

ARTÍCULOS DE REVISIÓN

Elementos móviles: ¿Ventaja evolutiva o parasitismo molecular?

Mobile elements: Positive evolution or molecular parasitism?

LÓPEZ, M.C.^{1*}; OLIVARES, M.¹; GONZÁLEZ, C.I.^{1#}; MARTÍN, F.¹; ;
GARCÍA PÉREZ, J.L. ¹ Y THOMAS, M.C¹.

(¹) Instituto de Parasitología y Biomedicina “López Neyra”, Consejo Superior de Investigaciones Científicas. Granada. España

(¹) Department of Immunology, Windeyer Institute of Medical Sciences. University College, London

([#]) Departamento de Microbiología. Universidad Salud, Bucaramanga, Colombia

(*) Correspondencia: Instituto de Parasitología y Biomedicina “López Neyra”, Consejo Superior de Investigaciones Científicas. Calle Ventanilla nº 11, 18001 Granada, España. Teléfono, 958 805190. Fax: 958 203323.

e-mail: mclopez@ipb.csic.es

RESUMEN

Las secuencias de ADN repetidas-dispersas constituyen una elevada proporción del genoma de organismos eucariotas y procariotas. La primera descripción de movilidad de estas secuencias fue realizada por Barbara McClintock en los años 50, la cual, estudiando la herencia del color y distribución de la pigmentación del maíz, demostró la existencia de elementos genéticos con capacidad de transponerse. En el presente trabajo se realiza una revisión de las características moleculares, filogenéticas y funcionales de los elementos móviles, centrada especialmente en los elementos LINE. El análisis de la presencia de elementos móviles en los organismos hospedadores, muestra que estos evolucionan coincidentemente con el hospedador para evitar o mitigar el efecto deletéreo de su inserción e incluso para proporcionarle beneficio. Actualmente, se está empezando a considerar que estos elementos juegan un importante papel en la evolución de los organismos, probablemente como resultado del desarrollo de una relación simbiótica con su hospedador.

PALABRAS CLAVES: ADN repetido-disperso, elementos LINE, transcripción, transposición, integración, estructura, papel biológico.

ABSTRACT

The repetitive-interspersed DNA sequences constitute a large part of the genome of eukaryotic and prokaryotic organisms. The mobility of these sequences was first described by Barbara MacClintock in the fifties who, by studying the inheritance of the colour and distribution of the pigmentation in maize, showed the existence of genetic elements with the ability to transpose along the genome. In the present study we review some of the molecular, phylogenetic and functional characteristics of these mobile elements with special emphasis on LINE elements. The analysis of the presence of mobile elements in host organisms indicates that they co-evolve with the host to avoid or mitigate the deleterious effect of their insertion and even to provide potential beneficial features. Presently, these elements are beginning to be considered as playing important roles in the evolution of the organisms within a symbiotic framework.

KEY WORDS: repetitive-interspersed DNA, LINE elements, transcription, transposition, integration, structure, biological role.

Recibido: 26-1-1999

Aceptado: 15-2-1999

Bibli [0004-2927(1999) 40:1, 5-25]

1.- INTRODUCCIÓN

Dado que las secuencias de ADN repetidas-dispersas constituyen una elevada proporción del genoma de organismos eucariotas (Jelinek, 1982) y procariotas (Lupski y Weinstock, 1992), es fácil postular para las mismas un papel universal, a pesar de que sus implicaciones funcionales no son bien conocidas. Mientras que fenómenos de recombinación justifican la existencia de ordenamientos en tandem, la presencia de las secuencias repetidas-dispersas, o elementos transponibles se explica por mecanismos de transposición y retrotransposición (Eickbush, 1992; Charlesworth y col. 1994). Una poderosa fuente de cambio tanto en los genomas de eucariotas como de procariotas está dada por los referidos elementos transponibles, genéricamente llamados elementos móviles (Weiner y col., 1986; Schmid y Maraia, 1992; Okada y Ohshima, 1995). A diferencia de otros procesos involucrados en la reestructuración del genoma, la transposición no cuenta con ninguna relación entre las secuencias donadoras y receptoras, ya que los elementos transponibles se limitan a moverse por sí mismos, o con la ayuda de secuencias auxiliares del propio genoma.

La primera descripción de la existencia de elementos de ADN móviles fue realizada por Barbara McClintock al inicio de los años 50 (McClintock, 1951). Estudiando la herencia del color y distribución de la pigmentación del maíz demostró la existencia de elementos genéticos que pueden transponerse. McClintock observó que la regulación de la activación y desactivación de ciertos genes implicados en la coloración del maíz, podía alterarse como consecuencia de la acción de distintas unidades genéticas con capacidad de transposición, a las cuales llamó *elementos controladores*. Desde ese momento un extenso número de elementos móviles, con capacidad de trasladarse de un lugar a otro del genoma que hospedan, han sido descritos tanto en organismos procariotas como en eucariotas. Así, estas secuencias llegan a ocupar el 50% del genoma del maíz (SanMiguel y col, 1996), el 35% del genoma de humanos, o el 10% del genoma de *Drosophila* localizándose principalmente en regiones de la heterocromatina (Pimpinelli y col, 1995).

En *Trypanosoma cruzi*, agente causante de la enfermedad de Chagas la cual afecta a unos 18

millones de personas de Centro y Sudamérica, hecho evidenciado por estudios de cinética de reasociación de ADN genómico, las secuencias de ADN repetidas-dispersas representan hasta el 30% de su genoma (Castro y col., 1981; Lanar y col., 1981). Basándose en la estructura y organización de tales secuencias repetidas, éstas se han relacionado con la plasticidad del genoma de *T. cruzi*, parásito en el que se han descrito tanto secuencias repetidas en tandem (mini- y microsatélites), como elementos dispersos (transposones y retrotransposones) (Donelson, 1996), habiéndose sugerido que los mismos pueden estar implicados en fenómenos de reordenamiento y control de la expresión de génica (Hasan y col. 1984; Murphy y col. 1987; Requena y col 1994). La primera secuencia de ADN repetida descrita en *T. cruzi* fue el ADN satélite (Sloof y col., 1983; Frash y col., 1983; González y col., 1984), secuencia de 195 bp presente en unas 120.000 copias, distribuidas por la mayoría de los cromosomas. Además, el genoma de *T. cruzi* contiene varias familias de secuencias repetidas-dispersas que comparten características comunes con las secuencias SINE (*elemento corto de ADN dispersos*) de eucariotas. Una de estas secuencias es el elemento E13 de 1 kb (Requena y col., 1992), el cual está repetido unas 10^5 veces en el genoma, representando aproximadamente el 7% del ADN nuclear del parásito. Se han encontrado homologías de E13 con otras regiones del genoma del parásito, tales como con el pseudogen 24S α DNAr (Vieira de Arruda y col., 1990), con zonas intergénicas del gen gp72 (Cooper y col., 1991), o del gen H2A (Puerta y col., 1994). Otro elemento repetido de *T. cruzi* es el denominado SIRE de 428 bp, el cual se encontró insertado en la región intergénica del gen P2b (Vazquez y col., 1994). Este elemento SIRE presenta además una homología del 95% con las primeras 122 bp del elemento E13, así como con las regiones 3'UTR de algunas de las unidades génicas que conforman el locus H2A de *T. cruzi* (Marañón y col., 1998). Otras dos putativas secuencias SINE altamente repetidas en *T. cruzi* son las llamadas E12 y E22 (Requena y col., 1993). El número de copias por genoma de E12 y E22 es de 5600 y 7200, respectivamente. Interesantemente, se observó que el elemento E12 se encontraba formando parte del extremo 3' del mensajero correspon-

diente a una nueva secuencia repetida-dispersa de tipo LINE (*elemento largo de ADN disperso*) (Requena y col., 1994), la cual se denominó L1Tc (Martín y col., 1995). Este elemento LINE de aproximadamente 5 kb, se transcribe activamente y se encuentra repetido en el genoma de *T. cruzi* unas 2.500 veces (Martín y col., 1995). Su extremo 5' presenta una homología significativa de secuencia con el elemento ribosomal RIME, descrito por Hasan y col., (1984), el cual asimismo se encontró asociado al retrotransposón ingi o TRS-1 de *T. brucei* (Kimmel y col., 1987; Murphy y col., 1987). Se ha sugerido que las secuencias RIME encontradas en los genes VSG (*glicoproteína variable de superficie*) activos de *T. brucei*, pero no en los silentes, pueden condicionar la activación de potenciales sitios de expresión (Pays y col., 1989).

Durante décadas, las secuencias repetidas-dispersas se han calificado como elementos egoístas, en el sentido de que se mantienen en el genoma solamente por su posibilidad de replicarse junto a él. Sin embargo, estos elementos repetidos con capacidad para duplicarse, para moverse de un lugar a otro del genoma y para provocar reordenamientos, posibilitan sin duda alguna, una

importante flexibilidad y capacidad evolutiva a los genomas hospederos. Se ha demostrado que estas secuencias interaccionan activamente con el genoma que hospedan y pueden desempeñar diversas funciones, entre las que se incluye la regulación de la expresión de genes. Así, en *Drosophila* los retrotransposones HeTA y TART son responsables de mantener la integridad de los telómeros (Jurka, 1998). Numerosos ejemplos muestran que la inserción de elementos transponibles actúa como una fuente de variación que conduce a cambios útiles positivamente seleccionados, que se mantienen a lo largo de la evolución. Es el caso de las secuencias Alu de humanos que intervienen en la regulación de la transcripción génica de diversos genes (Schmid, 1998), o el caso referido anteriormente del elemento SIRE, cuya inserción en la región 3' no traducida del gen que codifica para la proteína H2A de *T. cruzi* da lugar a un nuevo mensajero de mayor tamaño, el cual es regulado post-transcripcionalmente de forma más eficiente que el mensajero que no lo contiene (Marañón y col., 1998; Marañón y col., 1999).

2.- CLASIFICACIÓN DE LOS ELEMENTOS TRANSPONIBLES

La clasificación de los elementos transponibles ha estado basada en el tipo de hospedador (procariótico o eucariótico), en las estructuras terminales que los flanquean (repeticiones terminales invertidas o repeticiones terminales directas) y en la similitud de secuencias a genes homólogos, como transposasas. En general se pueden distinguir, de acuerdo a sus diferentes modos de transposición, dos tipos principales de elementos (Finnegan, 1992): los elementos de clase I, que se transponen por transcripción inversa de un ARN intermediario (mediante un mecanismo ADN-ARN-ADN), tales como los retrotransposones LTR: copia de *Drosophila melanogaster* (Mount, 1985), y Ty en *Saccharomyces cerevisiae* (Clare, 1985) o el elemento LINE ingi en *T. brucei* (Kimmel, 1987); y los elementos de clase II, tales como los elementos P de *D. melanogaster* (Engels, 1989) o Ac/Ds de *Zea mays* (Kunce, 1996), que se transponen directamente de ADN a ADN. Recientemente se ha descrito un nuevo tipo de elementos transponibles denominados MITEs, los cuales comparten propiedades de las dos anteriormente

citadas clases de elementos (Wessler y col., 1995). Estos elementos transponibles MITE, con repeticiones invertidas mínimas, son secuencias cortas (100-400 pb) a las que no se les ha encontrado capacidad codificante. Están presentes en un alto número de copias (3000 – 10000) por genoma, y en plantas se ha observado que se insertan preferentemente en las secuencias TAA o TA. Los elementos MITEs están asociados con frecuencia a las regiones reguladoras de los genes de distintas plantas florales. El elemento Tourist en el maíz y el elemento Stowaway en el sorgo (Wessler y col., 1995) se encuentran localizados, con una significativa mayor frecuencia, en las regiones 5' y 3' no codificantes de genes de dichas plantas. También se han descrito, elementos transponibles con propiedades similares en *Xenopus* (Unsal y Morgan, 1995), humanos (Morgan, 1995; Smit y Rigas, 1996) y en el mosquito transmisor de la fiebre amarilla, *Aedes aegypti* (Tu, 1997), sin embargo la mayoría de los elementos MITE presentes en genomas de eucariotas son defectivos y han perdido la capacidad de transponerse independientemente, aunque pueden ser reconocidos

como substratos para transposición por las enzimas producidas por los transposones funcionales del mismo grupo.

Dentro de los elementos de clase I o retroposones (Rogers, 1985; Weiner y col, 1986), en los que el flujo de información genética, denominado retroposición, ocurre de ARN a ADN, se distinguen dos grupos de elementos móviles. Un primer grupo que incluye a aquellos que no tienen capacidad de codificar las enzimas necesarias para su transposición, donde encontramos los elementos SINE, y un segundo grupo denominado retrotransposones, con capacidad para codificar las referidas enzimas, tales como la transcriptasa inversa, que utilizan para la replicación de sus propias secuencias. Los elementos SINE fueron en principio detectados gracias a los estudios de renaturalización del ADN cromosómico, los cuales mostraban que la mayoría de los organismos poseían secuencias cortas (menores de 1 kb), moderadamente repetidas, separadas por largos fragmentos de ADN de varias kilobases (Jelinek, 1982). En 1982 Singer encuadró como secuencias SINE aquellas que son típicamente menores de 500 pares de bases, se repiten unas 10^5 veces y se encuentran dispersas a lo largo del genoma que hospedan. Los ejemplos más representativos de este tipo de secuencias son las familias Alu de humanos y B1 de roedores. Otros grupos de secuencias SINE importantes son, entre otras, las ID de ratas, B2 de ratones y las repeticiones C de conejo. Las características más destacadas de las secuencias SINE son: a) la presencia de un promotor interno funcional capaz de llevar a cabo la transcripción del elemento completo mediante la ARN polimerasa III (Jelinek, 1982; Rogers, 1985); b) las duplicaciones directas de 4-10 nucleótidos encontradas en ambos extremos del elemento; y c) la presencia de una cola de poli (A) en el extremo 3'. Todas estas características reflejan el hecho de que la práctica totalidad de las secuencias SINE son retroposones, es decir son elementos cuyo origen es una molécula de ARN procesada, copiada a ADN complementario e integrada de nuevo en el genoma. Sin embargo, tal como referimos anteriormente, al ser secuencias pequeñas no contienen la capacidad para codificar las proteínas implicadas en el proceso de transposición, así que éstas han de ser suplidas en *trans*, probablemente por la maquinaria enzimática de los elementos LINE que cohospedan en el

genoma (Eickbush, 1992; Okada y col., 1997). Recientes estudios han mostrado que las secuencias de inserción del genoma humano, generadas por el elemento SINE Alu y por el elemento LINE L1, son mayoritariamente idénticas, lo que constituye un fuerte soporte de que la maquinaria del elemento L1 es la utilizada para la transposición del elemento Alu, e indicando que la integración de estos elementos no ocurre al azar (Jurka, 1997; Cost y Boeke, 1998).

El origen de estos elementos se encuentra en tránscritos de la ARN polimerasa III, tales como el ARN 7SL y ARNts (Weiner, 1986). De hecho, la secuencia de los elementos de la familia Alu de humanos y B1 de roedores son prácticamente idénticas a la del ARN 7SL, variando únicamente en delecciones internas tanto en Alu como en B1. Además, la mayoría de las secuencias SINE, tales como las familias ID, B2 y C, parecen tener como origen diversos ARNts. Esto al menos es lo que parecen demostrar las homologías detectadas entre ambos grupos de elementos (Rogers, 1985). Por lo tanto, las secuencias SINE se pueden considerar como resultado de un proceso evolutivo en el que los tránscritos de ARNts o ARN 7SL (o algunos otros ARNs pequeños), fueron copiados por una transcriptasa inversa a ADN complementario e integrados en un lugar diferente del genoma, utilizando algún mecanismo aún no bien conocido. Dado que el promotor es interno, la nueva copia del gen tendría una independencia transcripcional relativamente importante. Nuevas duplicaciones, siguiendo este mecanismo, se pudieron ir produciendo siempre y cuando las inserciones fueran neutrales o ventajosas para el organismo en cuestión. Dichas duplicaciones sólo requerirían la transcripción del elemento a ARN y que se mantuvieran las secuencias reconocidas por las proteínas involucradas en la transcripción inversa e integración, dando lugar por tanto a que las mutaciones se vayan acumulando en las zonas irrelevantes para la transposición. Como resultado de este proceso, las secuencias de ADN repetidas-dispersas solamente conservan (en la mayoría de los casos) la región promotora y algunas otras regiones de los genes de donde provienen. Esta es la razón de que solo recientemente se hayan detectado homologías entre elementos SINE y determinados ARNts, si bien la similitud entre el promotor de los elementos SINE y de los ARNts fue obvia desde mucho

antes (Rogers, 1985).

El otro grupo importante de secuencias de ADN repetidas-dispersas lo constituyen los retrotransposones, que son elementos móviles que teóricamente tienen la capacidad de codificar las proteínas necesarias para su propia transposición (Boeke, 1989). Estos elementos, basándose en su estructura, pueden ser divididos en retrotransposones LTR (*terminaciones largas repetidas*) y en retrotransposones no-LTR o elementos LINE. Los elementos LINE, fueron descritos como secuencias de ADN de un tamaño de aproximadamente 5 kb de longitud y presentes en un número de copias mayor de 10^4 veces por genoma. Inicialmente se establecieron diferentes grupos, tales como la familia KpnI de primates (Adams, 1980) o la MIF-1 de roedores (Cheng, 1980). Posterior análisis mostraron que en realidad estos dos grupos de elementos eran en realidad miembros de una única familia (Singer, 1983), que estaba ampliamente representada en mamíferos (Burton, 1986). A esta familia se le denominó LINE-1 o L1 (Rogers, 1985). De forma genérica se propuso una terminación de dos letras para especificar el origen del elemento L1, indicando género y especie. A partir de entonces se han caracterizado elementos con estructura similar a la familia L1 de mamíferos en organismos tan distantes como

insectos, protozoos, hongos, aves, plantas, anfibios y moluscos (Hutchison, 1989; Xiong, 1990). La denominación de estos elementos, ha sido a veces discutida, dado que el nombre alude únicamente al hecho de que sean secuencias largas y que se encuentren repetidas y dispersas. Sin embargo, se han descrito en el genoma de diferentes organismos la presencia de secuencias con gran homología a los elementos LINE, tanto en estructura como en secuencia, pero que se encuentran en bajo número de copias e incluso son de localización específica. Por ello, actualmente se acepta la clasificación de un elemento como perteneciente a esta familia de elementos LINE, en función de poseer propiedades comunes encontradas en todos los elementos (Hutchison, 1989). Estas propiedades serían la de ser elementos móviles activos capaces de codificar las proteínas involucradas en su propio proceso de transposición, al menos la actividad transcriptasa inversa (RT), poseer en el extremo 3' una cola de poli (A), contener repeticiones directas flanqueantes de entre 5-10 nucleótidos, ser activamente transcritos y presentarse como ARN maduros.

3. ESTRUCTURA DE LOS RETROTRANSPOSONES

Los retrotransposones LTR, como los elementos copia de *Drosophila*, Ty de levadura y elemento THE de humanos, son muy similares a retrovirus. Presentan en sus extremos 5' y 3' largas repeticiones terminales (LTR) que contienen señales para la iniciación y terminación de la transcripción del ARN intermediario, jugando un papel relevante en el mecanismo de transposición de los mismos (Boeke y Corces, 1989; Sandmeyer y col., 1990). Los retrotransposones LTR contienen normalmente 2 marcos abiertos de lectura (ORFs) (figura 1). El primer ORF codifica proteínas de la cápsida (similares a las proteínas *gag* de retrovirus) y el segundo ORF, similar a los genes *pol* de retrovirus, codifica las actividades proteasa, integrasa, transcriptasa inversa y RNasa H. El elemento *gypsy* de *Drosophila* presenta un tercer ORF muy similar a los genes *env*, codificantes para las proteínas de la envoltura de retrovirus, habiéndose descrito en deter-

minadas condiciones la formación de partículas virales con capacidad infectiva (Kim y col., 1994; Song y col., 1994; Song y col., 1997).

Los retrotransposones no-LTR o elementos LINE, originalmente descubiertos en humanos, son también llamados elementos poli-A⁺, dado que, tal como referimos anteriormente presentan en su extremo 3' una región rica en adeninas. Estos elementos pueden integrarse en una diana específica del ADN genómico hospedador (elementos sitio-específicos) como R1 y R2 de insectos (Jakubczac y col., 1991) y CRE1, SLACS y CZAR de tripanosomátidos (Gabriel y col., 1990; Aksoy y col., 1990; Villanueva y col., 1991), o pueden encontrarse insertados de manera dispersa a lo largo del genoma (elementos de sitio no-específico) entre los que destaca el elemento L1 de humanos con unas 10^5 copias distribuidas por todos los cromosomas (Fanning y Singer, 1987). Como resultado de la integración

de la nueva copia en el genoma se produce la duplicación del sitio de inserción lo que da lugar a que el elemento quede flanqueado por unas secuencias repetidas directas de entre 7 y 20 nucleótidos. Carecen de LTR y la mayoría de ellos contienen dos marcos abiertos de lectura, situados en fases diferentes que solapan en un corto segmento (figura 1). En algunos elementos estos ORFs están en la misma fase de lectura, separados por un codón de terminación (Fawcett y col., 1986; Skowronski y col., 1988), y en otros, solo existe un único marco de lectura (Burke y col., 1987; Gabriel y col., 1990). El primer ORF de los elementos LINE codifica para un polipéptido con un tamaño de entre 300-500 aminoácidos, contiene algunos motivos de cisteínas cerca de su extremo carboxi-terminal con una separación de residuos de cisteína e histidinas similar al encontrado en los genes *gag* de los retrovirus. Una excepción a esta estructura sería la presente en el ORF1 del elemento L1 humano, el cual posee motivos del tipo cremallera de leucinas (*leucine zipper*) (Holnes y col., 1992), que codifica para la proteína p40 que se une específicamente al tránsrito de L1 cerca del extremo 5' del ORF2 y juntos son transportados al núcleo (Hohjoh y Singer, 1996). Se cree, por analogía a los genes *gag*, que para el resto de los elementos LINE descritos, este ORF codifica para una proteína de unión a ácidos nucleicos la cual estaría actuando en la unión del ARN intermedio en el proceso de transcripción inversa. Así, el ORF1 del elemento I factor de *Drosophila* codifica para una proteína con capacidad de unión a ácidos nucleicos (ARN y ADNs de cadena sencilla y cadena doble), pero a diferencia de su homólogo humano L1, dicha unión es inespecífica (Dawson y col., 1997). El segundo ORF codifica para un polipéptido de alrededor de 1000 aminoácidos, presentando homología con los genes *pol* de retrovirus, especialmente en lo que respecta al dominio transcriptasa inversa, dominio conservado en todos los retroelementos desde retrovirus a secuencias internas del grupo II de intrones (Michel y Lang, 1985). En la mayoría de los casos este dominio está localizado en la zona central del ORF, pero en algunos elementos puede estar cerca de la zona amino-terminal (Kimmel y col., 1987) o en la carboxilo-terminal (O'Hare y col., 1991). En el segundo marco abierto de lectura de los retrotransposones LTR, además de la actividad transcriptasa inversa se

han detectado de forma similar a retrovirus, la existencia de actividades enzimáticas RNasa H, proteasa e integrasa (Doolittle, 1989). Sin embargo, en el caso de los elementos LINE no se han descrito la presencia de dominios proteasa y sólo en un pequeño número de elementos se ha descrito pequeñas homologías de secuencia con dominios integrasa (Gabriel y col., 1990; Villanueva y col., 1991) o RNasa H (Fawcett y col., 1986; McClure, 1991).

A diferencia de la mayoría de los retrotransposones no-LTR descritos, los cuales poseen dos ORFs, el elemento LINE de *T. cruzi* L1Tc, está organizado en tres ORFs en diferentes fase de lectura (figura 1). El ORF2 presenta homología con los genes *pol* de retrovirus y retrotransposones LTR, manteniendo conservados los 7 dominios consenso de las transcriptasas inversas (Xiong y Eickbush, 1990). Recientemente, ha sido detectada actividad transcriptasa inversa en lisados totales de *T. cruzi* (González y col., 1997a) y la proteína recombinante expresada a partir de un fragmento del ORF2 del elemento L1Tc, que contiene los dominios consenso de las transcriptasas inversas, posee dicha actividad enzimática medida sobre homopolímeros sintéticos de ARN, y ensayos de RT-PCR sobre ARN heterólogo (González y col., 1997b). El ORF3, con homología a los genes *gag* de retrovirus y retrotransposones LTR, presenta además homología con los dominios integrasa de retrovirus, así como dos motivos *Zinc finger* del tipo CCHH similares a los descritos en factores de transcripción de eucariotas (Martín y col., 1995). La mayoría de los retrotransposones presentan dominios *Zinc finger* del tipo CCHC en el ORF1. La estructura CCHH fue encontrada en los retrotransposones descritos en la familia *Tripanosomatidae* (SLACS, CZAR e ingi) además de en el elemento R2Bm del insecto *Bombyx mori* (Burke, 1987). Sin embargo, sólo los elementos ingi y L1Tc presentan estos dominios en el ORF3. Interesantemente el ORF1 de L1Tc presentó una homología significativa con miembros de la familia de endonucleasas del tipo AP clase II (Martín y col., 1995), proteínas que intervienen en la reparación del ADN (Demple y Harrison, 1994). Estudios de homología de secuencia y de estructura han evidenciado que la referida proteína codificada por el ORF1 del elemento L1Tc contiene todos los residuos activos des-

critos para la exonucleasa III y el 80% de los residuos activos de la DNaseI humana y bovina (Martín y col., 1996), homología extensiva a la región amino terminal del ORF2 de todos los elementos LINE descritos (Martín y col., 1996; Olivares y col., 1997c), a excepción del

elemento Dong. Este elemento contiene el primer motivo activo (QET) de la zona N-terminal, pero pierde los demás motivos, sugiriendo que la copia secuenciada de Dong probablemente presenta una delección interna del ORF2 (Feng y col., 1996).

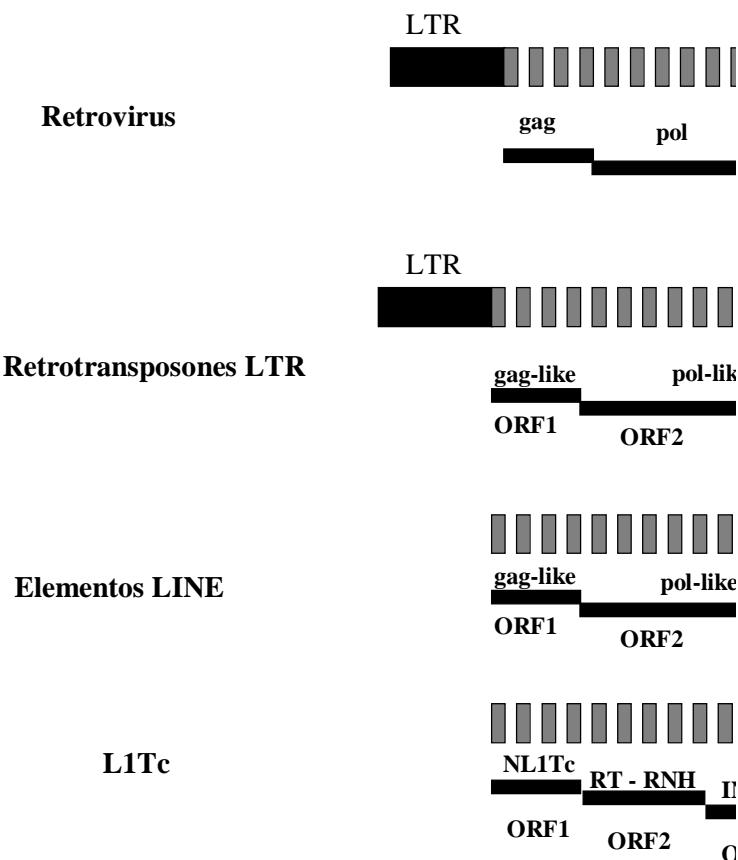


Figura 1: Esquema comparativo de la estructura de los retrovirus, retrotransposones LTR, elementos LINE y del elemento LINE de *T. cruzi* L1Tc. LTR, terminaciones largas repetidas. NL1Tc, endonucleasa con actividad hidrolítica de sitios apurínicos-apirimidínicos. RT, transcriptasa inversa. RNH, RNaseH. IN, dominios homólogos a las integrasas retrovirales. FT, dominios Zinc finger homólogos a factores de transcripción del tipo TFIII.

4.- FILOGENIA DE LOS RETROTRANSPOSONES

La mayoría de los elementos LINE se encuentran presentes en una amplia variedad de organismos (mamíferos, anfibios, plantas, insectos, hongos y protistas), localizados de forma dispersa a lo largo del genoma que hospedan. Sin embargo, hay un número significativo de elementos que se encuentran localizados, como se mencionó anteriormente, en sitios específicos del genoma. Así, los elementos R1 y R2 se encuentran insertados en una específica fracción de los genes ribosomales 28S de la mayoría de los

insectos (Jakubczak y col., 1991), al igual que el elemento R4 presente en nematodos (Burke y col., 1995). Tx1 de *Xenopus laevis* se inserta en Tx1d, un segundo elemento móvil presente en la misma especie (Garret y col., 1989). CRE1, SLACS y CZAR se encuentran localizados en el exon del *spliced leader* (SL) de algunas especies de la familia *Trypanosomatidae* (Gabriel y col., 1990; Aksoy y col., 1990; Villanueva y col., 1991). DRE de *Dictyostelium discoideum* se inserta adyacente a los genes del ARN de transfe-

rencia (Marschalek y col., 1992). Recientemente se han descrito nuevos retrotransposones no-LTR, tales como el elemento TRAS1 (Okasaki y col., 1995), el elemento SART1 de *Bombyx mori* (Takahashi y col., 1997), o el elemento HetA de *Drosophila melanogaster* (Biessmann y col., 1990), los cuales aparecen insertados en las secuencias repetidas del telómero.

Comparando la secuencia de aminoácidos del dominio transcriptasa inversa, Eickbush estableció una probable relación filogenética entre los retroelementos LTR, no-LTR, retrovirus y un grupo de elementos genéticos adicionales que codifican para transcriptasas inversas (Eickbush, 1992). Estos elementos incluyen virus ADN, como hepadnavirus de animales y caulimovirus de plantas (Toh y col., 1983), intrones mitocondriales del grupo II de hongos (Michel y Lang, 1985) y los elementos de inserción bacterianos los cuales producen múltiples copias de ADN de cadena sencilla (Inouye y col., 1989; Lim y Mass, 1989). De esta forma, basándose en la similitud de secuencia de los siete dominios conservados en más de 80 transcriptasas inversas y tomando como raíz del árbol filogenético las ARN polimerasas dependientes de ARN, todos los retroelementos pueden ser divididos en dos ramas principales (figura 2). Una rama contiene los retrotransposones no-LTR, elementos de bacterias y el grupo II de intrones. Todos los elementos de esta rama carecen de repeticiones terminales ya sean directas o indirectas. En la segunda rama del árbol se localizan retroelementos cuya replicación implica repeticiones terminales. En este segundo grupo se encuentran virus (retrovirus, caulimovirus y hepadnavirus) y los retrotransposones LTR. Dentro de los elementos LTRs se distinguen 2 grupos: el grupo Ty1-copia y el Ty3-gypsy. La principal diferencia estructural entre estos dos grupos de elementos LTRs es el orden en que se encuentran los dominios transcriptasa inversa e integrasa en el ORF homólogo al gen *pol* de retrovirus (Mount y Rubin, 1985).

La localización de todos los retrotransposones no-LTR en una misma rama sugiere que tienen un origen común y que no parecen proceder de elementos LTR que han perdido las repeticiones terminales. Además, salvo algunas excepciones, la relación filogenética de estos elementos no parece correlacionarse con las relaciones

filogenéticas de los organismos que habitan. Este tipo de relación indica que los retrotransposones no-LTR (o sus dominios transcriptasa inversa) se han transmitido de una forma horizontal entre los principales grupos taxonómicos o que el origen de estos elementos precede a la evolución de los metazoarios (Eickbush, 1992).

Si nos basamos en la organización estructural y en la conservación de dominios específicos, los retrotransposones sitio específicos descritos en tripanosomátidos parecen pertenecer a una misma familia de elementos móviles (figura 2). Cada uno de los elementos contiene uno o dos marcos de lectura abiertos. El análisis del ORF2 muestra que además de poseer los siete dominios conservados de la transcriptasa inversa, contienen un motivo conservado cisteína-histidina en la zona 5', similar al identificado para las integrasas de retrovirus (Villanueva y col., 1991). Mientras que el elemento CRE1 presente en *Critidhia* tiene un único marco abierto de lectura similar al ORF2 de los elementos SLACS y CZAR presentes en *T. brucei* y *T. cruzi* respectivamente, estos contienen un ORF adicional, situado en 5' respecto al anterior. Este ORF1 tiene un motivo conservado de cisteína-histidinas, similar al dominio *gag* de los retrovirus. A pesar de las similitudes entre los tres elementos citados, que hacen pensar en ellos como una familia de retrotransposones evolutivamente relacionados, SLACS y CZAR presentan una mayor similitud entre ellos que con CRE1, lo cual estaría de acuerdo con la distancia evolutiva existente entre los organismos en los cuales están presentes. Estudios filogenéticos realizados con genes mitocondriales revelan que *T. brucei* y *T. cruzi* están evolutivamente más cercanos entre sí, que con *Critidhia* (Lake y col., 1988). Además ambos tripanosomas son digénicos, requiriendo en su ciclo de vida un hospedador vertebrado y un insecto vector, mientras que *Critidhia* es monogénica realizando todo su ciclo biológico dentro del insecto. Es por ello, que los referidos retrotransposones sitio específicos de tripanosomátidos pueden ser evolutivamente tan lejanos como sus hospedadores, y las diferencias entre ellos serían reflejo de la evolución que han tenido en cada genoma.

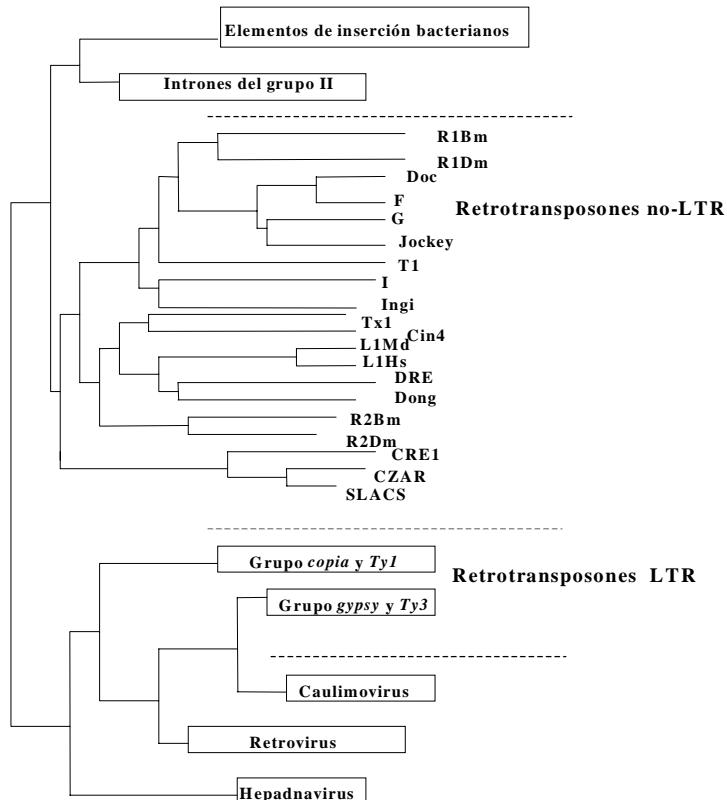


Figura 2: Árbol filogenético de los retroelementos, basado en el estudio comparativo de las secuencias de aminoácidos de los dominios consenso de las transcriptasas inversas, según el estudio realizado por Xiong y Eickbush, 1990.

5 .- EVIDENCIAS DE TRANSPOSICIÓN DE LOS ELEMENTOS LINE

La demostración de una actividad transcriptasa inversa codificada por un elemento LINE se ha descrito en el elemento jockey de *Drosophila melanogaster* (Ivanov y col., 1991), CRE1 de *Critchidia fasciculata* (Gabriel y Boeke, 1991) y L1 de humanos (Mathias y col., 1991). En 1991 se demostró experimentalmente, que dicha transposición ocurre vía ADN-ARN-ADN en el elemento I factor de *Drosophila*, L1 de ratón, R2Bm de *Bombyx mori* y más recientemente en L1 de humanos (Jensen y Heidmann, 1991; Evans y

Palmiter, 1991; Luan y col., 1993; Moran y col., 1996).

El elemento L1 ocupa aproximadamente el 15% del genoma humano. Aunque existen unas 100.000 copias de dicho elemento, aproximadamente el 95% de las mismas son truncadas o están internamente modificadas, de tal forma que son inactivas para la retrotransposición. Por ejemplo, se estima que de las 30 a 60 copias intactas existentes en el genoma humano del subgrupo Ta - L1, sólo 7 de estos elementos serían activa-

mente competentes para su retrotransposición autónoma (Kazazian y Moran, 1998). La evidencia más patente de la movilidad, y en suma de la actividad de los elementos L1, está relacionada con la aparición de mutaciones inducidas por su inserción *de novo* en algún lugar del genoma humano. En 1988 se describió por primera vez al elemento móvil L1 como agente causal de una enfermedad al comprobarse que la inserción de L1 en el gen del factor VIII daba lugar a un proceso patológico hemofílico del tipo A (Kazazian y col., 1988). Otras inserciones del elemento L1 se han descrito en los últimos años, tales como inserciones en el gen del factor VIII (Woods-Samuels y col., 1989), en el gen de la *bglobina*, en el gen APC (Miki y col., 1992), o más recientemente la descripción de varias inserciones en el gen DMD (gen de la distrofina) en pacientes con distrofia muscular (Narita y col., 1993; Holmes y col., 1994). Mientras que la mayoría de las inserciones descritas ocurren en la línea germinal o durante el desarrollo temprano, la inserción en el gen APC se encontró en células de cáncer de colon, pero no en otras células del paciente, indicando que la transposición no sólo puede ocurrir en células germinales sino también en células somáticas (Kazazian y col., 1998). El análisis de cada uno de estos elementos revela que cada secuencia es diferente, sugiriendo que cada una proviene de un elemento progenitor distinto. Recientemente se ha publicado el primer ejemplo de una mutación debida a la recombinación no homóloga entre elementos L1 en el gen que codifica la subunidad de la fosforilato kinasa (PHKB), lo cual ha llevado a la delección de una región que incluye el exón 8. Esta mutación dio lugar a una patología caracterizada por un almacenamiento de glucógeno motivado por la deficiencia de la fosforilato kinasa (Burwinkel y Kilimann, 1998).

Por otro lado se han observado otras estrecha relación entre la expresión de elementos LINE

con procesos cancerosos, ha sido evidenciada en numerosas ocasiones. En un carcinoma humano de mama ligado a una reorganización específica de un locus *c-myc* y a la amplificación de otro alelo *c-myc*, se ha identificado la inserción de un elemento LINE dentro del segundo intrón del locus reorganizado. Esta reorganización, así como la inserción, están presentes solamente en el tejido de mama malignizado y no en el tejido de la mama no malignizada de la misma paciente (Morse y col., 1988). En el tumor venéreo de perro, en el que se encuentra involucrado el mismo oncogen *c-myc*, se ha encontrado un elemento LINE insertado corriente arriba del primer exón del citado oncogen (Katzir y col., 1985). Otro ejemplo de inserción en *c-myc* asociado a un proceso tumoral, es el caso del inmunocitoma de rata donde se ha demostrado que *c-myc* se recomienda con una secuencia LINE dentro del agrupamiento de los genes de las inmunoglobulinas (Pear y col., 1988). Se ha observado también que la expresión de los elementos LINE está favorecida en células tumorales, tanto de origen germinal como epitelial (Singer y col., 1993). Se ha descrito una expresión generalizada del elemento L1 en tumores de células germinales de testículo humano, observándose también su expresión en las células metastatizadas (Bratthauer y Fanning, 1992). Estudios epidemiológicos muestran que al menos el 10% de los cánceres de células germinales de testículo expresan el elemento L1. Asimismo este elemento se expresa en un 5% de células de tumor ovárico y en un 30% de tumores extragonales (Bratthauer y Fanning, 1993). Por ello se postula, que la inserción de los elementos móviles L1 origina mutaciones que pueden jugar un papel importante en la etiología de algunas neoplasias. Además, existen claras evidencias de que la expresión de estos elementos puede contribuir al origen y progresión de algunos cánceres humanos de mama (Asch y col., 1993; Bratthauer y col., 1994).

6.- MECANISMO DE TRANSCRIPCIÓN Y TRANSPOSICIÓN DE LOS ELEMENTOS LINE

Los retrovirus y los retrotransposones LTR presentan promotores convencionales dirigidos por ARN polimerasa II localizados en el extremo 5' de la secuencia LTR, iniciándose la transcripción corriente abajo de referido promotor (Varus y Brown, 1989). La secuencia promotora,

la cual no está presente en el extremo 5' del ARNm resultante, es regenerada durante la transcripción inversa hacia el ADN intermedio, a través de la secuencia LTR localizada en el extremo 3'. En el caso de que en los retrotransposones no-LTR o elementos LINE existiera un promoto-

tor similar controlado por ARN polimerasa II, se plantearía el grave problema de cómo regenerar la secuencia promotora, en la nueva copia del elemento, ya que en estos elementos carentes de secuencias LTR no existe duplicación de la putativa secuencia promotora en el extremo 3' del ARNm intermedio.

Una solución al problema de cómo restablecer el promotor en los retrotransposones no-LTR podría ser resuelto por el hecho de que la inserción de estos elementos ocurre en específicos sitios del genoma. De esta forma, los elementos R1 y R2 insertados en los genes 28S de *Bombyx mori*, así como los elementos CRE1, SLACS y CZAR de tripanosomátidos insertados en los exones del SL, están posicionados en la correcta orientación para ser co-transcritos por el promotor del gen donde se insertan. Por tanto una posible explicación, del porqué estos elementos han evolucionado hacia una especificidad de sitio y orientación de inserción dentro de determinados genes, sería asegurar que de esta forma las nuevas copias insertadas estén localizadas corriente abajo de un promotor adecuado. En los elementos LTR, que no requieren de promotores externos, el número de elementos con especificidad de inserción es mucho menor. El uso de un promotor externo difícilmente resuelve el problema de los elementos no específicos de sitio de inserción. De hecho, los elementos LINE transcripcionalmente activos han sido encontrados en diferentes sitios del genoma sin un promotor conocido que pueda dirigir su transcripción, sugiriendo por tanto la existencia de un promotor interno (Fawcett y col., 1986; Mizrokhi y col., 1988; Dombroski y col., 1991). La demostración directa de la presencia de un promotor interno en estos elementos fue establecida para el elemento jockey de *D. melanogaster* (Mizrokhi y col., 1988) y para el elemento L1 de *Rattus norvegicus* (Nur y col., 1988). Mediante ensayos de transfección transitoria, se demostró que los únicos nucleótidos requeridos para la transcripción del elemento jockey estaban corriente abajo del inicio de la transcripción (Mizrokhi y col., 1988), siendo los 13 primeros nucleótidos esenciales para la misma. Otros experimentos con los elementos L1 (Swergold, 1990), I (Mclean y Finnegan, 1991) y F (Minichiotti y Di Nocera, 1991), han indicado que su transcripción está también controlada por un promotor interno localizado cerca del extremo 5' del ele-

mento. Para estos elementos LINE los componentes críticos del promotor se localizan en las primeras 100 bp del elemento. Sin embargo, toda la región controladora de la transcripción se distribuye en algunos cientos de pares de bases del extremo 5'. El inicio de la transcripción corriente arriba del promotor en los retrotransposones no-LTR o elementos LINE, permite la preservación de las regiones promotoras en el curso de la replicación vía transcripción inversa. Mientras que todos estos datos hacen pensar en un promotor dirigido por ARN polimerasa III, la ausencia de secuencias correspondientes a las cajas A y B típicas de este tipo de promotores, la poliadenilación del ARN mensajero y el hecho, observado en alguno de estos elementos, de que la transcripción de los mismos sea sensible a la aamanitina, son argumentos a favor de la participación de una polimerasa de tipo II (Mizrokhi y col., 1988). La localización de la región promotora de tipo II corriente abajo del sitio de inicio de la transcripción tiene algunos precedentes. Por ejemplo en algunos genes de desarrollo de *D. melanogaster* se han encontrado secuencias esenciales para la transcripción por ARN polimerasa II corriente abajo del sitio de inicio (Thummel, 1989; Winslow y col., 1989), pero en ningún caso todas las secuencias necesarias para la transcripción se encuentran localizadas corriente abajo.

En el caso del elemento L1 de humanos se ha hablado de una posible combinación de polimerasas (Kurose y col., 1995). El análisis de los tránscritos correspondientes al elemento L1 de humanos mostró tránscritos de diferentes tamaños algunos de los cuales eran más largos de la unidad y se encontraban ligados a secuencias no relacionadas. Estos datos sugieren que dichos tránscritos podrían provenir de la transcripción regulada por promotores de genes vecinos (Skowronski, 1985).

La transcripción de los genes en *T. cruzi* tiene la singularidad de ser mayoritariamente policistrónica y requiere de un específico mecanismo denominado *trans-splicing*, necesario para procesar los tránscritos primarios en ARN mensajeros maduros individuales. Este mecanismo implica la adición de una secuencia de 39 nucleótidos (*mini-exon* o SL) en el extremo 5' de todos los mensajeros maduros, y una cola de poli-A en el extremo 3'. El proceso involucra la producción de un ARNm quimérico resultante de la transcripción de dos segmentos de ADN

que se encuentran separados en el genoma (DeLange y col., 1984; Parsons y col., 1984). Interesantemente, ensayos de RT-PCR y extensión de cebador en los mensajeros maduros del retrotranspón L1Tc de *Trypanosoma cruzi* no han permitido detectar la presencia del referido *spliced leader* asociado en 5' al transcripto del elemento L1Tc (datos de laboratorio). Resultado similar al descrito, ha sido reportado para el elemento *ingi* (Vasella y col., 1996), con el cual L1Tc comparte una alta homología de secuencia y organización estructural. Recientemente, mediante ensayos de transformación estable no integrativa con diferentes vectores recombinantes, construidos a partir del vector de transformación pTEX descrito por Kelly y col., (1992), que contenían como putativa región promotora la región 5' del elemento L1Tc seguida de un adecuado gen informador, se evidenció la capacidad de la referida región 5' de L1Tc de dirigir y activar la transcripción del gen informador (Olivares y col., 1997b). La zona 5' de L1Tc, como referimos anteriormente, es homóloga a la secuencia RIME descrita por Hasan y col., (1982) en *T. brucei*, la cual se ha implicado en procesos de activación génica (Pays y col., 1989). Estos resultados podrían estar evidenciando la presencia de un promotor interno en L1Tc, similar al descrito para otros elementos LINE (Mizrokhi y col., 1988; Nur y col., 1988; Swergold, 1990; McClean y Finnegan., 1991; Minchietti y DiNocera, 1991). Por otra parte, dado que en *T. cruzi* no se han descrito promotores clásicos, no es posible excluir que el elemento L1Tc, repetido más de 2500 veces a lo largo de todo el genoma (Martín y col., 1995), podría actuar como promotor de grupos génicos, iniciando la transcripción de policistrones específicos o actuar como activador génico (*enhancers*) de promotores cercanos atípicos.

El mecanismo de transposición de los elementos LTR es bien conocido desde hace bastantes años (Wang y Seeger, 1993). La transcripción inversa del ARN en los elementos LTR tiene lugar en el citoplasma, donde se usa normalmente como cebador un ARNt para la síntesis de la primera cadena de ADN y una molécula de ARN de poli-purinas para la síntesis de la segunda cadena, tras la hidrólisis del ARN que forma el híbrido ARN-ADN por la acción del enzima RNasaH. La síntesis de ambas cadenas es completada a través de eventos de transferencia de cadena que involucran las redundancias termina-

les de las secuencias LTR. El ADN de doble cadena pre-procesado migra entonces al núcleo donde una integrasa codificada por el elemento se une a las secuencias LTR y las procesa eliminando dos nucleótidos del extremo 3'. Finalmente la propia integrasa rompe el ADN genómico diana y une los extremos del ADN procesado con los originados en el ADN genómico (figura 3). A pesar de la similitud de los elementos LINE con los elementos LTR y retrovirus, el mecanismo de transposición e integración de estos elementos LINE, es sin duda alguna diferente al anteriormente referido dada la ausencia de las secuencias LTR esenciales para dichos procesos, y no está bien conocido. Además, la falta de caracterización de actividades enzimáticas asociadas a las putativas proteínas codificadas por estos elementos LINE, dificulta aún más su conocimiento. Se ha postulado que la integración tiene lugar en sitios donde el ADN cromosómico ha sido previamente cortado por productos codificados por el hospedador, como podría ocurrir durante la reparación del ADN o recombinación (Eickbush, 1992). En el elemento sitio específico R2Bm se describió la presencia de una actividad nucleasa que corta específicamente dentro de la secuencia del gen del ARNr 28s, lugar de inserción de dicho elemento (Luan y col., 1993). La demostración de la existencia de una actividad endonucleasa en la proteína recombinante codificada por el ORF1 del elemento L1Tc, la cual reconoce sitios apurínicos/apirimidínicos (Olivares y col., 1997a), así como en la proteína codificada por la región amino-terminal del ORF2 del elemento L1 humano (Feng y col., 1996), ha permitido explicar el primer paso de la integración de estos elementos LINE en el genoma. Esta proteína con actividad endonucleasa, codificada por el propio elemento LINE, cortaría en una de las cadenas del ADN generando un extremo 3'-OH libre que sería usado por la transcriptasa inversa como cebador para la síntesis de la primera cadena de ADN complementaria al ARN molde. Además este modelo permite explicar la distribución dispersa de estos elementos en el genoma, sin la necesidad de que para ello deban existir roturas previas en el ADN genómico. Interesantemente, recientes resultados han mostrado que la proteína recombinante codificada en el ORF2 del elemento L1Tc, obtenida por expresión *in vitro* en un sistema de *E. coli*, tiene actividad polimerasa dependiente de cationes

divalentes (González y col., 1997b), pudiendo utilizar como substrato ARN y ADN de cadena sencilla (datos de laboratorio). Esta proteína, presenta además asociada una actividad RNasaH con requerimientos iónicos similares a los descritos para el enzima homólogo de retrovirus y por tanto diferentes a los requeridos por el enzima de *E. coli* (datos de laboratorio). La sustitución del segundo ácido aspártico del dominio activo de la

transcriptasa inversa (YXDD) conduce a la abolición de la actividad transcriptasa inversa, no afectándose la actividad RNasaH asociada, cuyos dominios catalíticos están localizados en la zona carboxilo terminal de la proteína recombinante, corriente abajo de los dominios consenso descritos para las transcriptasas inversas.

Estos resultados, unidos a otros recientemente publicados por otros autores (Dawson y col.,

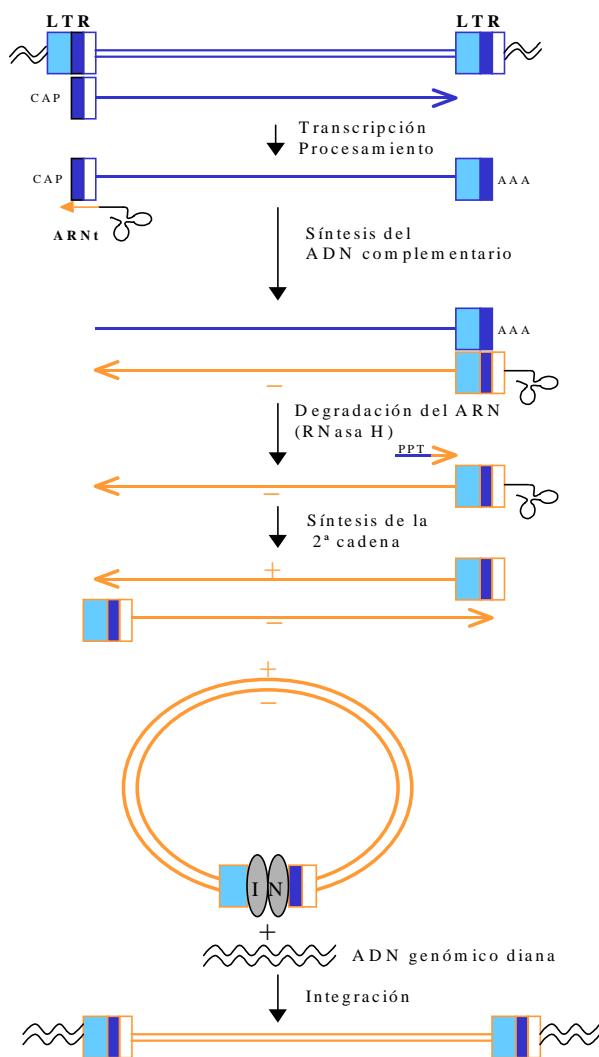


Figura 3: Esquema del mecanismo de integración y transposición de los elementos LTR según el modelo propuesto por Wang y Seeger, 1993. LTR, terminaciones largas repetidas. CAP, caperuza. PPT, tracto de polipurinas. IN, integrasa.

1997; Dhellin y col., 1997; Clements y col., 1998), permiten establecer un nuevo modelo para explicar el mecanismo de transposición e integración de estos elementos LINE en el genoma que los llevan, el cual queda esquematizado en la figura 4. Inicialmente se formaría el transcripto del elemento completo, con la participación de un promotor interno situado en el extremo 5'. Seguidamente se produciría la traducción de los distintos ORFs presentes en el mensajero intermedio transcripto y consecuentemente la obtención de las proteínas que codifican (endonucleasa, transcriptasa inversa, RNasaH, putativa integrasa y *zinc fingers*). Probablemente en este punto sea necesario un procesamiento posttraduccional de los referidos productos, sin embargo hasta el momento no se han descrito en los elementos LINE dominios que pudieran codificar para una putativa actividad proteolítica. Posteriormente los componentes del complejo de integración, *zinc fingers* e integrasa se unirían respectivamente, a la región 3' del transcripto completo, para facilitar su transporte al núcleo, y a la región del ADN genómico reconocido por la endonucleasa codificada por el propio elemento LINE, la cual habría

generado en la cadena negativa del ADN (c-) un extremo 3'-OH libre que podría ser usado como iniciador por la transcriptasa inversa. Este enzima utilizaría inicialmente como molde los primeros nucleótidos de la cadena positiva del ADN localizados en el 5' al punto de corte de la C-, seguidamente copiaría en dirección 3'®5' la cadena complementaria al ARN intermedio. Finalmente, la transposición e integración en el genoma de un nuevo elemento LINE, quedaría completada por la acción del enzima RNasaH que produciría la degradación de la cadena de ARN molde que forma el híbrido ARN-ADN, y por una actividad ADN polimerasa que copiaría la segunda cadena de ADN generándose las secuencias de duplicación. Esta actividad polimerasa probablemente pudiera ser ejercida por la propia enzima transcriptasa inversa codificada por el elemento LINE, dado que tal como describimos anteriormente, al menos el enzima presente en el elemento L1Tc de *T. cruzi*, es capaz de utilizar moldes de ADN como substratos (González y col., 1997b).

7.- PAPEL BIOLÓGICO DE LOS ELEMENTOS LINE

Los elementos transponibles se han involucrado directa o indirectamente en varios procesos de regulación de la expresión génica. Se han descrito casos en los que estas secuencias se encuentran como elementos activos tanto en procesos de regulación transcripcional como post-transcripcional. Ya en 1969 se postulaba que las secuencias repetidas podrían jugar un papel en las regiones 5' implicadas en regulación génica reguladoras de genes (Britten y Davidson, 1969), evidenciándose posteriormente que las variaciones en regiones de regulación eran fundamentales para la evolución (Britten y Davidson, 1971), lo que llevó a argumentar que las secuencias repetidas podrían ser una importante fuente de variabilidad evolutiva (Britten, 1997). Recientes estudios han mostrado que cierto número de elementos móviles son frecuentemente insertados en el ADN y pudiendo desempeñar funciones estrechamente ligadas a la regulación de la expresión génica. La secuencia intensificadora *enhancer* del gen de la apolipoproteína A en humanos está localizada dentro de un elemento

LINE (Yang y col., 1998). La transcripción de un elemento repetido que se encuentra insertado en el gen del receptor de eritropoyetina de ratón inhibe profundamente la transcripción del receptor (Youssoufian y Lodish, 1993). La señal de poliadenilación del gen de la timidilato sintasa se creó por la inserción de un elemento LINE (Harendza y Johnson, 1990). En lo que respecta a tripanosomátidos, se ha sugerido que el elemento ribosomal RIME, presente en los sitios de expresión activos de los genes VSG de *T. brucei*, puede condicionar la activación de potenciales sitios de expresión (Pays y col., 1989). El elemento SIRE de *T. cruzi* proporciona un nuevo sitio aceptor de *splicing* en ciertos genes de la proteína ribosomal P2 (Vazquez y col., 1994). Además, tal como referimos en el apartado de introducción, la inserción de parte de este elemento en la región 3' no traducida de algunas unidades génicas del locus H2A de *T. cruzi*, da lugar a un cambio en el sitio de poliadenilación y consecuentemente a un mensajero maduro de mayor tamaño, cuyo nivel citoplasmático está

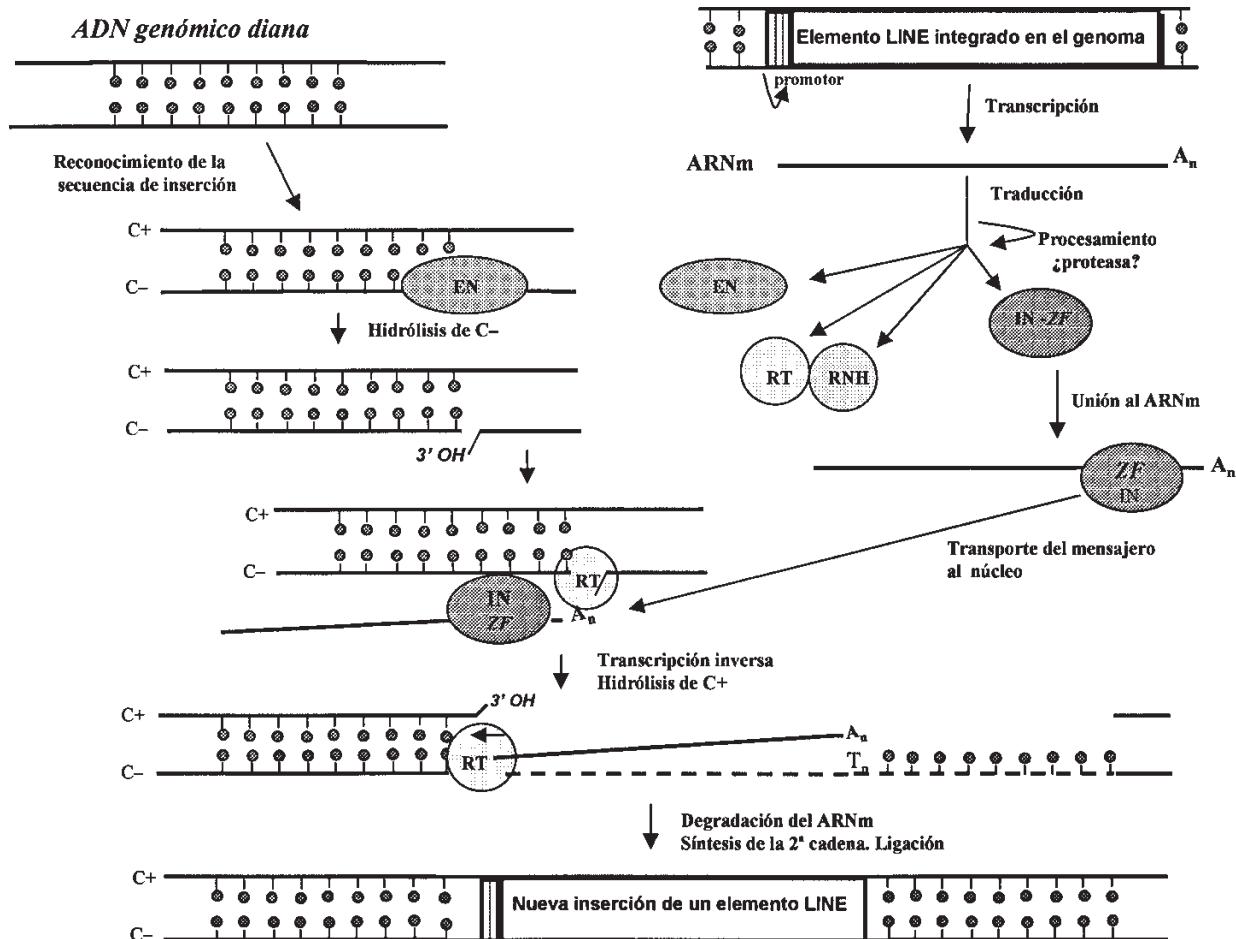


Figura 4: Esquema del modelo propuesto para el mecanismo de integración y transposición de los elementos LINE. EN, endonucleasa; C+ y C- cadenas positiva y negativa del ADN respectivamente; FT, Integrasa-Zinc fingers, Zinc fingers. RT, transcriptasa inversa.

regulado de forma más eficiente que el mensajero que no lo contiene (Marañón y col., 1997; Marañón y col., 1999). Por otra parte, dado que existen secuencias de ADN repetidas-dispersas en varias localizaciones del genoma, un proceso de reordenamiento por recombinación homóloga entre dos de estas secuencias, podría colocar en un entorno favorable para la expresión génica a genes que anteriormente no eran transcritos. Este parece ser el caso de los procesos que dan lugar a la activación o inactivación de los sitios de expresión de las VSG de *T. brucei* (Gottesdiener, 1991; Lodes, 1993).

Otro posible papel biológico de los retrotransposones, sería su implicación en los procesos de reparación del ADN cromosómico. Así, en ensayos con células de *Saccharomyces cerevisiae* en las cuales se anula la posibilidad de reparación por recombinación homóloga, me-

canismo normalmente utilizado en levaduras para reparar la rotura del ADN cromosómico, se observó la inserción de fragmentos del retrotransposón Ty1 en el punto específico de corte generado por la endonucleasa HO. En el caso de células *wild type* también se encontraron inserciones del elemento Ty1 aunque la frecuencia era significativamente menor (Moore y Haber, 1996). En estudios paralelos, Teng y col., (1996) mostraron igualmente que roturas en el ADN cromosómico de *S. cerevisiae* podían ser reparadas, en ausencia de recombinación homóloga, por la inserción de ADN complementarios marcadores generados por la presencia de las transcriptasas inversas codificadas por los elementos L1 de humanos, CRE1 de *Crithidia* o por el propio Ty1 de *S. cerevisiae*. Estos resultados estarían indicando una importante relación entre retrotransposición y reparación del ADN

en eucariotas, evidenciando por tanto un posible nuevo mecanismo de relevancia biológica para la reparación cromosómica en células que contienen estos elementos.

Dado que estos elementos transponibles pueden potencialmente actuar como fuertes agentes mutagénicos, y por tanto ejercer una acción deletérea para el hospedador, la movilidad de los mismos debe de estar fuertemente regulada, tanto a nivel de frecuencia como de especificidad de inserción. Por su parte, estos elementos móviles deben de mantener un cierto nivel de actividad que les permita asegurar su propagación y supervivencia en el genoma. Por todo ello el nivel de transposición es probablemente el resultado de un balance entre los intereses del elemento transponible y los intereses del hospedador que habitan. La mayoría de los retrotransposones de eucariotas se movilizan de forma esporádica en el genoma. Los elementos se mueven raramente y de forma imprevisible, sugiriendo la existencia de una fuerte regulación de su transposición. Una excepción a esta regla es el caso del elemento I factor el cual presenta una alta frecuencia de transposición durante la disgénesis híbrida en *Drosophila melanogaster* (Bucheton, 1990). A pesar de que los mecanismos de control de la movilidad de los elementos LINE por el hospedador no son muy conocidos, muy probablemente deben de estar actuando a múltiples niveles, tales como: transcripción del elemento, expresión específica de tejido, procesamiento del ARN y proteínas codificadas, ensamblaje de partículas virales, transcripción inversa y la propia integración de la doble cadena de ADN. Algunos de estos mecanismos se han descrito en elementos LTR, y dada la similitud entre estos y los elementos LINE, se postula que algunos de ellos pueden ser comunes. Solo recientemente ha sido descrito un mecanismo específico de control de la transposición de un elemento LINE, el cual afecta al promotor interno. Así, Chaboissier y col (1998) han puesto en evidencia que la actividad del promotor interno del elemento I factor de *Drosophila* está afectada de forma inversamente proporcional al número de copias existentes en el elemento de los primeros 186 nucleótidos, los cuales constituyen la región 5' no traducida del mismo.

Para un elemento móvil es beneficioso transponerse en aquellos tejidos que le aseguren la transmisión vertical a la siguiente generación del

hospedador. Sin embargo la transposición en tejidos somáticos probablemente sólo resultaría en una pérdida de funcionalidad en el hospedador, sin ningún beneficio para el elemento. Se piensa que la transcripción del elemento L1 de humanos se limita a células indiferenciadas: células germinales tempranas y células tumorales. Esto puede ser debido al significativo menor nivel de metilación existente en células indiferenciadas. De hecho, se ha sugerido que uno de los propósitos de la metilación de CpG es reducir la expresión de elementos móviles en células diferenciadas (Yoder y col., 1997).

Otro mecanismo de control que protege de mutaciones deletéreas es la especificidad de inserción en zonas no codificantes del genoma o incluso en elementos preexistentes. De las 100 nuevas inserciones del elemento Ty1 observadas en el cromosoma III, sólo el 3% fueron insertadas en ORFs (Ji y col., 1993). En el maíz aproximadamente el 50% del genoma está constituido por retroelementos (SanMiguel y col., 1996), entre ellos la familia Opie supone el 10-15%, pero la mitad de los elementos estudiados están insertados dentro de otros elementos. Interesantemente, la endonucleasa codificada por el elemento L1, encargada de generar extremos 3'-OH libres necesarios para la integración del elemento, es específica de estructuras de ADN de doble cadena formadas por secuencias $T_n-N_2-A_n$ especialmente inusuales en el genoma humano (Cost y Boeke, 1998), las cuales además, están al menos tres veces más representadas en regiones no codificantes del genoma que en regiones codificantes. Estos autores apuntan a la posible existencia de una evolución del referido enzima del elemento L1, con selección positiva para el reconocimiento de las mencionadas estructuras de ADN, con el fin de evitar inserciones deletéreas. Como referimos anteriormente la endonucleasa del elemento L1 presenta una alta homología de secuencia con la endonucleasa codificada por el elemento L1Tc de *T. cruzi* y consecuentemente con las enzimas de actividad nucleasa tipo AP (Martín y col., 1996), reconociendo sitios apurínicos/apirimidínicos (Olivares y col., 1999), que a priori podrían generarse en cualquier región del genoma. Por otra parte se ha sugerido que el procesamiento de los elementos transponibles a partir de secuencias exónicas ha evolucionado como un medio para disminuir sus efectos deletéreos en el hospedador. Se han des-

crito casos en *Drosophila*, *Caenorhabditis elegans* y en plantas donde el procesamiento del ARN lleva a la eliminación del elemento y a una reversión parcial de la mutación causada por la inserción original (Fridell y col., 1990; Purugganan y Wessler, 1992; Rushforth y Anderson, 1996). Fenotipos mutantes asociados con inserciones de retroelementos son enmascarados por alelos de genes supresores que actúan como reguladores en *trans* de la expresión del elemento (McDonald, 1995). En *Drosophila* hay evidencias de la presencia de genes del hospedador con función supresora en poblaciones naturales (Csink y McDonald, 1990).

A nivel de transcripción se ha comprobado que una sobre-expresión del elemento Ty1 es necesario y suficiente para conseguir unos mayores niveles de transposición. Esto sugiere la existencia de un umbral en los niveles de ARN o proteínas codificadas por el elemento, necesario para la movilización de una nueva copia del mismo. El caso del elemento gypsy de *D. melanogaster* es muy llamativo. La movilización de gypsy es muy poco frecuente, sin embargo en determinadas condiciones los niveles de ARN de gypsy se incrementan entre 10 y 20 veces y se procesa un nuevo marco de lectura que parece codificar una proteína similar a las de la envoltura retrovirales.

En resumen, dado que los elementos

transponibles se propagan por inserción en nuevas localizaciones en el genoma que habitan, cabría pensar que la movilización de estos elementos no reportaría ningún beneficio al organismo hospedador y que por ello serían inactivados para finalmente desaparecer. Sin embargo, como se ha comentado anteriormente, existen evidencias que demuestran la actividad de estos elementos y más aún, se han descrito numerosos casos que muestran que la inserción de un elemento no es necesariamente deletérea, sino que puede desempeñar importantes funciones e incluso dar lugar a un nuevo patrón de expresión génica que, si es fijado, podría ofrecer una ventaja evolutiva. Así, el análisis, desde el punto de vista evolutivo, del efecto de las transposiciones de los elementos móviles sobre los organismos hospedadores, muestra cómo estos elementos en algunos casos evolucionan coincidentemente con el hospedador, para evitar o mitigar el efecto deletéreo de las inserciones y cómo existe una selección positiva de los elementos que han evolucionado para dar algún beneficio a su hospedador (Kidwell y Lisch, 1997). Es decir, se está empezando a considerar a los elementos transponibles, como fuentes de variabilidad con un papel importante en la evolución de los organismos, como resultado del desarrollo de una especie de relación simbiótica entre los elementos transponibles y su hospedador.

AGRADECIMIENTOS

Los estudios, realizados en nuestro laboratorio, mostrados en el presente trabajo de revisión, han sido financiados por la DGESIC (Plan Nacional I+D, Programa de Biotecnología, CICYT Proyecto nº BIO96-0468 y PGC Proyecto nº PB96-0829). La doctora MC Thomas está financiada por el Fondo de Investigaciones Sanita-

rias, Beca de Perfeccionamiento Instituto Carlos III nº exped. 97/4207, Ministerio de Sanidad. M. Olivares está financiada por el Plan Andaluz de Investigación, Beca predoctoral de FPI, Junta de Andalucía. JL. García-Pérez está financiando con una Beca predoctoral de la Fundación Areces.

BIBLIOGRAFÍA

- ADAMS, J.W., KAUFMAN, R.E., KRETSCHMER, P. J., HARRISON, M., NIENHUIS, A.W. (1980) *Nucleic Acids Res.*, **8(6)**, 613-6128.
- AKSOY, S., WILLIAMS, S., CHANG, S., RICHARDS, F.F. (1990) *Nucleic. Acids Res.* **18**: 785-792.
- ASCH, B.B., ASCH, H.L., STOLER, D.I., ANDERSON, G.R. (1993) *Int. J. Cancer.*, **54**, 813-819.
- BIESSMANN, H., MASON, J.M., FERRY, K., D'HULST, M., VALGEIRSDOTTIR, K., TRAVERSE, K. I., PARDUE, M.L. (1990). *Cell*, **61**: 663-673.
- BOEKE, J. D., CORCES, V. G. (1989) *Annu. Rev. Microbiol.*, **43**, 403-433.
- BRATTHAUER, G. L. CARDIFF, R. D., FANNING, T. G. (1994) *Cancer* **73(9)**, 2333-2336
- BRATTHAUER, G. L., FANNING, T. G. (1992) *Oncogene* **7**, 507-510.

- BRATTHAUER, G. L., FANNING, T. G. (1993) *Cancer* **71**, 2383-2386.
- BRITTEN, R. J., DAVIDSON, E. H. (1969) *Science* **165**, 349-358.
- BRITTEN, R. J., DAVIDSON, E. H. (1971) *Q. Rev. Biol.* **46**, 111-138.
- BRITTEN, R. J. (1997) *Gene* **205**, 177-182.
- BUCHETON A. (1990) *Trends Genet.* **6**, 16-21.
- BURKE, W. D., CATALANG, C. C., EICKBUSH, T. H. (1987) *Mol. Cell. Biol.* **7**, 2221-2230.
- BURKE, W. D., MÜLLER, F., EICKBUSH, T. H. (1995) *Nucleic Acids Res.* **23**, 4628-4634.
- BURTON, F.H., LOEB, D.D., VOLIVA, C.F., MARTÍN, S.L., EDGELL, M.H., HUTCHISON, C.A. (1986) *J. Mol. Biol.* **187**, 291-304.
- BURWINKEL, B., KILIMANN, M. W. (1998) *J. Mol. Biol.* **277**, 513-517.
- CASTRO, C., CRAIG, S. P., CASTAÑEDA, M. (1981) *Mol. Biochem Parasitol.* **4**, 273-282.
- CHABOISSIER, M-C., BUCHETON, A., FINNEGAN, D.J. (1998) *Proc. Natl. Acad. Sci.* **95**, 11781-11785.
- CHARLESWORTH, B., SNIEGOWSKI, P., STEPHAN, W. (1994) *Nature* **371**, 215-220.
- CLARE, J., FARABAUGH, P. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2829-2833.
- CHENG, S.M., SCHILDKRAUT, C.L. (1980) *Nucleic Acids Res.* **8**, 4075-4090.
- CLEMENTS, A. P., SINGER, M. F. (1998) *Nucleic Acids. Res.* **26**: 3528-3535.
- COOPER, R., INVERSO, J., ESPINOSA, M., NOGUEIRA, N., CROSS, G. (1991) *Mol. Biochem. Parasitol.* **49**, 45-60.
- COST, G.J., BOEKE, J.D. (1998) *Biochemistry*, **37**, 18081-18093.
- CSINK, A., MCDONALD, J. F. (1990) *Genetics* **126**, 375-385.
- DAWSON, A., HARTSWOOD, E., PATERSON, T., FINNEGAN, D.J. (1997) *EMBO J.*, **16(14)**, 4448-4455.
- DE LANGE, T., MICHELS, P. A. M., VEERMAN, H. J. G., CORMELISSEN, W. C. A., BORST, P. (1984) *Nucleic Acids Res.* **12**, 3777-3790.
- DEMPLE, B., HARRISON, L. (1994) *Annu. Rev. Biochem.* **63**, 915-948.
- DHELLIN, O., MAESTRE, J., HEIDMANN, T. (1997) *EMBO J.* **16**: 6590-6602.
- DOMBROSKI, B., MATHIAS, S. L., NANTHAKUMAR, E., SCOTT, A.F., KAZAZIAN H. H. (1991) *Science* **254**, 1805-1808.
- DONELSON, J. E. (1996) *Curr. Opin. Genet. Dev.* **6**, 699-703.
- DOOLITTLE, R., FENG, D. F., JOHNSON, M. S., MCCLURE, M. A. (1989) *Q. Rev. Biol.* **64**, 1-30.
- EICKBUSH, T. (1992) *The New Biologist* **4**, 430-440.
- ENGELS, W. R. (1989) In: Mobile DNA, eds. Berg, D. E., Howe, M. M. Washington DC: An. Soc. Microbiol.
- EVANS, J. P., PALMITER, R. D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8792-8795.
- FANNIN, T., SINGER, M. (1987) *Nucleic Acids Res.* **15**, 2251-2260.
- FAWCETT, D., LISTER, C. K., KELLETT, E., FINNEGAN, D. J. (1986) *Cell* **47**, 1007-1015.
- FENG, Q., MORAN, J. V., KAZAZIAN, H. H., BOEKE, J. D. (1996) *Cell* **87**, 905-916.
- FINNEGAN, D. J. (1992) *Curr. Opin. Genet. Dev.* **2**, 861-867.
- FRASCH, A.C.C., CARRASCO, A.E., GOIJMAN, S.G., SANCHEZ, D.O. (1983) *Mol. Biochem. Parasitol.* **8**, 227-239.
- FRIDELL, R. A., PRET, A. M., SEARLES, L. L. (1990) *Genes Dev.* **4**, 559-566.
- GABRIEL, A., BOEKE, J. D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9794-9798.
- GABRIEL, A., YEN, T.J., SCHWARTZ, D.C., SMITH, C.L., BOEKE, J.D., SOLLNERWEBB, B., CLEVELAND, D.W. (1990) *Mol. Cell. Biol.* **10**, 615-624.
- GARRET, J., KNUTZON, D. S., CARROL, D. (1989) *Mol. Cell. Biol.* **9**, 3018-3027.
- GONZÁLEZ, A., PREDIGER, E., HUECAS, M.E., NOGUEIRA, N. LIZARDI, P.M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3356-3360.
- GONZÁLEZ, C. I., THOMAS, M.C., MARTÍN, F., ALCAMI, J., ALONSO, C., LÓPEZ, M. C. (1997) *Acta Tropica* **63**, 117-126.
- GONZÁLEZ, C.I., THOMAS, M.C., MARTÍN, F., LÓPEZ, M.C. (1997) *Ars Pharmaceutica*, **38(2-3)**, 221-227.
- GOTTESDIENER, K.M., GORIPARTHI, L., MASUCCI, J.D., VAN DER PLOEG, L.H.T. (1991) *Mol. Cell. Biol.* **11**, 2467-2480.
- HARENDAZ, C. J., JOHNSON, L. F. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2531-2535.
- HASAN, G., TURNER, M. J. AND CORDINGLEY, J. S. (1982) *Nucleic. Acids. Res.* **10**, 6747-6761.
- HASAN, G., TURNER, M. J., CORDINGLEY, J. S. (1984) *Cell* **37**, 333-341.
- HOHJOH, H., SINGER, M. F. (1996) *EMBO J.* **15**, 630-639.
- HOLMES, S.E., SINGER, M.F., SWERGOLD, G.D. (1992) *J. Biol. Chem.*, **267**, 19765-19768.
- HOLMES, S. E. DOMBROWSKI, B. A., KREBS, C. M., BOEHM, C. D., KAZAZIAN, H. H., JR. (1994) *Nature Genet.* **7**, 143-148.
- HUTCHISON, C. A. III, HARDIES, S. C., LOEB, D. D., SEHEE, W. R., EDGELL M. H. (1989) In: Mobile DNA. Eds Berg, D. E., Howe, M. M. Washington D. C.: American Society for Microbiology, pp 593-617.
- INOUE, S., HSU, M-Y., EAGLE, S., INOUYE, M. (1989) *Cell* **56**, 709-717.
- IVANOV, V.A., MELNIKOV, A.A., SIUNOV, A.V., FODOR II, ILLYN, Y.V (1991) *EMBO J.* **10**, 2489-2495.
- JAKUBZAC, J., BURKE, W. D., EICKBUSH, T. H. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 3295-3299.
- JELINEK, W. R. (1982) *Annu. Rev. Biochem.* **51**, 813-844.
- JENSEN, S., HEIDMANN, T. (1991) *EMBO J.* **10**, 1927-1937,
- JI, H., MOORE, D. P., BLOMBERG, M. A., BRAITERMAN, L. T., BOYTAS, D. F., NATSOULIS, G., BOEKE, J. D.

- (1993) *Cell* **73**, 1007-1018.
- JURKA, J. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 1872-1877.
- JURKA, J. (1998) *Curr. Opin. Struct. Biol.* **8**, 333-337.
- KATZIR, N., RECHAVI, G., COHEN, J.B., UNGER, T., SIMONI, F., SEGAL, S., COHEN, D., GIVOL, D. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1054-1058
- KAZAZIAN, H. H., JR, WONG, C., YOUSSEFIAN, H., SCOTT, A.F., PHILLIPS, D.G., ANTONARAKIS, S.E. (1988) *Nature* **332**, 164-166.
- KAZAZIAN, H.H., MORAN J. V. (1998) *Nature Genetics* **19**, 19-24.
- KELLY, J. M. WARD, H. M., MILES, M. A., KENDALL, G. (1992) *Nucleic Acids Res.* **20**, 3963-3969.
- KIDWELL M. G., LISCH D. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 7704-7711.
- KIM, A., TERZIAN, C., SANTAMARÍA, P., PELISSON, A., PRUD'HOMME, N., BUCHETON, A. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 1285-1289.
- KIMMEL, B. E., OLE-MOYROI, O. K., YOUNG, J. R. (1987) *Mol. Cell. Biol.* **7**, 1465-1475.
- KUNCE, R. (1996) *Curr. Top. Microbiol. Immunol.* **204**, 161-194.
- KUROSE, K., HATA, K., HATTORI, M., SAKAKI, Y. (1995) *Nucleic Acids Res.* **23(18)**, 3704-3709.
- LAKE, J. A., DE LA CRUZ, P. C. G., MOREL, C., SIMPSON, L. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 4779-4783.
- LANAR, D. E., LEVY, S. C., MANNING, J. E. (1981) *Mol. Biochem. Parasitol.* **3**, 327-341.
- LIM, D., MASS, W. K. (1989) *Cell* **56**, 891-904.
- LODES, M. J., SMILEY, B. J., STADNYK, A. W., BENNET, J. L., MYLER, P. J., STUART, K. (1993) *Mol. Cell. Biol.* **13**, 7036-7044.
- LUAN, D., KORMAN, M. H., JAKUBCZAK, J. L., EICKBUSH, T. (1993) *Cell* **72**, 595-605.
- LUPSKI, J. R., WEINSTOCK, G. M. (1992) *J. Bacteriol.* **174**, 4525-4529.
- MARSCHALEK, R., HOFMANN, J., SCUMANN, G., GOSSERINGER, R., DINGERMANN, T. (1992) *Mol. Cell. Biol.* **12**, 229-239.
- MARAÑÓN C., PUERTAS, C., LÓPEZ, M.C. (1997) *Ars Pharmaceutica*, **38(2-3)**, 229-235.
- MARAÑÓN, C., PUERTAS, C., ALONSO, C. LÓPEZ, M.C. (1998) *Mol. Biochem. Parasitol.*, **92**, 313-324.
- MARAÑÓN, C., THOMAS, M.C., PUERTAS, C., ALONSO, C. LÓPEZ, M.C. (1999) Sometido a publicación.
- MARTÍN, F., MARAÑÓN, C., OLIVARES, M., ALONSO, C., LÓPEZ, M. C. (1995) *J. Mol. Biol.* **247**, 49-59.
- MARTÍN, F., OLIVARES, M., ALONSO, C., LOPEZ, M. C. (1996) *TIBS* **21**, 283-285.
- McCLINTOCK, B. (1951) *Cold Spring Harbor Sym. Quant. Biol.* **16**, 13-47.
- MATHIAS, S. L., SCOTT, A. F., KAZAZIAN, H.H. BOEKE, J. D., GABRIEL, A. (1991) *Science* **254**, 1808-1810.
- MCLEAN, C. FINNEGAR, D. (1991) *J. Cell. Biochem.* **15D**, 74.
- MC CLURE, M. (1991) *Mol. Biol. Evol.* **8**, 835-856.
- MC DONALD, J. F. (1995) *Trends Ecol. Evol.* **100**, 123-126.
- MICHEL, F., LANG, B. F. (1985) *Nature* **316**, 641-643.
- MIKI, Y., NISHISHO I., HORII, A., MIYOSHI, Y., UTSUNOMIYA, J., KINZLER, KW., VOGELSTEIN, B., NAKAMURA, Y. (1992) *Cancer Res.* **52**, 643-645.
- MINCHIOTTI, G., DINOCERA, P.P.(1991) *Mol. Cell. Biol.* **11**, 5171-5180.
- MIZROKHI, L., GEORGIEVA, S. G., LLYIN, Y. V. (1988) *Cell* **54**, 685-691.
- MOORE, J. K., HABER, J. E. (1996) *Nature* **383**, 644-646.
- MORAN, J., HOLMES, S.E., NAAS, T. P., DEBERARDINIS, R. .., BOEKE, J.D., KAZAZIAN, JR., H.H. (1996) *Cell* **87**, 917-927.
- MORGAN, J. T. (1995) *J. Mol. Biol.* **248**, 812-823.
- MORSE, B. ROTHERG, PG., SOUTH, VJ., SPANDORFER, JM., ASTRIN, SM. (1988) *Nature* **333**, 87-90.
- MOUNT, S. M., RUBIN, G. M. (1985) *Mol. Cell. Biol.* **5**, 1630-1638.
- MURPHY, N. B PAYS A, TEBABI P, COQUELET H, GUYAUX M, STEINERT M, PAYS E. (1987) *J. Mol. Biol.* **195**, 855-871.
- NARITA, N., NISHIO H, KITOH Y, ISHIKAWA Y, ISHIKAWA Y, MINAMI R, NAKAMURA H, MATSUO M. (1993) *J. Clin. Invest.* **91**, 1862-1867.
- NUR, I., PASCALE, E., FURANO, A. V. (1988) *Nucleic. Acids. Res.* **16**, 9233-9251.
- O'HARE, K., ALLEY, M. R. K., CULLINGFORD, T. E., DRIVER, A., SANDERSON, M. J. (1991) *Mol. Gen. Genetic.* **225**, 17-24.
- OKADA, N., OHSHIMA, K. (1995). In: R. J. Maraia (Ed.), *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*. R. G. Landes Company, Austin: 61-79.
- OKADA, N., HAMADA, M., OGAWA, I., OHSHIMA, K. (1997) *Gene* **205**, 229-243.
- OKASAKI, S., ISHIKAWA, H., FUJIWARA, H. (1995) *Mol. Cell. Biol.* **15**, 4545-4552.
- OLIVARES, M., ALONSO, C., LOPEZ, M. C. (1997a) *J. Biol. Chem.* **272**, 25224-25228.
- OLIVARES, M., GONZÁLEZ, C.I., THOMAS, M.C., PORRAS, R., MARAÑÓN, C., GARCÍA-PÉREZ, J.L., MARTÍN, F., LOPEZ, M.C. (1997b) In: *Avances en cáncer, Jornadas sobre investigación en medicina*. Unidad Mixta de Investigaciones Médicas, 86-97.
- OLIVARES, M., MARTÍN, F., MARAÑÓN, C., LÓPEZ, A., LOPEZ, M. C. (1997c) *Ars Pharmaceutica*, **38(2-3)**, 209-220.
- OLIVARES, M., THOMAS, M.C., ALONSO, C., LOPEZ, M. C. (1999), en prensa. *J. Biol. Chem.*

- PARSONS, M., NELSON, R. G., WATKINGS, K. P., AGABIAN, N. (1984) *Cell* **38**, 309-316.
- PAYS, E., COQUELET, H., PAYS, A., TEBABI, P., STEINERT, M (1989) *Mol. Cell. Biol.* **9**, 4018-4021.
- (PAYS, E., STEINERT, M. (1988) *Annu. Rev. Genet.* **22**, 107-126.
- PEAR, W. S., NELSON, S.F., AXELSON, H., WAHLSTROM, G., BAZIN, H., KLEIN, G., SUMEGI, J. (1988) *Oncogen* **2**, 499-507
- PIMPINELLI, S., BERLOCO, M., FANTI, L., DIMITRI, P., BONACCORSI, S., MARCHETTI, E., CAIZZI, R., CAGGESSE, C., GATTI, M. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 3804-3808.
- PUERTA, C., MARTÍN, J., ALONSO, C., LÓPEZ, M. C. (1994) *Mol. Biochem. Parasitol.* **64**, 327-336.
- PURUGGANAN, M., WESSLER, S. (1992) *Genetic*. **86**, 295-303.
- REQUENA, J.M., JIMENEZ-RUIZ A., SOTO, M., LOPEZ M.C., ALONSO, C. (1992) *Mol. Biochem. Parasitol.* **51**, 271-280.
- REQUENA J.M., MARTÍN, F., SOTO, M., LÓPEZ, M.C., ALONSO, C. (1994) *Gene* **146**, 245-250.
- REQUENA J.M., SOTO, M., ALONSO, C. (1993) *Biol. Res.* **26**, 11-18.
- ROGERS, J. H. (1985) *Int. Rev. Cytol.* **93**, 231-279.
- RUSHFORTH, A., Mc ANDERSON, P. (1996) *Mol. Cell. Biol.* **16**, 422-429.
- SANDMEYER, S. B., HANSEN, L. J., CHALKER, D. L. (1990) *Annu. Rev. Genet.* **24**, 491-518.
- SANMIGUEL, P., TIKHONOV, A., JIN, YK., MOTCHOULSKAIA, N., ZAKHAROV, D., MELAKE-BERHAN, A., SPRINGER, PS., EDWARDS, KJ., LEE, M., AVRAMOVA, Z., BENNETZEN, JL. (1996) *Science* **274**, 765-768.
- SCHMID, C.W., MARAIA, R. (1992) *Curr. Opin. Genet. Dev.* **2**, 874-882.
- SCHMID, C.W. (1998) *Nucleic Acids Res.*, **26**, 4541-4550.
- SINGER, M. F., KREK, V., MCMILLAN, JP., SWERGOLD, GD., THAYER, RE. (1993) *Gene* **135**, 183-188.
- SKOWRONSKI, J., SINGER, M. F. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6050-6054.
- SKOWRONSKI, J., FANNING, T. G., SINGER, M. F. (1988) *Mol. Cell. Biol.* **8**, 1385-1397.
- SLOOF, P. BOS, J.L., KONINGS, A.F., MENKE, H.H., BORST, P., GUTTERIDGE, W.E., LEON, W. (1983) *J. Mol. Biol.* **167**, 1-21.
- SMIT, A. F. A., RIGGAS, A. D. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 1443-1448.
- SONG, S.U., GERASIMOVA, T. I., KURKULOS, M., BOEKE, J. D., CORCES, V.G. (1994) *Genes Dev.* **8**, 2046-2057.
- SONG, S.U., KURKULOS, M., BOEKE, J.D., CORCES, V.G. (1997) *Development* **124**, 2789-2798
- SWERGOLD, G. (1990) *Mol. Cell. Biol.* **10**, 6718-6729.
- TENG, S-C., KIM, B., GABRIEL, A. (1996) *Nature* **383**, 641-644.
- TOH, H, HAYASHIDA, H., MIYATA, T. (1983) *Nature* **305**, 827-829.
- THUMMEL C. S. (1989) *Genes Dev.* **3**, 782-792.
- TU, Z. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 7475-7480.
- VASELLA, E., RODITI, I. BRAUN, R. (1996) *Mol. Biochem. Parasitol.* **82**, 131-135.
- VAZQUEZ, M. P., SCHIJMAN, A. G., LEVIN, M.J. (1994) *Mol. Biochem. Parasitol.* **64**, 327-336.
- VIEIRA DE ARRUDA, M. REINACH, F.C., COLLI, W., ZINGALES, B. (1990) *Mol. Biochem. Parasitol.* **40**, 35-42.
- VILLANUEVA, M. S. WILLIAMS, S. P., BEARD, C. B., RICHARDS, F.F., AKSOY, S. (1991) *Mol. Cell Biol.* **11**, 6139-6148.
- WANG, H., SEEGER, C. (1993) *J. Virol.*, **67**, 6507-6512.
- WEINER, A. M., DEININGER, P. L., EFSTRATIADIS, A. (1986) *Annu. Rev. Biochem.* **55**, 631-661.
- WESSLER, S., BUREAU, T. E., WHITE, S. E. (1995) *Curr. Opin. Genet. Dev.* **5**, 814-821.
- WINSLOW, G. M. HAWASHI, S. KRASNOW, M. HOHNESS, D. S. SCOTT, M. P. (1989) *Cell* **57**, 1017-1030.
- WOODS-SAMUELS, P., WONG, C., MATHIAS, S.L., SCOTT, A.F., KAZAZIAN, H.H., ANTONARAKIS, S.E. (1989) *Genomics* **4**, 290-296.
- (XIONG, Y., EICKBUSH T. H. (1990) *EMBO J.* **9**, 3353-3362.
- YANG, Z., BOFFELLI, D., BOONMARK, N., SCHWARTZ, K., LAWN, R. (1998) *J. Biol. Chem.* **273**, 891-897.
- YODER, J. A., WALSH, C. P., BESTOR, T. H. (1997) *Trends Genet.* **13**, 335-340.
- YOUSSOUFIAN, H., LODISH, H. F. (1993) *Mol. Cell. Biol.* **13**, 98-104.