejar secar se fija durante 2 minutos con luz ultravioleta y se guarda hasta su uso. Para el revelado de la digoxigenina la membrana se bloquea durante 30 minutos en tampón-2 en agitación y se incuba durante otros 30 minutos en tampón-

2 con antidigoxigenina conjugada con fosfatasa alcalina (1:5000). Posteriormente se realizan dos lavados en tampón-1 de 15 minutos cada uno y un lavado de 5 minutos en tampón-3. Por último, la membrana se introduce en una pequeña bolsa de plástico que se sella una vez añadida la solución de color (20 µl de NBT/BCIP en 1,5 ml de tampón-3). Tras unos minutos comienza a aparecer señal, y podemos establecer la concentración aproximada de ADN marcado por comparación entre la intensidad de color de las diluciones del ADN sonda y del control (Figura 4).

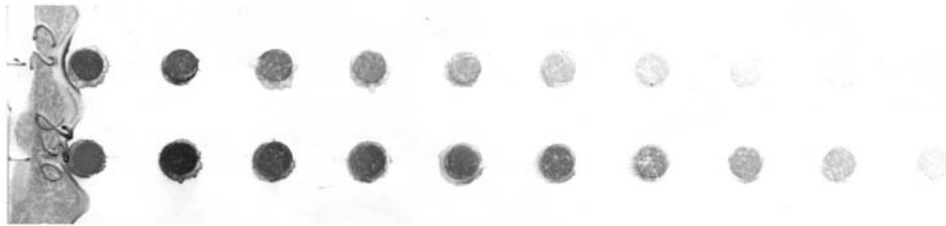


Figura 4. Dot-blot usado para comprobar el grado de marcaje de las sondas con digoxigenina. En la fila de arriba se han depositado las diluciones del ADN control y en la fila de abajo las diluciones de la sonda preparada. Tras el revelado de la digoxigenina se puede estimar la concentración de la sonda comparando la señal de ésta con la del ADN control.

4.2.9. Selección de plásmidos recombinantes mediante dot-blot

Para la búsqueda de clones positivos, se depositan aproximadamente 100 ng de cada plásmido sobre una membrana de nylon Hybond-N⁺ (Amersham). Los plásmidos han sido previamente desnaturalizados durante 10 en agua hirviendo. El ADN se fija a la membrana durante dos minutos con luz ultravioleta.

Para la selección de los plásmidos recombinantes se utiliza como sonda una muestra de la banda de ADN extraída de la agarosa y usada para la clonación. Este ADN se marca con DIG-11-dUTP con el Kit de marcaje por "random priming" de Roche. El ADN se desnaturaliza durante 10 minutos en agua hirviendo y se enfría rápidamente en hielo.

La membrana de nylon con el ADN fijado se incuba durante dos horas en un horno de hibridación a la temperatura deseada (debe de ser la misma que la temperatura de hibridación), en este caso 55°C, con la solución de prehibridación

(solución de hibridación sin el ADN sonda). Se elimina la solución de prehibridación de la membrana y se añade la que contiene el ADN sonda desnaturalizado (a una concentración aproximada de 10-25 ng/ml), para incubar a la temperatura de hibridación (55°C) durante toda la noche. A la mañana siguiente se retira la sonda y se realizan lavados de posthibridación para eliminar la sonda que se haya unido inespecíficamente, en este caso, con 2xSSC, 0,1% SDS a temperatura ambiente durante 30 minutos y 1xSSC, 0,1% SDS a 55°C durante 30 minutos.

Posteriormente se realiza un lavado de 5 minutos en agitación con tampón-1 y a temperatura ambiente, para después bloquear la membrana durante una hora en tampón-2. A continuación se incuba durante otra hora con anti-digoxigenina conjugada con fosfatasa alcalina en tampón-2 (1:5000). Se realizan dos lavados con tampón-1 durante 15 minutos para eliminar los restos de anticuerpo no unidos, y un lavado de 5 minutos en tampón-3. La membrana se introduce en una pequeña bolsa de plástico que se sella una vez añadida la solución de color (20 µl de NBT/BCIP en 1,5 ml de tampón-3). Una vez que aparecen las señales de hibridación, se lava la membrana con agua destilada y se deja secar a temperatura ambiente.

4.2.10. Amplificación de ADN por PCR

Se llevaron a cabo dos tipos diferentes de PCR, una sobre ADN genómico para amplificar las secuencias localizadas entre las secuencias HrufAlu de 800 pb y otra para amplificar el inserto de los fragmentos de ADN clonados en los plásmidos pUC19 o pGEM-T.

Para la amplificación de las secuencias localizadas entre las secuencias HrufAlu de 800 pb se diseñaron dos cebadores a partir de las secuencias de 800 pb, localizados en los extremos de estas y dirigidos hacia fuera (HrufA-1 5'-CTCTCACCAGAGATAGGAG y HrufA-2 5'-AAAATATCAAGATCTCCTGT). Las PCRs se llevaron a cabo en un volumen total de 50 µl, con 50 ng de ADN genómico, dNTPs 0,5 mM, 50 pmol de los cebadores y una unidad de Taq polimerasa (Biotaq, Bioline). El programa utilizado fue el siguiente: 92°C 1' 30", 30 ciclos (92°C 20", 50°C 1' y 72°C 2'), 72°C 5', 4°C. Los productos de PCR son migrados en un gel de agarosa al 1% y purificados mediante el kit NucleoSpin Extract II (Machery-Nagel).

Para amplificar los insertos de clonados en los plásmidos pUC19 o pGEM-T se usaron los cebadores universales M13F-pUC (-40) 5'-GTTTTCCCAAGTCACGAC y M13R-pUC (-40) 5'-CAGGAAACAGCTATGAC, usando como ADN molde 20 ng de plásmido. Las condiciones de la PCR fueron las mismas que para la amplificación de las secuencias HruFAlu.

4.2.11. Reacciones de ligado de los productos de PCR

Los productos de PCR fueron clonados en el vector pGEM-T easy vector (Promega). La reacción de ligado se llevo a cabo con 12 µl del producto de PCR purificado de la agarosa, 0,5 µl de vector, 1,5 µl de tampón (10x) para la ligasa T4, de la que se añadió 1 µl llegando así al volumen final de 15 µl. La reacción de ligado se dejó a 4 °C durante 12 h.

4.2.12. Selección de plásmidos recombinantes mediante lisis rápida

Para los clones producidos con el vector pGEM-T, y cuyo inserto es un producto de PCR, la selección se ha realizado mediante lisis rápida. Para ello se recoge con ayuda de un asa de siembra estéril un aspa (de las dibujadas al sembrar las colonias en forma de cruz en las placas con LB) en 50 µl de EDTA (10 mM, pH=8). Se añaden 50 µl de solución tampón de lisis (NaOH 10 M, SDS 10%, sacarosa 40%), se incuba a 70°C durante 5 minutos y se deja enfriar a temperatura ambiente. Posteriormente se añade 1,5 µl de KCl 4M y 0,5 µl de azul de bromofenol. Se agita en vortex durante 30 segundos y se incuba en hielo 5 minutos. Como paso final se centrifuga a 14.000 rpm a 4°C y el sobrenadante se carga en un gel de agarosa al 1% para realizar una electroforesis y observar los resultados. También se carga en el gel una muestra del vector sin inserto para comparar el tamaño de los plásmidos sin inserto con aquellos que han incorporado el fragmento que se desea clonar.

4.2.13. Conservación de los clones

Para la conservación los clones que se han obtenido, se realiza un cultivo de cada clon en 10 ml de medio Luria-Bertani líquido con ampicilina hasta que se llega

a la fase exponencial. Una vez crecido el cultivo se centrifuga durante 10 minutos a 2000 g, se elimina el sobrenadante y se lava el precipitado con medio LB sin antibiótico. Las bacterias se resuspenden con ayuda de un agitador y se añaden 500 µl de glicerol:medio LB (1:1) estéril. Se mezclan homogéneamente con un agitador y se guardan a -80°C.

4.2.14. Preparación de ADN plasmídico para secuenciación

Los plásmidos para secuenciación se obtuvieron a partir de 5 ml de cultivo líquido. El ADN plasmídico se extrajo mediante con el kit NucleoSpin Plasmid (Machery-Nagel) siguiendo las instrucciones del proveedor y fueron enviados al servicio de secuenciación del CNIO (Centro Nacional de Investigaciones Oncológicas).

4.2.15. Southern-blot del ADN genómico digerido con enzimas de restricción

El ADN genómico se digiere con enzimas de restricción, y los fragmentos resultantes son separados en un gel al 2% y transferidos a una membrana Hybond-N⁺ (Amersham). El Southern-blot se realizan siguiendo la técnica descrita por Sambrook et al. (1989).

El gel de agarosa se depurina con HCl 0,2 N durante 10 minutos, tras lo cual se lava en agua desionizada durante otros 10 minutos. Después, se realiza una desnaturalización mediante una solución de NaCl 1,5 M, NaOH 1,5 M durante 45 minutos, y por último se restablece el pH con una solución de Tris-HCl 1 M, NaCl 1,5 M, pH=7,4 durante 30 minutos. Esta última solución se elimina y se añade una nueva durante otros 15 minutos. La transferencia del ADN se realiza por capilaridad a una membrana de nylon Hybond-N⁺ (Amersham) durante toda la noche con una solución de 20xSSC. Posteriormente se deja secar la membrana y se fija el ADN con luz ultravioleta durante 2 minutos.

Los insertos de los clones, previamente amplificados por PCR, fueron utilizados como sondas, para lo cual se marcaron con DIG-11-dUTP con el Kit de marcaje por "random priming" de Roche. La hibridación y el revelado de la

membrana se realizan tal como se ha explicado anteriormente para la “Selección de plásmidos recombinantes mediante dot-blot”.

4.2.16. Análisis de las secuencias

Las secuencias obtenidas se alinearon utilizando el programa CLUSTALW, y fueron analizadas y comparadas con las bases de datos GenBank/NCBI usando el programa BLAST y con la base de datos EMBL usando el programa FASTA (Altschul et al. 1997, Pearson & Lipman 1998). Las variaciones de las secuencias se analizaron utilizando el programa DnaSP (Rozas & Rozas 1999). Para determinar la naturaleza de las secuencias se usó el CDD (Conserved Domain Database) del NCBI (Marchler-Bauer et al. 2013).

La composición del ADN, la búsqueda de repeticiones directas e invertidas, u ORFs y otros análisis fueron llevados a cabo mediante programas disponibles en el portal Mobyly del Instituto Pasteur (<http://mobyle.pasteur.fr/cgi-bin/portal.py?%23welcome%20%20#welcome>). El programa FoldDNA (<http://mfold.rna.albany.edu/?q=mfold/DNA-Folding-Form>, Zuker 2003) fue utilizado para la predicción de estructuras secundarias.

4.2.17. Composición y preparación de los reactivos utilizados

Tampón-TAE (50x)

Trizma base	242 g
Ácido acético glacial	57,1 ml
EDTA 0,5 M (pH=8)	100 ml
Agua destilada	hasta un litro

Medio de cultivo Luria-Bertani

Líquido:

Extracto de levadura	5 g/l
Triptona	10 g/l
NaCl	10 g/l
NaOH 1M	160 µl/l

Sólido:

Extracto de levadura	5 g/l
Triptona	10 g/l
NaCl	10 g/l
NaOH 1M	160 µl/l
Agar-agar	15 g/l

Solución stock de ampicilina (1000x)

Ampicilina	100 mg/ml
Agua destilada	c.s.

Tampones para la detección de la digoxigenina

Solución de prehibridación

5xSSC	
N-laurosilsarcosina	0,1%
SDS	0,02%
Agente de bloqueo	2%

Tampón-1 (pH=7,5).

Ácido maleico	0,1 M
NaCl	0,15 M

Tampón-2.

Agente de bloqueo (Roche)	2%
Tampón-1	c.s.

Tampón-3 (pH=9,5).

Tris-HCl	100 mM
NaCl	100 mM
MgCl ₂	50 mM

Tampón-4 (pH=8).

Tris-HCl	10 mM
EDTA	1 Mm

5. RESULTADOS Y DISCUSIÓN

El ADN genómico de *Heliotaurus ruficollis* fue digerido con una batería de enzimas de restricción. Los fragmentos de ADN resultantes se separaron mediante electroforesis en geles de agarosa al 2%. La mayor parte de las enzimas utilizadas genera la aparición de una o varias bandas (**Figura 5**). Estas bandas se originan por la existencia de un ADN repetitivo que es cortado por dicha enzima de restricción. El tamaño de estos fragmentos es variable dependiendo del enzima utilizado y oscila entre 400 pb y 4 Kb, como ocurre en las digestiones con *Bgl*II.

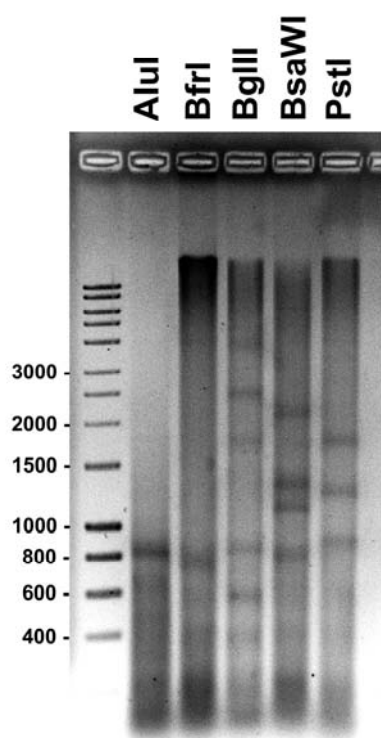


Figura 5. Electroforesis en gel de agarosa al 2 % de la restricción con varios enzimas de restricción del ADN genómico de *Heliotaurus ruficollis*. Se observa la aparición en todas las digestiones de bandas de ADN repetido. Los números de la izquierda indican el tamaño de los fragmentos de ADN en pares de bases.

El género *Heliotaurus* pertenece a la familia Tenebrionidae, subfamilia Alleculinae. La familia Tenebrionidae ha sido modelo de estudio para el ADN altamente repetitivo, especialmente del ADN satélite (ADNsat). En las especies estudiadas hasta el momento se ha observado la existencia de grandes cantidades de ADNsat, pudiendo coexistir en un mismo genoma varias familias (Palomeque & Lorite 2008, Plohl 2010). Menos se conoce sobre ADN repetitivo interdisperso, pero los datos que se han obtenido de la secuenciación del genoma de *Tribolium castaneum* muestran que el genoma de estos insectos también presenta cantidades importantes de elementos transponibles, tanto de la clase I o ADN transposones así

como de retrotransposones (Wang et al. 2008a, Richards et al. 2008). El complejo patrón de bandas obtenidas en las digestiones del ADN genómico de *H. ruficollis* y el tamaño variable de las mismas parece indicar la posible existencia en su genoma de varias familias de ADN repetitivos, lo que indicaría que en la subfamilia Alleculinae presenta un comportamiento similar al observado en otras subfamilias de tenebriónidos.

En *H. ruficollis* ninguna de las enzimas de las que se han usado origina un patrón en escalera típico de los ADNsats. Esto no descarta que alguna de estas bandas corresponda a algún ADNsat organizado en tándem, ya que a veces estos no generan patrones en escalera, debido, generalmente, a la existencia de varias dianas para el mismo enzima y/o a que el fragmento generado sea parte de un monómero de tamaño mayor. Este tipo de patrones de restricción complejos también ha sido observado en otras especies de coleópteros, como en *Chrysolina carnifex* o en el escarabajo de la patata (Palomeque et al. 2005, Lorite et al. 2013).

De los diferentes ADN repetitivos encontrados en este trabajo se analiza el ADN puesto de manifiesto mediante la digestión con *AluI* ya que presenta un patrón sencillo apareciendo una única banda de aproximadamente 800 pb.

Esta banda de 800 pb se extrajo de la agarosa y posteriormente se clonó en un plásmido pUC19. Los clones obtenidos se denominaron HrufAlu, Hruf de *Heliotaurus ruficollis* y Alu, para diferenciar este ADN repetitivo del generado con otros enzimas de restricción. Parte del ADN purificado de la banda de 800 pb fue marcado con digoxigenina y usado para la selección de clones positivos. La selección de clones se realizó mediante un dot-blot (**Figura 6**). De los clones que mostraban más intensa hibridación se seleccionaron cuatro que fueron posteriormente secuenciados; HrufAlu-10, -17, -28 y -38.

En la **Figura 7** se muestra el alineamiento de las secuencias de los distintos clones obtenidos. Para el alineamiento de secuencias y obtención de una secuencia consenso se utilizó el programa CLUSTALW. El análisis de estas secuencias indica que el tamaño real del fragmento clonado es de 811 pb. De las 811 posiciones, 769 son invariables lo que indica la existencia de un elevado grado de conservación

entre las secuencias. Las variaciones de la secuencia consenso predominantes entre las especies son los cambios puntuales de nucleótidos. Se han detectado sustituciones en determinadas posiciones en una única secuencia (por ejemplo, una A en la posición 58 de la secuencia HrufAlu-28), y sustituciones que se repiten en varias secuencias (por ejemplo, una G en la posición 22 de las secuencias HrufAlu-10 y HrufAlu-28. Estas mutaciones compartidas entre secuencias se consideran que han sido generadas por un evento único de mutación y que posteriormente se han extendido a otras copias de la secuencia de forma gradual (Strachan *et al.* 1985). A diferencia de lo que ocurre en otros ADN repetitivos el número de inserciones y deleciones observadas es especialmente bajo, de hecho no se observa ninguna inserción en las secuencias analizadas y solamente dos deleciones de un solo nucleótido en dos de los clones, una deleción de C en la posición 67 del clon HrufAlu-38 y una deleción de G en la posición 805 del clon HrufAlu-10.

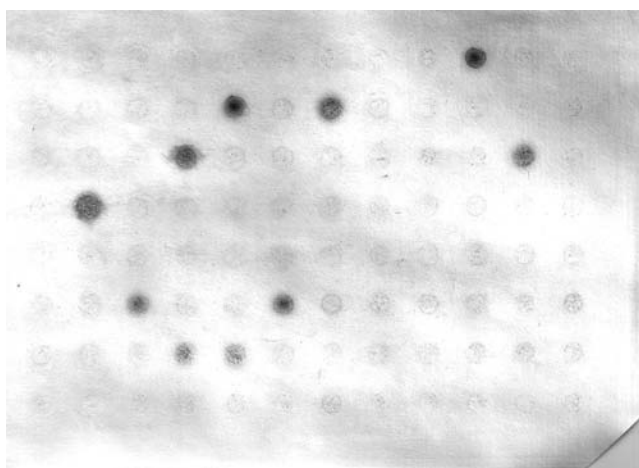


Figura 6. Dot-blot utilizado para la selección de los clones positivos generados a partir de la banda de 800 pb. La sonda para revelarlos se obtuvo a partir de la misma banda de 800 pb marcada con digoxigenina.

La riqueza en AT de este ADN repetitivo es del 47.22%, lo cual es inusual en insectos, cuyo genoma es rico en AT y especialmente el ADNsat (revisión en Palomeque & Lorite 2008). En los tenebriónidos analizados hasta el momento el contenido en AT del ADNsat es especialmente alto, llegando a superar el 79% como ocurre en la familia TDEST 140-bp de *Tribolium destructor* (Mravinac *et al.* 2004). Si bien no es común, se han descrito casos de ADNsat en tenebriónidos cuyo contenido es bajo en AT, como por ejemplo en la especie *Alphitobius diaperinus*, cuyo ADNst tiene una riqueza en AT del 49% (Plohl & Ugarkovic 1994, Bruvo *et al.* 1995).

```

      10      20      30      40      50      60      70      80      90     100
consenso  CTCGTACAGGAGATCTTGATATTTGTCCGTTTTCTCTCTCTTCAGGACTATGTTGGTCTCCGCCGGAGCAGAAAACTCGATGACATACATAGTCCTG
HrufAlu-10 .....G.....
HrufAlu-17 .....G.....
HrufAlu-28 .....A.....
HrufAlu-38 .....C.....
      TGCCTCTAGAACTATAAAA - 5' HrufA-2
      110     120     130     140     150     160     170     180     190     200
consenso  GTGTGTTTGTTCGAGGAGGACAATGTCAGGCTTGTGGCGTGAATAAGCCTGACAGTCGGAAATGAATAGTCCAGAAAACTGGCATCGCTCATTCCCGA
HrufAlu-10 .....A.....
HrufAlu-17 .....T.....
HrufAlu-28 .....C.....
HrufAlu-38 .....C.....
      210     220     230     240     250     260     270     280     290     300
consenso  CTACGGACTCAATATCTCCCGGTAAAGTAAGGAGGCACTGGCTCTTGGTCTATCTCATACGAGTGCCTCAGGTGGTAAATAGAGCACTCGTAACGCCGCGTT
HrufAlu-10 .....A.....
HrufAlu-17 .....C.....
HrufAlu-28 .....C.....
HrufAlu-38 .....G.....
      310     320     330     340     350     360     370     380     390     400
consenso  ATGTCGATGCGGTGTACGCCGAGGCAGCCTTCATCGGACAGCCGACAACAATGCATCAGCGTCTCCGGTGCCCGGCGACAAGCCCTGCGAGTTGACATCA
HrufAlu-10 .....G.....
HrufAlu-17 .....G.....
HrufAlu-28 .....G.....
HrufAlu-38 .....G.....
      410     420     430     440     450     460     470     480     490     500
consenso  GGGACTTGCACACGCATCACGCGACTTCGATAGGCCAGGGTGTGATAACGCCATCTTGGCAGGCCGCAAGAACCCTTCGGTCTCCGACATGAGCCCGG
HrufAlu-10 .....A.....
HrufAlu-17 .....A.....
HrufAlu-28 .....A.....
HrufAlu-38 .....T.....
      510     520     530     540     550     560     570     580     590     600
consenso  CTGACTTAAGATAGAGGAAAGTCAGCCTTGTGACAGACCCAGCGTCTCGATGTTTCTATAGAAAACACCATGCATGGGTCTGTCAACATGCAGCCTAAA
HrufAlu-10 .....C.....
HrufAlu-17 .....GT.....
HrufAlu-28 .....G.....
HrufAlu-38 .....A.....
      610     620     630     640     650     660     670     680     690     700
consenso  CAAGCGATCTTGTTCCGGCTGCCCTTATTGGGCTCTTAGGCGATTGCGCTCCAACCTCGGTAATCAACTCAGGGTTGACACCTCGCCGGCTCCTCGCCTA
HrufAlu-10 .....G.....
HrufAlu-17 .....G.....
HrufAlu-28 .....G.....
HrufAlu-38 .....T.....
      710     720     730     740     750     760     770     780     790     800
consenso  AACTCGAAGCGGAGACCCAGTTCACTGGCGGATCTAACAGCCTCTTCAATATAAACGACCCACACCTGCAATCTCGTGATCTCTCACCAGAGATAGGA
HrufAlu-10 .....G.....
HrufAlu-17 .....G.....
HrufAlu-28 .....G.....
HrufAlu-38 .....C.....
      810
consenso  GGGGGTCCGAG
HrufAlu-10 .....T.....
HrufAlu-17 .....T.....
HrufAlu-28 .....T.....
HrufAlu-38 .....T.....
      HrufA-1 5' - CTCTCACCAGAGATAGGAG

```

Figura 7. Alineamiento múltiple de todas las secuencias de los clones obtenidos de la clonación de la banda de 800 pb generada mediante digestión con *AluI* y secuencia consenso obtenida a partir de ellos. Se indica en la consenso las secuencias diana para algunos enzimas de restricción: *HaeIII* (GGCC), *RsaI* (GTAC), *ScrFI* (CCNGG), *PstI* (CTGCAG). También se indica la localización de los cebadores usados en la PCR.

Una característica que sí comparte el ADN HrufAlu con el ADNsat de tenebriónidos y otros insectos es la presencia de repeticiones invertidas y secuencias palindrómicas. Estas secuencias pueden dar lugar a plegamientos sobre sí mismas por apareamientos intracatenarios originando estructuras secundarias cruciformes. La aplicación de programas capaces de detectar posibles estructuras

secundarias (mfold, Zuker 2003) pone de manifiesto que estas pueden formarse prácticamente en la totalidad de la secuencia de 811 pb (**Figura 8**). Se ha indicado que estas estructuras podrían actuar como señales en el posicionamiento de nucleosomas en el repliegamiento de la cromatina y/o actuar como puntos de reconocimiento de proteínas (Bigot et al. 1991, Brázda et al. 2011).

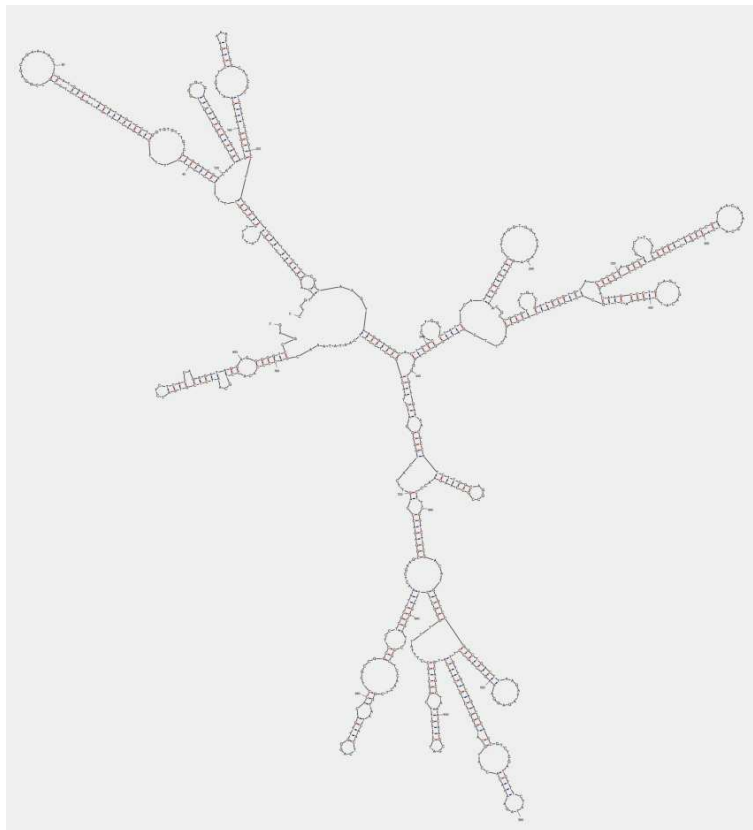


Figura 8. Estructura secundaria de la secuencia de 811 pb obtenida con el programa mfol. Pueden observarse las numerosas regiones donde es posible la formación de estructuras cruciformes.

Para determinar el modelo de organización de las secuencias HrufAlu de 811 pb y como se relacionan entre sí, si están organizadas en tándem o dispersas, se analizó el ADN genómico mediante hibridación Southern-blot. Para ello se digirió ADN genómico con *AluI* así como con otros enzimas de restricción cuya diana estaba presente en la secuencia de 811 pb. Como sonda se utilizó el ADN de los clones HrufAlu-10 y 17 previamente amplificado por PCR. En la **Figura 9** se muestra el resultado obtenido. En el ADN genómico digerido con *AluI* solo es posible observar una única banda correspondiente al fragmento clonado. Si la secuencia de 811 pb fuese la unidad de repetición de un ADN organizado en tándem cabría esperar la aparición de una escalera de bandas. Estas escaleras se

originan por mutaciones en las secuencias diana de los enzimas de modo que el enzima no puede cortar en todas las repeticiones generando fragmentos que contienen dos, tres o más copias de la secuencia repetida en tándem. Sin embargo la ausencia de esta escalera no puede por sí sola descartar que se trate de un ADN organizado en tándem. A veces la elevada conservación de las secuencias hace que la mayor parte de ellas conserven la diana para el enzima y no se generan escaleras visibles en los geles de agarosa o en las hibridaciones Southern, tal como se ha descrito en el escarabajo de la patata (Lorite et al. 2013).

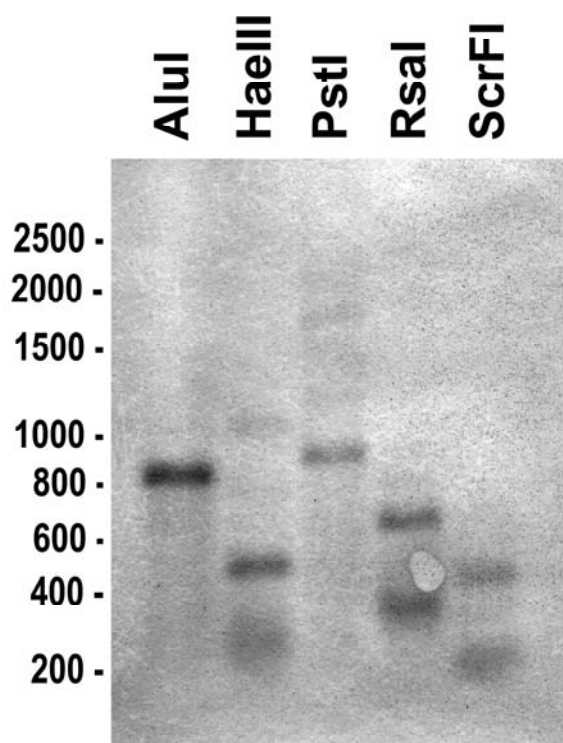


Figura 9. Hibridación Southern-blot realizada a partir de ADN genómico digerido con las enzimas que se indican en la parte superior. A la izquierda los números indican el tamaño en pares de bases

Aunque las secuencias HrufAlu-800 pb son altamente conservadas, no parece que esta secuencia sea el monómero de repetición de un ADN organizado en tándem, tal como se puede deducir de los resultados obtenidos con los otros enzimas de restricción. La aparición de dos bandas de menos de 800 pb para los enzimas *HaeIII*, *RsaI* y *ScrFI* es un resultado esperable puesto que la secuencias HrufAlu-800 pb tiene al menos dos dianas para cada una de ellas (**Figura 7**). Especialmente importante es el resultado obtenido con *PstI*. Para este enzima la secuencia de 811 pb solo tiene una diana, de modo que cabría esperar el mismo resultado que con *AluI* o incluso la aparición de una escalera de 800 pb si la diana

estuviese mutada en algunas copias. Sin embargo, la banda que aparece en el Southern-blot tras la digestión con *Pst*I es de mayor tamaño, aproximadamente de 870 pb). Por tanto el fragmento de 811 pb clonado debe ser parte de un ADN repetitivo de mayor tamaño. Para comprobar esta hipótesis y si este ADN se encuentra repetido en tándem o disperso por el genoma se llevó a cabo PCR con cebadores que se unen a la secuencia de 811 pb y dirigidos hacia afuera (**Figura 10**). Si la secuencia repetida se encuentra organizada en tándem la PCR generaría una única banda discreta correspondiente a la amplificación de la región comprendida entre dos de las secuencias de 811 pb, amplificándose por tanto la unidad de repetición. También sería posible la amplificación de dímeros, trímeros, etc, siempre y cuando el tamaño del fragmento tuviese un tamaño factible de ser amplificado por la Taq polimerasa. Si por el contrario la secuencia repetida estuviese dispersa en el genoma el resultado de la PCR podría ser muy variable, dependiendo de su frecuencia en el genoma y por tanto de la probabilidad de que dos de ellas estuviesen lo suficientemente cerca para poder ser amplificadas por PCR, no guardando el tamaño de los productos ningún tipo de relación. En el caso de secuencias organizadas en tándem la diferencia de tamaño entre las bandas que se obtuviesen sería el de la secuencia que se repite en tándem.

Cuando se realizó la PCR usando los cebadores basados en la secuencia de 811 pb se obtuvieron dos bandas, de aproximadamente 400 pb y 3 kb respectivamente (**Figura 11**). El ADN de estas bandas fue extraído de la agarosa y clonado en vector pGEM-T. En este caso la selección de los plásmidos con inserto se realizó mediante lisis rápida de los clones bacterianos obtenidos. En la electroforesis del DNA obtenido de las bacterias es posible diferenciar plásmidos que no contienen inserto de aquellos que han incorporado el ADN purificado a partir de la PCR (**Figura 12**). De los plásmidos obtenidos se seleccionaron los clones HrufAlu-419, -420, -421 y -422 de los que incluían el fragmento de 400 pb, y los clones HrufAlu-325, -328 y -335 de los clones que incluían el fragmento de 3 kb.

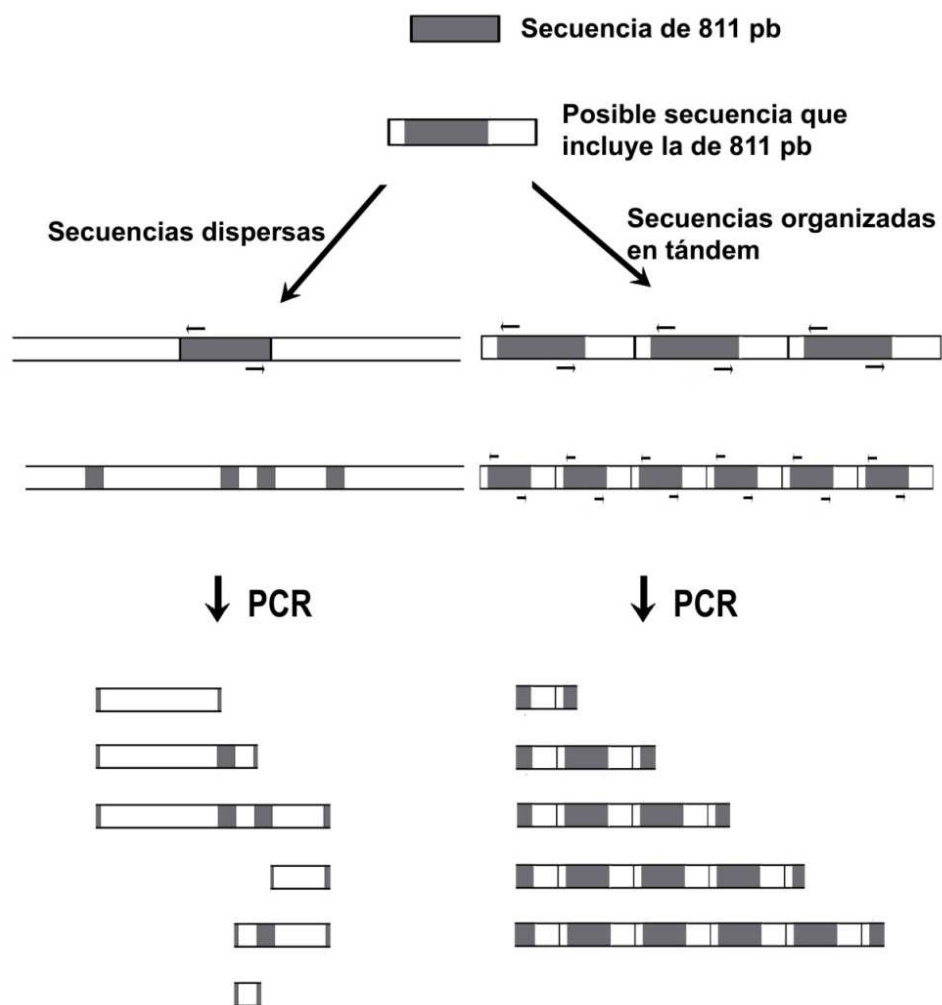


Figura 10. Esquema del procedimiento basado en la PCR para determinar la organización de las secuencias HrufAlu de 800 pb. Explicación en el texto.

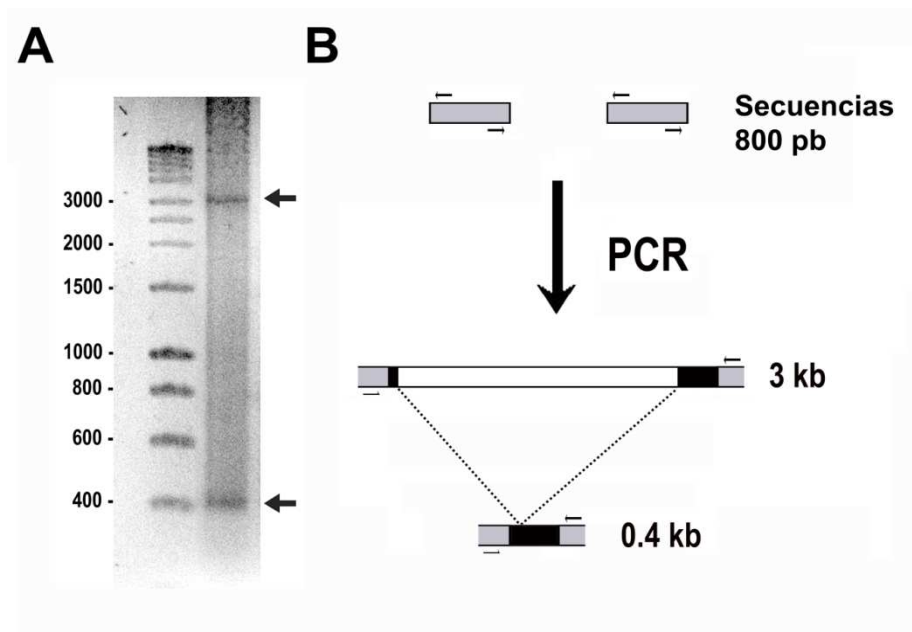


Figura 11. (A) Resultado de la PCR sobre DNA genómico de *H. ruficollis* usando cebadores basados en la secuencias de 811 pb. Se observa la amplificación de solamente dos bandas, una de aproximadamente 400 pb y otra de 3 kb. (B) Esquema del resultado obtenido por PCR. Los fragmentos de 400 pb y de 3 kb son similares entre sí, los fragmentos de 400 pb presentan una delección interna de aproximadamente 2.6 kb con respecto a los de 3 kb.

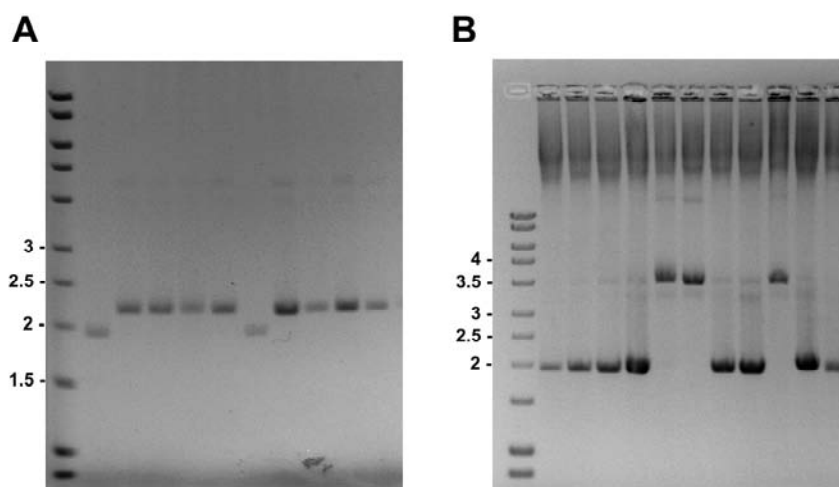


Figura 12. Selección de plásmidos en la clonación de los fragmentos de 400 pb (A) y de 3 kb (B) obtenidos en la PCR. El primer carril en (A) y (B) corresponde a la escalera de peso molecular. Se indica el tamaño de algunas bandas en kb. El segundo carril en ambos corresponde a un plásmido sin inserto. En (A) se observa que la mayoría de los plásmidos aislados contienen el fragmento de 400 pb, excepto el del carril 8 que tiene el mismo tamaño que el plásmido sin inserto. En (B) se observan 3 plásmidos que han incorporado la banda de 3 kb, correspondiente a los carriles 6, 7 y 10.

10 20 30 40 50 60 70 80 90 100

HrufA-1 5'- CTCTCACCAGATAGGAG
HrufAlu-419 CTCTCACCAGATAGGAGGGGGTCCGAGCTCTCAAGATCTTATACGCCATG-----
HrufAlu-420 CTCTCACCAGATAGGAGGGGGTCCGAGCTCTTAAGATCTTA-----
HrufAlu-421 CTCTCACCAGATAGGAGGGGGTCCGAGCTCTCAAGATCTTATACGCCATGTG-----
HrufAlu-422 CTCTCACCAGATAGGAGGGGGTCCGAGCTCTCAAGATCTTATACGCCATG-----
HrufAlu-325 CTCTCACCAGATAGGAGGGGGTCCGAGCTCTCAAGATCTTATACGCCATACCCAAACCACTCTGCCGTGCAAACTCCAGGCTCAGGAGTCCCT
HrufAlu-328 CTCTCACCAGATAGGAGGGGGTCCGAGCTCTCAAGATCTTATACGCCATACCCAAACCACTCTGCCGTGCACTCCAGGCTCAGGAGTCCCT
HrufAlu-335 CTCTCACCAGATAGGAGGGGGTCCGAGCTCTCAAGACCTTATACGCCATACCCAAACCACTCTGCCGTGCACTCCAGGCTCAGGAGTCCCT

110 120 130 140 150 160 170 180 190 200

HrufAlu-325 ACCTCCTTGATATCGTGGAGGTACAATCGAGGAACGGAGCATCTGGGATGTAAACTGCGTTGCATGTGCAACATTTTACGAGTGTCCCTATCGAGTTGG
HrufAlu-328 ACCTCCTTGATATCGTGGAGGTACAATCGAGGAACGGAGCATCTGGGATGTAAACTGCGTTGCATGTGCAACATTTTACGAGTGTCCCTATCGAGTTGG
HrufAlu-335 ACCTCCTTGATATCGTGGAGGTACAATCGAGGAACGGAGCATCTGGGATGTAAACTGCGTTGCATGTGCAACATTTTACGAGTGTCCCTATCGAGTTGG

210 220 230 240 250 260 270 280 290 300

HrufAlu-325 CGCAACTCGTCAACTGTCCACTTGACTACTCCTCAAGAGTGTAGAGCAGAAATGGAAACGGCAAGTAAATAGTGGCAGACACCTTGTCTGCCCGGAAAGTT
HrufAlu-328 CGCAACTCGTCAACTGTCCACTTGACTACTCCTCAAGAGTGTAGAGCAGAAATGGAAACGGCAAGTAAATAGTGGCAGACACCTTGTCTGCCCGGAAAGTT
HrufAlu-335 CGCAACTCGTCAACTGTCCACTTGACTACTCCTCAAGAGTGTAGAGCAGAAATGGAAACGGCAAGTAAATAGTGGCAGACACCTTGTCTGCCCGGAAAGTT

310 320 330 340 350 360 370 380 390 400

HrufAlu-325 CAGAAGACCAAATCTTTTGGACCAAGGCGCTGTACTTGGCGGGAGATCCTCTTTACCTGGGGGGCATCATGGGTGTTATGTTGCTCAATCCTCAGTTA
HrufAlu-328 CAGAAGACCAAATCTTTTGGACCAAGGCGCTGTACTTGGCGGGAGATCCTCTTTACCTGGGGGGCATCATGGGTGTTATGTTGCTCAATCCTCAGTTA
HrufAlu-335 CAGAAGACCAAATCTTTTGGACCAAGGCGCTGTACTTGGCGGGAGATCCTCTTTACCTGGGGGGCATCATGGGTGTTATGTTGCTCAATCCTCAGTTA

410 420 430 440 450 460 470 480 490 500

HrufAlu-325 TCTATAGGCTCTCCAGCATCAAGATGTCTGAAGATGCTCCGTCGGTAAGTTGCACGCCCTCAGGGAACTCTGCAGATGCTCTACCCTTAGATGGAAC
HrufAlu-328 TCTATAGGCTCTCCAGCATCAAGATGTCTGAAGATGCTCCGTCGGTAAGTTGCACGCCCTCAGGGAACTCTGCAGATGCTCTACCCTTAGATGGAAC
HrufAlu-335 TCTATAGGCTCTCCAGCATCAAGATGTCTGAAGATGCTCCGTCGGTAAGTTGCACGCCCTCAGGGAACTCTGCAGATGCTCTACCCTTAGATGGAAC

510 520 530 540 550 560 570 580 590 600

HrufAlu-325 ACTGCGCACTTGTCCAACCCCAATTCATTCCCAGCTCCCGTGTATCTCTGGACCAGCTTGAGAGACTGATGGAGATCTCTCTCCCGGAGGCATATA
HrufAlu-328 ACTGCGCACTTGTCCAACCCCAATTCATTCCCAGCTCCCGTGTATCTCTGGACCAGCTTGAGAGACTGATGGAGATCTCTCTCCCGGAGGCATATA
HrufAlu-335 ACTGCGCACTTGTCCAACCCCAATTCATTCCCAGCTCCCGTGTATCTCTGGACCAGCTTGAGAGACTGATGGAGATCTCTCTCCCGGAGGCATATA

610 620 630 640 650 660 670 680 690 700

HrufAlu-325 GCTTGAGATCATCCATGTAAACAAATGGGTGATCTCATGCTCCGGTTACTGGGTGGTCCGCATCTGTATCCACGGGTACGCTCTCAGGGCAATGGAAAG
HrufAlu-328 GCTTGAGATCATCCATGTAAACAAATGGGTGATCTCATGCTCCGGTTACTGGGTGGTCCGCATCTGTATCCACGGGTACGCTCTCAGGGCAATGGAAAG
HrufAlu-335 GCTTGAGATCATCCATGTAAACAAATGGGTGATCTCATGCTCCGGTTACTGGGTGGTCCGCATCTGTATCCACGGGTACGCTCTCAGGGCAATGGAAAG

710 720 730 740 750 760 770 780 790 800

HrufAlu-325 CGGCAGCAGTGAGATGCAAAACAGAACGGGACTCAGGGAACTCTCCCTGGAAACACCTCTTCTGAACGTACCAACTCTGTAGTAATTCGCCGAGTGCCG
HrufAlu-328 CGGCAGCAGTGAGATGCAAAACAGAACGGGACTCAGGGAACTCTCCCTGGAAACACCTCTTCTGAACGTACCAACTCTGTAGTAATTCGCCGAGTGCCG
HrufAlu-335 CGGCAGCAGTGAGATGCAAAACAGAACGGGACTCAGGGAACTCTCCCTGGAAACACCTCTTCTGAACGTACCAACTCTGTAGTAATTCGCCGAGTGCCG

810 820 830 840 850 860 870 880 890 900

HrufAlu-325 GATGTTATGGCAAAACGAGTTTTCAGAGTGGCAACAGGCGTTCACTAGCACCCTACCAAGCGAGCAGGTACCGCAAGGCAAGTTGAGCAAGTGTAAAAAGT
HrufAlu-328 GATGTTATGGCAAAACGAGTTTTCAGAGTGGCAACAGGCGTTCACTAGCACCCTACCAAGCGAGCAGGTACCGCAAGGCAAGTTGAGCAAGTGTAAAAAGT
HrufAlu-335 GATGTTATGGCAAAACGAGTTTTCAGAGTGGCAACAGGCGTTCACTAGCACCCTACCAAGCGAGCAGGTACCGCAAGGCAAGTTGAGCAAGTGTAAAAAGT

910 920 930 940 950 960 970 980 990 1000

HrufAlu-325 GCTCATGAGATGTTGAGTCAAATGCCCTTGGGTAATCAATCCAGGCCATGGAGAGGTTGCGCTTGTACTGGATGGAATCCTGGGTGACCACACGGTCAAT
HrufAlu-328 GCTCATGAGATGTTGAGTCAAATGCCCTTGGGTAATCAATCCAGGCCATGGAGAGGTTGCGCTTGTACTGGATGGAATCCTGGGTGACCACACGGTCAAT
HrufAlu-335 GCTCATGAGATGTTGAGTCAAATGCCCTTGGGTAATCAATCCAGGCCATGGAGAGGTTGCGCTTGTACTGGATGGAATCCTGGGTGACCACACGGTCAAT

1010 1020 1030 1040 1050 1060 1070 1080 1090 1100

HrufAlu-325 CACCAGGTTCTCGCGACAACCCCTCCATATCCC - TTTTCCCCACGCTGCTGCTGATCTCTTGTAGTACCGGCTGATCGATCTCTCCAGCCTGGCATA
HrufAlu-328 CACCAGGTTCTCGCGACAACCCCTCCATATCCC - TTTTCCCCACGCTGCTGCTGATCTCTTGTAGTACCGGCTGATCGATCTCTCCAGCCTGGCATA
HrufAlu-335 CACCAGGTTCTCGCGACAACCCCTCCATATCCC - TTTTCCCCACGCTGCTGCTGATCTCTTGTAGTACCGGCTGATCGATCTCTCCAGCCTGGCATA

1110 1120 1130 1140 1150 1160 1170 1180 1190 1200

HrufAlu-325 TAAGATTGGGTTGAAGAGCTTATAGCAGGTGTTGAGACAGGTGATGCTCGGTAGTTCTTGGGGTCAAGACAGATCCCTTTCTTGGGGAGCAGCAGATG
HrufAlu-328 TAAGATTGGGTTGAAGAGCTTATAGCAGGTGTTGAGACAGGTGATGCTCGGTAGTTCTTGGGGTCAAGACAGATCCCTTTCTTGGGGAGCAGCAGATG
HrufAlu-335 TAAGATTGGGTTGAAGAGCTTATAGCAGGTGTTGAGACAGGTGATGCTCGGTAGTTCTTGGGGTCAAGACAGATCCCTTTCTTGGGGAGCAGCAGATG

1210 1220 1230 1240 1250 1260 1270 1280 1290 1300

HrufAlu-325 CGCCCTTCGACAGCCAGTTCTGGAATAGGCGCTCCCCATTAGCCACAGGATCATAAACTCTGGCAAGACGGCGATGGGTGGTGGAAACTTTTTCCACC
HrufAlu-328 CGCCCTTCGACAGCCAGTTCTGGAATAGGCGCTCCCCATTAGCCACAGGATCATAAACTCTGGCAAGACGGCGATGGGTGGTGGAAACTTTTTCCACC
HrufAlu-335 CGCCCTTCGACAGCCAGTTCTGGAATAGGCGCTCCCCATTAGCCACAGGATCATAAACTCTGGCAAGACGGCGATGGGTGGTGGAAACTTTTTCCACC

1310 1320 1330 1340 1350 1360 1370 1380 1390 1400

HrufAlu-325 ACCAGTTGTTAATGCCGTCTGGTCCGGGGAGGCGAAGTTCTTGGCGCCAGAGAGTGCTGCCGAACCTTCTCCTCGCTGACTGCCTCGCACTCTTCGAC
HrufAlu-328 ACCAGTTGTTAATGCCGTCTGGTCCGGGGAGGCGAAGTTCTTGGCGCCAGAGAGTGCTGCCGAACCTTCTCCTCGCTGACTGCCTCGCACTCTTCGAC
HrufAlu-335 ACCAGTTGTTAATGCCGTCTGGTCCGGGGAGGCGAAGTTCTTGGCGCCAGAGAGTGCTGCCGAACCTTCTCCTCGCTGACTGCCTCGCACTCTTCGAC

1410 1420 1430 1440 1450 1460 1470 1480 1490 1500

HrufAlu-325 TTGTGGCTCTAAAGCCGTTGGCAAAAGAGTCAAGGTGTACTAGCGCTGGAGTGTCTAGACGGAAGTGGACGCGGATGGTATAGACTCCCTTCCAGAAG
HrufAlu-328 TTGTGGCTCTAAAGCCGTTGGCAAAAGAGTCAAGGTGTACTAGCTGGAGTGTCTAGACGGAAGTGGACGCGGATGGTATAGACTCCCTTCCAGAAG
HrufAlu-335 TTGTGGCTCTAAAGCCGTTGGCAAAAGAGTCAAGGTGTACTAGCGCTGGAGTGTCTAGACGGAAGTGGACGCGGATGGTATAGACTCCCTTCCAGAAG

1510 1520 1530 1540 1550 1560 1570 1580 1590 1600

HrufAlu-325 TTTTCGACTTCTCTGCGGTTGGCGGATTATCCACGCGTGGTGGTGGAGCCTCAAAACAGCCGAGACGGGTTCTACGAAACAAAGCGGTTCTCCTTGATAC
HrufAlu-328 TTTTCGACTTCTCTGCGGTTGGCGGATTATCCACGCGTGGTGGTGGAGCCTCAAAACAGCCGAGACGGGTTTTCGCAAAACAAAGCGGTTCTCCTTGATAC
HrufAlu-335 TTTTCGACTTCTCTGCGGTTGGCGGATTATCCACGCGTGGTGGTGGAGCCTCAAAACAGCCGAGACGGGTTTTCGCAAAACAAAGCGGTTCTCCTTGATAC

1610 1620 1630 1640 1650 1660 1670 1680 1690 1700

HrufAlu-325 GCTTCACCCGCTGCGGAGACCCCTTTTTCACCACCACTAGCATTCGATCTATCCGACAGTGTGCGCGATGGTCTGCAGCAGTACAGTTTAGAGT
HrufAlu-328 GCTTCACCCGCTGCGGAGACCCCTTTTTCACCACCACTAGCATTCGATCTATCCACAGTGTGCGCGATGGTCTGCAGCAGTACAGTTTAGAGT
HrufAlu-335 GCTTCACCCGCTGCGGAGACCCCTTTTTCACCACCACTAGCATTCGATCTATCCGACAGTGTGCGCGATGGTCTGCAGCAGTACAGTTTAGAGT

1710 1720 1730 1740 1750 1760 1770 1780 1790 1800

HrufAlu-325 GTGATGCTGGCGGCGTAACATGGCAGCTACGCGCTTGATCTTGGGCGTAACGCTCTGCCGCCAACAAATGACTGCATCACACTCTGGATCCGAGAAGCG
HrufAlu-328 GTGATGCTGGCGGCGTAACATGGCAGCTACGCGCTTGATCTTGGGCGTAACGCTCTGCCGCCAACAAATGACTGCATCACACTCTGGATCCGAGAAGCG
HrufAlu-335 GTGATGCTGGCGGCGTAACATGGCAGCTACGCGCTTGATCTTGGGCGTAACGCTCTGCCGCCAACAAATGACTGCATCACACTCTGGATCCGAGAAGCG

1810 1820 1830 1840 1850 1860 1870 1880 1890 1900

HrufAlu-325 TGTGCTGTCGCAACCGAACTTGGCTCAACCTGTTGTAGGCGCCGTTCAAGTTTCCCAACCAAGGTCAGACGGCTCTCACGCCGCTCTCCGGT--AC

HrufAlu-328 TGTGCTGTCGCAACCGAACTTGGCTCAACCTGTTGTAGGCGCCGTTCAAGTTTCCCAACCAAGGTCAGACGGCTCTCACGCCGCTCTCCGGT--TAC

HrufAlu-335 TGTGCTGTCGCAACCGAACTTGGCTCAACCTGTTGTAGGCGCCGTTCAAGTTTCCCAACCAAGGTCAGACGGCTCTCACGCCGCTCTCCGGTGTATC

1910 1920 1930 1940 1950 1960 1970 1980 1990 2000

HrufAlu-325 GAATCGCCTCGCGCGCCATATACGGCTCGTTGATCTGCGCAGAGGTCGCTTCTTCCGGAAGTTTCATCGTCTATTAAGCGGTCAATTTCCGCCGCTTCT

HrufAlu-328 GAATCGCCTCGCGCGCCATATACGGCTCGTTGATCTGCGCAGAGGTCGCTTCTTCCGGAAGTTTCATCGTCTATTAAGCGGTCAATTTCCGCCGCTTCT

HrufAlu-335 GAATCGCCTCGCGCGCCATATACGGCTCGTTGATCTGCGCAGAGGTCGCTTCTTCCGGAAGTTTCATCGTCTATTAAGCGGTCAATTTCCGCCGCTTCT

2010 2020 2030 2040 2050 2060 2070 2080 2090 2100

HrufAlu-325 TGTGGTTTTAGGGGACTCTGATGTTAATGCGGTCACCCCTTGCCCTCAAGTTAACTTCGGTGTCTGAACCAAGGTCAGTCAGCAGACCTGTTACGCC

HrufAlu-328 TGTGGTTTTAGGGGACTCTGATGTTAATGCGGTCACCCCTTGCCCTCAAGTTAACTTCGGTGTCTGAACCAAGGTCAGTCAGCAGACCTGTTACGCC

HrufAlu-335 TGTGGTTTTAGGGGACTCTGATGTTAATGCGGTCACCCCTTGCCCTCAAGTTAACTTCGGTGTCTGAACCAAGGTCAGTCAGCAGACCTGTTACGCC

2110 2120 2130 2140 2150 2160 2170 2180 2190 2200

HrufAlu-325 GCACCGGAGGTGTTGGGGGTGCGGCTGTTGAGGATCAGCTGTTGAAGGTCGCTCTAATGTAACCGACCCCTTCGGAGATACCAACCTGATCTCTCAG

HrufAlu-328 GCACCGGAGGTGTTGGGGGTGCGGCTGTTGAGGATCAGCTGTTGAAGGTCGCTCTAATGTAACCGACCCCTTCGGAGATACCAACCTGATCTCTCAG

HrufAlu-335 GCACCGGAGGTGTTGGGGGTGCGGCTGTTGAGGATCAGCTGTTGAAGGTCGCTCTAATGTAACCGACCCCTTCGGAGATACCAACCTGATCTCTCAG

2210 2220 2230 2240 2250 2260 2270 2280 2290 2300

HrufAlu-325 CTTT-GTTGAGAGAAGCGCGGTGTTGGGATGCAATTTACACCAAAGTGAGTGCAACCGCGCCATGATATCGCTCCGGGCGGCTCGCTGCGCTGATAG

HrufAlu-328 AGCTTTGTTGAGAGAAGCGCGGTGTTGGGATGCAATTTACACCAAAGTGAGTGCAACCGCGCCATGATATCGCTCCGGGCGGCTCGCTGCGCTGATAG

HrufAlu-335 CTTTGTGTTGAGAGAAGCGCGGTGTTGGGATGCAATTTACACCAAAGTGAGTGCAACCGCGCCATGATATCGCTCCGGGCGGCTCGCTGCGCTGATAG

2310 2320 2330 2340 2350 2360 2370 2380 2390 2400

HrufAlu-325 CAGTCCAGTATGCTGTTCCGAAGTCTGCTCGTCCACTTAAAGCGTCCGTGTCATGCGCAAGGTCGTCATCGTGGGACCATCCGGTATGGCCCTCTCA

HrufAlu-328 CAGTCCAGTATGCTGTTCCGAAGTCTGCTCGTCCACTTAAAGCGTCCGTGTCATGCGCAAGGTCGTCATCGTGGGACCATCCGGTATGGCCCTCTCA

HrufAlu-335 CAGTCCAGTATGCTGTTCCGAAGTCTGCTCGTCCACTTAAAGCGTCCGTGTCATGCGCAAGGTCGTCATCGTGGGACCATCCGGTATGGCCCTCTCA

2410 2420 2430 2440 2450 2460 2470 2480 2490 2500

HrufAlu-325 CCTGTCCTGTTCTCATCCGACGAGGAAGTATGTTGGGAGGCCCTTCGATATTTGCGGTAACAAATATGTTGTTCTACCGGGGCGTTAACCCGGCAGA

HrufAlu-328 CCTGTCCTGTTCTCATCCGACGAGGAAGTATGTTGGGAGGCCCTTCGATATTTGCGGTAACAAATATGTTGTTCTACCGGGGCGTTAACCCGGCAGA

HrufAlu-335 CCTGTCCTGTTCTCATCCGACGAGGAAGTATGTTGGGAGGCCCTTCGATATTTGCGGTAACAAATATGTTGTTCTACCGGGGCGTTAACCCGGCAGA

2510 2520 2530 2540 2550 2560 2570 2580 2590 2600

HrufAlu-325 CACGGCATTGTTGTTTAAAGGGTACCAGGCCGATTTAAACCCCGATTTTATCGAGGTTGTCTCTCTACGCTCGGGAATGTGGTGGCACTACTCACCA

HrufAlu-328 CACGGCATTGTTGTTTAAAGGGTACCAGGCCGATTTAAACCCCGATTTTATCGAGGTTGTCTCTCTACGCTCGGGAATGTGGTGGCACTACTCACCA

HrufAlu-335 CACGGCATTGTTGTTTAAAGGGTACCAGGCCGATTTAAACCCCGATTTTATCGAGGTTGTCTCTCTACGCTCGGGAATGTGGTGGCACTACTCACCA

2610 2620 2630 2640 2650 2660 2670 2680 2690 2700

HrufAlu-419 -----TTTTTTTTTTTTTTTTTTCGCTACAAACT-----GGTCGGGGCACTGCGCGTTCCGAC

HrufAlu-420 -----ATCTATTATTTTTTTTCGATTTTTTTTCCCTACAAACT-----GGTCGGGGCACTGCGCGTTCCGAC

HrufAlu-421 -----TTTTTTTTTTTTTTTTTTCGCTACAAACT-----GGTCGGGGCACTGCGCGTTCCGAC

HrufAlu-422 -----TTTTTTTTTTTTTTTTTTCGCTACAAACT-----GGTCGGGGCACTGCGCGTTCCGAC

HrufAlu-325 ATGAACAAGCTAAAGTTTGTCTCACTCTGCTATTTCTTTTTTTTTTTTTTTTTTTCGCTACAAACT-----GGTCGGGGCACTGCGCGTTCCGAC

HrufAlu-328 ATGAACAAGCTAAAGTTTGTCTCACTCTGCTATTTCTTTTTTTTTTTTTTTTTTTCGCTACAAACT-----GGTCGGGGCACTGCGCGTTCCGAC

HrufAlu-335 ATGAACAAGCTAAAGTTTGTCTCACTCTGCTATTTCTTTTTTTTTTTTTTTTTTTCGCTACAAACT-----GGTCGGGGCACTGCGCGTTCCGAC

2710 2720 2730 2740 2750 2760 2770 2780 2790 2800

HrufAlu-419 ACCATCGCAGCGACCGCGCCAGAAATCTAAAGGTGTTAAATAACTGCGACTACTCTCTCCACCGGTCGCTTAAAGTAAACCGAGGTTGCCGAATCGTGTGC

HrufAlu-420 ACCATCGCAGCGACCGCGCCAGAAATCTAAAGGTGTTAAATAACTGCGACTACTCTCTCCACCGGTCGCTTAAAGTAAACCGAGGTTGCCGAATCGTGTGC

HrufAlu-421 ACCATCGCAGCGACCGCGCCAGAAATCTAAAGGTGTTAAATAACTGCGACTACTCTCTCCACCGGTCGCTTAAAGTAAACCGAGGTTGCCGAATCGTGTGC

HrufAlu-422 ACCATCGCAGCGACCGCGCCAGAAATCTAAAGGTGTTAAATAACTGCGACTACTCTCTCCACCGGTCGCTTAAAGTAAACCGAGGTTGCCGAATCGTGTGC

HrufAlu-325 ACCATCGCAGCGACCGCGCCAGAAATCTAAAGGTGTTAAATAACTGCGACTACTCTCTCCACCGGTCGCTTAAAGTAAACCGAGGTTGCCGAATCGTGTGC

HrufAlu-328 ACCATCGCAGCGACCGCGCCAGAAATCTAAAGGTGTTAAATAACTGCGACTACTCTCTCCACCGGTCGCTTAAAGTAAACCGAGGTTGCCGAATCGTGTGC

HrufAlu-335 ACCATCGCAGCGACCGCGCCAGAA--CAAAGGTGCTGATTAATGCGGCTACTCTCTCCACCGGTCGCTTAAAGTAAACCGAGGTTGCCGAATCGTGTGC

2810 2820 2830 2840 2850 2860 2870 2880 2890 2900

HrufAlu-419 CCTAAGTAAACGAAGCGAACCAGAACGACCGCTTCTGATCCGGCATAGTAGGCTCTCCGCAAGTTCTTACAAGCGGAATGGCTTTGAGTTACAGAC

HrufAlu-420 CCTAAGTAAACGAAGCGAACCAGAACGACCGCTTCTGATCCGGCATAGTAGGCTCTCCGCAAGTTCTTACAAGCGGAATGGCTTTGAGTTACAGAC

HrufAlu-421 CCTAAGTAAACGAAGCGAACCAGAACGACCGCTTCTGATCCGGCATAGTAGGCTCTCCGCAAGTTCTTACAAGCGGAATGGCTTTGAGTTACAGAC

HrufAlu-422 CCTAAGTAAACGAAGCGAACCAGAACGACCGCTTCTGATCCGGCATAGTAGGCTCTCCGCAAGTTCTTACAAGCGGAATGGCTTTGAGTTACAGAC

HrufAlu-325 CCTAAGTAAACGAAGCGAACCAGAACGACCGCTTCTGATCCGGCATAGTAGGCTCTCCGCAAGTTCTTACAAGCGGAATGGCTTTGAGTTACAGAC

HrufAlu-328 CCTAAGTAAACGAAGCGAACCAGAACGACCGCTTCTGATCCGGCATAGTAGGCTCTCCGCAAGTTCTTACAAGCGGAATGGCTTTGAGTTACAGAC

HrufAlu-335 CCTAAGTAAACGAAGCGAACCAGAACGACCGCTTCTGATCCGGCATAGTAGGCTCTCCGCAAGTTCTTACAAGCGGAATGGCTTTGAGTTACAGAC

2910 2920 2930 2940 2950 2960 2970 2980 2990 3000

HrufAlu-419 AGCAGCATTGGACGGACCCACCAAGAACCCGATAACGAGGACCAAGCAGCAGCTTTATGGCTTGGGTACAGCAGCGGAGCTCGTACAGGAGATCTT

HrufAlu-420 AGCAGCATTGGACGGACCCACCAAGAACCCGATAACGAGGACCAAGCAGCAGCTTTATGGCTTGGGTACAGCAGCGGAGCTCGTACAGGAGATCTT

HrufAlu-421 AGCAGCATTGGACGGACCCACCAAGAACCCGATAACGAGGACCAAGCAGCAGCTTTATGGCTTGGGTACAGCAGCGGAGCTCGTACAGGAGATCTT

HrufAlu-422 AGCAGCATTGGACGGACCCACCAAGAACCCGATAACGAGGACCAAGCAGCAGCTTTATGGCTTGGGTACAGCAGCGGAGCTCGTACAGGAGATCTT

HrufAlu-325 AGCAGCATTGGACGGACCCACCAAGAACCCGATAACGAGGACCAAGCAGCAGCTTTATGGCTTGGGTACAGCAGCGGAGCTCGTACAGGAGATCTT

HrufAlu-328 AGCAGCATTGGACGGACCCACCAAGAACCCGATAACGAGGACCAAGCAGCAGCTTTATGGCTTGGGTACAGCAGCGGAGCTCGTACAGGAGATCTT

HrufAlu-335 AGCAGCATTGGACGGACCCACCAAGAACCCGATAACGAGGACCAAGCAGCAGCTTTATGGCTTGGGTACAGCAGCGGAGCTCGTACAGGAGATCTT

CTATAAAA - 5' HrufA-2

HrufAlu-419 GATATTTT

HrufAlu-420 GATATTTT

HrufAlu-421 GATATTTT

HrufAlu-422 GATATTTT

HrufAlu-325 GATATTTT

HrufAlu-328 GATATTTT

HrufAlu-335 GATATTTT

Figura 13. Alineamiento de las secuencias obtenidas por PCR a partir de los cebadores basados en las secuencias HrufAlu de 811 pb.

Los resultados de PCR parecen indicar que las secuencias HrufAlu-800 pb no se encuentran dispersas por el genoma y que dos de estas secuencias están a su vez separadas entre sí por 400 pb en unos casos o por 3 kb en otros. La secuenciación de los fragmentos amplificados por PCR muestran la existencia de similitud entre ambos tipos de secuencias, ya que la secuencia de 3 kb correspondería a la de 400 pb con una secuencia insertada de aproximadamente 2.6 kb (**Figura 13**).

Así pues se puede deducir que en el genoma de *H. ruficollis* hay dos tipos de secuencias repetidas diferentes, una formada por una secuencia de 3 kb rodeada por dos de 811 y otras de 400 pb rodeadas por dos de 811, que denominaremos HrufAlu-4.6 kb y HrufAlu-2kb, respectivamente. Probablemente el origen de ambas sea el mismo y una se haya originado a partir de la otra por procesos de inserción o delección. Debido a esto el análisis se hará solo de la secuencia de 3 kb.

El alineamiento de las tres secuencias obtenidas (HrufAlu-325, -328 y -335) se deduce una secuencia consenso de 2989 pb. Entre estas secuencias existe una elevada similitud, en torno al 97%. Al igual que ocurre con las secuencias de 811 pb, la variación entre las secuencias de 3 kb son debidas sobre todo a cambios puntuales de nucleótidos. Además es posible observar pequeñas inserciones o delecciones de hasta 4 nucleótidos, como la presente en la secuencia HrufAlu-335 en la posición 2672 y sobre todo una delección de 63 nucleótidos también en esta secuencia en la posición 2132. La secuencia de 3 kb es aún menos rica en AT (46%) que la secuencia de 811 pb y al igual que ocurre con ella presenta secuencias palindrómicas y repeticiones invertidas que le permitirían originar estructuras cruciformes, tal como pone de manifiesto el programa mfold (**Figura 14**).

De los resultados obtenidos hasta el momento podemos concluir que en el genoma de *H. ruficollis* tiene una familia de ADN repetitivo formado por dos tipos de secuencias, HrufAlu-2 kb y HrufAlu-4.6 kb. Estas secuencias son relativamente abundantes en el genoma ya que han sido puestas de manifiesto mediante digestión con enzimas de restricción generando una banda prominente en geles de agarosa (**Figura 5**). Sin embargo los datos no permiten aún determinar cómo se organizan estas secuencias en el genoma, existiendo varias posibilidades, tal y

como se recoge en la **Figura 15**; o bien se trata de un ADN repetitivo disperso o bien están organizadas en tándem, tal como hacen los ADNs satélite. Estos tándems podrían estar formados de tres modos diferentes, en primer lugar podrían estar formados por uno solo de los tipos de secuencias (**Figura 15A**) o estar mezclados, bien de forma alternante (**Figura 15C**) o podrían estar organizados aleatoriamente (**Figura 15D**). Si bien no se puede concluir cual es la organización de estas secuencias, algunos de estos modelos podrían en principio ser descartados a tenor de los resultados obtenidos en la PCR. Como se ha comentado, usando cebadores basados en la secuencia de 811 pb solo se amplifican bandas de 0.4 o 3 kb. Este resultado es compatible con un modelo de secuencias dispersas en el genoma, tal como se expone en la **Figura 15A**.

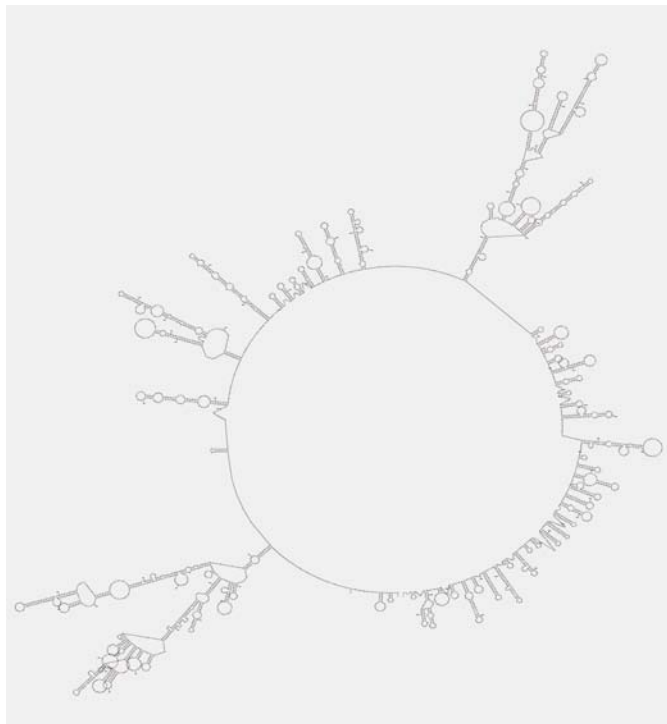


Figura 14. Estructura secundaria de la secuencia de 3 kb obtenida con el programa mfol. Se observa como existen numerosas regiones a lo largo de la secuencia que permiten la generación de estructuras cruciformes.

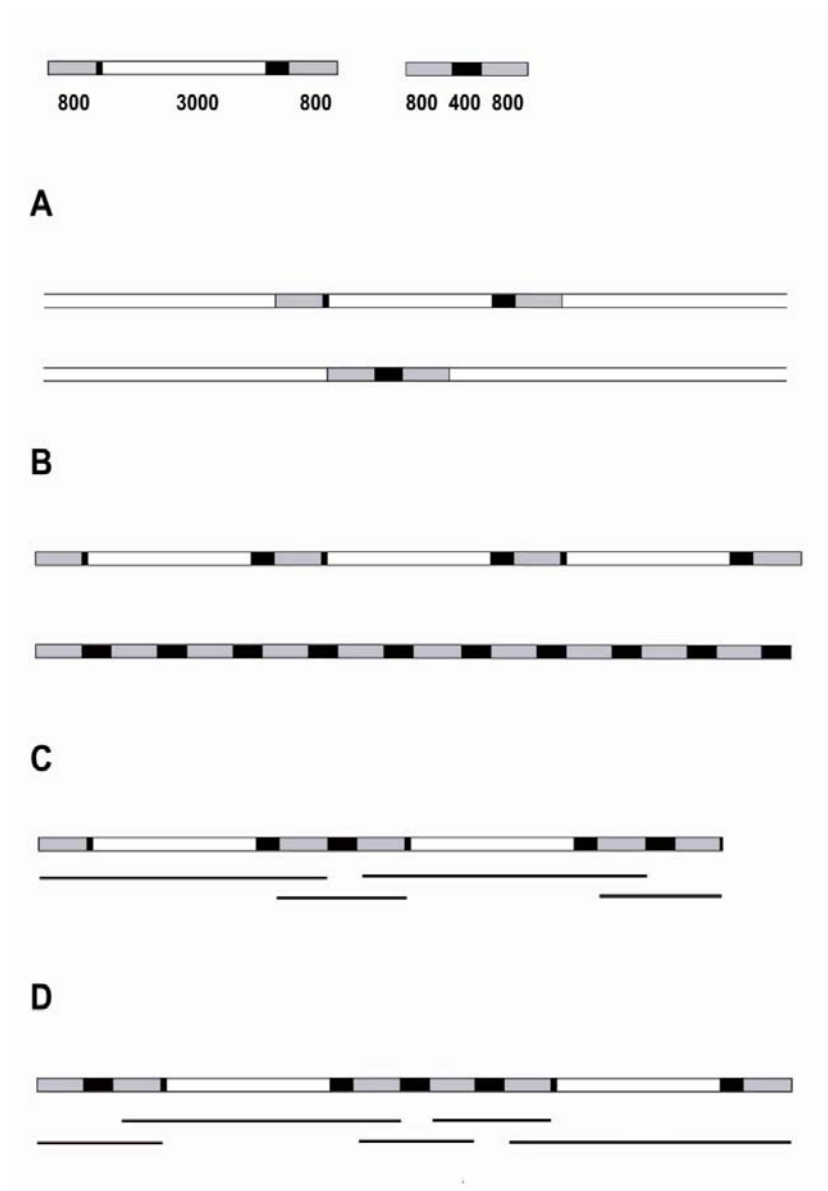


Figura 15. Posible organización de las secuencias HrufAlu-4.6 kb y HrufAlu-2 kb en el genoma de *H. ruficollis*: dispersas por el genoma (A) o bien organizadas en tándem (B, C, D).

En los modelos expuestos en las **Figuras 15B y 15C**, se asume la existencia de dos o más secuencias HrufAlu-2 kb consecutivas. Si esto es así mediante PCR podría generarse además de banda de 400 pb otras bandas cuyo tamaño difiriese en 1.2 kb (**Figura 16**), lo cual no ocurre. La no amplificación de dos secuencias HrufAlu-4.6 kb es esperable ya que el fragmento que se generaría tendría un tamaño de 6.8 kb que posiblemente escapa de la capacidad de amplificación de la Taq polimerasa normal.

El modelo con HrufAlu-2 kb y HrufAlu-4.6 kb alternadas tampoco parece compatible con el resultado de la PCR. La amplificación conjunta de ambas secuencias generaría una banda de 3.4 kb que tampoco aparece en el producto de PCR. Por tanto, y a la espera de nuevos experimentos que permitan aclarar la organización genómica de las secuencias HrufAlu-2 kb y HrufAlu-4.6 kb, el modelo de secuencias dispersas por el genoma parece el modelo más plausible.

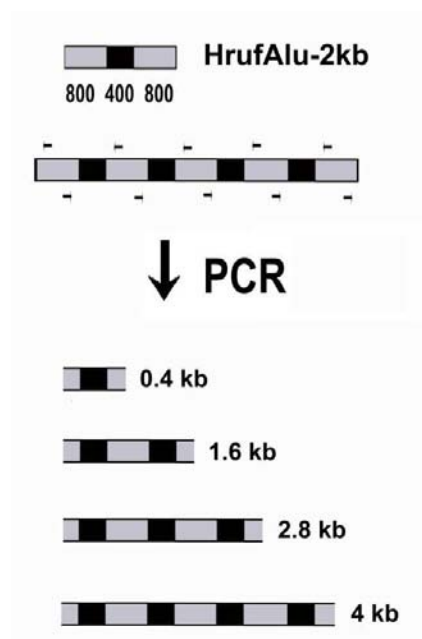


Figura 16. Esquema en el que se indica cual sería el resultado de la PCR en caso de la presencia de dos o más secuencias HrufAlu-2 kb consecutivas (Modelo de la figura 15B, 15D).

Se ha realizado una búsqueda en las bases de datos con el programa BLAST a fin de determinar la posible existencia de secuencias con similitud con las aisladas en *H. ruficollis*. Para ello se ha usado la secuencia formada por el fragmento de 3 kb amplificada por PCR rodeado de dos fragmentos de 811 pb (HrufAlu-4.6 kb). Los resultados muestran la existencia de varias regiones con similitud del 65-71% con una región del genoma de la avispa parasitoide *Cotesia congregata* (GenBank: FM212912.1), que a su vez corresponde a un provirus de un *Bracovirus* presente en el genoma de esta especie (GenBank: HF586473.1) (Bézier et al. 2009, Louis et al 2013). Estos resultados sugieren que la familia de ADN repetitivo HrufAlu estudiado en este trabajo puede tener un origen vírico.

A fin de determinar la naturaleza de las secuencia HrufAlu-4.6 kb se ha usado el CDD (Conserved Domain Database) del NCBI (Marchler-Bauer et al. 2013). A

partir de una secuencia el programa la traduce en las 6 posibles pautas de lectura y las correspondientes secuencias de aminoácidos se comparan con los modelos de la base de datos. Los resultados obtenidos (**Figura 17**) muestran que en una de las pautas de lectura es posible encontrar un dominio propio de una reverso transcriptasa (RT). La presencia de un gen de la RT es generalmente indicativo de un elemento transponible como un retrotransposón o un retrovirus. La organización de la secuencia HrufAlu-4.6 kb hace pensar que pueda tratarse de un LTR-retrotransposón, donde las secuencias de 811 representarían las repeticiones terminales directa y una región central codificante.

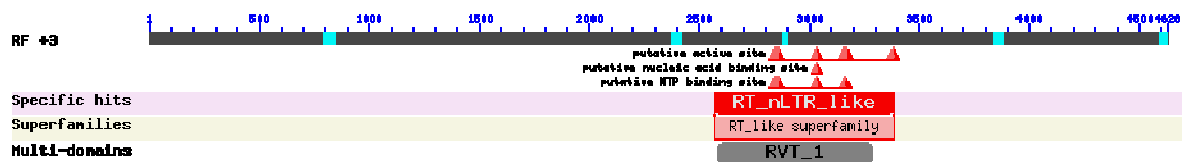


Figura 17. Resultado de la búsqueda de dominios en la secuencia HrufAlu-4.6 kb. Se muestra la existencia de un dominio perteneciente a la familia de las reverso transcriptasas.

Los LTR-retrotransposones son similares a los retrovirus excepto que, generalmente, carecen del gen *env*, si bien puede estar presente en algunos de ellos. Contiene el gen *gag* que codifica para proteínas de la cápsida y el gen *pol* que codifica varios dominios proteicos: proteasa, reverso transcriptasa, ARNasaH e integrasa. Las secuencias HrufAlu-4.6 kb solo conservan el dominio de la RT por lo que todo hace pensar que se trata de un retrotransposón no funcional. Esto no es un resultado inesperado ya que muchos transposones funcionales dan lugar a elementos no autónomos (Wicker et al. 2007). Generalmente los elementos no autónomos han perdido la capacidad codificante por mutaciones. En otros casos contienen algunos genes pero carecen de otros, como ocurre con los elementos BARE2 en plantas de la tribu Triticeae que han perdido por delección el gen *gag* (Tanskanen et al. 2006). La participación de los procesos de delección en la inactivación de este retrotransposón en *H. ruficollis* se ve apoyado por la existencia

de las secuencias HrufAlu-2 kb, donde el proceso de delección ha sido más intenso, conservándose solamente 400 pb entre las repeticiones terminales.

La búsqueda de ORFs en la secuencia HrufAlu-4.6 kb muestra la existencia de tres pequeñas ORFs. Dos de ellas se encuentran dentro de las repeticiones terminales de 811 pb (**Figura 18**). La interna da lugar a una secuencia de 498 aa. Esta secuencia se usó para buscar secuencias similares en la base de datos usando el programa blastp. El resultado obtenido muestra que esta secuencia de aa es la que contiene el motivo de la RT, así como su similitud con otras RT, algunas de ellas pertenecientes a otros insectos como el piojo humano, *Pediculus humanus* (GenBank: XP_002431918.1) o el mosquito *Anopheles gambiae* (GenBank: BAC82595.1).

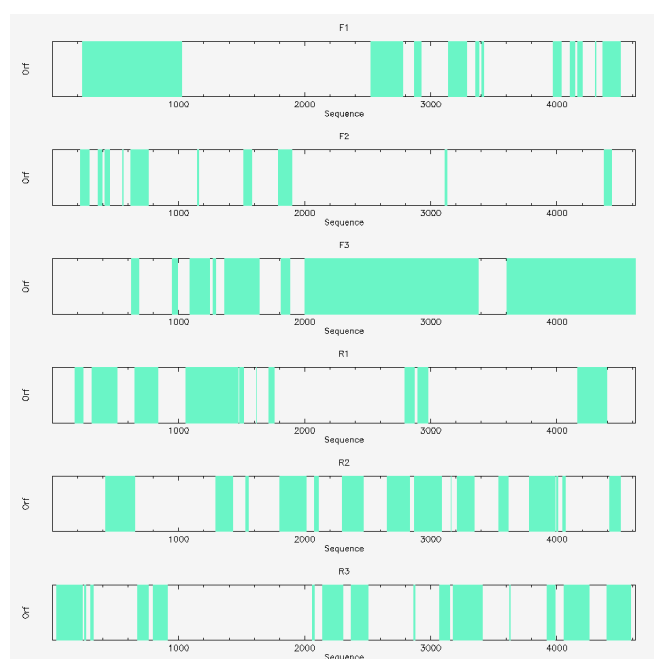


Figura 18. Resultado de la búsqueda de ORFs en la secuencia HrufAlu-4.6 kb. Las más importantes aparecen en la primera y la tercera pauta de lectura de la secuencia (F1 y F3). La ORF en F1 y la segunda de F3 están dentro de las secuencias de 800 pb. La primera ORF de F3 está dentro de la región de 3 kb.

6. BIBLIOGRAFÍA

- Altschul S F, Stephen F, Thomas L, Madden L, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402.
- Arensburger P, Hice RH, Zhou L, Smith RC, Tom AC, Wright JA, Knapp J, O'Brochta DA, Craig NL, Atkinson PW (2011). Phylogenetic and functional characterization of the *hAT* transposon superfamily. *Genetics* 188: 45-57.
- Barceló F, Gutierrez F, Barjau I, Portugal J (1998). A theoretical perusal of the satellite DNA curvature in tenebrionid beetles. *J Biomol Struct Dyn* 16: 41-50.
- Barceló F, Pons J, Petitpierre E, Barjau I, Portugal J (1997). Polymorphic curvature of satellite DNA in three subspecies of the beetle *Pimelia sparsa*. *Eur J Biochem* 244: 318-324.
- Beeman RW, Thomson MS, Clark JM, DeCamillis MA, Brown SJ, Denell RE (1996). Woot, an active *gypsy*-class retrotransposon in the flour beetle, *Tribolium castaneum*, is associated with a recent mutation. *Genetics* 143: 417-426.
- Benfante R, Landsberger N, Maiorano D, Badaracco G (1990). A binding protein (*p82 protein*) recognizes specifically the curved heterochromatic DNA in *Artemia franciscana*. *Gene* 94: 217-222.
- Bézier A, Annaheim M, Herbinière J, Wetterwald C, Gyapay G, Bernard-Samain S, Wincker P, Roditi I, Heller M, Belghazi M, Pfister-Wilhem R, Periquet G, Dupuy C, Huguet E, Volkoff AN, Lanzrein B, Drezen JM (2009). Polydnviruses of braconid wasps derive from an ancestral nudivirus. *Science* 323: 926-930.
- Bigot Y, Hamelin MH, Periquet G (1990) Heterochromatin condensation and evolution of unique satellite-DNA families in two parasitic wasp species: *Diadromus pulchellus* and *Eupelmus vuilleti* (Hymenoptera). *Mol Biol Evol.* 7: 351-364.
- Booth RG, Cox ML, Madge RB (1990). Guides to insects of importance to man 3. Coleoptera. International Institute of Entomology. The Natural History Museum of London, London.
- Braquart C, Royer V, Bouhin H. (1999). DEC: a new miniature inverted-repeat transposable element from the genome of the beetle *Tenebrio molitor*. *Insect Mol Biol* 8: 571-574.
- Brázda V, Laister RC, Jagelská EB, Arrowsmith C (2011). Cruciform structures are a common DNA feature important for regulating biological processes. *BMC Mol Biol* 12: 33.
- Britten RJ, Davidson EH (1971). Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. *Q Rev Biol* 46: 111-138.

- Brown TA (2002). The repetitive DNA content of genomes. En "Genomes", 2^a ed. Oxford: Wiley-Liss. pp 59-64.
- Bruford MW, Wayne RK (1993). Microsatellites and their application to population genetic studies. *Curr Opin Genet Dev* 3: 939-943.
- Bruvo B, Plohl M, Ugarkovic D (1995). Uniform distribution of satellite DNA variants on the chromosomes of tenebrionid species *Alphitobius diaperinus* and *Tenebrio molitor*. *Hereditas* 123: 69-75.
- Bruvo-Madaric B, Plohl M, Ugarkovic D (2007). Wide distribution of related satellite DNA families within the genus *Pimelia* (Tenebrionidae). *Genetica* 130: 35-42.
- Burgtorf C, Bünemann H (1993). A telomere-like satellite (GGGTCAT)_n comprises 4% of genomic DNA of *Drosophila hydei* and is located mainly in centromeric heterochromatin of all large acrocentric autosomes. *Gene* 137: 287-291.
- Cavalier-Smith T (1985). Eukaryotic gene numbers, non-Coding DNA, and genome size. En "The evolution of genome size." T. Cavalier-Smith, John Wiley and Sons (Ed). Chichester, U.K. pp. 69–103.
- Charlesworth B, Sniegowski P, Stephan W (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371: 215-220.
- Duncan L, Bouckaert K, Yeh F, Kirk DL (2002). Kangaroo, a mobile element from *Volvox carteri*, is a member of a newly recognised third class of retrotransposons. *Genetics* 162: 1617–1630.
- Edwards A, Civitello A, Hammond HA, Caskey CT (1991). DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am J Hum Genet* 49: 746–756.
- Eickbush T. (1992). Transposing without ends: the non-LTR retrotransposable elements. *New Biol* 4: 430-440.
- Elgar G, Vavouri T (2008). Tuning in to the signals: noncoding sequence conservation in vertebrate genomes. *Trends Genet* 24: 344–352.
- Fattorini S, Ulrich W (2012). Spatial distributions of European Tenebrionidae point to multiple postglacial colonization trajectories. *Biol J Linn Soc* 105: 318-329.
- Finnegan D (1992). Transposable elements. *Curr Opin Genet Dev* 2: 861-867.
- Fry K, Salser W (1977). Nucleotide sequences of HS-alpha satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell* 12: 1069-1084.
- Galián J, Vogler AP (2003). Evolutionary dynamics of a satellite DNA in the tiger beetle species pair *Cicindela campestris* and *C. maroccana*. *Genome* 46: 213-223.
- Goodwin TJ, Poulter RT, Lorenzen MD, Beeman RW. (2004). DIRS retroelements in arthropods: identification of the recently active TcDir1 element in the red flour beetle *Tribolium castaneum*. *Mol Genet Genomics* 272: 47-56.

- Gur-Arie R, Cohen CJ, Eitan Y, Shelef L, Hallerman EM, Kashi Y (2000) Simple sequence repeats in *Escherichia coli*: abundance, distribution, composition, and polymorphism. *Genome Res* 10: 62–71.
- Hartley SE, Davidson WS (1994). Characterization and distribution of genomic repeat sequences from arctic char (*Salvelinus alpinus*). En "Genetics and Evolution of Aquatic Organisms." Beamont AR (Ed). Chapman and Hall, London. pp 271-279.
- Heinze J, Gadau J, Hölldobler B, Nanda I, Schmid M, Scheller K (1994). Genetic variability in the ant *Camponotus floridanus* detected by multilocus fingerprinting. *Naturwissenschaften* 81: 34-36.
- Herrera J (1988). Pollination relationships in Southern Spanish Mediterranean shrublands. *J Ecol* 76: 274-287.
- Jonstrup A T, Thomsen T, Wang Y, Knudsen BR, Koch J, Andersen AH (2008). Hairpin structures formed by alpha satellite DNA of human centromeres are cleaved by human topoisomerase IIa. *Nucleic Acids Res* 36: 6165–6174.
- Juan C, Pons J, Petitpierre E (1993). Localization of tandemly repeated DNA sequences in beetle chromosomes by fluorescent in situ hybridization. *Chromosome Res* 1: 167-174.
- Kit S (1961). Equilibrium sedimentation in density gradients of DNA preparations from animal tissues. *J Mol Biol* 3: 711-716.
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. (2001) Initial sequencing and analysis of the human genome. *Nature* 409: 860-921.
- Lorite P, Torres MI, Palomeque T (2013). Characterization of two unrelated satellite DNA families in the Colorado potato beetle *Leptinotarsa decemlineata* (Coleoptera, Chrysomelidae). *Bull Entomol Res* 103: 538-546.
- Louis F, Bézier A, Periquet G, Ferras C, Drezen JM, Dupuy C (2013). The Bracovirus genome of the parasitoid wasp *Cotesia congregata* is amplified within 13 replication units, including sequences not packaged in the particles. *J Virol* 87: 9649-9660.
- Marchler-Bauer A, Zheng C, Chitsaz F, Derbyshire MK, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, Zhang D, Bryant SH (2013). CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Res* 41(Database issue): D348-D352.
- Mestrovic N, Plohl M, Mravinac B, Ugarkovic D (1998) Evolution of satellite DNAs from the genus *Palorus* – Experimental evidence for the library hypothesis. *Mol Biol Evol* 15: 1062–1068.
- Miklos GL, Gill AC (1982). Nucleotide sequences of highly repeated DNAs; compilation and comments. *Genet Res* 39: 1-30.

- Miller WJ, Capy P (2004). Mobile genetic elements as natural tools for genome evolution. *Meth Mol Biol* 260: 1-20
- Mravinac B, Plohl M (2007). Satellite DNA junctions identify the potential origin of new repetitive elements in the beetle *Tribolium madens*. *Gene* 394: 45–52.
- Mravinac B, Plohl M (2010). Parallelism in evolution of highly repetitive DNAs in sibling species. *Mol Biol Evol* 27: 1857–1867.
- Mravinac B, Plohl M, Mestrovic N, Ugarkovic D (2002). Sequence of PRAT satellite DNA "frozen" in some coleopteran species. *J Mol Evol* 54: 774–783.
- Mravinac B, Plohl M, Ugarkovic D (2004). Conserved patterns in the evolution of *Tribolium* satellite DNAs. *Gene* 332: 169-177.
- Mravinac B, Ugarkovic D, Franjevic D, Plohl M (2005). Long inversely oriented subunits form a complex monomer of *Tribolium brevicornis* satellite DNA. *J Mol Evol* 60: 513–525.
- Oromí P (1982). En "Los tenebriónidos de las Islas Canarias". Instituto de Estudios Canarios. Santa Cruz de Tenerife.
- Palomeque T, Lorite P (2008). Satellite DNA in insects: a review. *Heredity* 100: 564–573.
- Palomeque T, Muñoz-López M, Carrillo JA, Lorite P (2005). Characterization and evolutionary dynamics of a complex family of satellite DNA in the leaf beetle *Chrysolina carnifex* (Coleoptera, Chrysomelidae). *Chromosome Res* 13: 795-807.
- Pathak D, Ali S (2012). Repetitive DNA: a tool to explore animal genomes/transcriptomes. En "Functional Genomics", Meroni G, Petrera F (Eds). pp 155-280. <http://www.intechopen.com/books/functional-genomics/repetitive-dna-a-tool-to-explore-animal-genomes-transcriptomes>
- Patience C, Takeuchi Y, Weiss RA (1997). Infection of human cells by an endogenous retrovirus of pigs. *Nat Med* 3: 276-282.
- Pearson WR, Lipman DJ (1998). Improved tools for biological sequence comparison. *Proc Natl Acad Sci USA* 85: 2444-2448.
- Petitpierre E, Gatewood JM, Schmid CW (1988). Satellite DNA from the beetle *Tenebrio molitor*. *Experientia* 44: 498-499.
- Plasterk, RHA, Van Luenen HG (2002). The Tc1/mariner family of transposable elements. En "Mobile DNA II" Craig, N.L., Craigie, R., Gellert, M. y Lambowitz, A.M. (Eds). American Society for Microbiology Press, Washington DC. pp 519-532.
- Plohl M, Borstnik B, Lucijanac-Justic V, Ugarkovic D (1992) Evidence for random distribution of sequence variants in *Tenebrio molitor* satellite DNA. *Genet Res* 60: 7–13.
- Plohl M, Borstnik B, Ugarkovic D, Gamulin V (1990). Sequence-induced curvature of *Tenebrio molitor* satellite DNA. *Biochimie* 72: 665–670.

- Plohl M, Bruvo B, Mestrovic N, Mravinac B, Pet-Rovic V, Durajlija-Zinic S, Ugarkovic D (2004). Satellite DNA sequences in centromeric heterochromatin. *Period Biol* 106: 95–102.
- Plohl M, Mestrovic N, Bruvo B, Ugarkovic D (1998). Similarity of structural features and evolution of satellite DNAs from *Palorus subdepressus* (Coleoptera) and related species. *J Mol Evol* 46: 234-239.
- Plohl M, Ugarkovic D (1994). Characterization of two abundant satellite DNAs from the Mealworm *Tenebrio obscurus*. *J Mol Evol* 39: 489-495.
- Plohl M (2010). Those mysterious sequences of satellite DNAs. *Period Biol* Vol. 112, No 4: 403–410.
- Pons J (2004). Cloning and characterization of a transposable-like repeat in the heterochromatin of the darkling beetle *Misolampus goudoti*. *Genome* 47: 769–774.
- Pons J, Bruvo B, Juan C, Petitpierre E, Plohl M, Ugarkovic D (1997). Conservation of satellite DNA in species of the genus *Pimelia* (Tenebrionidae, Coleoptera). *Gene* 205: 183-190.
- Pons J, Bruvo B, Petitpierre E, Plohl M, Ugarkovic D, Juan C (2004). Complex structural features of satellite DNA sequences in the genus *Pimelia* (Coleoptera: Tenebrionidae): random differential amplification from a common 'satellite DNA library'. *Heredity* 92: 418-427.
- Pons J, Petitpierre E, Juan C (1993). Characterization of the heterochromatin of the darkling beetle *Misolampus goudoti*: cloning of two satellite DNA families and digestion of chromosomes with restriction enzymes. *Heredity* 119: 179-185.
- Pons J, Petitpierre E, Juan C (2002). Evolutionary dynamics of satellite DNA family PIM357 in species of the genus *Pimelia* (Tenebrionidae, Coleoptera). *Mol Biol Evol* 19: 1329-1340.
- Reed KM, Phillips RB (1995). Molecular characterization and cytogenetics analysis of highly repeated DNAs of lake trout, *Salvelinus namaycush*. *Chromosoma* 104: 242-251.
- Richards S, Gibbs RA, Weinstock GM, Brown SJ, Denell R, Beeman RW, et al. (2008). The genome of the model beetle and pest *Tribolium castaneum*. *Nature* 452: 949-955.
- Rozas J, Rozas R (1999). DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15: 174-175.
- Sambrook J, Fritsch EF, Maniatis T (1989). *Molecular cloning: A laboratory manual*. Cold Spring Harbour Laboratory.
- Saitoh Y, Miyamoto N, Okada T, Gondo Y, Showguchi-Miyata J, Hadano S, Ikeda JE (2000). The RS447 human megasatellite tandem repetitive sequence encodes a novel deubiquitinating enzyme with a functional promoter. *Genomics* 67: 291-300.
- Slamovits CH, Rossi MS (2002). Satellite DNA: agent of chromosomal evolution in mammals. *Mastozoología Neotropical* 9: 297-308.

- Strachan T, Webb D, Dover GA (1985). Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. EMBO J 4: 1701–1708.
- Tanskanen JA, Sabot F, Vicient C, Schulman AH (2006). Life without GAG: The *BARE-2* retrotransposon as a parasite's parasite. Gene 390: 166–174.
- Tautz D (1993). Notes on the definition and nomenclature of tandemly repetitive DNA sequences. En "DNA Fingerprinting: State of the science". Pena SD (Ed.) Birkhäuser Basel pp. 21–28.
- Toth G, Gaspari Z, Jurka J (2000). Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res 10: 967–981.
- Ugarkovic D, Durajlija S, Plohl M (1996). Evolution of *Tribolium madens* (Insecta, Coleoptera) satellite DNA through DNA inversion and insertion. J Mol Evol 42: 350–358.
- Ugarkovic D, Plohl M (2002). Variation in satellite DNA profiles-causes and effects. EMBO J 21: 5955-5959.
- Ugarkovic D, Podnar M, Plohl M (1996). Satellite DNA of the red flour beetle *Tribolium castaneum* – comparative study of satellites from the genus *Tribolium*. Mol Biol Evol 13: 1059–1066.
- Venter JC et al. (2001). The sequence of the human genome. Science 291: 1304-1351.
- Wang J, Du Y, Wang S, Brown SJ, Park Y (2008b). Large diversity of the piggyBac-like elements in the genome of *Tribolium castaneum*. Insect Biochem Mol Biol 38: 490-498.
- Wang S, Lorenzen MD, Beeman RW, Brown SJ (2008a). Analysis of repetitive DNA distribution patterns in the *Tribolium castaneum* genome. Genome Biology 9: R61.
- Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007). A unified classification system for eukaryotic transposable elements. Nat Rev Genet. 8: 973-982.
- Wilfinger WW, Mackey K, Chomczynski P (1997) Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. Biotechniques 22: 474-476, 478-481.
- Zuker M (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406-3415.