

REVISIÓN

Impacto de la biología molecular, genómica y proteómica en el estudio de los Kinetoplastida

José Luis Ramírez*

En este trabajo se presenta una visión de la evolución de la biología molecular en el estudio de la biología de los Kinetoplastida, orden de protozoarios que incluye especies patogénicas al hombre, plantas y animales. El ensayo está centrado en los nuevos desarrollos generados a partir de los proyectos genoma de *Leishmania major*, *Trypanosoma cruzi* y *Trypanosoma brucei*, identificando aquellos puntos en donde se puede derivar información para luchar contra estos parásitos. Se discuten además las tendencias que predominarán en la parasitología moderna como son las aproximaciones sistémicas a la relación parásito-hospedador, tratando de desviar nuestra atención de la superficie de los parásitos en donde estos disponen de sistemas altamente evolucionados o surfomas que resultan difícil de atacar.

Palabras claves: genoma, proteómica, aproximaciones sistémicas, interacción parásito- hospedador.

INTRODUCCIÓN

Desde hace más de 30 años se han venido aplicando las herramientas de la biología molecular en el estudio de protozoarios patógenos del Orden Kinetoplastida. En los primeros tiempos estos estudios estaban centrados en el análisis de ADN (o material genético) tanto nucleares como extra-cromosomales. Así se logró determinar que la mayoría de estos organismos presentan peculiaridades en la organización de sus genomas tales como la ausencia de condensación cromosomal durante los procesos mitóticos y genomas extra-cromosomales con organizaciones caprichosas como el kinetoplasto. Con el tiempo se logró clonar y producir in vitro un gran número de proteínas del metabolismo central y de la superficie de estos parásitos y se identificaron

secuencias repetidas que han brindado información sobre la heterogeneidad poblacional de los mismos. En 1984 se introdujo la técnica de electroforesis en campo eléctrico pulsado (Pulse Field Electrophoresis) (Schwartz & Cantor, 1984) la cual permitió visualizar por primera vez las bandas cromosomales de muchos protozoarios. En 1985 se incorpora la poderosa herramienta de la reacción en cadena de polimerasa o PCR (Saiki, *et al.*, 1985), como método de detección de agentes infecciosos en muestras biológicas. Más recientemente se ha completado el secuenciamiento los genomas de tres tripanosomatidos patógenos, y se ha iniciado la era de estudios postgenómicos que incluyen: la genómica, la proteómica, el interactoma, el metaboloma y la biología de sistemas.

En esta breve revisión discutiré cada tecnología empleada, su evolución e impacto en las enfermedades causadas por estos protozoarios. La revisión no es exhaustiva sino didáctica y dirigida a trabajadores de la salud no especialistas en ciencias básicas.

He diseñado gráficos para las tecnologías menos convencionales, pasando por alto a aquellas que ya son de uso común en muchos laboratorios.

Centro de Biotecnología, Instituto de Estudios Avanzados Ministerio de Ciencia y Tecnología e Instituto de Biología Experimental, Facultad de Ciencias, UCV - Caracas, Venezuela.

Dirección postal Apartado 47525, Los Chaguaramos, Caracas 1041

*Autor de correspondencia: ramjoseluis@gmail.com

Las tecnologías mas usadas

Los primeros ensayos usando el material genético de protozoarios patógenos se remontan a los años 70 cuando se aplicaron técnicas de gradientes de flotación en cloruro de cesio que permitían determinar la composición de bases y contenido de ADN. Era común entonces, siguiendo estudios seminales en bacterias, el tratar de establecer correlaciones entre cantidad de ADN, contenido de bases y estatus taxonómico (Chance *et al.*, 1971). Para la misma época se examinaba el material genético de parásitos mediante técnicas de microscopía electrónica (Simpson & Da Silva, 1981) y electroforesis en geles. Aunque todos estos estudios tuvieron una utilidad práctica limitada, ayudaron a cimentar las bases para estudios posteriores. De estos tiempos han quedado conocimientos como: la gran variabilidad en contenido de ADN por genoma, aún dentro de miembros de una misma especie (Lanar *et al.*, 1978), la organización atípica de sus genomas mitocondriales con dos tipos de ADN circulares (Simpson & Da Silva, 1981) uno mayor o maxi-círculo el cual codifica para la mayoría de las proteínas y ARNs de origen mitocondrial, y otro pequeño en la forma de miles de pequeños círculos concatenados o mini-círculos que codifican para ARNs llamados guías (ver mas abajo); la fragmentación de los ARN ribosomales (Villalba & Ramirez, 1982), la ubicación telomérica o subtelomérica de genes de contingencia en *Trypanosoma brucei* (Johnson & Borst 1986) y el clonamiento de muchos genes básicos.

Concomitante a la aplicación de estas tecnologías comenzaron a aparecer los estudios de RFLP (restriction fragment length polymorphisms) (Botstein *et al.*, 1980). Las enzimas de restricción descubiertas en el año 1972 (Meselson *et al.*, 1972) son endonucleasas bacterianas que tienen la propiedad de reconocer sitios específicos dentro del ADN, y a diferencia de otras enzimas que degradan ADN, producen patrones específicos en vez de digestiones al azar. Siguiendo la tendencia en los estudios poblacionales en otros organismos, se comenzó a utilizar esta herramienta para distinguir entre especies, clones y variantes. Aún hoy en día, los RFLPs son empleados en conjunción con PCR para detectar e identificar parásitos. Uno de los logros más importantes de esta tecnología junto con el uso de isoenzimas (Tibayrenc *et al.*, 1991), fue la de revelar que la estructura poblacional en la mayoría de los

protozoarios patógenos era de naturaleza clonal con poca o ninguna recombinación de tipo sexual. En paralelo al desarrollo de los RFLPs, investigadores de varias partes del mundo lograron clonar un alto número de genes de superficie, con miras a una posible inmunoterapia o inmunoprevención, o genes del metabolismo central en la búsqueda de blancos quimioterapéuticos específicos.

Estudios sobre la expresión de genes (transcripción) en los Kinetoplastida han revelado características únicas en estos organismos, por ejemplo el proceso de transcripción es poco regulado a nivel primario ya que todos los genes (exceptuando los genes PARP y variantes de superficie en *T. brucei*) se expresan como largos ARN policistrónicos (Martínez-Calvillo *et al.*, 2003) (conteniendo la información para muchas proteínas), los cuales posteriormente son procesados mediante un mecanismo de trans-empalme que logra fragmentar el ARN policistrónico en mRNAs mensajeros maduros añadiéndole una señal de entrada al ribosoma y una cola de poli adeninas (poliA). En otras palabras, probablemente la regulación de la expresión genética se ejerce a nivel postranscripcional y postraduccional (después de ser sintetizadas las proteínas). A fines de los 80 (Feagin, *et al.*, 1988) se descubre en kinetoplastidos un proceso de edición de ARN circunscrito al genoma mitocondrial con el cual los mensajeros primarios producidos a partir del maxi-círculo son editados mediante la adición o delección de uracilos para producir el mensajero final que será traducido a proteínas. La plantilla que sirve de base al proceso de edición está en los ARN guía producidos a partir de los mini-círculos. Desde esta época y dado lo exclusivo de este proceso, se ha tratado de desarrollar drogas contra el mismo.

A partir de los años 90 se perfeccionaron los protocolos de transformación de Kinetoplastida, y se desarrollaron vectores para clonamiento y expresión de genes. Estos vectores permitieron avanzar hacia estudios de lo que hoy se llama genética en reverso o búsqueda de la importancia de un gen mediante: disrupción (knock-out o reemplazamiento) (Fig. 1A); neutralización de su expresión (por ARN interferencia, sentido y antisentido) (Fig. 1B y 1C) y mutagénesis sitio dirigida (aislando un gen, cambiando su secuencia in vitro y re-insertándolo de vuelta en el parásito). Sin embargo dentro de las tecnologías más importantes como el ARNi (ARN interferencia) y

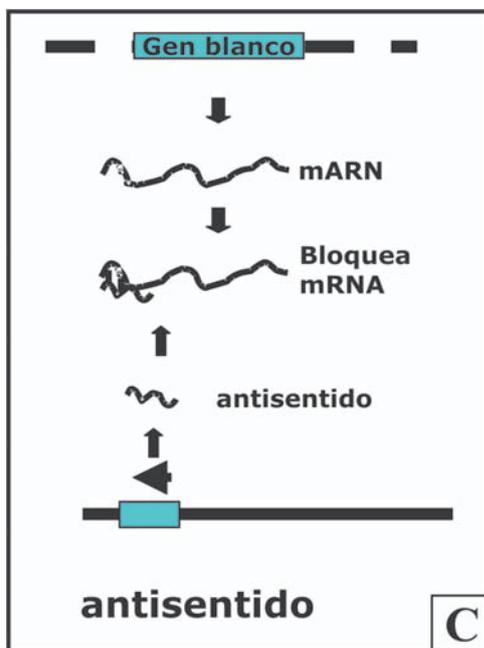
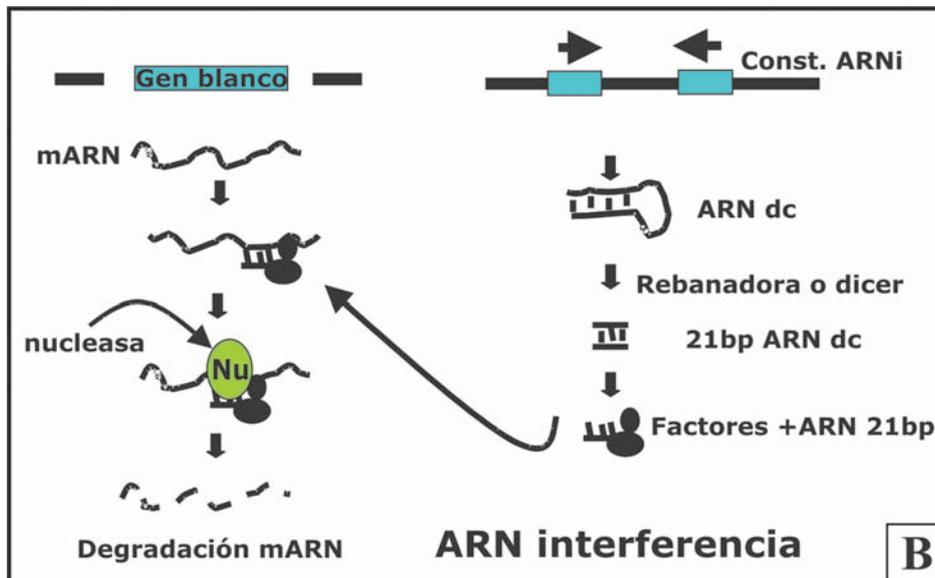
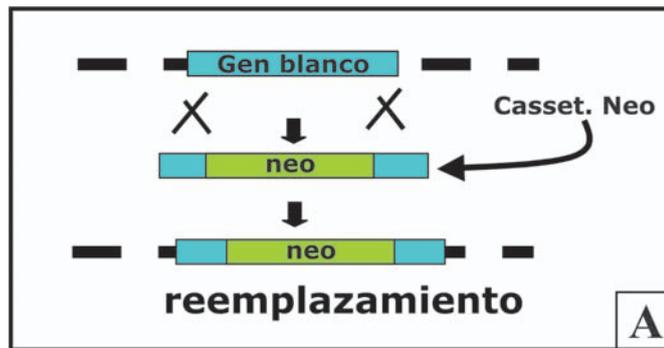


Fig. 1. Técnicas de silenciamiento genético.

A. Reemplazamiento de un gen por un casete hecho por ingeniería genética, en el cual el mismo gen clonado está interrumpido por un marcador que confiere resistencia a un antibiótico.

B. Silenciamiento por medio de la cascada de ARN de interferencia. Un fragmento clonado del gen que se quiere inactivar es puesto en un casete en dos direcciones con una secuencia de relleno en el medio. Al colocarlo dentro del organismo cuyo gen se quiere inactivar, se produce un ARN doble cadena. Este ARN es reconocido por la maquinaria de ARNi, en donde se dispara una degradación continua de los mensajeros codificados por ese gen.

C. ARN antisentido, en este caso se expresa un fragmento del gen de forma invertida produciendo un ARN parcialmente complementario al mensajero, se cree que estos al interactuar impiden que el mensajero pueda ser traducido a proteínas.

el ARN antisentido, nos encontramos con algunas sorpresas, la tecnología de ARNi solo opera en *T. brucei*, y no así en *Leishmania* o *Trypanosoma cruzi*, y el ARN antisentido se ha utilizado con éxito en *T. cruzi* y *T. brucei* pero no parece funcionar en *Leishmania*. La explicaciones que se han dado para estas diferencias son parciales, en el caso de ARNi no funciona porque la maquinaria de procesamiento de esta ruta degradativa no existe en *Leishmania* y *T. cruzi*, mientras que si existe en *T. brucei*. Para el ARN antisentido aún no se conoce la razón por la cual no ha funcionado en *Leishmania*.

La época de los genomas

Entre los días 14 y 15 de Abril de 1994 (Degrave *et al.*, 1994) tuvo lugar una importante reunión en la Fundación Oswaldo Cruz (Fig. 2) destinada a coordinar varias iniciativas de secuenciación de los genoma de Kinetoplastida. En esta reunión se conciliaron las iniciativas del CYTED-D motorizada por un grupo de Latinoamericanos y Españoles, y las del National Institute of Health, el Instituto TIGR de los Estados Unidos, y el Sanger Center en el Reino Unido.

Al final de esta reunión fueron escogidos los genomas de *Leishmania major* Friedlin, *T. cruzi* CL Brener y *T. brucei*, para ser secuenciados. Las

estrategias asumidas para el genoma de *Leishmania* incluía transferir todos los marcadores genéticos que se habían mapeado (ubicar en los cromosomas del organismo) en las diferentes cepas o especies a *L. major*, y la tarea se ejecutó repartiendo un cierto número de cromosomas para cada uno de los centros participantes. Para *T. brucei* se decidió secuenciar solo los cromosomas mayores dejando a un lado los mini-cromosomas (pequeños cromosomas supernumerarios en donde residen copias extras de los genes de variantes de superficie). En el caso de *T. cruzi*, el cual contaba con menos recursos se optó por una estrategia de “salto de BACs” que consistía en secuenciar los extremos de los recombinantes obtenidos en cromosomas artificiales bacterianos (siglas en Inglés BAC), para luego tratar de ensamblar el genoma asumiendo una gran proximidad y un alto grado de sintenia (u orden lineal de los genes en los cromosomas) con *T. brucei* y *Leishmania*.

Los proyectos de *Leishmania* y *T. brucei* marcharon casi en paralelo, mientras que *T. cruzi* quedó rezagado hasta el año 2003 cuando el Instituto TIGR decidió aplicar una aproximación “Shot-gun” en la cual, a diferencia de la de secuenciar cromosoma por cromosoma, se fragmenta todo el ADN del parásito, se clonan estos fragmentos, se secuencian, y luego las computadoras se encargan de ensamblar todo el genoma.

Fig. 2. Fotografía de los investigadores participantes de la reunión preparatoria de la Red de Genoma de tripanosomatídeos. Fundación Oswaldo Cruz , Rio de Janeiro, Brasil del 14 ald 15 de Abril 1994.



El día 15 de Julio del 2005 salieron publicados en el volumen 309 de la revista Science de ese año los resultados, en una primera lectura o borrador, de los tres genomas, siendo el resultado para *L. major* bien acabado y completo (Ivens *et al.*, 2005), en segundo lugar *T. brucei* (Berriman, *et al.*, 2005) al cual aún le faltan las secuencias de los minicromosomas, y finalmente *T. cruzi* (El-Sayed *et al.*, 2005) cuyas secuencias no pudieron ser organizadas en cromosomas individuales. El secuenciamiento del genoma de *T. cruzi* fue bastante difícil por: su alto contenido de ADN; presencia de un gran número de bandas cromosomales que no pueden ser bien resueltas por electroforesis PFG; alto contenido de secuencias repetidas que dificultan el ensamblaje in silico (mediante computadoras); y por último, pero no menos importante, la cepa CL Brener resultó ser un híbrido celular de dos linajes de *T. cruzi*, o también mal llamados haplotipos. Es decir en esta cepa coexisten dos genomas algo distintos los cuales no se han integrado.

Algunas de las conclusiones importantes de estos estudios son: i) el gran número de datos generados por los proyectos genoma permiten hacer un análisis global de todo el repertorio de proteínas de superficie, y en este sentido, las perspectivas para una vacuna para el mal de Chagas lucen remotas teniendo en cuenta que en *T. cruzi* se han descubierto nuevas y más complejas familias de proteínas de superficie.; ii) se han revelado rutas únicas en la reparación del ADN, su replicación, y los procesos de segregación; iii) una ausencia de muchos de los receptores descritos para otras células eucarióticas; iv) una serie de fosfatasa y quinasas únicas que podrían servir como nuevo blancos quimioterapéuticos.

Para los países en vías de desarrollo los resultados de estos proyectos genomas proveen información de gran utilidad, por ejemplo en la actualidad para clonar cualquier gen o región del genoma solo basta con identificar la secuencia en las bases de datos, diseñar los cebadores o primers oligonucleotídicos alrededor de la secuencia, hacer su amplificación por PCR, e insertar el producto de amplificación en un plásmido vector. Las bases de datos también permiten hacer búsquedas focalizadas (datamining) de genes susceptibles a una droga en particular. En este sentido, es de suma importancia el preparar recursos humanos en bioinformática para llevar a cabo una investigación in silico la cual no solo

es barata, requiriendo computadoras y programas de libre acceso, sino que tiene un gran potencial.

Del genoma a las ómicas

Los proyectos genoma permitieron visualizar muchas características importantes de los Kinetoplastida, pero la gran cantidad de datos generados necesitaban ser examinados por tecnologías y programas de computación especiales. Esta aproximación usando técnicas de gran formato (en Inglés High Output) que permiten estudiar simultáneamente muchos genes o muchas proteínas es lo que da el nombre a las ómicas, en otras palabras en vez de estudiar gen por gen o producto por producto, se estudian todos en conjunto o cuerpo entero.

Genómica

La genómica es el estudio simultáneo de la expresión (o no) de todos los genes de un organismo en distintos estadios y condiciones. Tan pronto como comenzaron a salir las primeras secuencias completas de los proyectos genoma en el 2003, empezaron las primeras experiencias con las tecnologías de la genómica. El tipo de experimento más común en genómica utiliza micromatrices, o también llamados microarreglos en los cuales se imprimen, mediante técnicas de microrrobótica, en un portaobjeto de vidrio todos los genes de un organismo, o los ADN copia producidos desde los mensajeros codificados por esos genes (Schena *et al.*, 1995). También se pueden arreglan puntos de proteínas o anticuerpos. El arreglo de estos puntos está registrado en la computadora, en donde también está almacenada la información de secuencia para cada punto. Un experimento típico de genómica, consiste en hibridar simultáneamente una micromatriz con los ADNc (sintetizados a partir de los ARN Mensajeros o ADN copia) de un estadio marcándolos previamente con un fluoróforo (rodamina), y los ADNc aislados de otro estadio marcados con verde CYBER. Luego de la hibridación, el microarreglo se lee en un lector láser que tiene varias longitudes de onda de excitación, y así podemos cuantificar la cantidad de un mensajero producido por un determinado gen, en un determinado estadio (Fig. 3).

En el caso de los Kinetoplastida este tipo de experimentación no ha dado buenos resultados, ya que como se ha mencionado anteriormente, los

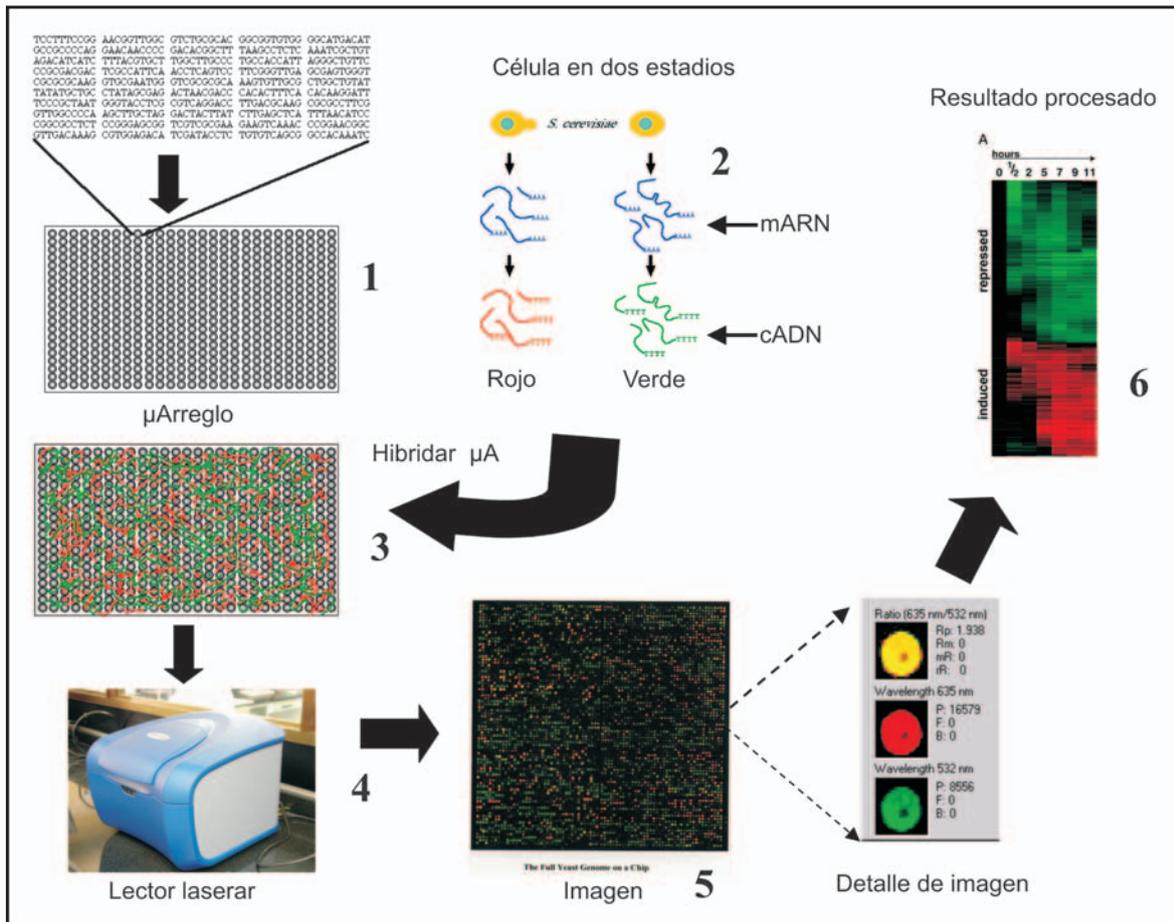


Fig. 3. Uso de Micromatrices en el estudio de la expresión genética: 1. Microarreglo de secuencias u oligonucleótidos de un organismo; 2. Preparación de ADN copia marcado con fluoróforos a partir de los ARN mensajeros de una célula en dos situaciones distintas; 3. Hibridación de los ADN copia a la micromatriz; 4. Lectura de la micromatriz hibridada con los ADN copia en un lector láser de varias longitudes de onda; 5. Conversión y agrupación de las señales en grupos de genes con igual nivel de expresión.

genes de un estadio son transcritos todos al mismo tiempo en largos mensajeros policistrónicos, de ahí que no se observen claramente en cuales genes sube o baja su expresión, y las diferencias entre estadios no han sido tan marcadas ni han conducido a ningún descubrimiento importante. Una alternativa que examina los productos post-transcripcionales es la de aislar polisomas (o ribosomas unidos a mensajeros) y hacer ADNc a partir de dichos mensajeros, así estaríamos estudiando mensajeros que han sido procesados (como se indicó arriba con su señal para el ribosoma y su cola de poliA) y que están

siendo activamente traducidos a proteínas. Usando esta aproximación Avila *et al.* (2003) han logrado observar diferencias estadio-especifica importantes. El factor limitante para esta aproximación es la pequeñísima cantidad de mensajeros que se puede recuperar de las formas metacíclicas sanguíneas, y las formas amastigotas intracelulares, las cuales a su vez vienen contaminadas con los ARN mensajeros del hospedador vertebrado.

Existe otra modalidad de micromatrices en las cuales se sintetizan sobre un chip (un cristal

de germanio de aproximadamente 1 centímetro cuadrado) oligonucleótidos correspondientes a partes de los genes mediante un proceso de fotolitografía (www.affymetrix.com/) la detección se hace por vía electrónica. Sin bien las micromatrices no han sido de mucha utilidad en el estudio de la expresión génica en Kinetoplastida, prometen ser importantes instrumentos para el diagnóstico del futuro, de hecho ya existen técnicas y equipos aprobados por la FDA (Food and Drug Administration) de los Estados Unidos para detectar las distintas variantes alélicas de los genes de Citocromo P450 genes involucrado en la degradación de drogas. (http://www.roche-diagnostics.com/media/pdf/presskit/final_technology_behind_cyp450.pdf).

Proteómica

Proteómica es el estudio a gran escala de las proteínas de una célula, particularmente sus estructuras y funciones. El término fue acuñado por analogía con la genómica, pero es mucho más complicada que esta última. Mientras el genoma es una entidad más bien constante, el proteoma difiere de célula a célula, y experimenta cambios a través de sus interacciones bioquímicas con el genoma y el ambiente. Así, un organismo tendrá un perfil de expresión de proteínas radicalmente diferente en diferentes partes de su cuerpo, en diferentes estadios del ciclo celular y bajo diferentes condiciones ambientales. Las técnicas de la proteómica se están aplicando ampliamente en Kinetoplastida.

La técnica más utilizada en la proteómica es la electroforesis en dos dimensiones, separando mezclas de proteínas por tamaño molecular en la primera dimensión, y por carga en la segunda (Fig. 4). Las nuevas técnicas de 2D, como se abrevia esta tecnología, utilizan geles más grandes con mayor poder de resolución, además de usar marcaje diferencial entre las proteínas que vienen de un estadio u otro o en presencia de una droga y su control. Los marcajes pueden hacerse con fluoróforos o con isótopos, los cuales son leídos y analizados por lectores especiales. La segunda tecnología de más amplio uso es la espectrometría de masas acoplada a sistemas de ionización de péptidos producto de la digestión (Cristea *et al.*, 2004) de una proteína con proteasas específicas, y el análisis de estos péptidos, los cuales son ionizados y analizados hasta llegar a la determinación de su secuencia primaria. Con esta tecnología se están logrando identificar y validar la

mayoría de los genes putativos registrados en los proyectos genoma, permitiendo también identificar las distintas formas de una proteína que se producen luego de modificaciones post-traduccionales.

El interactoma

Recientemente se ha llegado a la conclusión que para poder entender como operan los sistemas complejos de los sistemas eucarióticos, no basta con conocer como se expresan los genes, o cuantas proteínas distintas se expresan bajo diferentes condiciones, si no que es necesario conocer como interactúan las proteínas entre si. Estas interacciones entre proteínas no son fáciles de predecir basándonos solamente en estructuras, modificaciones o secuencias. Por ejemplo a veces una proteína ejerce una influencia sobre otra no directamente sino a través de una tercera proteína. También una misma proteína interviene en procesos distintos asociándose con distintas proteínas.

Las tecnologías de uso más frecuentemente para detectar interacciones entre proteínas son: i) cromatografía de afinidad en la cual se adhiere una proteína a una matriz insoluble, se hace pasar otra proteína y se determina si hay alguna asociación entre ellas por la retención de la segunda proteína; ii) técnicas de Resonancia Plasmón (nombre del aparato BIACORE, <http://www.biacore.com/lifesciences/index.html>), en las cuales la interacción entre proteínas se hace casi de la misma forma que en con cromatografía de afinidad, pero en este caso la proteína inmovilizada está adherida a un cristal, cuya parte posterior está cubierta de oro. Un rayo láser que incide sobre la superficie del oro causa una onda llamada plasmón. Al pasar otra proteína que interactúa con la primera ocurren cambios en la onda plasmón que son detectadas por el instrumento de detección. Esta técnica es muy sensible y poderosa; iii) sistema de interacción dos híbridos (2H), esta prueba es de naturaleza genética y está basada en sistemas de expresión desarrollados para levaduras. En la misma un activador de la expresión génica (transcripción) ha sido dividido en dos partes. Para que este activador funcione se necesita que las dos partes estén juntas sobre la región promotora de un gen reportero. Las dos partes que han sido separadas se fusionan mediante técnicas de ADN recombinante a dos genes de proteínas que queremos probar. Cuando estas proteínas fusionadas son co-expresadas dentro de la célula de la levadura, si las partes fusionadas

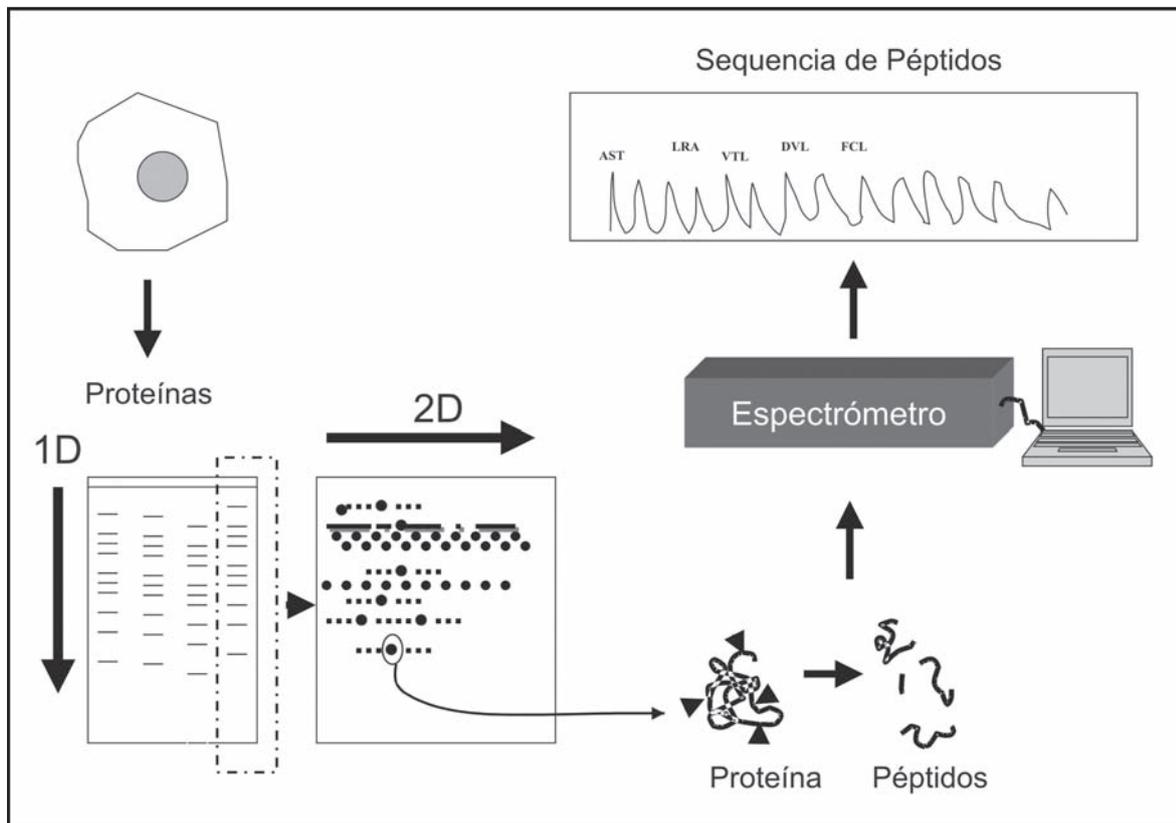


Fig. 4. Experimentos clásicos de proteómica. La mezcla de proteína se resuelve en una dimensión (por tamaño), luego se corta uno de los carriles y se corre en una segunda dimensión separando por carga. Los puntos de proteínas revelados por tinción pueden ser recuperados, digeridos con proteasas, y los péptidos producidos estudiados por espectrometría de masas hasta llegar a la secuencia de los mismos.

interactúan entre sí, se recompone el activador y se da la expresión del gen reportero acoplado a este activador (Fig. 5). Utilizando tecnología 2H se ha logrado hacer un primer borrador del interactoma de *Plasmodium falciparum* (La Count *et al.*, 2005) (Fig. 6); iv.) otra manera de descubrir interacciones múltiples entre proteínas es a través de la búsqueda bibliográfica masiva de artículos, a esto se le llama en Inglés “data mining” y para hacer este tipo de búsqueda existen varios programas de computación de libre acceso en la WWW (por ejemplo www.chilibot.net). Los resultados de los estudios de interactoma, nos permiten ver múltiples y noveles interacciones que no son detectables por otras vías. Los resultados del interactoma son diagramas de redes complejas en las cuales se evidencian aquellas proteínas que son nodos importantes dentro de las redes. Este tipo de estudio puede ser de mucha importancia para

identificar blancos sobre las cuales deben diseñarse drogas o aproximaciones de regulación.

El metaboloma

Siguiendo las mismas pautas de las descripciones anteriores, en el caso del metaboloma se trata de descubrir a partir de los resultados de los proyectos genoma, del proteoma y el interactoma, aquellas relaciones entre genes y proteínas que tienen que ver con las rutas metabólicas de un organismo y las conexiones entre distintas rutas. Haciendo uso de importantes bases de datos públicas como la Kegg (<http://www.genome.ad.jp/kegg/pathway/map/map01100.html>), así como de programas de computación, puede llegarse a predecir nuevas rutas antes no detectadas, y determinar la ausencia de una enzima en una ruta metabólica. Estos estudios van

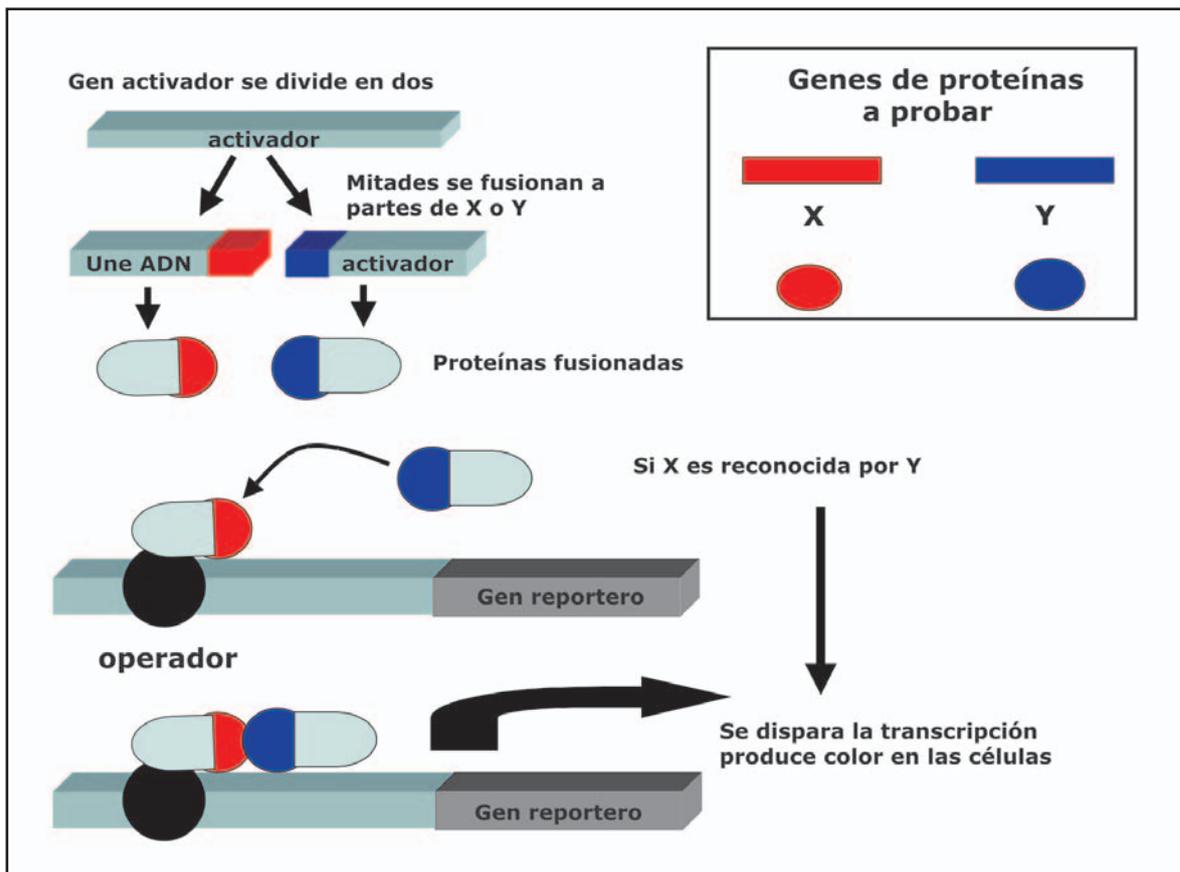


Fig. 5. Sistema de dos híbridos (2H) para estudiar interacciones entre proteínas. Los genes clonados correspondientes a dos proteínas que se quiere saber si interactúan entre si (X y Y) se fusionan con una parte de una proteína activadora que se une a un promotor de la transcripción de levadura. Para que ocurra la transcripción del gen reportero, es necesario que las dos proteínas (o partes de ellas) se reconozcan y ayuden así a reconstruir al activador, sobre la región operadora del gen. Usualmente se detecta esta interacción en la célula de levadura receptora por cambio de la coloración producida por el gen reportero.

acompañados de otros bioquímicos y estructurales. En Kinetoplastida se ha hecho un estudio comparativo de su metaboloma (Merriman *et al.*, 2005) en el mismo se confirman blancos quimioterapéuticos descritos en la era pre-genómica tales como: i) la síntesis de ergosterol en *Leishmania* y *T. cruzi*, ii) la ruta metabólica del tripanotión, iii) la ausencia de una ruta para sintetizar grupos hemo y iv) las enzimas de las vías glicolíticas. Pero aparecen nuevos y promisorios blancos como: i) los transportadores de hexosas fosfatos en *T. cruzi*; ii) la síntesis de aminoetilfosfonato, una alternativa al ancla GPI, exclusiva de *T. cruzi*. Aún es prematuro afirmar que la información del metaboloma de los Kinetoplastida proveerá una salida inmediata a la

quimioterapéutica de las enfermedades que ellos producen, pero los investigadores tienen en sus manos para estudiar alrededor del 50% de los genes putativos de los cuales aún no se les conoce ninguna función.

La biología de sistemas

Los métodos masivos de estudio usados en la actualidad hacen que la biología pase de ser una ciencia pobre en datos a una rica en ellos, tal como lo es la física. Por ello es necesario recurrir a nuevas aproximaciones integrales u holísticas para poder analizar esta enorme cantidad de información a fin de que se convierta en conocimiento útil.

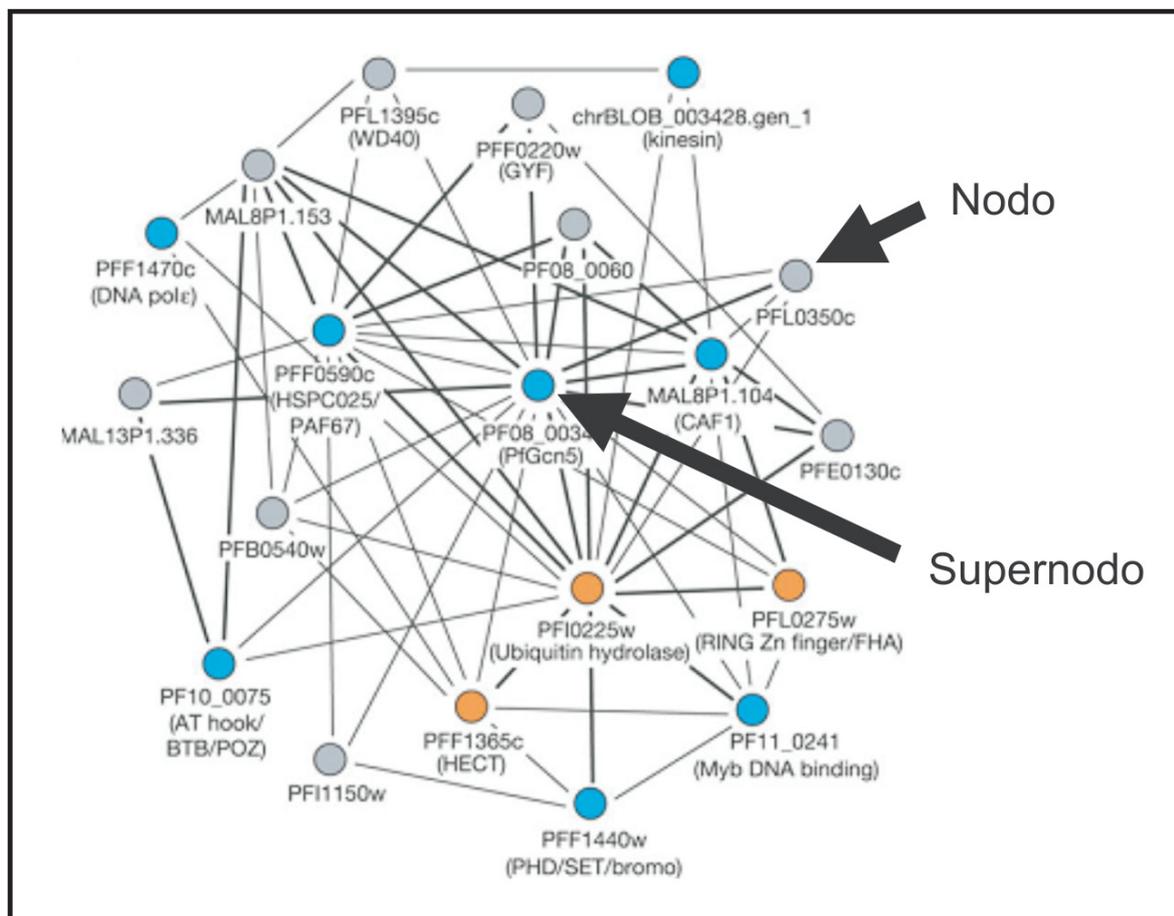


Fig. 6. Una red del interactoma de *Plasmodium falciparum* (La Count et al., 2005) correspondiente a genes que en su mayoría están asociados a remodelación de la cromatina, se indican los nodos muy bien conectados o Hubs y los nodos mas simples.

El área de la nueva biología que se encarga de desarrollar nuevos algoritmos y métodos de estudio para racionalizar y entender los datos generados por los distintos proyectos se le llama biología de sistemas o biología sistémica, término que fue acuñado en Inglés como “Systems Biology” (Kitano, 2001).

El conocimiento y manejo de sistemas complejos es familiar y común a los ingenieros, por ejemplo a aquellos que estudian y diseñan aeroplanos. Estos aparatos tienen muchos de los atributos de los sistemas biológicos tales como la de ser robustos, y la redundancia de funciones para alcanzar esta cualidad. La biología de sistemas nos permitirá abordar la relación parásito-hospedador en una forma integral, teniendo una visión no constreñida a ventanas de información que no se traducen a información útil o conocimiento.

Las distintas interacciones que se describen a partir de las ómicas pueden ser graficadas, como fue mencionado para el interactoma de *Plasmodium*, como conexiones entre un punto (nodos) (proteína o gen) y los demás (Fig. 6). La biología de sistemas analiza las estructuras de las redes para determinar su arquitectura organizándolas en redes o super-redes con distintos grados de jerarquía. También trata de descifrar las leyes que regulan su organización e identificar a los nodos mas ricos en conexiones, a los que se llaman Hubs o supernodos, estos supernodos constituyen puntos clave para la regulación de los sistemas biológicos, y por ello pueden ser de gran importancia para diseño de drogas, o como puntos de diagnóstico del funcionamiento de un determinado sistema. En el ejemplo del interactoma de *P. falciparum* una alternativa a la búsqueda de antígenos superficiales para desarrollo de vacunas,

sería la de atacar el super-nodo PfGcn5 que centra las actividades de remodelación de la cromatina de este parásito (Fig. 6). Si extrapolamos este mismo razonamiento a *T. cruzi*, los resultados del proyecto genoma revelan una complejidad en la organización, contenido, y composición de las proteínas de superficie que sugieren un sistema que podríamos llamar surfoma, en el cual se expresan de manera coordinada, dependiendo del estadio del ciclo celular y del ambiente, familias y subfamilias de genes cuya dinámica serían importante dilucidar. Este hecho nos lleva a pensar que al enfocarnos en la superficie de este parásito para tratar contrarrestarlo caemos en su juego de evasión, el cual es el producto de miles de años de evolución. Tal vez este tipo de investigaciones deberían de estar dirigidas hacia otros tantos blancos identificados y por identificar en el proyecto genoma.

CONCLUSIONES

Aunque aun resulta difícil medir el impacto de las tecnologías más recientes en la lucha contra la *Leishmaniasis*, el mal de Chagas y la enfermedad del sueño, los proyectos genoma de sus agentes etiológicos han brindado un caudal de información que necesita ser sistematizado y aprovechado. En especial los países en vías de desarrollo deberíamos entrenar investigadores en como explotar y obtener información útil de estos proyectos los cuales son de dominio público. Por otro lado, el cúmulo de resultados también obliga a la aplicación de aproximaciones sistémicas para poder estudiar la complejidad de los datos, dando una visión biológica integral más en consonancia con las complejas interacciones existentes en la relación parásito-hospedador. En algunos parásitos como *Plasmodium falciparum* (Gardner *et al.*, 2002) el ciclo de genomas interactuando se ha cerrado con la publicación del genoma humano; (Venter J.C. *et al.*, 2001), (Lander *et al.*, 2001) y el de *Anopheles gambiae* (Holt *et al.*, 2002), pero todavía no contamos con las herramientas bioinformáticas adecuadas para detectar a gran escala el diálogo genético que se establece entre el parásito y el organismo que lo alberga.

El abordaje sistémico nos lleva también a dejar de un lado el enfoque reduccionista de la biología molecular del siglo XX que nos permitía ver partes pero no el todo. El entender la dinámica

de las redes complejas de los componentes genéticos de estos parásitos, nos ayudarán a orientar nuestras investigaciones a la búsqueda de los puntos más débiles de estas redes y así ejercer un control sobre ellos.

Impact of molecular biology, genomics and proteomics in the study of Kinetoplastida

SUMMARY

Kinetoplastida is an Order of Protozoa that includes important pathogens to humans, animals and plants. Here, a historical recounting of the impact of molecular biology on Kinetoplastida biology is made. This review is mainly focused on the new developments originated after the publication of the genome sequences of *Leishmania major*, *Trypanosoma cruzi*, and *Trypanosoma brucei*, emphasizing those subjects where useful information can be derived to fight these parasites. In addition, the new tendencies that may be relevant to parasitology in the near future are discussed, such as the application of systems biology to study the host-parasite interaction. Systemic considerations are made to direct the attention to new metabolic or genetic hubs, de-emphasizing the attack of the complex parasite system of surface proteins or surfome. This review is aimed to non-molecular biology specialists.

Key words: Parasite genome, proteomics, systemic approaches, host-parasite interaction

REFERENCIAS

- Avila A. R., Dallagiovanna B., Yamada-Ogatta S. F., Monteiro-Goes V., Fragoso S. P., Krieger M. A. & Goldenberg S. (2003). Stage-specific gene expression during *Trypanosoma cruzi* metacyclogenesis. *Genet. Mol. Res.* **31**: 159-68.
- Berriman M., Ghedin E., Hertz-Fowler C., Blandin G., Renauld H., Bartholomeu D. C. *et al.* (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Science*. **309**: 416-435.
- Botstein D., White R. L., Skolnick M. & Davis R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32**: 314-3.

- Chance M. L., Schnur L. F., Thomas S. C. & Peters W. (1971). The biochemical and serological taxonomy of *Leishmania* from the Ethiopian zoogeographical region of Africa. *J. Mol. Biol.* **56**: 443-473.
- Cristea I. M., Gaskell S. J. & Whetton A. D. (2004). Proteomics techniques and their application to haematology. *Blood*. **103**: 3624-3634.
- Degrave W., Levin M. J., Franco Da Silveira J. & Morel, C.M. (1994). Parasite Genome Projects and the *Trypanosoma cruzi* Genome Initiative. Mem. Inst. Oswaldo Cruz. **92**: 859-862.
- El-Sayed N. M., Myler P. J., Bartholomeu D. C., Nilsson D., Aggarwal G., Tran A., *et al.* (2005). The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science*. **309**: 409-415.
- Feagin J. M., Abraham J. E. & Stuart K. (1988). Extensive editing of the cytochrome C oxidase III transcript in *Trypanosoma brucei*. *Cell*. **53**: 413-422.
- Gardner M. J., Hall N., Fung E., White O., Berriman M., Hyman R. W. *et al.* (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*. **419**: 498-511.
- Holt R. A., Subramanian G. M., Halpern A., Sutton G. G., Charlab R., Nusskernet D. R. *et al.* (2002). The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science*. **298**: 129-49.
- Ivens A. C., Peacock C. S., Worthey E. A., Murphy L., Aggarwal G., Berriman M. *et al.* (2005). The genome of the kinetoplastid parasite *Leishmania major*. *Science*. **309**: 436-442.
- Johnson P.J. & Borst P. (1986). Mapping of VSG genes on large expression-site chromosomes of *Trypanosoma brucei* separated by pulsed-field gradient electrophoresis. *Gene*. **43**: 213-220.
- Koza J. R., Myldlowec W., Lanza G., Yu J. & Keane M. A. (2001). *Automated reverse engineering of metabolic pathways from observed data by means of genetic programming*. 95-117. En: Foundations of Systems Biology. Ed. Kitano H. Primera edición. The MIT Press, USA y London, UK.
- LaCount D. J, Vignali M. , Chettier R., Phansalkar A., Bell R., Hesselberth J. R., *et al.* (2005). A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature*. **438**: 103-107.
- Lanar D.E, Levy L.S & Manning J.E..(1978). Complexity and content of the DNA and RNA in *Trypanosoma cruzi*. *Ann. Trop. Med. Parasitol.* **72**: 533-42.
- Lander E., Linton L. M., Birren B., Nusbaum C., Zody M. C., Baldwin J. *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature*. **409**: 860-921. .
- Martinez-Calvillo S., Yan S., Nguyen D., Fox M., Stuart K. & Myler P. J. (2003). Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. *Mol Cell*. **11**: 1291-1299.
- Meselson M., Yuan R. & Heywood J. (1972). Restriction and Modification of DNA. *Ann. Rev. Biochem.* **41**: 447-466.
- Saiki R. K., Scharf S., Faloona F., Mullis K. B., Horn G. T., Erlich H. A. & Arnheim N. (1985). Enzymatic Amplification of β -Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia. *Science*. **230**: 1350-1354
- Schena M., Shalon D., Davis R. W. & Brown P. O. (1995). Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. *Science*. **270**: 467-470.
- Schwartz D. C. & Cantor C. R. (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell*. **37**: 67-75.
- Simpson L. & Da Silva A. (1981). Isolation and characterization of kinetoplast DNA from *Leishmania tarentolae*. *Mol. Biochem. Parasitol.* **3**: 327- 41.
- Tibayrenc M., Kjellberg F. & Ayala F. J. (1991). The Clonal Theory of Parasitic Protozoa. *BioScience*. **41**: 767-774.

Venter J. C., Adams M. D., Myers E. W., Li P. W., Mural R. J., Sutton G. G. *et al.* (2001). The sequence of the human genome The sequence of the human genome. *Science*. **291**: 1304-51.

Villalba E. & Ramírez J. L.(1982). Ribosomal RNA of *Leishmania brasiliensis*: number of ribosomal copies and gene isolation. *J. Protozool.* **29**: 438-441.

Recibido el 14/07/2006
Aceptado el 08/08/2006
