

Identificación de proteínas reguladoras de la expresión génica en tripanosomátidos

Elizabeth Ruiz¹, Cesar A. Ramírez¹, Paola Nocua¹, José María Requena², Concepción J. Puerta^{1,*}

¹ Laboratorio de Parasitología Molecular, Departamento de Microbiología, Facultad de Ciencias, Pontificia Universidad Javeriana, Bogotá, Colombia.

² Centro de Biología Molecular Severo Ochoa (CSIC-UAM), Universidad Autónoma de Madrid, Madrid, España.

Artículo de posesión para el ingreso como miembro correspondiente a la Academia Colombiana de Ciencias Exactas, Físicas y Naturales el 8 de Noviembre de 2018

Resumen

Los tripanosomátidos son parásitos causantes de patologías de reconocido impacto en salud pública como la enfermedad de Chagas, la enfermedad del sueño y la leishmaniasis. Estos microorganismos divergieron tempranamente de la línea evolutiva de los eucariotas y se caracterizan por poseer mecanismos peculiares de regulación génica finamente orquestados, tan eficaces que han asegurado su transmisión al permitirles adaptarse a ambientes inhóspitos y dispares como los de sus huéspedes invertebrados y mamíferos. Como consecuencia de su peculiar organización genómica, los tripanosomátidos han apostado por regular la expresión de sus genes a través de mecanismos posteriores a la transcripción, mediados principalmente por la acción de proteínas de unión a ARN (*RNA-binding proteins*, RBP), que reconocen su mensajero blanco gracias a la presencia de elementos reguladores en *cis* y se asocian con el ARN formando complejos ribonucleoprotéicos. De esta manera, las células establecen redes reguladoras en las que una misma RBP puede actuar sobre centenares de ARN mensajeros y el destino de cada uno de estos es dictado por la combinación de RBP con las que interactúa. Si bien mediante herramientas de bioinformática se han predicho cerca de un centenar de proteínas con capacidad de unión al ARN en tripanosomátidos, son pocas las que se han caracterizado y, sin duda, son muchas las que están aún por descubrir. En este artículo, se presentan las estrategias seguidas para la identificación y caracterización de proteínas reguladoras de la expresión génica en tripanosomátidos durante la última década en nuestro grupo de investigación, especialmente de las proteínas RBP directamente implicadas en la regulación posterior a la transcripción de los genes *HSP70* de *Leishmania braziliensis*. © 2018. Acad. Colomb. Cienc. Ex. Fis. Nat.

Palabras claves: Estrategia; Expresión génica; Proteínas de unión a ARN; Regulación génica; Tripanosomátidos.

Identification of proteins regulating gene expression in trypanosomatids

Abstract

Trypanosomatids are parasites that cause pathologies with renowned impact on public health such as Chagas disease, the sleeping sickness, and leishmaniasis. These eukaryotes are characterized by having diverged early from their evolutionary path developing regulatory mechanisms that are efficient and finely orchestrated. Mechanisms which have ensured their transmission by allowing their adaptation to inhospitable and disparate environments such as those encountered in their invertebrate and mammal hosts. As a consequence of their peculiar genome organization, trypanosomatids have opted for regulating their gene expression mainly through post-transcriptional mechanisms, which are mediated through the action of RNA-binding proteins (RBP). These proteins recognize trypanosomatids target messengers due to the presence of *cis* elements and they link to the corresponding RNA forming ribonucleoprotein complexes. Thus, cells establish regulatory networks where a single RBP can act over hundreds of RNA messengers, and the destiny of any given RNA is dictated by the combination of the RBP with which it interacts. Around 100 RNA-binding proteins have been predicted by bioinformatics tools in trypanosomatids, but few of them have been characterized and there is no doubt that many are to be discovered still. In this article, we present the strategies used for the identification and characterization of gene-expression regulatory proteins in trypanosomatids over the past decade in our research group, particularly of RBPs directly implicated in the post-transcriptional regulation of the *HSP70* genes of *Leishmania braziliensis*. © 2018. Acad. Colomb. Cienc. Ex. Fis. Nat.

Key words: Gene regulation; Pipeline; RNA binding protein; Trypanosomatids.

*Correspondencia:

Concepción J. Puerta, cpuerta@javeriana.edu.co, **Recibido:** 15 de junio de 2018, **Aceptado:** 25 de julio de 2018, **Editor:** John Mario González

Introducción

Los tripanosomátidos son parásitos protozoos flagelados de importancia médica que se transmiten al hombre y a los animales a través de insectos vectores y causan diversas enfermedades de reconocido impacto en salud pública como la enfermedad de Chagas, la enfermedad del sueño y la leishmaniasis. Estos microorganismos eucariotas se caracterizan por haber divergido tempranamente de la línea evolutiva, de tal manera que poseen mecanismos de regulación génica tremendamente eficaces que pueden calificarse como *sui generis*, y que les han permitido enfrentar los sistemas de defensa de sus huéspedes y completar su complejo ciclo biológico tras millones de años. Así, para lograr adaptarse a ambientes tan disímiles como los de sus huéspedes invertebrados y mamíferos, adoptan a lo largo de su ciclo de vida diversas formas que son el resultado de una orquestada modulación de la expresión de sus genes, lo cual redundo no solo en la adquisición de una morfología propia, sino de un potencial bioquímico y una capacidad de adaptación diferentes (Puerta, 2007). A diferencia de lo que ocurre en la mayoría de eucariotas, la expresión génica en los tripanosomátidos se regula después de la transcripción gracias, precisamente, a la organización génica peculiar que presentan (Clayton, 2016). Sus cromosomas contienen grandes agrupaciones génicas, algunas conformadas por más de un centenar de genes, a menudo sin ninguna relación funcional, pero sí con la misma orientación de la transcripción. Estas agrupaciones se denominan unidades de grupos de genes policistrónicos (*polycistronic gene clusters*, PGC) y están diseñadas para ser transcritas como grandes policistrones (Berriman, et al., 2005; El-Sayed, et al., 2005; Ivens, et al., 2005). En cada cromosoma existen varias unidades PGC; las unidades de transcripción vecinas tienen una orientación convergente o divergente y están separadas por las denominadas regiones de cambio de hebra. Una vez se genera el transcrito policistrónico, debe ser procesado y convertido a ARNm monocistrónicos antes de ser traducido, lo cual es mediado por dos procesos íntimamente acoplados: el procesamiento en *trans* o *trans-splicing* y la poliadenilación (Requena, 2011). En el primero, se adiciona en el extremo 5' una secuencia conocida como *spliced leader* (SL) de 39 nucleótidos, la cual ya porta la caperuza o estructura cap (cap-4 en tripanosomátidos) propia de los ARN mensajeros, cuya función, al igual que sucede en el resto de eucariotas, es proteger los ARNm y facilitar su interacción con la maquinaria de traducción (Agabian, 1990; Bangs, et al., 1992). Por otra parte, en el proceso de poliadenilación se adiciona una cola de adeninas en el extremo 3' de los mensajeros tras el corte endonucleotídico, la cual, en general, le otorga estabilidad a la molécula (LeBowitz, et al., 1993).

La peculiar organización génica y la transcripción policistrónica de estos organismos explican por qué la tasa de transcripción es similar en la mayoría de sus genes. En consecuencia, dada la pérdida de este punto de control, los

tripanosomátidos han apostado por regular la expresión de sus genes a través de mecanismos posteriores a la transcripción, entre ellos el procesamiento del pre-ARNm y el transporte, la localización, la estabilidad y la traducción de los mensajeros monocistrónicos maduros (Araujo & Teixeira, 2011). Todos estos procesos están mediados por la acción de proteínas de unión a ARN (*RNA binding protein*, RBP), que reconocen su mensajero blanco gracias a la presencia de elementos *cis* localizados en las sus regiones 5' y 3' no traducidas (*untranslated region*, UTR) y se asocian con dicho ARN formando un complejo ribonucleoproteico cuya composición va a dictar el destino de los mensajeros (Cassola & Frash, 2009; Romaniuk, et al., 2016). De esta manera, las RBP desempeñan un papel central en la regulación de la expresión génica, y ya que no puede haber una RBP específica para cada transcrito, las células han establecido redes reguladoras en las que una misma RBP puede actuar sobre centenares de ARN mensajeros (regulones u operones de ARN), en tanto que el destino de cada ARN es dictado por la combinación de RBP con las que interactúa (Lapointe, et al., 2015; Muller-McNicoll & Neugebauer, 2013).

Si bien ya se conocen cerca de un centenar de proteínas con capacidad de unión al ARN en tripanosomátidos (Kramer & Carrington, 2011; Romaniuk, et al., 2016), sin duda son muchas las que están aún por descubrir. Dada su divergencia evolutiva, la mayoría de las proteínas de estos parásitos carecen de ortólogos reconocibles en otros eucariotas. Además, muchas RBP carecen de dominios de unión a ARN típicos. En este artículo, se presentan las estrategias seguidas para la identificación de proteínas reguladoras de la expresión génica en tripanosomátidos durante la última década en el Laboratorio de Parasitología Molecular del Departamento de Microbiología de la Facultad de Ciencias de la Pontificia Universidad Javeriana, trabajos que se han desarrollado en colaboración con el Laboratorio de Regulación de la Expresión Génica en *Leishmania* del Centro de Biología Molecular "Severo Ochoa" de la Universidad Autónoma de Madrid. Entre los estudios llevados a cabo, se describen con mayor detalle y consignando resultados ilustrativos aquellos dedicados a la identificación y caracterización de proteínas RBP directamente implicadas en la regulación posterior a la transcripción de los genes *HSP70* de *Leishmania braziliensis*.

Concepto y punto de partida

Aunque la investigación propuesta es de carácter básico y su objetivo es aumentar nuestro conocimiento sobre los mecanismos de control de la expresión génica en tripanosomátidos desentrañando las redes de interacción entre las RBP y sus mensajeros blancos, no podemos olvidar que estos organismos son patógenos que actúan en humanos. Por ello, como punto de partida se propuso el estudio de la expresión de proteínas críticas para el parásito o de factores de virulencia, cuyos mensajeros (más concretamente, las

regiones UTR) se usaron como cebos para la captura de las RBP. En este sentido, el primer paso del estudio consistió en la elección del gen y del parásito a estudiar. Si bien en nuestro grupo se ha trabajado con otros tripanosomátidos, se escogió *Leishmania* como el tripanosomátido objeto de estudio por las siguientes razones: *Leishmania* es el agente causal de la leishmaniasis, parasitosis asociada a la pobreza con un amplio espectro de manifestaciones clínicas que van desde lesiones cutáneas autocurativas (LC), pasando por lesiones mucocutáneas (LMC) causantes de desfiguración hasta afectaciones viscerales (LV), siendo estas últimas fatales si no se las trata adecuadamente (Bañuls, et al., 2007). Se estima que aproximadamente 12 millones de personas sufren actualmente leishmaniasis en un centenar de países de todo el mundo, y al menos 500.000 nuevos casos de LV y 1,5 millones de casos de LC grave se reportan anualmente (Alvar, et al., 2012). En Latinoamérica, la leishmaniasis es endémica en 22 países, y *Leishmania (Viannia) braziliensis* es el principal agente causante de la forma mucocutánea. En Colombia, la leishmaniasis es endémica en casi todo el territorio nacional, y *L. braziliensis* es una de las seis principales especies causantes de leishmaniasis cutánea y mucocutánea en el país (Instituto Nacional de Salud - INS, 2011).

Una vez elegida *L. braziliensis*, se pensó en estudiar un factor clave como es la proteína de choque térmico de 70 kDa (HSP70) con base a las siguientes características: (i) durante la transición de promastigote (insecto-vector) a amastigote (huésped mamífero), se requiere de la actividad chaperona de las proteínas de choque térmico, fundamentalmente de la HSP70 (Requena, 2012); (ii) a partir de un determinado gen del locus *HSP70* (el *HSP70-II*) de *Leishmania infantum* la expresión de dicha proteína es necesaria para la supervivencia de los promastigotes en fase estacionaria y de los amastigotes tanto *in vitro* como *in vivo* (Folgueira, et al., 2008); (iii) la depleción del gen *HSP70-II* disminuye la capacidad infectiva del parásito y, por ende, la enfermedad que ocasiona según diferentes modelos animales (Carrion, et al., 2011; Folgueira, et al., 2008), y (iv) induce una fuerte respuesta inmunitaria humoral y celular (Requena, et al., 2012).

El factor de virulencia

Dado que las RBP gobiernan la regulación posterior a la transcripción a través de su interacción con elementos en *cis*, localizados principalmente en las regiones 5' y 3' UTR de los mensajeros blanco (Kramer & Carrington, 2011), en la primera etapa del desarrollo de la propuesta experimental, es necesario conocer la arquitectura de los genes y de los mensajeros del factor de virulencia seleccionado.

En este caso, se encontró que *L. braziliensis*, al igual que otras especies del género *Leishmania*, posee dos tipos de genes *HSP70*, el *HSP70-I* y el *HSP70-II*, los cuales se organizan en el cromosoma 28 en una agrupación conformada por seis copias de los genes *HSP70-I* y una de los *HSP70-II* (Ramírez, et al., 2011). Cabe mencionar que estos genes se diferencian solamente en la secuencia y la longitud de su región 3' UTR (Ramírez, et al., 2011), pero dichas diferencias han resultado fundamentales en los mecanismos moleculares de adaptación del parásito, pues le permiten tolerar el estrés térmico al que es sometido durante su ciclo de vida: 26 °C en el insecto vector y 35 a 37 °C en el huésped mamífero (Folgueira, et al., 2005). Para definir y caracterizar la región 5' UTR, común a ambos mensajeros, así como la 3' UTR específica para cada uno de ellos, se hicieron ensayos de RT-PCR a partir del ARN de formas promastigotas del parásito; a continuación, los ADN complementarios (ADNc) obtenidos se clonaron y secuenciaron (Ramírez, 2012). Para el caso de la región 5' UTR (número de acceso al GenBank JF449363, posiciones 1-198), se encontró que, además de la secuencia SL de 39 nucleótidos (nt), esta comprende los 198 nt corriente arriba del codón de inicio de la traducción. Se observó que las regiones 3' UTR abarcan 936 (UTR-I, número de acceso al GenBank JF449364, posiciones 143-1078) o 932 (UTR-II, número de acceso al GenBank JF449365, posiciones 143-1074) nucleótidos corriente abajo del codón de parada de los genes respectivos (Figura 1). Teniendo en cuenta el papel de los elementos *cis* presentes en las UTR como mediadores de la interacción ARN-proteína, se procedió a

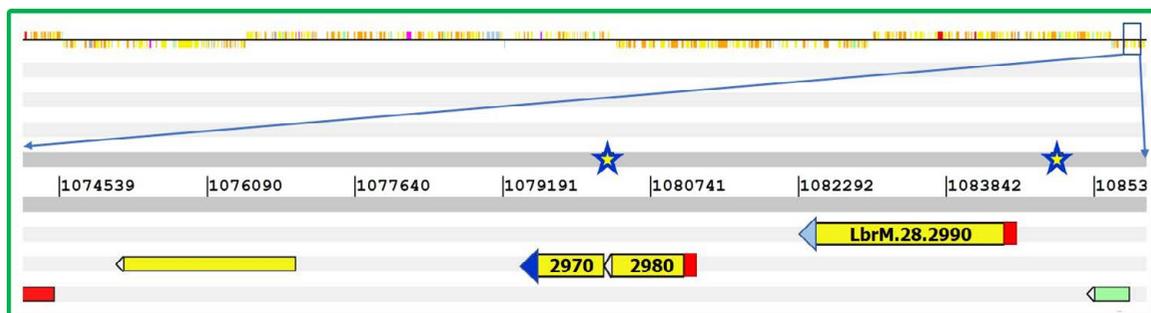


Figura 1. Ubicación de las regiones UTR determinadas para los genes *HSP70* sobre el cromosoma 28 en el genoma publicado de *L. braziliensis*. El gen LbrM.28.2990 (2990), codificante para la proteína HSP70, es portador de la región 3' UTR-I (triángulo azul claro) mientras que el gen LbrM.28.2970 (2970), también codificante para la proteína HSP70, tiene la región 3' UTR-II (triángulo azul oscuro). Los genes LbrM.28.2990 y LbrM.28.2980 (2980) comparten la región 5' UTR. Cada estrella simboliza un hueco en la secuencia genómica actualmente disponible en la base de datos (GeneDB.org).

la identificación de secuencias que potencialmente pudieran funcionar como tales elementos. Para ello, se implementó una aproximación que combinó hallazgos previos consignados en la literatura con el análisis computacional de secuencias conservadas y la predicción de la estructura secundaria del ARN mediante la cual se identificaron dos secuencias ricas en pirimidinas sobre la 3' UTR-I (Ramírez, 2012). Una de ellas, ubicada en la posición 808-842 (5' GCGCTCCCCCTCCTGCCGCCTGTCTTCCTCCCCTT 3'), contiene tramos de cadena sencilla y tiene un 53,8 % de identidad con un elemento rico en pirimidinas de la región 3' UTR del gen *HSP83* de *L. amazonensis* relacionado con la traducción preferencial de estos genes en condiciones de choque térmico (David, et al., 2010). Otro elemento *cis* putativo (5' ACCGACTCTCTTTCCTCCCCCTACCCGTA 3') se encontró en una región altamente conservada entre los nt 489 y 605 de la 3' UTR-I, posición que concuerda con la localización de una secuencia reguladora identificada en la región 3' UTR de los genes *HSP70-I* de *L. infantum*, asociada con la acumulación de estos mensajeros en respuesta al choque térmico (Quijada, et al., 2000). Al hacer la predicción de la estructura secundaria de estas regiones en *L. braziliensis*, se observó que el tramo rico en pirimidinas presentó una identidad del 89,7 % con la misma región en *L. infantum* y se ubicó en una zona estructurada en forma de cadena sencilla (Ramírez, 2012). La conformación del ARN en cadena sencilla es una condición importante para que las proteínas de unión a ARN puedan interactuar con este, dado que cuando el ARN adopta una estructura de doble hélice, el surco mayor es demasiado estrecho e impide el establecimiento de puentes de hidrógeno entre las bases nitrogenadas del ARN y los residuos de aminoácidos de la proteína implicada (Baumeister, 1997). Por su parte, en la 3' UTR-II del mensajero de la *HSP70* se encontraron dos motivos, ambos en regiones de cadena sencilla (Nocua, et al., 2017; Ramírez, 2012). Uno de ellos, localizado en la posición 252-302, tiene forma de horquilla y se relaciona con los elementos ARE, conocidos por desestabilizar a los ARNm que los contienen (Di Noia, et al., 2000): 5' ATGTCTTTTATTTTTTTGTGTGTGTTTATATTTTTCTCCTTCGTAATAA 3'. El otro motivo es una región rica en citosinas en las posiciones 338-361 (5' CCCCTCCTCCCCTCTCCCCCTCCC 3'). Según la literatura, las proteínas que se unen a secuencias ricas en citosinas participan en el control posterior a la transcripción de la expresión génica y tienen un amplio rango de funciones que van desde la estabilización del ARNm y activación de la traducción hasta su silenciamiento (Makeyev & Liehaber, 2002).

Detección de las proteínas

Para identificar las RBP asociadas con la regulación de la expresión del gen seleccionado, se utiliza un ensayo de *pull-down* o pesca en el cual las regiones UTR de los mensajeros de interés, en este caso los que codifican para

la proteína *HSP70*, se usan como cebos para capturar las proteínas RBP (Ramírez, et al., 2013). Para ello, las regiones UTR previamente identificadas deben clonarse en un vector apropiado para dirigir la transcripción *in vitro* y así obtener una cantidad suficiente de moléculas de ARN que servirán como cebos en el ensayo. Como molde para la transcripción *in vitro* también se pueden utilizar las regiones UTR amplificadas con un oligonucleótido quimérico hacia adelante que contenga en su extremo 5' la secuencia del promotor, el cual se usará seguido de una secuencia complementaria de la UTR. Una vez transcritas las UTR, deben ser unidas a un soporte sólido antes de proceder con el siguiente paso experimental. Para ello, se utilizan perlas magnéticas recubiertas de estreptavidina, a las cuales se une el cebo de ARN mediante un oligonucleótido puente marcado con biotina cuya secuencia es complementaria de uno de los extremos del cebo. En este caso, en el experimento de captura con la región 5' UTR se utilizó la secuencia complementaria de la región SL marcada con biotina en el extremo 3' como oligonucleótido puente (Figura 2) y, para las regiones 3' UTR, un oligo d(T) marcado en su extremo 5'. Paralelamente, se prepararon extractos de proteínas totales del parásito a partir de los cuales se pescaron las RBP de interés. En este estudio se usaron extractos proteicos provenientes de lisados de las formas promastigotas de la cepa MHOM/BR/75/M2904 de *L. braziliensis* cuyo contenido proteico se verificó cualitativamente mediante electroforesis en gel de poliacrilamida con dodecil sulfato de sodio (*sodium dodecyl sulfate polyacrylamide gel electrophoresis*, SDS-PAGE), y cuantitativamente con el método de Bradford, antes de usarlas. A continuación, cada cebo de ARN del complejo de unión se incubó con la mezcla de proteínas del parásito; en el medio de incubación es preciso añadir inhibidores tanto de ARNasas como de proteasas para evitar la degradación de los componentes de la mezcla, y de ARN competidores (como ARNt de levadura, ARN de *Escherichia coli* o de células Vero), para disminuir uniones inespecíficas de proteínas a las moléculas cebo. Posteriormente, los complejos de proteína y ARN cebo se recuperaron mediante atracción magnética ejercida sobre las perlas, y se lavaron varias veces para retirar las proteínas de unión inespecífica. Por último, la resolución de los complejos obtenidos y la identificación de las proteínas que los conforman se hicieron mediante técnicas de separación como la electroforesis en geles bidimensionales o la cromatografía líquida de alta eficiencia (*high performance liquid chromatography*, HPLC), seguidas de análisis de espectrometría de masas (EM). Los datos sobre las huellas peptídicas obtenidas y sus masas se buscan en las bases de datos de proteínas de la secuencia genómica de *L. braziliensis* para identificar las proteínas presentes en los complejos obtenidos. En el caso de la electroforesis en geles bidimensionales se requiere de un paso adicional: la disgregación de las proteínas del complejo, las cuales son luego separadas mediante electroforesis en gel, primero en función de su punto isoeléctrico (pI)

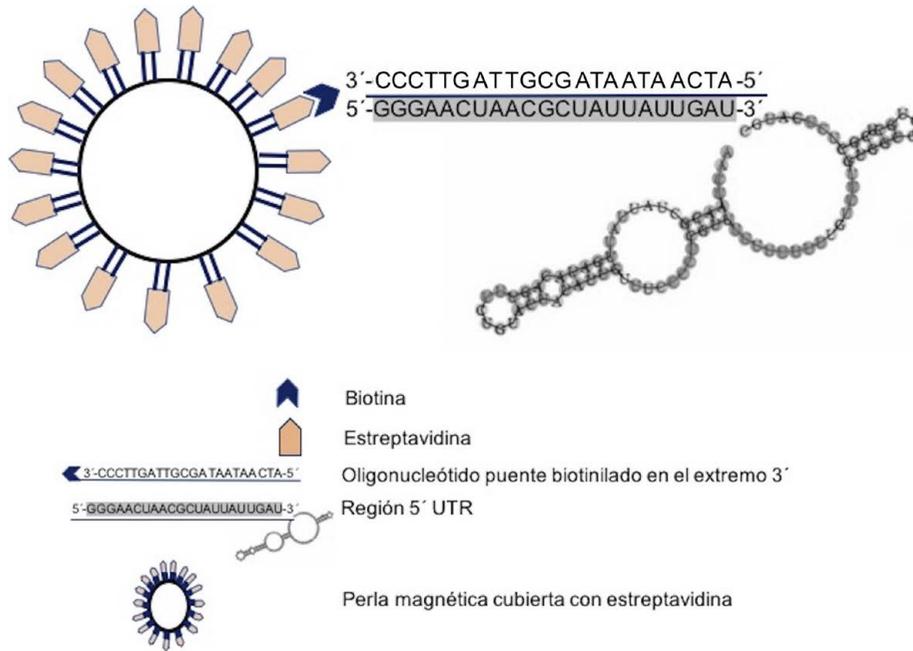


Figura 2. Representación esquemática de la unión de los cebos (5' UTR) a las perlas magnéticas en el ensayo de *pull-down*

y, después, por su tamaño molecular. Una vez obtenidas las bandas correspondientes a las proteínas, estas se analizan mediante espectrometría de masas, lo que en este caso se hizo con las técnicas de MALDI/TOF o MALDI/TOF/TOF. En este procedimiento es importante incluir un cebo de control, es decir, una secuencia irrelevante para los propósitos del ensayo que, por comparación con los resultados de los cebos problema, garantice su especificidad. También es primordial hacer réplicas biológicas de cada cebo de ARN, con el fin de obtener resultados más sólidos y confiables y configurar todos los parámetros para la identificación según los análisis de separación y espectrometría de masas llevados a cabo. En el ejemplo que nos ocupa, se utilizaron como cebo de control las regiones UTR de los mensajeros de la proteína estructural del citoesqueleto de la tubulina alfa y se hicieron dos réplicas biológicas. Los parámetros utilizados para el análisis de las imágenes de los geles y de los espectros de los péptidos se detallan en **Ramírez, et al.** (2013). Además, dado que se ha demostrado que en *L. infantum* los mensajeros del *HSP70-I* se traducen tanto a 26 como a 35 °C, en tanto que los mensajeros del *HSP70-II* se traducen preferencialmente bajo choque térmico (**Folgueira, et al.**, 2005), los ensayos de pesca o *pull-down* se realizaron para cada cebo usando lisados proteicos de parásitos que habían crecido tanto a 26 °C como a 35 °C. De esta forma, se obtuvieron 52 proteínas distintas usando los cebos de las regiones UTR de los mensajeros de la proteína HSP70, todas ellas ausentes en los ensayos de pesca con el cebo de control. De estas, 26 se asociaron exclusivamente con la región 5' UTR, cinco con la región 3' UTR-I y seis con la 3' UTR-II. Además, tan solo cuatro proteínas interactuaron con todos los tres cebos y 11 lo hicieron con dos de ellos: cuatro proteínas con la

5' UTR y la 3' UTR-I, cinco con las 3' UTR-I y II y dos con la 5' UTR y la 3' UTR-I. Estos resultados muestran la especificidad de la unión ARN-proteína, lo cual se explica, muy seguramente, tanto por la secuencia como por la estructura secundaria particular de cada una de las UTR usadas como cebo (**De Gaudenzi, et al.**, 2005). Ahora bien, teniendo en cuenta la longitud de los cebos, fue un tanto sorprendente que la región con mayor número de interacciones fuera la 5' UTR, ya que su longitud es casi cinco veces menor que las de la 3' UTR. Sin embargo, estos hallazgos podrían indicar que la regulación sobre los transcritos de HSP70 opera fundamentalmente a nivel de la iniciación de la traducción. De hecho, entre las proteínas asociadas con la región 5' UTR se encontraron proteínas que participan tanto en el inicio como en la elongación de la traducción. Por otra parte, en concordancia con la temperatura de crecimiento del parásito, la gran mayoría de proteínas (41) se identificaron a los 26 °C (**Ramírez, et al.**, 2013), lo cual podría obedecer, en parte, a un sesgo de la metodología *in vitro*, la cual solo opera para la interacción de las proteínas libres presentes en el lisado y descarta aquellas que participan en los complejos que se forman durante el choque térmico entre las proteínas reguladoras y los mensajeros endógenos de HSP70.

Análisis de las proteínas

Una vez obtenido el listado de proteínas que interactúan con los distintos cebos, es preciso ponderar el valor de su presencia en el proteoma obtenido. Así, las proteínas identificadas inicialmente se analizan según su categoría funcional utilizando la *Kyoto Encyclopedia of Genes and Genomes* (KEGG: <http://www.genome.jp/tools/blast/>) (**Kanehisa, et al.**, 2006) y la *Protein ANalysis THrough Evolutionary Relationships*,

version 7 (Panther: <http://www.pantherdb.org/pathway/>) (Mi, et al., 2010), así como datos previamente reportados en la literatura y en las bases de datos de tripanosómátidos. En nuestro ejemplo, las categorías funcionales más representadas en las 52 proteínas identificadas fueron, como se esperaba, las de metabolismo del ARN (27 %) y las de traducción (7 %), para un total de 34 %, seguidas por las de proteínas citoplasmáticas relacionadas con el metabolismo energético (21 %) y el metabolismo de proteínas (19 %) que, en principio, no guardan relación con los procesos de regulación génica. La presencia de estas proteínas podría explicarse como un ruido de fondo de la metodología, que arrastra proteínas irrelevantes para el estudio gracias a que se encuentran en mayor cantidad en la célula, o a que la condición *in vitro* posibilita interacciones que *in vivo* se encuentran limitadas por factores como la

compartimentación subcelular. Sin embargo, es importante no perder de vista que en numerosos estudios en organismos distintos y alejados a lo largo de la escala evolutiva, se ha descrito la existencia de proteínas pluriempleadas (*moonlight*) que, siendo enzimas metabólicas, son capaces de realizar otras funciones bioquímicas, como unirse al ARN y cumplir un papel en la regulación de la expresión génica, entre otros (Jeffery, 2018). Ahora bien, para efectos de la metodología aquí propuesta, se recomienda que mientras se amplían las bases de proteínas multitarea o *moonlight* (Multitasking Proteins DataBase: <http://wallace.uab.es/multitask/>), debe enfocarse la atención en aquellas relacionadas con el metabolismo del ARN y la traducción. En la Tabla 1 se presentan las 19 proteínas identificadas en estas categorías en nuestro estudio. De estas, es importante descartar las que tienen localización mitocondrial (3), ya que

Tabla 1. Lista de proteínas identificadas con los cebos de las regiones UTR de los mensajeros de la proteína HSP70 pertenecientes a las categorías funcionales del metabolismo del ARN y la traducción

| Identificación en el GeneDB | Cebo | Temperatura (°C) | Función putativa | Clasificación |
|-----------------------------|---------------------------------|------------------|---|--------------------------|
| LbrM.17.0090 | 5' UTR 3' UTR-I | 26 | Factor de elongación EF-1 α | Trasporte de ARN |
| LbrM.34.3010 | 5' UTR | 26 | Helicasa ATP-dependiente, putativa | Actividad ARN helicasa |
| LbrM.35.0270 | 5' UTR 3' UTR-II | 26/35 26 | Factor de elongación 2 | Traducción |
| LbrM.32.0470 | 5' UTR 3' UTR-I 3' UTR-II | 26 | Helicasa ATP-dependiente, putativa | Actividad ARN helicasa |
| LbrM.20.0320 | 5' UTR | 26 | Factor 5 de inicio de la traducción eucariota, putativo | Trasporte de ARN |
| LbrM.32.2400 | 5' UTR | 26 | Subunidad k del factor 3 de inicio de la traducción | Traducción |
| LbrM.20.2140 | 5' UTR | 26 | Proteína hipotética conservada | <i>Splicing</i> del ARNm |
| LbrM.23.0850 | 5' UTR 3' UTR-I | 26/35 35 | Proteína mitocondrial de unión a ARN, putativa | Unión a ARN |
| LbrM.27.1360 | 5' UTR | 35 | Proteína hipotética conservada | Transporte de ARN |
| LbrM.35.2250 | 5' UTR | 35 | Precursor mitocondrial, chaperonina HSP60 | Degradación de ARN |
| LbrM.15.0320 | 5' UTR | 35 | Precursor de la ribonucleoproteína p18 mitocondrial, putativa | Unión a ARN |
| LbrM.25.0080 | 3' UTR-I | 26 | Proteína 3 de unión a poli(A), putativa | Metabolismo del ARN |
| LbrM.01.0740 | 3' UTR-I | 26 | Factor 4A eucariótico de inicio de la traducción, putativo | Transporte de ARN |
| LbrM.25.2210 | 3' UTR-II | 26 | Proteína hipotética (SCD6.10) | Catabolismo de ARN |
| LbrM.03.0170 | 3' UTR-I 3' UTR-II | 26 | Proteína ABC transportadora, putativa | Traducción |
| LbrM.34.4130 | 3' UTR-I 3' UTR-II | 26 | Proteína 2 de unión a poli(A), putativa | Metabolismo de ARN |
| LbrM.17.0700 | 5' UTR | 26 | Proteína hipotética conservada | No se pudo asignar |
| LbrM.31.2460 | 5' UTR | 26 | Proteína hipotética de función desconocida | No se pudo asignar |
| LbrM.30.3080 | 3' UTR-II | 26 | Proteína hipotética conservada | No se pudo asignar |

muy seguramente fueron detectadas gracias a la naturaleza *in vitro* del estudio y no a las condiciones fisiológicas *in vivo*. En cuanto a las restantes, en este estudio se detectaron tres grupos de proteínas: las previamente identificadas en el parásito con un rol ya establecido (2); las identificadas en el parásito con funciones putativas asignadas (7), y las hipotéticas, cuya existencia y función en el parásito se desconocen (7). En el presente caso, en el primer grupo se identificaron dos componentes del proceso de elongación de la traducción: el factor de elongación 1 α (LbEF1 α , LbrM.35.0270), asociado a las regiones 5' UTR y 3' UTR-I a 26°C y, el factor de elongación 2 (LbEF2, LbrM.35.0270), aislado por su unión tanto a la 5' UTR como a la 3' UTR-II a 26°C. La LbEF1 α es una proteína abundante que participa en otros procesos celulares como el transporte nuclear, el control de calidad de las proteínas y la degradación concomitante con la traducción, entre otros (Mateyak & Kinzy, 2010; Mittal, *et al.*, 2013). La LbEF2, con un 61 % de identidad con la proteína ortóloga humana y con la de *Saccharomyces cerevisiae*, se conoce por su función en la translocación de los ribosomas a lo largo del ARNm durante el proceso de traducción (Justice, *et al.*, 1998; Noble & Song, 2008). Acerca del segundo y tercer grupo de proteínas, primero debe comprobarse tanto *in vitro* como *in vivo* su interacción con el ARN, para luego confirmar o elucidar su función específica. Por ello, en el presente estudio la atención se centró en estas proteínas, especialmente las hipotéticas.

Teniendo en cuenta que las herramientas bioinformáticas pueden predecir la presencia de dominios o motivos secuenciales o estructurales relacionados con la interacción entre ARN y proteína, estas fueron útiles para seleccionar con cuáles de las siete proteínas hipotéticas continuar el estudio (Nocua, *et al.*, 2014). Para ello, se usaron programas como el InterProScan (<http://www.ebi.ac.uk/Tools/pfa/iprscan/>), el Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) y el Modeller (<http://salilab.org/modeller/>), y se hizo una revisión de la literatura. En el ejemplo que se presenta, gracias a la presencia de los dominios Lsm y NTF2 presentes en familias de proteínas relacionadas con la represión de la traducción, o la acumulación de mensajeros en los cuerpos P y los gránulos de estrés (Lsm) y en el transporte de ARN (NTF2), se seleccionaron las proteínas codificadas por los genes LbrM.25.2210 y LbrM.30.3080 (Nocua, *et al.*, 2017).

Análisis estructural y modelamiento de las proteínas de interés

Con el ánimo de ahondar en la estructura de las proteínas codificadas por los genes LbrM.25.2210 y LbrM.30.3080, se recurrió a diversas predicciones bioinformáticas que, en primer lugar, mostraron que tenían ortólogos tanto en otras especies de leishmanias como de tripanosomas y de otros organismos evolutivamente mucho más distantes, como *S. cerevisiae*, *Xenopus laevis* y *Homo sapiens* (De Gaudenzi, *et al.*, 2005; Kruger, *et al.*, 2013; Nocua, 2017).

Tras estos análisis de homología de secuencia, se consideró conveniente cambiar el nombre de las proteínas codificadas por los genes LbrM.25.2210 y LbrM.30.3080 a LbSCD6 y LbRBP42, respectivamente. Para el caso de LbSCD6, la predicción de dominios permitió localizar en los primeros 79 aminoácidos de la proteína un dominio Lsm, que ha sido implicado en la formación de cuerpos P y gránulos de estrés (Kruger, *et al.*, 2013). En la parte central de la molécula se identificó el dominio DFDF (DFDF-x(6)-F), caracterizado por una alternancia de aminoácidos no polares y cargados, responsables de la interacción con el ARN (Anantharaman & Aravind, 2004). Además, en el extremo carboxilo terminal, se identificaron las cajas FFD (Y-x-K-x(3)-FFD-x-[IL]-S) y TFG ([RKH]-x(2-5)-Ex(0-2)-[RK]-x(3-4)-[DE]-TFG), y, flanqueando el dominio DFDF y la caja TFG, se encontró un motivo de aminoácidos RGD, el cual ha sido implicado en la mediación de interacciones entre ARN y proteína y entre proteína y proteína, así como en la formación de gránulos de estrés (Kruger, *et al.*, 2013; Nocua, *et al.*, 2017; Rajyaguru & Parker, 2012). En el modelamiento tridimensional de la proteína, se encontró que estos dominios adoptaron las estructuras y plegamientos previamente descritos en la literatura, con excepción de la estructura de hélice α presente en el extremo N-terminal del dominio Lsm (Nocua, *et al.*, 2017). Sin embargo, al comparar la estructura de LbSCD6 con la EDC3 humana y la TraI de *Drosophila melanogaster*, se encontró que la LbSCD6 encajaba perfectamente en los segmentos conservados Sm1 y Sm2 que hacen parte de la estructura típica de barril beta de este dominio (Nocua, *et al.*, 2017). En cuanto a la RBP42, se identificó el dominio NTF2 (Ribbeck, *et al.*, 1998), relacionado con procesos de transporte entre núcleo y citoplasma de ARN, y que en esta proteína está conformado por los primeros 114 aminoácidos del extremo N-terminal y también adopta el plegamiento típico consistente en tres hélices α alineadas contra una hoja β conformada por cinco cadenas. Así mismo, los 73 aminoácidos del extremo C-terminal de RBP42 conforman un dominio RRM (*RNA recognition motif*), que es uno de los dominios de unión a ARN más conservados a lo largo de la escala filogenética. No obstante, a diferencia de la mayoría de estos dominios, la LbRBP42 solo posee una de las dos marcas conservadas, RNP1, consistente en ocho residuos (con abundancia de aminoácidos aromáticos o cargados), al parecer responsable de la unión al ARN (Maris, *et al.*, 2005). Cabe destacar que la proteína ortóloga de humanos, la G3BP, la cual, además de poseer ambos dominios como la proteína LbRBP42, también presenta tres repeticiones del motivo PxxP y ha sido implicada en la localización de esta proteína en los gránulos de estrés (Tourriere, *et al.*, 2003).

Clonación, expresión y obtención de anticuerpos contra las proteínas de interés

Una vez conocidos los dominios y la estructura tridimensional de las proteínas promisorias, se procede a su clonación y expresión como proteínas recombinantes, con el fin

de obtenerlas en estado nativo y producir sus anticuerpos. Dichas proteínas constituyen herramientas muy útiles en los experimentos encaminados a descifrar su funcionalidad. En nuestro ejemplo, las proteínas se clonaron en el sistema de expresión pQE30 y, tras ensayar diversas condiciones de solubilización y renaturación, finalmente se purificaron y se obtuvieron en su estado nativo, en tanto que los anticuerpos se obtuvieron mediante inmunización en conejos (Nocua, *et al.*, 2017).

Confirmación de la capacidad de unión a ARN tanto *in vitro* como *in vivo* de las proteínas de interés

Teniendo en cuenta que la técnica de pesca o *pull-down* empleada es una tamización del lisado de proteínas totales del parásito, el siguiente paso metodológico son los ensayos confirmatorios, los cuales combinan los ensayos de *pull-down* y visualización tras el *dot blotting*, utilizando exclusivamente las proteínas específicas de interés y los cebos de ARN, con el fin de confirmar su capacidad de unión a este. En el presente caso, con esta metodología se corroboró que tanto la LbSCD6 como la LbRBP42 son proteínas capaces de unirse al ARN e interactuar de forma específica con la región 3' UTR-II de los mensajeros de la HSP70 (Nocua, *et al.*, 2017).

Después de demostrar la interacción *in vitro* de ambas proteínas (LbSCD6 y LbRBP42) con la región 3' UTR de los genes *HSP70-II*, se hicieron ensayos para delimitar posibles elementos *in cis* involucrados en dicha interacción. Para lograr este objetivo, se sintetizaron los dos elementos *in cis* previamente identificados en la región 3' UTR-II. El primero de ellos, correspondiente a la secuencia rica en adeninas y uridinas relacionada con los elementos ARE, se denominó 3' UTR-II.1 y, el segundo, correspondiente a la región rica en citosinas, se denominó 3' UTR-II.2. Para analizar la posible interacción entre las proteínas LbSCD6 y LbRBP42 con los motivos seleccionados, se hicieron ensayos de retraso en la movilidad electroforética (*electrophoretic mobility shift assay*, EMSA) utilizando las proteínas recombinantes, así como los oligorribonucleótidos marcados con digoxigenina en su extremo 5'. Los resultados mostraron una fuerte interacción entre el motivo 3' UTR-II.1 y cada una de las proteínas recombinantes *rLbSCD6* (Figura 3A) y *rLbRBP42* (Figura 3B) (Nocua, 2017). En cambio, en el ensayo con el motivo 3' UTR-II.2 no se observó interacción con las proteínas recombinantes (Nocua, *et al.*, 2017). Los ensayos adicionales de EMSA de competición confirmaron los resultados y evidenciaron que la interacción entre las proteínas recombinantes y el motivo 3' UTR-II.1 es específica y depende tanto de sus características estructurales como de su secuencia (Nocua, *et al.*, 2017).

Confirmación *in vivo* de la capacidad de unión al ARN de las proteínas de interés

Una vez confirmada la capacidad de estas proteínas de interactuar directamente con moléculas de ARN *in vitro*,

se determinó esta misma capacidad *in vivo*, con el objetivo adicional de conocer los transcritos de ARN que pueden unirse a ellas en promastigotes de *L. braziliensis* cultivados a 26 °C. Para ello, se hicieron ensayos de coimmunoprecipitación utilizando anticuerpos específicos para cada una de las proteínas, con el fin de obtener los complejos LbSCD6-ARN y LbRBP42-ARN. Estos mismos ensayos se realizaron utilizando el suero preinmunización correspondiente a cada anticuerpo (preSCD6 y preRBP42) y la información obtenida se utilizó como control de uniones no específicas, las cuales se descartaron de las lecturas obtenidas en las dos réplicas biológicas hechas para cada una de las proteínas estudiadas (Nocua, 2017). Los resultados en las dos proteínas confirmaron que estas interactúan *in vivo* con distintos transcritos del parásito, principalmente a través de sus regiones UTR (Nocua, 2017). Específicamente, para proteínas LbSCD6 se encontraron 155 transcritos que codifican, siendo la interacción con las regiones 3' UTR la más frecuente (66 %). Con el fin de categorizar estos genes según el proceso biológico en el que están implicados, se hizo un análisis de enriquecimiento con la herramienta Gene Ontology (<http://geneontology.org/page/go-enrichment-analysis>), y se

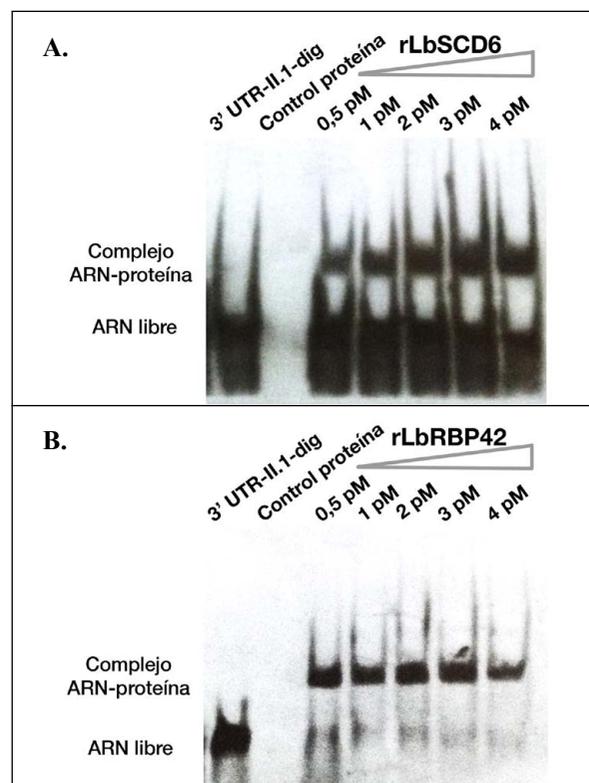


Figura 3. Interacción de las proteínas rLbSCD6 y rLbRBP42 con el motivo de ARN 3' UTR-II.1 detectada mediante ensayos EMSA. (A) Análisis de la capacidad de interacción de rLbSCD6 con el motivo 3' UTR-II.1 utilizando diferentes concentraciones de la proteína. (B) Análisis de la capacidad de interacción de rLbRBP42 con el motivo 3' UTR-II.1 utilizando diferentes concentraciones de la proteína

encontró que 98 (63 %) correspondían a genes anotados en la base de datos TriTryp con una función conocida (Aslett, *et al.*, 2010), en tanto que el 37 % restante correspondió a genes hipotéticos. En cuanto a la adscripción de los 98 transcritos a procesos biológicos, se encontró que, en su mayoría, estos pertenecen a proteínas del metabolismo celular (30,61 %), seguidos por proteínas asociadas a procesos celulares (11,28 %), al transporte celular (8,16 %) y a la proteólisis celular (7,14 %). Resultó interesante observar la interacción de la LbSCD6 con transcritos que codifican para la proteína DNAJ-like, indispensable en la maquinaria de activación y funcionamiento de la proteína HSP70 (Kampinga & Craig, 2010). Se encontraron 192 (74,13 % del total) transcritos que codifican asociados a la proteína LbRBP42 (Nocua, 2017). En cuanto a la localización de la interacción ARNm-proteína, se encontraron principalmente las interacciones con las regiones 3' UTR de los transcritos (57,8 %). De los 192 transcritos sometidos a coimmunoprecipitación, el 68,04 % correspondió a proteínas con anotación funcional en la base de datos TriTryp, en tanto que el 31,96 % correspondió a proteínas hipotéticas. Según su clasificación por asociación a procesos biológicos, se observó una gran proporción de proteínas asociadas al metabolismo celular (18,18 %), así como a diferentes procesos celulares, entre las cuales cabe resaltar las adscritas al ciclo celular (13,64 %). Asimismo, se encontraron proteínas asociadas a la respuesta a estímulos (7,58 %), entre ellas las asociadas a la respuesta al estrés. Cabe destacar que el transcrito correspondiente al gen *HSP70-II* se encontró asociado a la proteína LbRBP42, interacción localizada en la región 3' UTR de este transcrito.

Teniendo en cuenta que tanto la LbSCD6 como la LbRBP42 se encontraron interactuando *in vivo* con la región 3' UTR de más de cien transcritos y que *in vitro* se demostró la interacción de ambas proteínas con el motivo 3' UTR-II.1, relacionado con el elemento ARE, se hizo una búsqueda preliminar de motivos en la región 3' UTR de todos los transcritos asociados. Así, se encontró un motivo estructural asociado a las regiones que interactúan tanto con la proteína LbSCD6 (Figura 4A) como con la proteína LbRBP42 (Figura 4B). Puede observarse que, si bien la secuencia del motivo no corresponde a la de un elemento ARE típico, su estructura presenta una horquilla semejante a la observada en el motivo 3' UTR-II.1, lo que sugiere que la interacción entre los ARNm y las proteínas LbSCD6 y LbRBP42 puede estar guiada principalmente por motivos estructurales más que por una secuencia específica.

Estudios de aproximación a la función biológica de las proteínas de interés

Con el propósito de dilucidar los posibles procesos biológicos en los cuales podrían estar involucradas estas proteínas, se determinó, en primer lugar, el interactoma de cada una. Los resultados evidenciaron que los interactomas de LbSCD6 y LbRBP42 comparten una alta proporción de

proteínas, lo que sugiere algún tipo de relación funcional entre ellas; de hecho, las dos proteínas forman parte de los interactomas recíprocos. Es de resaltar que ambos interactomas están enriquecidos en proteínas asociadas a los procesos de traducción, al metabolismo del ARN y a la regulación de la expresión génica. Además, se encontraron varias proteínas cuyos homólogos se han reportado como constituyentes de gránulos de ARN, principalmente gránulos de estrés y cuerpos de procesamiento (Nocua, 2017); ambos son estructuras macromoleculares carentes de membrana en las que se acumulan ARN y proteínas. En eucariotas superiores los gránulos de estrés se forman transitoriamente en respuesta al estrés celular y están asociados con la traducción y estabilidad del ARNm. Bajo las condiciones apropiadas, los ARNm pueden salir de estos gránulos y asociarse de nuevo con polisomas (Buchan & Parker, 2009). En cuanto a los cuerpos de procesamiento, estos son estructuras constitutivas, las cuales contienen generalmente maquinaria relacionada con la degradación del ARNm (Parker & Sheth, 2007). Tanto los gránulos de estrés como los cuerpos de procesamiento interactúan entre sí y están en una estrecha asociación con el proceso de traducción (Decker & Parker, 2012).

Por último, para analizar la relevancia celular y funcional de LbSCD6 y LbRBP42 se generaron líneas mutantes de parásitos que sobreexpresan las proteínas fusionadas a la proteína fluorescente mCherry, y que denominamos LbSCD6-mCherry y LbRBP42-mCherry. La línea que sobreexpresó control (expresión de la proteína mCherry) de *L. braziliensis* pudo establecerse, pero luego de una semana de crecimiento, las líneas de sobreexpresión para cada una de estas proteínas perdieron viabilidad. Si bien estos resultados no son concluyentes, sugieren que la sobreexpresión de LbSCD6 y LbRBP42 en *L. braziliensis* afecta negativamente el desarrollo y la supervivencia del parásito. Como alternativa, se intentó con éxito la generación de estas líneas de sobreexpresión utilizando la especie *L. major*, lo que ha permitido vislumbrar importantes implicaciones de estas proteínas en el desarrollo de los parásitos, observándose

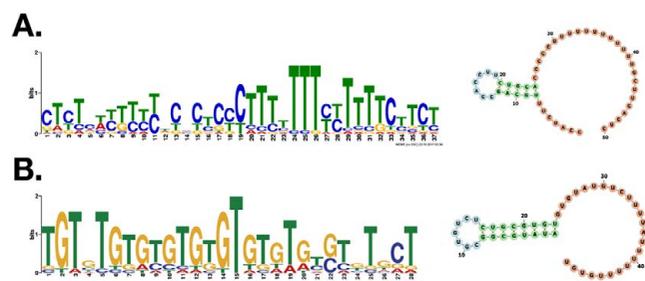


Figura 4. Motivos estructurales presentes en los transcritos asociados *in vivo* con las proteínas SCD6 y RBP42. Los posibles motivos asociados con las proteínas (A) SCD6 y (B) RBP42, se encontraron con la herramienta MEME, y las estructuras secundarias de un motivo representativo para cada proteína se obtuvieron con la herramienta RNAfold.

alteraciones morfológicas significativas y el retardo del crecimiento de los parásitos en condiciones fisiológicas de crecimiento, siendo el efecto mucho más pronunciado en condiciones de estrés térmico. Por otra parte, la capacidad infectiva *in vitro* de estos parásitos mutantes resultó también significativamente afectada en relación con la cepa silvestre (Nocua, 2017). Teniendo en cuenta todos los hallazgos encontrados, en la figura 5 se presenta un modelo que refleja el papel de las proteínas LbSCD6 y LbRBP42 en el parásito.

Conclusiones

Se concluyó que la LbSCD6 y la LbRBP42 son proteínas de unión al ARN implicadas en la regulación de la estabilidad o degradación no solo del transcrito de *HSP70-II* sino de un amplio grupo de transcritos en *Leishmania*.

En cuanto a la estrategia aquí propuesta, la cual se resume en la figura 6, los resultados evidencian su bondad y su potencial para la identificación de proteínas reguladoras de la expresión génica en tripanosomátidos.

Perspectivas de la estrategia. La metodología empleada comprende varias modificaciones y mejoras que pueden implementarse para subsanar algunas falencias detectadas. Por ejemplo, para simular las condiciones *in vivo*, es posible someter a los parásitos a la acción de la luz ultravioleta antes de la obtención del lisado proteico, con el fin de

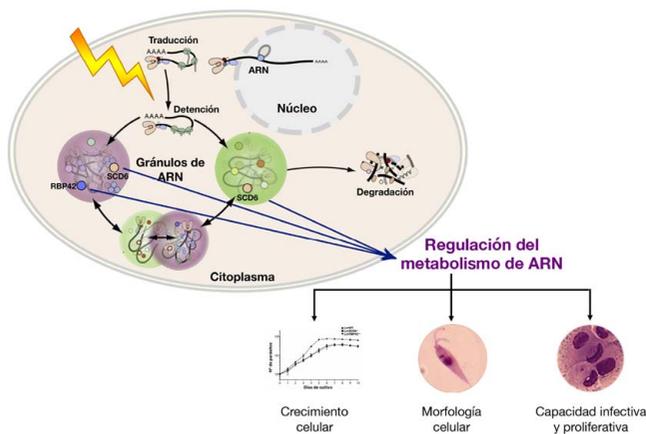


Figura 5. Modelo del rol de las proteínas LbSCD6 y LbRBP42 en la regulación de la expresión génica a nivel del ARN en el parásito. Teniendo en cuenta todos los hallazgos derivados de este trabajo, se sugiere que las proteínas SCD6 y RBP42 están implicadas en la regulación de transcritos de ARN debido a su capacidad de unión con este tipo de moléculas y a la asociación que presentan con constituyentes de gránulos ribonucleoproteicos, como los cuerpos de procesamiento y los gránulos de estrés, localizados en el citoplasma. La localización citoplasmática granular se observa claramente cuando los parásitos están sometidos a condiciones de estrés, particularmente de estrés térmico (Datos no mostrados). Los cambios en la expresión de estas proteínas afectan la regulación de la estabilidad y su capacidad de traducción de los ARN mensajeros (metabolismo de ARN), lo cual desencadena un efecto pleiotrópico en el parásito, observándose cambios en el crecimiento y morfología celular, así como en la capacidad infectiva y proliferativa del parásito.

inmovilizar los complejos de ARN y proteínas formados bajo condiciones fisiológicas. Asimismo, se puede trabajar con los componentes subcelulares de interés, así como usar técnicas más finas de separación de proteínas que en una misma reacción permitan elucidar las proteínas diferenciales en relación con el control no relevante utilizado.

Teniendo en cuenta que la estrategia aquí propuesta se focaliza en la búsqueda de RBP que modulen la expresión de factores de virulencia de los tripanosomátidos, que como se vio en el caso de las proteínas LbSCD6 y LbRBP42, no solo gobiernan la expresión de dicho factor, en este caso la proteína HSP70, sino de muchas otras proteínas; en consecuencia, se proponen estas RBP como nuevos blancos parasitarios para el desarrollo de medicamentos. En este sentido, no resulta extraño que, dado su papel preponderante en la desregulación asociada al desarrollo de las células cancerosas, hayan sido propuestas también como nuevos blancos para el desarrollo de alternativas terapéuticas frente a los procesos tumorales (Turner & Díaz-Muñoz, 2018).

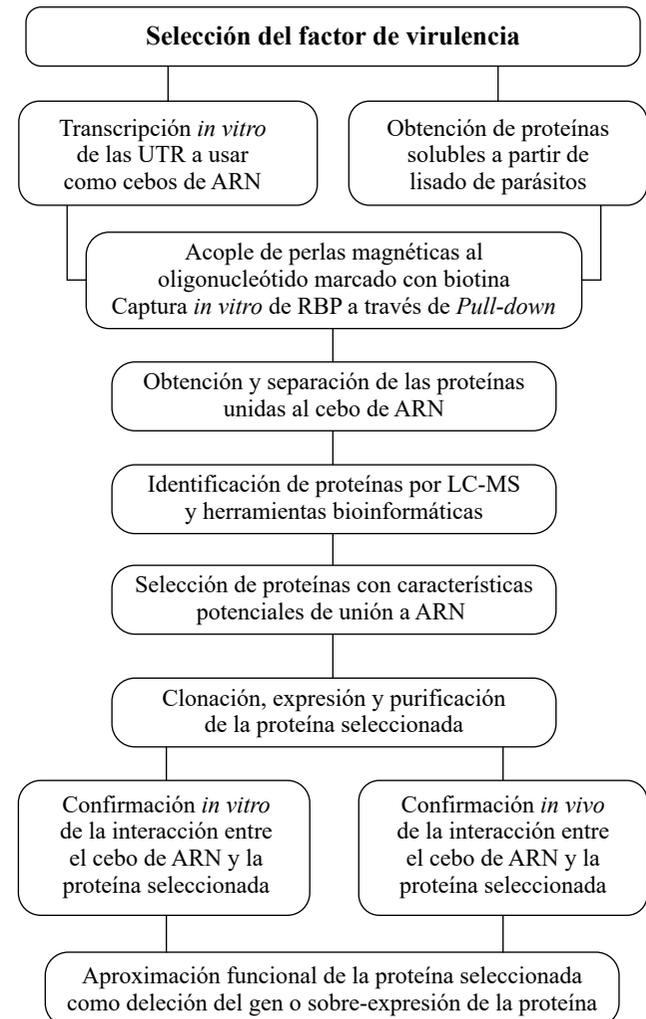


Figura 6. Esquema de resumen de la estrategia con el conjunto de pasos para la identificación de proteínas reguladoras de la expresión génica en tripanosomátidos.

Agradecimientos

Este trabajo no hubiese sido posible sin el respaldo incondicional de la Pontificia Universidad Javeriana, en cabeza de sus distintas autoridades de gobierno, a lo largo de toda una vida de labor académica y científica. Agradecemos, asimismo, a todas las entidades que han financiado nuestros proyectos, principalmente Colciencias y la Pontificia Universidad Javeriana.

Contribución de los autores

Concepción J. Puerta: concepción del manuscrito, participación en su escritura, revisión y corrección, y en el diseño y seguimiento de los experimentos. José María Requena: participación en su escritura, revisión y corrección, y en el diseño y seguimiento de los experimentos. Elizabeth Ruiz, Cesar A. Ramírez y Paola Nocua: realización de los experimentos, participación en la elaboración de figuras y en la revisión y corrección del manuscrito.

Conflicto de intereses

Los autores declaran no tener conflictos de intereses.

Referencias

- Agabian, N. (1990). Trans splicing of nuclear pre-mRNAs. *Cell*. **61**(7): 1157-1160. doi: 10.1016/0092-8674(90)90674-4
- Álvar, J., Vélez, I.D., Bern, C., Herrero, M., Desjeux, P., Cano, J., Jannin, J., den Boer, M. (2012). Leishmaniasis worldwide and global estimates of its incidence. *PLoS One*. **7** (5): e35671. doi: 10.1371/journal.pone.0035671
- Anantharaman, V., Aravind, L. (2004). Novel conserved domains in proteins with predicted roles in eukaryotic cell-cycle regulation, decapping and RNA stability. *BMC Genomics*. **5** (1): 45. doi: 10.1186/1471-2164-5-45
- Araújo, P., Teixeira, S. (2011). Regulatory elements involved in the post-transcriptional control of stage-specific gene expression in *Trypanosoma cruzi* - A review. *Mem Inst Oswaldo Cruz*. **106** (3): 257-266. doi: 10.1590/S0074-02762011000300002
- Aslett, M., Aurrecochea, C., Berriman, M., Brestelli, J., Brunk, B.P., Carrington, M., Depledge, D.P., Fischer, S., Gajria B., Gao, X., Gardner, M.J., Gingle, A., Grant, G., Harb, O.S., Heiges, M., Hertz-Fowler, C., Houston, R., Innamorato, F., Iodice, J., Kissinger, J.C., Kraemer, E., Li, W., Logan, F.J., Miller, J.A., Mitra, S., Myler, P.J., Nayak, V., Pennington, C., Phan, I., Pinney, D.F., Ramasamy, G., Rogers, M.B., Roos, D.S., Ross, C., Sivam, D., Smith, D.F., Srinivasamoorthy, G., Stoeckert, C.J. Jr., Subramanian, S., Thibodeau, R., Tivey, A., Treatman, C., Velarde, G., Wang, H. (2010). TriTrypDB: A functional genomic resource for the *Trypanosomatidae*. *Nucleic Acids Res*. **38** (Database issue): D457-462. doi: 10.1093/nar/gkp851
- Bangs, J.D., Crain, P.F., Hashizume, T., McCloskey, J.A., Boothroyd, J.C. (1992). Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. *J Biol Chem*. **267** (14): 9805-9815. PMID: 1349605.
- Bañuls, A.L., Hide, M., Prugnolle, F. (2007). *Leishmania* and the leishmaniasis: A parasite genetic update and advances in taxonomy, epidemiology and pathogenicity in humans. *Adv Parasitol*. **64**: 1-10. doi:10.1016/S0065308X(06)64001-3
- Baumeister, H. (1997). RNA-protein interactions. Oxford, Reino Unido: IRL Press at Oxford University Press.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., Böhme, U., Hannick, L., Aslett, M.A., Shallom, Marcello, L., Hou, L., Wickstead, B., Alsmark, U.C., Arrowsmith, C., Atkin, R.J., Barron, A.J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T.J., Churcher, C., Clark, L.N., Corton, C.H., Cronin, A., Davies, R.M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M.C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B.R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A.X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line A, Lord A, Macleod A, Mooney PJ, Moule S, Martin DM, Morgan G.W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C.S., Peterson, J., Quail, M.A., Rabbinowitsch, E., Rajandream, M.A., Reitter, C., Salzberg, S.L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A.J., Tallon, L., Turner, C.M., Tait, A., Tivey, A.R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M.D., Embley, T.M., Gull, K., Ullu, E., Barry, J.D., Fairlamb, A.H., Opperdoes, F., Barrell, B.G., Donelson, J.E., Hall, N., Fraser, C.M., Melville, S.E., El-Sayed, N.M. (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Science*. **309**: 416-422. doi: 10.1126/science.1112642
- Buchan, J.R., Parker, R. (2009). Eukaryotic stress granules: The ins and outs of translation. *Mol Cell*. **36** (6): 932-941. doi: 10.1016/j.molcel.2009.11.020
- Carrion, J., Folgueira, C., Soto, M., Fresno, M., Requena, J.M. (2011). *Leishmania infantum* HSP70-II null mutant as candidate vaccine against leishmaniasis: A preliminary evaluation. *Parasit Vectors*. **4**: 150. doi: 10.1186/1756-3305-4-150
- Colombia – Instituto Nacional de Salud. (2011). Boletín Epidemiológico semana 40. Instituto Nacional de Salud Colombia. Bogotá, Colombia. Pp.9.
- Cassola, A., Frasch, A. (2009). An RNA recognition motif mediates the nucleocytoplasmic transport of a *Trypanosoma* RNA-binding protein *J Biol Chem*. **284** (50): 35015-35028. doi: 10.1074/jbc.M109.031633
- Clayton, C. (2016). Gene expression in kinetoplastids. *Curr Opin Microbiol*. **32**: 46-51. doi: 10.1016/j.mib.2016.04.018
- David, M., Gabdank, I., Ben-David, M., Zilka, A., Orr, I., Barash, D., Shapira, M. (2010). Preferential translation of Hsp83 in *Leishmania* requires a thermosensitive polypyrimidine-rich element in the 3' UTR and involves scanning of the 5' UTR. *RNA*. **16** (2): 364-74. doi: 10.1261/rna.1874710
- De Gaudenzi, J., Frasch, A., Clayton, C. (2005). RNA-binding domain proteins in Kinetoplastids: a comparative analysis. *Eukaryo Cell*. **4** (12): 2106-2114. doi: 10.1128/EC.4.12.2106-2114.2005
- Di Noia J.M., D'Orso I., Sánchez D.O, Frasch A.C. (2000). AU-rich elements in the 3'-untranslated region of a new mucin-type gene family of *Trypanosoma cruzi* confers mRNA instability and modulates translation efficiency. *J Biol Chem*. **275** (14): 10218-10227. PMID: 9556557

- El-Sayed, N.M., Myler, P.J., Bartholomeu, D.C., Nilsson, D., Aggarwal, G., Tran, A.N., Ghedin, E., Worthey, E.A., Delcher, A.L., Glandin, G., Westenberger, S.J., Caler, E., Cerqueira, G.C., Branche, C., Haas, B., Anupama, A., Arner, E., Aslund, L., Attipoe, P., Bontempi, E., Bringaud, F., Burton, P., Cadag, E., Campbell, D.A., Carrington, M., Crabtree, J., Darban, H., da Silveira, J.F., de Jong, P., Edwards, K., Englund, P.T., Fazelina, G., Feldblyum, T., Ferella, M., Frasch, A.C., Gull, K., Horn, D., Hou, L., Huang, Y., Kindlund, E., Klingbeil, M., Kluge, S., Koo, H., Lacerda, D., Levin, M.J., Lorenzi, H., Louie, T., Machado, C.R., McCulloch, R., McKenna, A., Mizuno, Y., Mottram, J.C., Nelson, S., Ochaya, S., Osoegawa, K., Pai, G., Parsons, M., Pentony, M., Pettersson, U., Pop, M., Ramirez, J.L., Rinta, J., Robertson, L., Salzberg, S.L., Sanchez, D.O., Seyler, A., Sharma, R., Shetty, J., Simpson, A.J., Sisk, E., Tammi, M.T., Tarleton, R., Teixeira, S., Van Aken, S., Vogt, C., Ward, P.N., Wickstead, B., Wortman, J., White, O., Fraser, C.M., Stuart, K.D., Andersson, B. (2005). The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science*. **309**: 409-415. doi: 10.1126/science.1112631
- Decker, C.J., Parker, R. (2012). P-bodies and stress granules: Possible roles in the control of translation and mRNA degradation. *Cold Spring Harb Perspect Biol*. **4** (9): a012286. doi: 10.1101/cshperspect.a012286
- Folgueira, C., Carrion, J., Moreno, J., Saugar, J.M., Canavate, C., Requena, J.M. (2008). Effects of the disruption of the HSP70-II gene on the growth, morphology, and virulence of *Leishmania infantum* promastigotes. *Int Microbiol*. **11** (2): 81-89. PMID: 18645958. En PUBMED no aparece el doi, coloqué el PMID.
- Folgueira, C., Quijada, L., Soto, M., Abanades, D.R., Alonso, C., Requena, J.M. (2005). The translational efficiencies of the two *Leishmania infantum* HSP70 mRNAs, differing in their 3'-untranslated regions, are affected by shifts in the temperature of growth through different mechanisms. *J Biol Chem*. **280** (42): 35172-35183. doi: 10.1074/jbc.M505559200
- Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Adlem, E., Aert, R., Anupama, A., Apostolou, Z., Attipoe, P., Bason, N., Bauser, C., Beck, A., Beverley, S.M., Bianchetti, G., Borzym, K., Bothe, G., Bruschi, C.V., Collins, M., Cadag, E., Ciarloni, L., Clayton, C., Coulson, R.M., Cronin, A., Cruz, A.K., Davies, R.M., De Gaudenzi, J., Dobson, D.E., Duesterhoeft, A., Fazelina, G., Fosker, N., Frasch, A.C., Fraser, A., Fuchs, M., Gabel, C., Goble, A., Goffeau, A., Harris, D., Hertz-Fowler, C., Hilbert, H., Horn, D., Huang, Y., Klages, S., Knights, A., Kube, M., Larke, N., Litvin, L., Lord, A., Louie, T., Marra, M., Masuy, D., Matthews, K., Michaeli, S., Mottram, J.C., Müller-Auer, S., Munden, H., Nelson, S., Norbertczak, H., Oliver, K., O'neil, S., Pentony, M., Pohl, T.M., Price, C., Purnelle, B., Quail, M.A., Rabinowitsch, E., Reinhardt, R., Rieger, M., Rinta, J., Robben, J., Robertson, L., Ruiz, J.C., Rutter, S., Saunders, D., Schäfer, M., Schein, J., Schwartz, D.C., Seeger, K., Seyler, A., Sharp, S., Shin, H., Sivam, D., Squares, R., Squares, S., Tosato, V., Vogt, C., Volckaert, G., Wambutt, R., Warren, T., Wedler, H., Woodward, J., Zhou, S., Zimmermann, W., Smith, D.F., Blackwell, J.M., Stuart, K.D., Barrell, B., Myler, P.J. (2005). The genome of the kinetoplastid parasite, *Leishmania major*. *Science*. **309**: 436-442. doi: 10.1126/science.1112680
- Jeffery, C.J. (2018). Protein moonlighting: What is it, and why is it important? *Philos Trans R Soc Lond B Biol Sci*. **373** (1738). pii: 20160523. doi: 10.1098/rstb.2016.0523
- Justice, M.C., Hsu, M.J., Tse, B., Ku, T., Balkovec, J., Schmatz, D., Nielsen, J. (1998). Elongation factor 2 as a novel target for selective inhibition of fungal protein synthesis. *J Biol Chem*. **273** (6): 3148-51. doi: 10.1074/JBC.273.6.3148
- Kampinga, H.H., Craig, E.A. (2010). The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nat Rev Mol Cell Biol*. **11** (8): 579-592. doi: 10.1038/nrm2941
- Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K.F., Itoh, M., Kawashima, S., Katayama, T., Araki, M., Hirakawa, M. (2006). From genomics to chemical genomics: New developments in KEGG. *Nucleic Acids Res*. **34**: D354-7. doi: 10.1093/nar/gkj102
- Kramer, S., Carrington, M. (2011). Trans-acting proteins regulating mRNA maturation, stability and translation in trypanosomatids. *Trends Parasitol*. **27** (1): 23-30. doi: 10.1016/j.pt.2010.06.011
- Krüger, T., Hofweber, M., Kramer, S. (2013). SCD6 induces ribonucleoprotein granule formation in trypanosomes in a translation-independent manner, regulated by its Lsm and RGG domains. *Mol Biol Cell*. **24** (13): 2098-2111. doi: 10.1091/mbc.E13-01-0068
- Lapointe, C.P., Wilinski, D., Saunders, H.A., Wickens, M. (2015). Protein-RNA networks revealed through covalent RNA marks. *Nat Methods*. **12** (12): 1163-70. doi: 10.1038/nmeth.3651
- LeBowitz, J.H., Smith, H.Q., Rusche, L., Beverley, S.M. (1993). Coupling of poly(A) site selection and trans-splicing in *Leishmania*. *Genes Dev*. **7** (6): 996-1007. doi: 10.1101/gad.7.6.996
- Makeyev, A.V., Liebhaber, S.A. (2002). The poly(C)-binding proteins: A multiplicity of functions and a search for mechanisms. *RNA*. **8** (3): 265-278. doi: 10.1017/S1355838202024627
- Maris, C., Domínguez, C., Allain, F.H. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J*. **272** (9): 2118-2131. doi: 10.1111/j.1742-4658.2005.04653.x
- Mateyak, M.K., Kinzy, T.G. (2010). eEF1A: Thinking outside the ribosome. *J Biol Chem*. **285** (28): 21209-2113. doi: 10.1074/jbc.R110.113795
- Mi, H., Dong, Q., Muruganujan, A., Gaudet, P., Lewis, S., Thomas, P.D. (2010). PANTHER version 7: Improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. *Nucleic Acids Res*. **38**: D204-10. doi: 10.1093/nar/gkp1019
- Mittal, N., Subramanian, G., Butikofer, P., Madhubala, R. (2013). Unique posttranslational modifications in eukaryotic translation factors and their roles in protozoan parasite viability and pathogenesis. *Mol Biochem Parasitol*. **187** (1): 21-31. doi: 10.1016/j.molbiopara.2012.11.001
- Müller-McNicoll, M., Neugebauer, K.M. (2013). How cells get the message: Dynamic assembly and function of mRNA-protein complexes. *Nat Rev Genet*. **14** (4): 275-287. doi: 10.1038/nrg3434

- Noble, C.G., Song, H.** (2008). Structural studies of elongation and release factors. *Cell Mol Life Sci.* **65** (9): 1335-1346. doi: 10.1007/s00018-008-7495-6.
- Nocua, P. A.** (2017). Estudio de factores proteicos asociados a la regulación de los genes HSP70 en *Leishmania braziliensis*. Tesis Doctoral, Pontificia Universidad Javeriana. Bogotá, Colombia.
- Nocua, P.A., Ramírez, C.A., Barreto, G.E., González, J., Requena, J.M., Puerta, C.J.** (2014). *Leishmania braziliensis* replication protein A subunit 1: Molecular modelling, protein expression and analysis of its affinity for both DNA and RNA. *Parasit Vectors.* **7**: 573. doi: 10.1186/s13071-014-0573-8
- Nocua, P.A., Ramírez, C.A., Requena, J.M., Puerta, C.J.** (2017). *Leishmania braziliensis* SCD6 and RBP42 proteins, two factors with RNA binding capacity. *Parasit Vectors.* **10** (1): 610. doi: 10.1186/s13071-017-2557-y
- Parker, R., Sheth, U.** (2007). P bodies and the control of mRNA translation and degradation. *Mol Cell.* **25** (5): 635-646. doi: 10.1016/j.molcel.2007.02.011
- Puerta, C.** (2007). Consideraciones taxonómicas, polimorfismo y clasificación. En: C.J. Puerta. *Parásito, genoma y biología. Aproximación molecular al estudio de Trypanosoma rangeli y su relación con Trypanosoma cruzi.* (59-68). Bogotá, Colombia: Editorial Pontificia Universidad Javeriana.
- Quijada, L., Soto, M., Alonso, C., Requena, J.M.** (2000). Identification of a putative regulatory element in the 3'-untranslated region that controls expression of HSP70 in *Leishmania infantum*. *Mol Biochem Parasitol.* **110** (1): 79-91. doi: 10.1016/S0166-6851(00)00258-9
- Rajyaguru, P., Parker, R.** (2012). RGG motif proteins: Modulators of mRNA functional states. *Cell Cycle.* **11** (14): 2594-2599. doi: 10.4161/cc.20716
- Ramírez, C.A.** (2012). Caracterización molecular de los genes HSP70, a-tubulina y ND8 de *Leishmania braziliensis* y detección de factores proteicos asociados a la regulación de los genes HSP70. Tesis Doctoral, Pontificia Universidad Javeriana. Bogotá, Colombia.
- Ramírez, C.A., Dea-Ayuela, M.A., Gutiérrez-Blázquez, M.D., Bolas-Fernández, F., Requena, J.M., Puerta, C.J.** (2013). Identification of proteins interacting with HSP70 mRNAs in *Leishmania braziliensis*. *J Proteomics.* **94**: 124-137. doi: 10.1016/j.jprot.2013.09.008
- Ramírez, C.A., Requena, J.M., Puerta, C.J.** (2011). Identification of the HSP70-II gene in *Leishmania braziliensis* HSP70 locus: Genomic organization and UTRs characterization. *Parasit Vectors.* **4**: 166. doi: 10.1186/1756-3305-4-166
- Requena, J.M.** (2011). Lights and shadows on gene organization and regulation of gene expression in *Leishmania*. *Front Biosci.* **16**: 2069-2085. doi: 10.2741/3840
- Requena, J. M.** (2012). The stressful life of pathogenic *Leishmania* species. En: J.M. Requena. *Stress response in Microbiology.* (323-346). Norfolk, Reino Unido: Causer Academic press. ISBN: 978-1-908230-04-1
- Ribbeck, K., Lipowsky, G., Kent, H.M., Stewart, M., Gorlich, D.** (1998). NTF2 mediates nuclear import of Ran. *EMBO J.* **17** (22): 6587-6598. doi: 10.1093/emboj/17.22.6587
- Romaniuk, M.A., Cervini, G., Cassola, A.** (2016). Regulation of RNA binding proteins in trypanosomatid protozoan parasites. *World J Biol Chem.* **7** (1): 146-157. doi: 10.4331/wjbc.v7.i1.146
- Tourriere, H., Chebli, K., Zekri, L., Courselaud, B., Blanchard, J.M., Bertrand, E., Tazi, J.** (2003). The RasGAP-associated endoribonuclease G3BP assembles stress granules. *J Cell Biol.* **160** (6): 823-831. doi: 10.1083/jcb.200212128
- Turner, M., Díaz-Muñoz, M.** (2018). RNA-binding proteins control gene expression and cell fate in the immune system. *Nat Immunol.* **19**: 120-129. doi: 10.1038/s41590-017-0028-4