



Universidad Nacional Autónoma de México

Maestría en Ciencias Bioquímicas

Facultad de Química

Identificación de proteínas diferencialmente expresadas entre aislados de *Leishmania mexicana* de pacientes con formas clínicas difusa (LCD) y localizada (LCL)

TESIS

Para optar por el grado de:

Maestro en Ciencias

Presenta:

I.B.T. Said Alejandro Muñoz Montero

Tutor Principal

Dr. Ruy Pérez Montfort

Instituto de Fisiología Celular (IFC, UNAM)

Miembros del Comité Tutor

Dra. Claudia Rangel Escareño (INMEGEN)

Dr. Sergio Encarnación Guevara (CCG, UNAM)

Ciudad de México

Diciembre, 2018



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Tabla de contenido

AGRADECIMIENTOS	¡ERROR! MARCADOR NO DEFINIDO.
1. INTRODUCCIÓN	8
1.1 LEISHMANIASIS	8
1.2 MANIFESTACIONES CLÍNICAS DE LA LEISHMANIASIS	8
1.3 DISTRIBUCIÓN DE CASOS A NIVEL MUNDIAL	10
1.4 LEISHMANIASIS EN MÉXICO.....	11
1.5 <i>LEISHMANIA MEXICANA</i>	13
1.5.3. PROTEÓMICA DEL PARÁSITO	17
1.5.4. PROTEÓMICA DE <i>LEISHMANIA MEXICANA</i>	18
1.6 GENERALIDADES DE PROTEÓMICA.	19
1.7 CROMATOGRAFÍA LÍQUIDA Y ESPECTROMETRÍA DE MASAS (LC-MS/MS).....	19
1.8 PROTEÓMICA EN TANDEM TIPO <i>BOTTOM-UP</i>	20
1.8.1 CUANTIFICACIÓN LIBRE DE MARCAJE	21
2. JUSTIFICACIÓN	22
3. HIPÓTESIS	22
4. OBJETIVOS	23
4.1 OBJETIVO GENERAL.....	23
4.2 OBJETIVOS ESPECÍFICOS.....	23
5. MATERIALES Y MÉTODOS	24
5.1 PACIENTES.....	24
5.2 EXTRACCIÓN DE PROTEÍNAS	25
5.3 WESTERN-BLOT (WB).....	26
5.4 LC-MS/MS.....	27
5.5 BÚSQUEDA EN BASES DE DATOS E IDENTIFICACIÓN DE PROTEÍNAS	27
5.6 PARÁMETROS DE BÚSQUEDA	28
5.7 IDENTIFICACIÓN DE PROTEÍNAS.....	28
5.8 CUANTIFICACIÓN Y NORMALIZACIÓN	28
5.9 ANÁLISIS DE ENRIQUECIMIENTO	29
5.10 IDENTIFICACIÓN DE PROTEÍNAS NO ANOTADAS.....	29
5.11 GRÁFICOS.....	29
6. RESULTADOS	30
6.1 INTEGRIDAD DE LAS MUESTRAS.....	30
LA INTEGRIDAD DE LAS MUESTRAS ANTES DE SER PROCESADAS EN LOS ENSAYOS DE LC-MS/MS SE ANALIZÓ CORRIENDO LOS EXTRACTOS PROTÉICOS EN GELES SDS-PAGE DESNATURALIZANTES TEÑIDOS CON AZUL DE COOMASSIE. TODAS LAS MUESTRAS ESTABAN ÍNTEGRAS Y MOSTRABAN PESOS MOLECULARES ENTRE 10 Y 300 kDA (FIGURA 8).....	30
6.2 IDENTIFICACIÓN DE PROTEÍNAS A PARTIR DE LOS DATOS LC-MS/MS	31
6.3 PROTEÍNAS EXCLUSIVAS PARA LCL.....	32
6.4 PROTEÍNAS EXCLUSIVAS PARA LCD	33
6.5 EXPRESIÓN DIFERENCIAL	34
6.5.1 Distribución de la expresión de proteínas por muestra.	34

6.5.2 Gráfica de densidad	35
6.6 PROTEÍNAS DIFERENCIALMENTE EXPRESADAS.....	36
6.7 MAPA DE CALOR	44
6.8 ANÁLISIS DE ENRIQUECIMIENTO	46
6.9 PROTEOMICA	48
MEJORAMIENTO DE LA ANOTACIÓN DE <i>LEISHMANIA MEXICANA</i>	48
6.10 VALIDACIÓN	52
7. DISCUSIÓN	53
8. CONCLUSIONES	55
9. REFERENCIAS	57

El presente trabajo se realizó en el laboratorio de Bioquímica y Biología Estructural bajo la dirección del Dr. Ruy Pérez Montfort, en colaboración con el Departamento de Genómica Computacional del Instituto Nacional de Medicina Genómica y el Laboratorio de Inmunoparasitología de la Unidad de Investigación en Medicina Experimental de la Facultad de Medicina de la Universidad Nacional Autónoma de México.

Este trabajo fue apoyado por el Consejo Nacional de Ciencia y Tecnología (CONACyT) a través de los donativos No. 254694 y FINNOVA 242368. PAPIIT IN211418.

El alumno fue becado por CONACyT para los estudios de maestría con CVU 467639.

Lista de Abreviaciones y Símbolos

BALB/c	Cepa de ratón albino
BLASTP	Herramienta de Alineamiento Básico Local para Proteínas
CID	Disociación Inducida por Colisión
Contenido GC	Contenido de Guanina Citosina
ELISA	Ensayo por inmunoadsorción ligado a enzimas
FCh	Tasa de Cambio
FDR	Tasa de Descubrimiento Falso
GO	Ontología de Genes
InterPro	Plataforma con acceso a base de datos de familias, dominios y sitios funcionales de proteínas.
LC	Cromatografía Líquida
LCD	Leishmaniasis Cutánea Difusa
LCL	Leishmaniasis Cutánea Localizada
m/z	Relación masa carga
MALDI	Desorción / Ionización Láser Asistida por Matriz
MCL	Leishmania Mucocutánea
MPM	Marcador de Peso Molecular
MS	Espectrometría de Masas
NCBI	Centro Nacional (Estadounidense) para la Información Biotecnológica
NSAF	Factor de Abundancia espectral normalizada
OMS	Organización Mundial de la Salud
PAGE	Gel de electroforesis de poliacrilamida
PBS	Amortiguador fosfato salino
PSM	Masa espectral de proteínas
PVDF	Polifluoruro de vinilideno
RIPA	Amortiguador de ensayo de radio inmunoprecipitación
SDS	Dodecilsulfato sódico
	SDS-PAGE definir
SFB	Suero Fetal Bovino
SPE	Extracción en fase sólida
tBLASTn	Herramienta de Alineamiento Básico Local para Proteínas en una base de datos de nucleótidos
TBST	Amortiguador Tris salino y Polisorbato 20
UniProt	Base de datos <i>Universal Protein</i>
VL	Leishmaniasis Visceral
WB	Western blot

1. INTRODUCCIÓN

1.1 Leishmaniasis

La infección por parásitos protozoarios del género *Leishmania* produce un amplio espectro de formas clínicas conocidas como leishmaniasis. De acuerdo con la Organización Mundial de la Salud (OMS), esta enfermedad afecta a 98 países y 3 territorios en 5 continentes [WHO, 2013].

La leishmaniasis es una de las 17 enfermedades clasificadas por la OMS como desatendidas, es decir, al padecerlas personas con pocos recursos económicos no representan un mercado lucrativo para la industria farmacéutica. Hasta el momento no existen tratamientos adecuados dada la alta toxicidad que presentan, o su costo representa una barrera de entrada para los principales países afectados por la enfermedad [RutaN, 2015].

Por otro lado, la leishmaniasis es una de las 10 enfermedades consideradas en expansión y para las cuales no hay instrumentos efectivos para su control. El número de casos aumenta en todo el mundo y son considerados insuficientes los instrumentos de intervención en sus diferentes aspectos: diagnóstico, terapéutico y control del vector [DNDi (Drug for Neglected Diseases initiative), 2010].

1.2 Manifestaciones clínicas de la leishmaniasis

La leishmaniasis tiene manifestaciones clínicas muy diversas, desde formas viscerales hasta cutáneas; estas diferencias se indican en la Tabla 1.

Tabla 1. Formas clínicas de la leishmaniasis y sus características

Forma Clínica	Siglas	Características clínicas	Agente etiológico
Cutánea Localizada	LCL	Úlceras cutáneas localizadas que pueden persistir por meses, en ocasiones años. Las lesiones evolucionan típicamente de placas nodulares a úlceras, que se pueden cubrir por una costra. Algunas lesiones persisten como nódulos. Las lesiones generalmente son indoloras. El proceso de cicatrización típicamente resulta en huellas atróficas.	<u>Viejo Mundo:</u> <i>Leishmania tropica</i> , <i>L. major</i> , <i>L. aethiopica</i> , <i>L. infantum</i> y <i>L. donovani</i> . <u>Nuevo Mundo:</u> Complejo de especies de <i>L. mexicana</i> (<i>L. mexicana</i> , <i>L. amazonensis</i> y <i>L. venezuelensis</i>); subgénero <i>Viannia</i> (<i>L. [V.] braziliensis</i> , <i>L. [V.] guyanensis</i> , <i>L. [V.] panamensis</i> y <i>L. [V.] peruviana</i>).
Cutánea Difusa	LCD	Lesiones cutáneas diseminadas en pacientes deficientes en la respuesta inmune celular frente al parásito.	<u>Viejo Mundo:</u> <i>L. tropica</i> , <i>L. major</i> , <i>L. aethiopica</i> , <i>L. infantum</i> y <i>L. donovani</i> . <u>Nuevo Mundo:</u> Complejo de especies de <i>L. mexicana</i> (<i>L. mexicana</i> , <i>L. amazonensis</i> y <i>L. venezuelensis</i>); subgénero <i>Viannia</i> (<i>L. [V.] braziliensis</i> , <i>L. [V.] guyanensis</i> , <i>L. [V.] panamensis</i> , y <i>L. [V.] peruviana</i>).
Mucocutánea	MCL	Destrucción parcial o completa de las membranas mucosas de la nariz, la boca y la garganta.	<u>Viannia:</u> subgenus (<i>L. [V.] braziliensis</i> , <i>L. [V.] panamensis</i> y <i>L. [V.] guyanensis</i>); además de <i>L. (Leishmania) amazonensis</i> .
Visceral	VL	Se caracteriza por episodios irregulares de fiebre, pérdida de peso, hepatoesplenomegalia y anemia.	<i>Leishmania tropica</i> , <i>L. major</i> , <i>L. aethiopica</i> , <i>L. infantum</i> y <i>L. donovani</i> .

1.3 Distribución de casos a nivel mundial

La distribución de la leishmaniasis cutánea en el mundo (Figura 1), está asociada a la presencia del vector, una mosquita hembra del género *Lutzomyia* o *Phlebotomus*, en climas cálidos. La enfermedad es endémica en muchas regiones tropicales y subtropicales del mundo [WHO, 2013].

Anualmente, se estiman 1.3 millones de nuevos casos, de los cuales 300,000 son viscerales. EL90% de estos ocurre en: Bangladesh, Brasil, Etiopía, India, Nepal, Sudán del Sur y Sudán, y un millón corresponden a leishmaniasis cutánea QUE principalmente afecta a personas en: Afganistán, Algeria, Brasil, Colombia, Irán, Paquistán, Perú, Arabia Saudita, Siria y Túnez o mucocutánea, para la que los habitantes de Brasil, Perú y Bolivia, son los principalmente afectados.

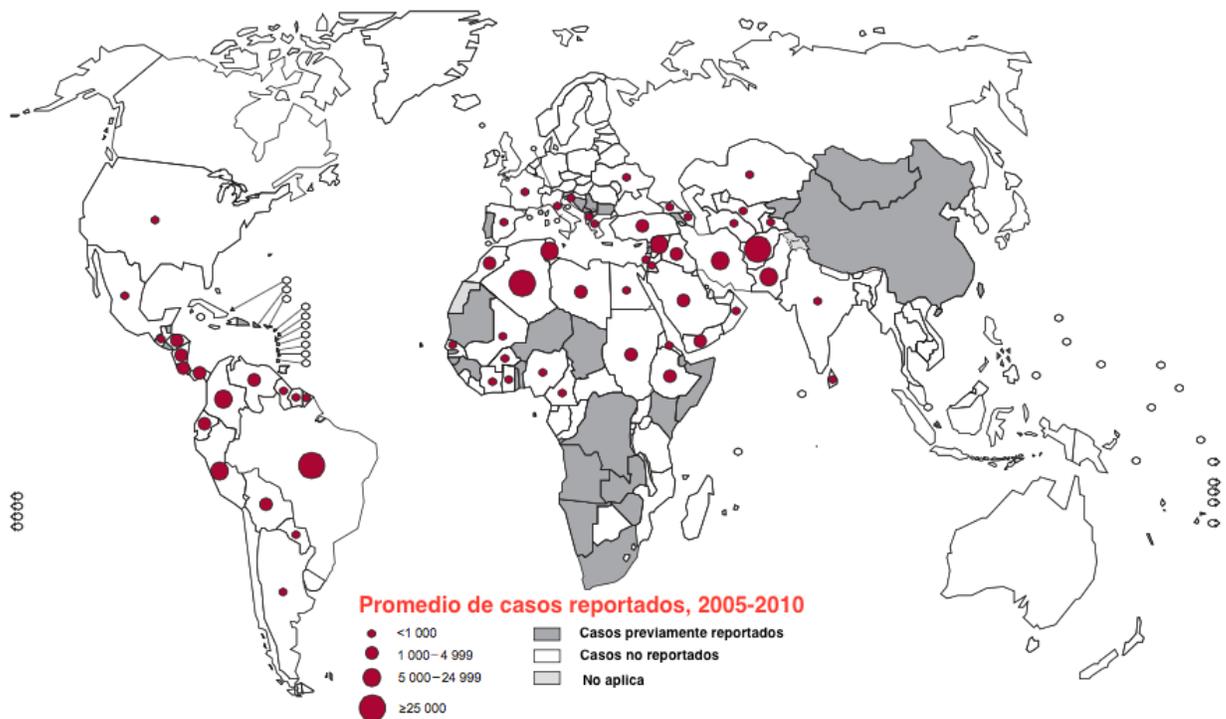


Figura 1. Distribución mundial de la leishmaniasis cutánea en el periodo 2005-2010 [WHO, 2013].

El reporte médico de la enfermedad es obligatorio en 33 de los 98 países afectados, pero el incremento real de los casos es una incógnita por el sesgo de

casos no reportados. A pesar de las predicciones, solamente se reportan oficialmente 600,000 casos [WHO, 2013].

1.4 Leishmaniasis en México

La leishmaniasis cutánea fue descrita por vez primera en la Península de Yucatán por Seidelin [Seidelin, 1912], quien la nombró "úlceras de los chicleros". En México, se han reportado las 4 formas clínicas de leishmaniasis: (1) Leishmaniasis cutánea localizada (LCL), leishmaniasis cutánea difusa (LCD), leishmaniasis mucocutánea (MC) y leishmaniasis visceral (LV), siendo la cutánea, la forma más común [Alvar et al., 2012].

La mayoría de los casos corresponden a LCL, principalmente en el sureste del país en los estados de Veracruz, Tabasco, Oaxaca, Chiapas, Nayarit, Yucatán, Quintana Roo y Campeche, lo que constituye el foco sur. También se han encontrado casos en Coahuila, Nuevo León, Tamaulipas, Hidalgo y San Luis Potosí (foco norte). Los casos de LCD tienen menor frecuencia y se localizan en ambos focos, con excepción de Yucatán y Quintana Roo (Figura 2). En los dos tipos de LCL la especie identificada ha sido *L. mexicana*. Los casos de MC se han reportado en los estados de Veracruz, Tabasco, Chiapas y Oaxaca, causados aparentemente por *L. braziliensis*. En cuanto a la LV, existe un foco en la cuenca del río Balsas, entre los estados de Guerrero, Puebla y Morelos, causados por *L. chagasi* y otro foco en Chiapas [Berrueta, 2015].

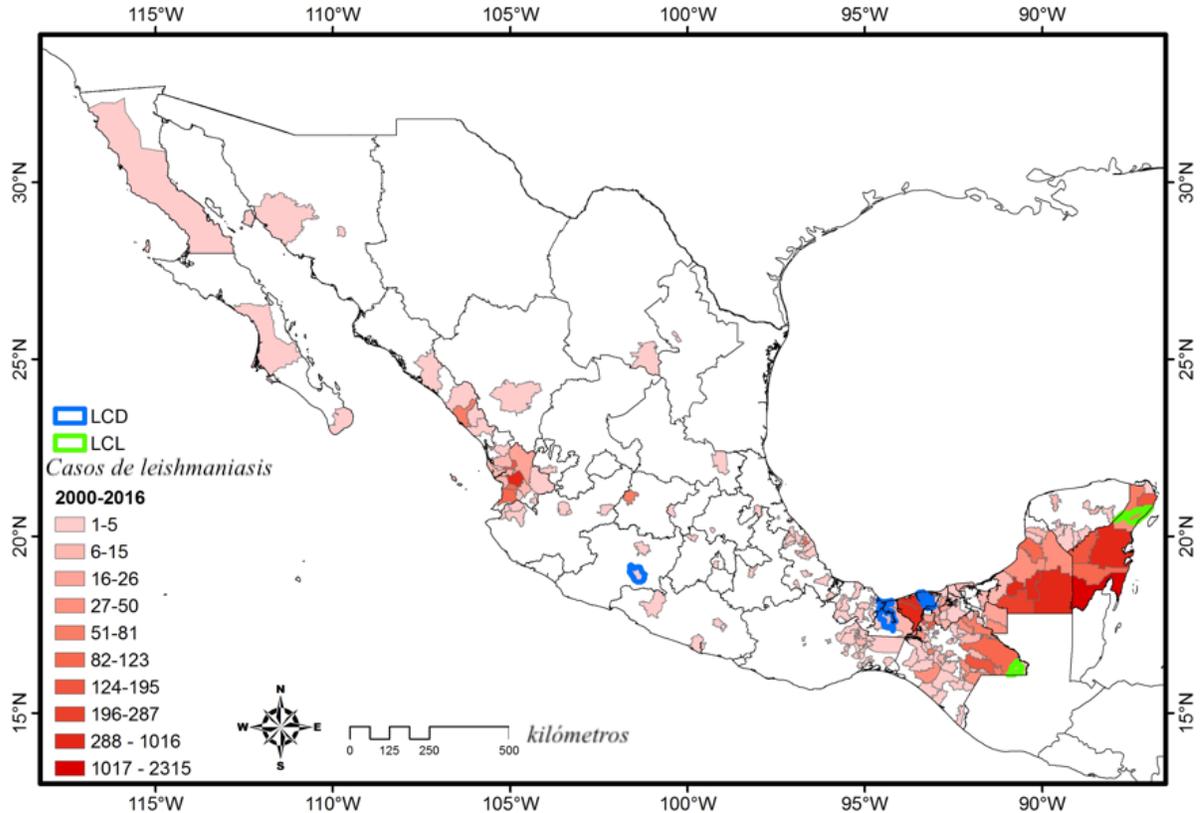


Figura 2. Distribución de la leishmaniasis cutánea en México (2000-2016). Datos tomados de la Secretaría de Salud. Mapa realizado por el Dr. César Ríos Muñoz (mapa no publicado).

En los Anuarios de Morbilidad de 1984-2013, emitidos por la Dirección General de Epidemiología de la Secretaría de Salud, se tienen registrados un total de 17,933 casos de leishmaniasis cutánea durante el periodo comprendido entre 1990-2013, careciendo de datos los años 1996 y 2000. La tasa de incidencia más elevada se registró en 1997 con 1.68 casos por cada 100,000 habitantes y la más baja en 1990 con 0.28 por cada 100,000 habitantes. La distribución estatal de los casos tiene el siguiente patrón: 36.17% (6,487) de los casos se registran en Tabasco, seguido por Quintana Roo 28.77% (5,159) y Campeche 12.67% (2,273), mientras que aquellos que solo tienen un reporte aislado dentro del periodo son: Coahuila, Hidalgo, México, Michoacán, Querétaro y Tamaulipas. A partir del año 2003 se comienza la notificación de los casos por sexo, registrándose 6,110 casos para el sexo masculino, comparados con 2,216 para el femenino, lo cual representa una proporción de 3:1.

1.5 Respuesta Inmune del hospedero

La respuesta inmune innata (RII) es la primera línea de defensa contra microorganismos patógenos, que es activada en minutos. El reconocimiento de los patógenos por parte de la RII se caracteriza por tener diversidad limitada, no tener memoria y tener una respuesta muy parecida frente a diferentes de agentes infecciosos. Sin embargo, la RII tiene una alta capacidad discriminatoria entre lo propio y lo extraño, como son los microorganismos patógenos [Abbas, 2004]. Los elementos esenciales de la inmunidad innata consisten en mecanismos de defensa como las barreras físicas y químicas, tales como los epitelios y sustancias antimicrobianas, el sistema del complemento, etc [Netea, 2011].

Las células que participan en la RII son las células fagocíticas (neutrófilos, monocitos, macrófagos), células dendríticas, leucocitos polimorfonucleares y las células NK (Natural Killer). Todas estas células reconocen patrones moleculares asociados a patógenos (PAMPs; Pathogen Associated Molecular Patterns) a través de receptores celulares denominados receptores de reconocimiento patrón (PRR, Pattern Recognition Receptors). La fagocitosis y la muerte de los microorganismos, son mecanismos efectores inespecíficos de la inmunidad natural del hospedero, y estos no inducen una memoria inmunológica [Netea, 2011].

Las interacciones hospedero-patógeno inician generalmente a través del reconocimiento de estructuras moleculares conservadas, esenciales para el ciclo de vida de los patógenos. El reconocimiento rápido de los PAMPs se lleva a cabo mediante PRRs que activan complejos de señalización que finalizan con la inducción de la inflamación mediada por varias citocinas y quimiocinas, lo que subsecuentemente trae como resultado la eliminación del patógeno. Diversos PRRs, como los TLRs, receptores tipo RIG-1(RLRs), receptores tipo NODs (NLRs) y los receptores de DNA, participan en este reconocimiento. Estos receptores son capaces de reconocer varias clases de moléculas como proteínas, lípidos, carbohidratos y ácidos nucleicos de diversos microorganismos [Fernández, 2013].

Los monocitos/macrófagos representan la principal diana del IFN- γ (producida entre otras células por linfocitos NK) [31]. El receptor del IFN- γ , que se encuentra en la membrana celular, reconoce al IFN- γ y tras la unión y la consiguiente oligomerización con las cinasas JAK 1 y 2, estas interaccionan entre sí y se activan por fosforilación recíproca y fosforilan determinados residuos de tirosina que se encuentran en la región intracitoplasmática de la cadena a del receptor. Dichos residuos son reconocidos por el dominio SH2 de la proteína STAT-1 α (Signal Transducer and Activator of Transcription) que es blanco de fosforilación por proteínas JAK. STAT-1 α se dimeriza dando lugar al complejo transcripcional GAF, que se transloca al núcleo para inducir la expresión de numerosos genes, entre ellos FcgR1 y factores de transcripción necesarios para la regulación de genes como IFN- α/β , MHC I y la iNOS.

1.5.1. Respuesta inmune a la infección por el parásito

La respuesta inmune contra la infección por *Leishmania* depende principalmente del tipo de célula del hospedero y de la especie de parásito. La resolución de la infección depende de una respuesta tipo Th1 junto con la producción de interleucina-2 (IL-2) e IFN- γ . Mientras que la susceptibilidad a la enfermedad se asocia con la falta de respuesta Th1, el desarrollo se asocia con una respuesta de Th2, caracterizada por la presencia de IL-4 e IL-10. El TNF- α producido por macrófagos activados y células NK también es importante para el control de las infecciones por *Leishmania*, ya que, junto con el IFN- γ , participa en la activación de macrófagos y en la producción de óxido nítrico e intermediarios reactivos del oxígeno que ayudan a la eliminación del parásito [Fernández, 2013].

La leishmaniasis cutánea se asocia con una respuesta inmune protectora, mientras que en la leishmaniasis visceral y la cutánea difusa la respuesta inmune no es capaz de eliminar al parásito. En el caso de la leishmaniasis mucocutánea, la intensa reacción inflamatoria genera un cuadro clínico muy grave, adicionalmente, la susceptibilidad a lesiones de la mucosa se relaciona con una respuesta Th1 hiperactiva, con un aumento de la producción de TNF- α e IFN- γ y una disminución en la producción de IL-10. Durante el desarrollo de la leishmaniasis mucocutánea, el mayor daño tisular se asocia con una baja expresión del receptor de IL-10, asociado a células citotóxicas CD4⁺ e inflamatorias, lo cual pudiera participar como posible mecanismo de progresión [Lupi, 2009].

Durante la infección por *Leishmania*, el sistema inmune innato de la piel desencadena dos vías de respuesta inmune complementarias: el reclutamiento de células proinflamatorias y la promoción e inducción de la respuesta inmune adaptativa incluyendo células T y la formación de anticuerpos [4, 5]. Diversas células del sistema inmune innato como los mastocitos, macrófagos y células dendríticas son las responsables de la detección de parásitos, mismas que, junto con la infiltración de neutrófilos, producen mediadores pro-inflamatorios que ayudan a contener y controlar al patógeno en los sitios de infección. Al mismo tiempo, estas células promueven la respuesta inmune adaptativa de células T, que se requieren para resolver eficazmente la infección de este parásito [Fernández, 2013].

Poco se sabe sobre los factores implicados en la modulación de la evolución de la leishmaniasis y entre los posibles factores se proponen los mediadores tempranos de la inflamación como las citocinas proinflamatorias y las quimiocinas que atraen a neutrófilos, cuya presencia se asocia con la inflamación excesiva y progresión de la enfermedad. La disminución del infiltrado por neutrófilos se correlaciona con un mejor control de la enfermedad [Villaseñor, 2008].

El papel facilitador de los neutrófilos está asociado con su capacidad para fagocitar a los parásitos y favorecer su transporte fuera del sitio de infección, evitando así los efectos tóxicos del complemento y de la respuesta inmune local. Estudios hechos en el modelo murino mostraron que durante la infección por *Leishmania* los neutrófilos entran en los vasos linfáticos favoreciendo la distribución de los parásitos hacia los tejidos periféricos [Fernández, 2013].

Los macrófagos constituyen el principal reservorio del parásito *in vivo* y a pesar de que no producen IL-12 después de la infección con *L. major*, son capaces de producir citocinas proinflamatorias como TNF- α en los sitios de infección y de producir NO mediante el cual participan en el control local de la infección [Von Stebut, 2007].

1.6 *Leishmania mexicana*.

1.6.1. Formas clínicas

L. mexicana puede causar dos formas clínicas de la enfermedad: LCL y LCD. La primera está caracterizada por úlceras en el sitio de inoculación del parásito (Figura 3A), mientras que la LCD, la forma más invasiva de la enfermedad, está caracterizada por macrófagos altamente parasitados y la formación de nódulos en todo el tegumento (Figura 3B). La causa de la progresión de la enfermedad es aún desconocida, pero se sospecha que los eventos de la respuesta inmune temprana durante el desarrollo de la enfermedad podrían establecer las condiciones que posiblemente determinan el resultado de la infección [Fernández, 2013].

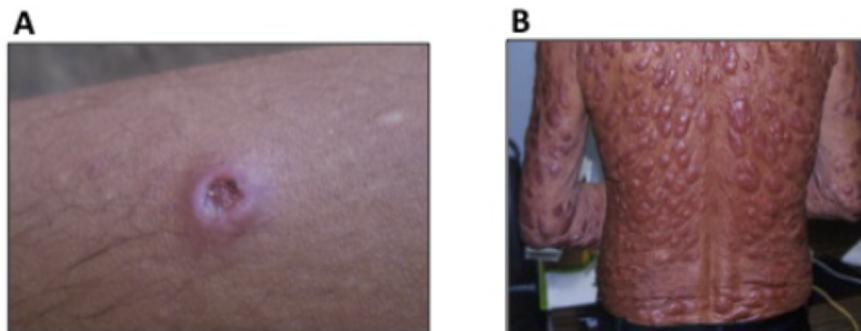


Figura 3. Diferentes formas clínicas de leishmaniasis cutánea generadas por *Leishmania mexicana*. A: Leishmaniasis cutánea localizada (LCL); B: Leishmaniasis cutánea difusa(LCD) [Fernández-Figueroa, 2013].

1.6.2. Factores de virulencia del parásito

Leishmania, al ser un organismo que vive dentro de células fagocíticas, necesita de mecanismos especializados para penetrar y sobrevivir en su célula blanco [14]. Los promastigotes de *Leishmania* están cubiertos por un glucocalix denso, compuesto por varias moléculas ancladas por glucosilfosfatidilinositol (GPI) como: proteofosfoglicanos (PPG), glucoinositolfosfoglicanos (GIPLS), fosfatasas ácidas de secreción, lipofosfoglicano (LPG) y la glicoproteína 63 (gp63). Esta última es una metaloproteasa dependiente de zinc, que entre otras funciones protege al parásito de la degradación por enzimas lisosomales del hospedero a través de la conversión proteolítica de C3b a C3bi en la superficie del parásito. El LPG y la gp63 son reconocidos por receptores de la célula hospedera, y logran inhibir diversos procesos proteolíticos del macrófago [15, 16].

El LPG es un importante factor de virulencia durante el ciclo de infección de *Leishmania* en mamíferos. Ha sido relacionado con la sobrevivencia del parásito dentro del vector y con la regulación de la adhesión e invasión celular dentro de la célula hospedera. LPG participa en la resistencia al complemento, evadiendo el complejo de ataque a membrana C5b-C9 [14].

1.5.3. Genoma del parásito

La especie *L. mexicana* produce leishmaniasis cutánea en el nuevo mundo, es la cuarta del género en haber sido secuenciada en 2011 y es representativa de 4 especies que forman un complejo. El genoma de *L. mexicana* tiene un tamaño de 32 Mb, con 34 cromosomas. El contenido de GC es aproximadamente de 59%. La cepa secuenciada fue *L. mexicana* MHOM/GT/2001/U1103, que se obtuvo del laboratorio de Byron Arana en Guatemala [Rogers, 2011].

Los genomas de *Leishmania spp.*, incluyendo a *L. mexicana*, están caracterizados por su alta densidad de genes, presentes en arreglos largos de genes policistrónicos, y la ausencia de intrones como se observa en la figura 4 [Siegel, 2011] [Ivens et al., 2005].

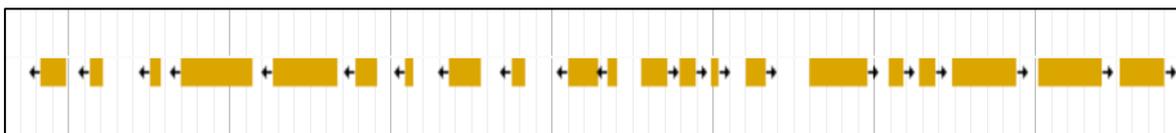


Figura 4. Arreglo divergente de genes en el cromosoma 1. 46,451- 100,150 3nt de *Leishmania mexicana*.

En la última actualización (9 de octubre de 2016) en la base de datos UniProt (<http://www.uniprot.org>) se reportaron 8,044 genes que codifican a proteínas anotados para *L. mexicana*, de los cuales 5,398 están marcadas como no caracterizadas.

1.5.3. Proteómica del parásito

Considerando que la expresión génica en tripanosomatidos está altamente regulada a nivel post-transcripcional, un análisis masivo y sistemático de proteínas es una buena aproximación para conocer mejor los mecanismos de estos patógenos [Myler and Fasel, 2008].

El primer mapa proteómico de *Leishmania* spp fue publicado a inicios de los años 1980s, mucho antes de que el término proteómica fuera acuñado. En total 31-35% del proteoma de *Leishmania* spp. predicho por el genoma ha sido identificado a la fecha [Paape and Aebischer, 2011].

Actualmente, se han desarrollado gran cantidad de análisis proteómicos de *Leishmania* spp., como, por ejemplo, la comparación entre el perfil proteómico de amastigotes y promastigotes de *Leishmania* spp., mostrando proteínas expresadas diferencialmente entre ambos estadios [Alcolea *et al.*, 2011] [Pawar *et al.*, 2012].

Por otro lado, esta reportado el análisis comparativo entre parásitos resistentes o susceptibles a antimoniales [Biyani *et al.*, 2011] y al estrés oxidativo [Sardar *et al.*, 2013]. Adicionalmente, se han identificado fosfoproteínas en promastigotes y amastigotes de *Leishmania donovani*, que muestran el nivel de fosforilación asociado a un estadio celular [Tsigankov *et al.*, 2013]. De la misma forma, se han identificado una gran cantidad de proteínas de membrana en promastigotes y amastigotes que regularmente serían difíciles de detectar y caracterizar por técnicas convencionales [Brotherton *et al.*, 2012].

1.5.4. Proteómica de *Leishmania mexicana*

En *L. mexicana* se ha comparado el perfil proteómico entre amastigotes y promastigotes utilizando geles de 2 dimensiones (masa molecular vs punto isoeléctrico) [Nugent, 2004]. Además, otro estudio realizado por Paape *et al.* en 2010, analizó el perfil en amastigotes usando LC-MS/MS, libre de geles. En este estudio se describen las principales proteínas identificadas en este estadio celular y puntualizan su aporte al conocimiento de las proteínas de *L. mexicana* en ese momento. En este trabajo se identificaron 1,764 proteínas, de las que 741 de ellas no habían sido descritas previamente (Figura 5).

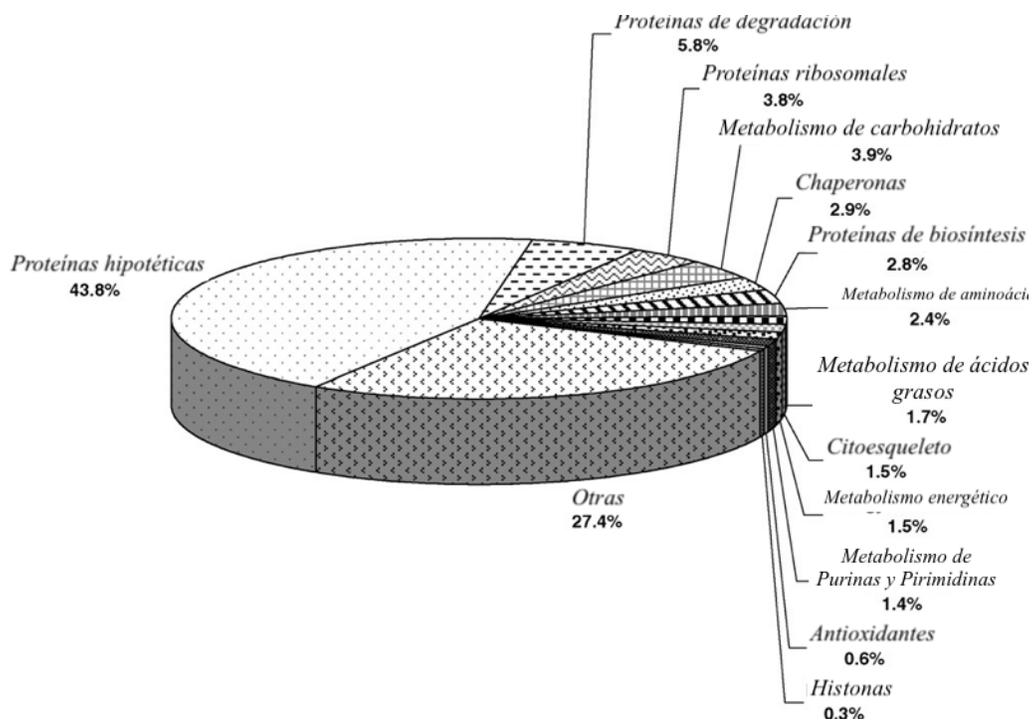


Figura 5. Clasificación funcional de las proteínas de amastigotes de *L. mexicana* identificadas por Paape *et al.*, 2010.

1.6 Generalidades de proteómica.

Las proteínas son importantes porque incluyen a las moléculas funcionales principales en un organismo vivo. El término proteómica fue acuñado de la unión de la palabra proteína con genómica en los años 90's [James, 1997].

Como una disciplina post-genómica, la proteómica dirige sus esfuerzos a la identificación y cuantificación de todas las proteínas incluyendo su expresión, localización celular, interacciones, modificaciones post-traduccionales, cambio de la función con respecto al tiempo, espacio y tipo celular. Todas estas variables hacen que el estudio completo de un proteoma sea mucho más demandante de analizar, que una secuenciación de genoma [Pawar et al., 2014].

El estudio de proteínas que se expresan a partir de un genoma, así como sus modificaciones (fosforilación, acetilación, ubiquitinación, etc.) es de importancia, pues en muchos casos no son los cambios en los genes los que dan lugar a algún fenotipo, sino las alteraciones a nivel de traducción de proteínas y las modificaciones post-traduccionales que determinan su función biológica [Fernández-Figueroa, 2013]. Esto es especialmente cierto para los tripanosomátidos que regulan la expresión y abundancia proteica en forma post-transcripcional [Pawar et al., 2014].

El progreso de la proteómica se ha basado en el desarrollo de nuevas tecnologías para la separación de péptidos y proteínas, análisis de espectrometría de masas, en el marcaje con isótopos para la cuantificación, y en el análisis bioinformático de los datos. La espectrometría de masas se ha establecido como una herramienta clave para el análisis a gran escala de proteínas, debido al rápido avance que ha tenido esta tecnología en la última década [Zhang et al., 2013].

1.7 Cromatografía líquida y Espectrometría de masas (LC-MS/MS)

La espectrometría de masas (MS) es una técnica analítica que permite medir con precisión la relación masa / carga (m/z) de una molécula. Las moléculas analizadas son transformadas en iones con carga en una fase gaseosa, que a su vez pueden separarse a través de campos magnéticos basadas en su relación m/z en el analizador.

Un espectrómetro de masas cuenta con tres principales componentes: una fuente de ionización que puede ser de electrospray o MALDI (Desorción/Ionización Láser Asistida por Matriz) que permite la carga y transferencia a una fase gaseosa de las moléculas; un analizador de masas que separa los péptidos ionizados, selecciona los iones para su fragmentación y determina los valores de m/z de los iones

precursores y productos de fragmentación; y un detector de iones que convierte la energía de las partículas que ingresan en una señal de corriente la cual posteriormente se transforma en los mapas espectrales para su análisis [Edwards, 2007].

Para determinar la estructura primaria de péptidos y proteínas, la cadena polipeptídica se somete a un aumento de la energía cinética mediante la aplicación de campos electromagnéticos. Este tipo de disociación es conocida como CID (*Collision Induced Dissociation*). El espectro resultante, llamado espectro de MS/MS (MS2), corresponde a los coeficientes de m/z de fragmentos correspondientes a la masa específica de los diferentes aminoácidos. [Guthals, 2012]

La cromatografía líquida acoplada en tandem a espectrometría de masas (LC-MS/MS) ha emergido como una herramienta fundamental para la proteómica, fundamentalmente para la caracterización e identificación de muestras complejas de proteínas derivadas de muestras biológicas. En esta técnica, los analitos derivados de proteínas se separan por cromatografía líquida y se ionizan por *electrospray* antes de entrar al espectrómetro de masas. Ahí pueden ser sujetos a dos fases del análisis de masas (MS1 y MS2) separados por una fase de selección, y una fase de fragmentación en una celda de colisión. Se han desarrollado tres estrategias para el análisis proteómico a partir de LC-MS/MS:

- *Bottom-up*
- *Middle-down*
- *Top-down*

1.8 Proteómica en tandem tipo *Bottom-up*

Los análisis proteómicos tipo *Bottom-up* se caracterizan por analizar péptidos derivados de una proteólisis de proteínas. Cuando un análisis de este tipo se lleva a cabo en una mezcla de proteínas, se le llama proteómica tipo *shotgun*, nombre acuñado por Yates en alusión a la técnica de secuenciación de ácidos nucleicos. En una corrida típica de proteómica *shotgun*, la mezcla de péptidos se fracciona y después se realiza un análisis de LC-MS/MS. [Zhang., 2013] [Yates III et al., 1995]

La identificación de péptidos se realiza comparando los patrones de fragmentación de los iones en los espectros MS2 con un espectro generado *in silico*, a partir de una base de datos de proteínas. Esta identificación puede ser de todos los productos hipotéticamente generados a partir de un genoma conocido o una base de datos conocida. Esto se hace automáticamente con programas como SEQUEST, X!Tandem o Mascot, entre otros. [Qian, 2015] [Bjornson et al., 2007], [Brosch, 2009]

La inferencia de las proteínas identificadas en el análisis está acompañada de la asignación de cada péptido a los modelos proteicos más parecidos. Debido a que un péptido puede ser asignado a una o varias proteínas, cada asignación debe ser ponderada [Zhang., 2013].

1.8.1 Cuantificación libre de marcaje

Hay dos grandes abordajes para la cuantificación libre de marcaje en proteínas:

1. Usando la intensidad de iones del péptido integrado de la primera fase (MS1).
2. Usando cuentas espectrales (contando el número de espectros MS/MS de los péptidos de una proteína en particular).

El primer método se basa en que la intensidad de la ionización del *electrospray* está correlacionada con la concentración del ion y que el área del pico cromatográfico, o la altura del pico, están linealmente correlacionados con la concentración de la proteína. Algunas herramientas como MassChroQ, MaxQuant, Skyline o Progenesis se han desarrollado para detectar, cuantificar y asignar los picos cromatográficos en MS1 o en MS2. [Zhang, 2013]

El segundo método, el conteo espectral, fue introducido por Liu *et al.*, 2004, mostrando que existe una correlación para las proteínas abundantes. Se basa en la suposición que la tasa de muestreo del MS/MS de un péptido en particular, está relacionada con la abundancia del péptido presentado por su ion precursor en la mezcla. La abundancia de un péptido es medida como la cantidad de péptidos pareados a un espectro PSMs (Protein Spectral Mass). La abundancia de la proteína se determina por el número total de espectros asociados al de todos los péptidos que se detectan [Liu., 2004].

Otro método como: NSAF (*Normalized Spectral Abundance Factor*) involucra la normalización por el tamaño de la proteína y número total de espectros MS/MS en un experimento. Esta normalización por tamaños se basa en que las proteínas más largas producirán más péptidos. Se define por la siguiente ecuación:

$$\text{NSAF} = \frac{(SC/L)}{\sum (SC/L)}$$

Donde SC, es el conteo espectral y L es la masa de la proteína en kDa.

2. JUSTIFICACIÓN

Leishmania mexicana en México desarrolla en el hospedero vertebrado dos cuadros clínicos cutáneos con características muy diferentes; LCL que desarrolla una respuesta inmune tipo Th1 la cual favorece la eliminación del parásito, y la LCD que desarrolla una respuesta inmune tipo Th2. Al ser una enfermedad parasitaria transmitida por un vector, cada uno de los agentes juega un papel importante dentro de la enfermedad, sin embargo, al estudiar al parásito en su forma infectiva se puede tratar de dilucidar el mecanismo por el cual el parásito genera los diferentes cuadros clínicos.

A nivel genómico, la primera referencia para *L. mexicana* se publicó en 2011. Esta misma cepa (U1103) se usa normalmente para el estudio del parásito, sin embargo, se ha observado una variabilidad genómica en aislados secuenciados que vienen directamente de biopsias de pacientes [Dawning et al., 2011]. Además, las herramientas bioinformáticas para la anotación de genomas han mejorado del 2011 a la fecha, lo que permite realizar una actualización y mejoramiento en los genes anotados en *L. mexicana*.

Al tener datos de diferentes plataformas ómicas (DNAseq, RNAseq y LC-MS/MS), se puede reportar una mejor referencia para el genoma de *Leishmania mexicana*, así como una mejor anotación; adicionalmente esto puede aumentar el número de proteínas encontradas en el estudio de LC-MS/MS del parásito.

El análisis genómico y proteómico permite tener una visión más amplia de los mecanismos de patogenicidad y supervivencia de *Leishmania*. En el presente proyecto se busca analizar el perfil proteico que permite a los diferentes aislados generar diferentes lesiones cutáneas en pacientes. Para este abordaje se utilizan promastigotes en fase estacionaria, asumiendo que la mayoría los mismos se encontraron en estadio metacíclico (fase infectiva).

3. HIPÓTESIS

Existe un perfil de expresión diferencial de proteínas entre promastigotes de aislados de *Leishmania mexicana* que clasifican los distintos cuadros clínicos: Difuso (LCD) y Localizado (LCL).

4. OBJETIVOS

4.1 Objetivo General

Analizar las proteínas diferencialmente expresadas en promastigotes de *Leishmania mexicana* que desarrollan cuadros clínicos distintos.

4.2 Objetivos específicos

1. Complementar la anotación con los datos de secuenciación de genoma y transcriptoma del parásito.
2. Obtener aislados de parásitos de lesiones de pacientes con LCD y LCL.
3. Realizar experimentos de proteómica tipo *bottom-up shotgun* para diferentes aislados de *Leishmania mexicana*.
4. Analizar el perfil de las proteínas diferencialmente expresadas entre los aislados de *Leishmania mexicana*.
5. Realizar un análisis proteogenómico.

5. MATERIALES Y MÉTODOS

5.1 Pacientes

Anterior a este trabajo, los pacientes fueron clínicamente diagnosticados como LCL o LCD de acuerdo con las lesiones presentes. Para corroborar el diagnóstico se hicieron frotis y se tiñeron con colorante Giemsa. Para los pacientes con LC se realizó la prueba de hipersensibilidad de Montenegro. El diagnóstico se confirmó por el ensayo por inmunoadsorción ligado a enzimas (ELISA, por sus siglas en inglés) específico para el género *Leishmania*. Los parásitos fueron aislados de lesiones activas de 10 pacientes en los estados de Tabasco, Veracruz, Michoacán, Chiapas y Quintana Roo: 5 LCD y 5 LCL (Figura 6).

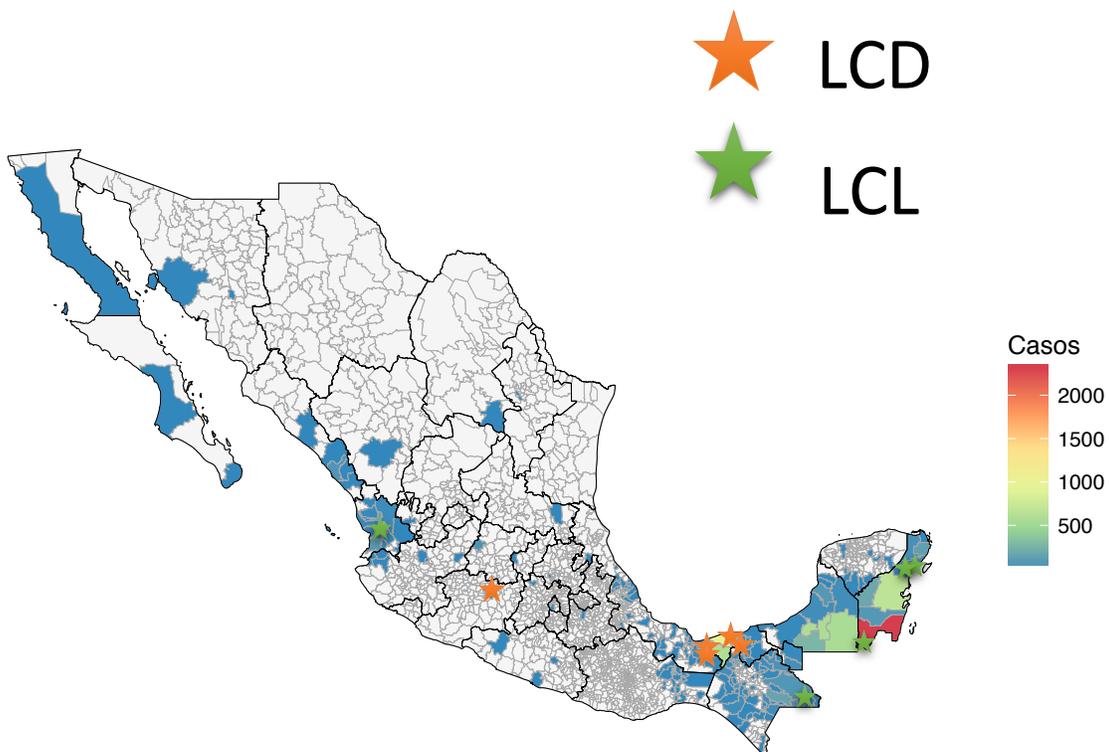


Figura 6. Distribución de los casos de pacientes con leishmaniasis utilizados en este estudio (las estrellas naranjas muestran los casos LCD y las verdes los casos LCL). A la derecha se observa la escala del número de casos por municipio.

5.2 Aislamiento y expansión de parásitos

Posterior al aislamiento de parásitos de las lesiones de pacientes, se infectaron ratones hembras de la cepa BALB/c con 1×10^7 parásitos. De las lesiones en cojinete se aislaron parásitos y se cultivaron en medio 199 suplementado con Suero Fetal Bovino (SFB) 10% para transformarlos a promastigotes.

Se realizaron curvas de crecimiento (Figura 7) y se cosecharon los parásitos en la fase estacionaria de crecimiento (un día antes de la fase de muerte celular). Se asume que en esta fase se encuentran el mayor número de promastigotes metacíclicos (infectivos) lavándolos 3 veces con amortiguador fosfato salino (PBS) frío 1X y centrifugándolos a $1300 \times g$ durante 12 minutos.

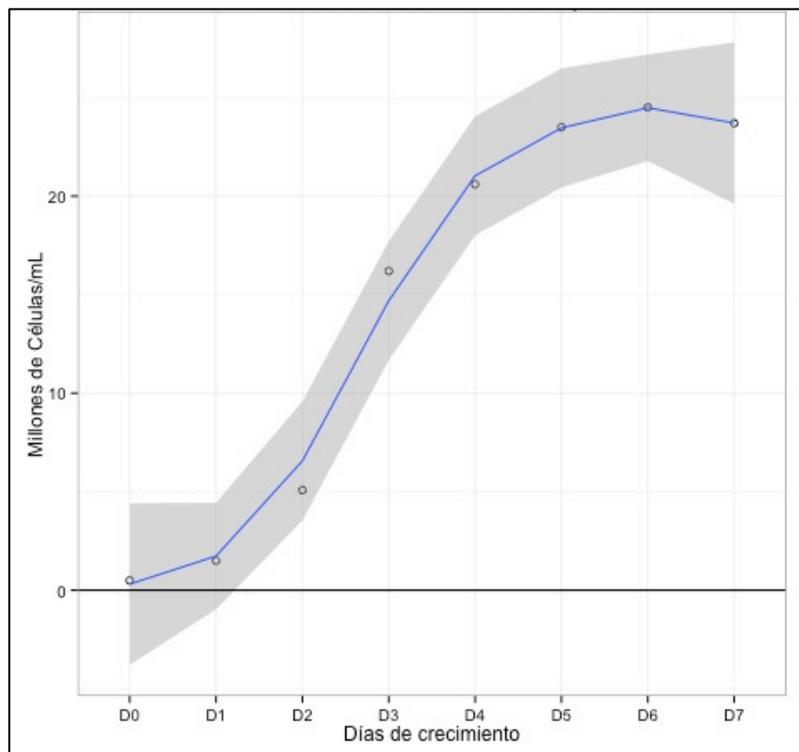


Figura 7. Curva de crecimiento representativa de los cultivos de parásitos de *L. mexicana* utilizados en este estudio.

Después de cosechados y lavados, los parásitos (1×10^8) se resuspendieron en amortiguador RIPA (*Radio Immunoprecipitation Assay*) que contenía inhibidores de proteasas.

5.3 Extracción de proteínas

Se incubaron en hielo por 10 minutos y se homogenizaron por sonicación durante 10 minutos a 4 °C. Las muestras se incubaron 30 minutos en hielo agitando en *vortex* cada 5 minutos durante 20 segundos. Se centrifugó a 1300 x g por 10 minutos a 4 °C y se recuperó el sobrenadante en tubos nuevos.

La concentración de proteína en las muestras se cuantificó utilizando un método compatible con detergentes (RC-BioRad), utilizando BSA como curva patrón.

Para verificar la integridad de las muestras se utilizaron 25 µg de proteína de cada una y se diluyeron 1:2 usando amortiguador Laemmli 2X. Las muestras fueron hervidas por 3 minutos a 98 °C e incubadas en hielo por 3 minutos. Posteriormente fueron separadas usando SDS-PAGE desnaturalizante al 10%. Los geles fueron teñidos con azul de Coomassie por 1 hora en agitación constante, se enjuagaron con agua destilada y se destiñeron por 2 horas.

5.4 Western-blot (WB)

Para el análisis de la expresión de la proteína PP2C en los extractos proteicos de los aislados de *Leishmania mexicana* se realizó un WB. Se utilizaron geles de acrilamida-bis-acrilamida al 10%. Las muestras se ajustaron a 20 µg de proteína, y se diluyeron amortiguador Laemmli 2x, se hirvieron por 3 min a 95 °C y se corrieron en los geles a 60 V por 30 min y a 100 V hasta terminar la corrida. Los geles se incubaron en amortiguador de transferencia durante 10 min y la transferencia se realizó en membranas de Polifluoruro de vinilideno (PVDF)-Immobilon de 0.45 µm (Millipore) activadas previamente con metanol durante 1 min. La transferencia se hizo en un equipo transblot semi-seco (Bio-Rad) durante 30 min a 25 V. Las membranas transferidas se incubaron durante 5 min en amortiguador TBST y se tiñeron con solución de rojo de Ponceau S (Sigma) para verificar que las proteínas se hayan transferido y se lavaron con TBST (amortiguador tris salino y polisorbato 20) hasta desteñir las membranas. Los sitios inespecíficos se bloquearon con leche al 5% en TBST durante 1 hora para posteriormente realizar 4 lavados con TBST durante 10 minutos cada uno. Se colocó el anticuerpo anti-PP2C a una dilución 1:1000 en BSA al 5% en TBST durante toda la noche a 4 °C. Al día siguiente se retiró el anticuerpo y se lavaron las membranas con TBST. Se agregó un segundo anticuerpo (anti-conejo) conjugado con peroxidasa (Cell Signaling Technology) a una dilución 1:5000 con 5% de leche en TBST durante 1 hr a temperatura ambiente. Se retiró el anticuerpo y las membranas se lavaron durante 2 horas con TBST. Las

membranas se incubaron por 3 min con sustrato Luminata Forte Western HRP (Millipore) y se revelaron por quimioluminiscencia en un equipo de BioRad.

5.5 LC-MS/MS

Para cada muestra, se utilizaron 100 µg de proteínas las cuales fueron reducidas con una solución de urea 8 M, 100 mM Tris. Posteriormente, las proteínas se alquilaron con iodoacetamida y se digirieron con tripsina. Después de la digestión se limpiaron las muestras por extracción en fase sólida (SPE). Los péptidos de cada muestra se solubilizaron en amortiguador de carga (3% acetonitrilo complementado con 0.1% ácido acético y 0.01% ácido trifluoroacético). Un total de 10 µL fueron inyectados al nano LC-MS/MS usando un sistema automático Easy-nLC 1000 acoplado a un espectrómetro de masas Q-Exactive plus (Thermo Fisher Scientific, San Jose, CA, USA). Para la cromatografía líquida se utilizó una pre-columna (20 mm x 75 µm-C18) y una columna analítica EASY-Spray (50 cm x 75 µm ID, PepMap C18, 2 µm de tamaño de partícula, 100 Å de tamaño de poro; Thermo Fisher Scientific) con una fase móvil A (0.1% ácido fórmico en agua) y una B (0.1% ácido fórmico en ACN). Los péptidos se separaron en condiciones estándar, con un flujo de 350 nL/minutos usando el gradiente: 2–40% de fase móvil B de 0-30 minutos, 40–98% fase móvil B de 30-170 minutos y 98% de fase móvil B de 170 - 240 min. Los espectros se procesaron en la modalidad adquisición dependiente de datos por 240 minutos utilizando los 10 mejores picos. Los datos se colectaron a una razón de 400-2000 m/z, con una resolución de 70,000, y Max IT 100 ms. Para MS/MS, se utilizó una resolución de 17,500 y Max IT 64 ms, con una ventana de 2 m/z y energía normalizada de colisión de 28.

5.6 Búsqueda en bases de datos e identificación de proteínas

Para la búsqueda en las bases de datos e identificación de las proteínas obtenidas se descargó la base de datos UniProt (*Leishmania mexicana* MHOM/GT/2001/U1103) la cual contiene 8,044 proteínas, misma que fue actualizada por última vez el 9 de julio de 2016. Se complementó esta base de datos con las secuencias resultantes de la predicción en el genoma y transcriptoma de *L. mexicana* desarrollado a partir de los 10 aislados en nuestro mismo grupo de trabajo (datos no publicados). El identificador de las secuencias provenientes de genómica-transcriptómica empieza con el código "LMEX". Para evitar redundancias en la base de datos se utilizó el algoritmo CD-HIT, descartando aquellas secuencias con un 90% de similitud o mayor. El resultado final es una base de datos con 8,924

proteínas, 155 contaminantes comunes y 9,079 Decoys (secuencias de proteínas y contaminantes en reversa).

5.7 Parámetros de búsqueda

Se utilizó el programa ProteomeDiscoverer v1.4 usando el algoritmo de Sequest. Los parámetros de búsqueda fueron:

- Precursores MS1 con masa mínima de 350 Da y máxima de 5,000 Da
- Razón de señal/ruido de 1.5; tolerancia del precursor de 50 ppm y tolerancia del fragmento de 0.6 Da
- Carbamidometilación de cisteínas (+57.021 Da) como modificación estática
- Oxidación de metioninas (+15.995 Da) como modificación dinámica.

Se utilizó la base de datos de *Leishmania mexicana* No Redundante con Contaminantes y *decoy* reversa (descrita anteriormente). Los espectros también se buscaron a través de X!Tandem bajo las mismas condiciones. La identificación de péptidos se aceptó si la probabilidad de asignación de Peptide Prophet [Keller, 2002] era mayor a 95% con la corrección por delta-masa de Scaffold.

5.8 Identificación de Proteínas

Scaffold (versión 4.7.2, Proteome Software Inc., Portland, OR) se utilizó para validar los patrones de fragmentación y la identificación de proteínas. Dichas identificaciones de proteínas fueron aceptadas si la probabilidad de asignación de Protein Prophet [Nesvizhskii, 2003] era mayor a 90% con al menos 2 péptidos identificados. Se consideraron como identificaciones válidas aquellas presentes en más de 2 aislados. Aquellas proteínas que estaban presentes en al menos 4 de las 5 muestras de un grupo y ausentes totalmente en el otro grupo fueron declaradas como exclusivas del grupo.

5.9 Cuantificación y normalización

Se utilizó la cuantificación por conteos espectrales y se normalizó por el factor de abundancia espectral (NSAF).

Para determinar la significancia estadística de la diferencia de medias entre los grupos LCD y LCL se utilizó una prueba estadística t-Student para cada proteína identificada.

5.10 Análisis de enriquecimiento

Se utilizó Blast2GO versión 4 Pro para hacer el mapeo de identificadores de InterPro y *Gene Ontology* (GO) para las 779 proteínas identificadas. Para el análisis de enriquecimiento, se seleccionaron las proteínas diferencialmente expresadas y sus respectivos identificadores de GO fueron contrastados con el perfil global de los GOs de las 779 proteínas.

5.11 Identificación de proteínas no anotadas

De las 8,924 proteínas de la base de datos generada en este trabajo, 8,044 pertenecen a las anotadas en UniProt y 880 son proteínas no anotadas para *L. mexicana*, o variantes de las que actualmente están en UniProt/NCBI.

Aquellas proteínas identificadas en Scaffold con un número de acceso comenzando por LMEX se analizaron por BLASTP para identificar proteínas homólogas y tBLASTn para identificar el marco de lectura abierto más cercano a la proteína en cuestión y verificar la especie de origen de dicha secuencia.

5.12 Gráficos

Se utilizó R versión 3.2.1 para realizar los diagramas de Venn, y el mapa de calor. Utilizando los paquetes VennDiagram y gplots.

6. RESULTADOS

6.1 Integridad de las muestras.

La integridad de las muestras antes de ser procesadas en los ensayos de LC-MS/MS se analizó corriendo los extractos protéicos en geles SDS-PAGE desnaturizantes teñidos con azul de Coomassie. Todas las muestras estaban íntegras y mostraban pesos moleculares entre 10 y 300 kDa (Figura 8).

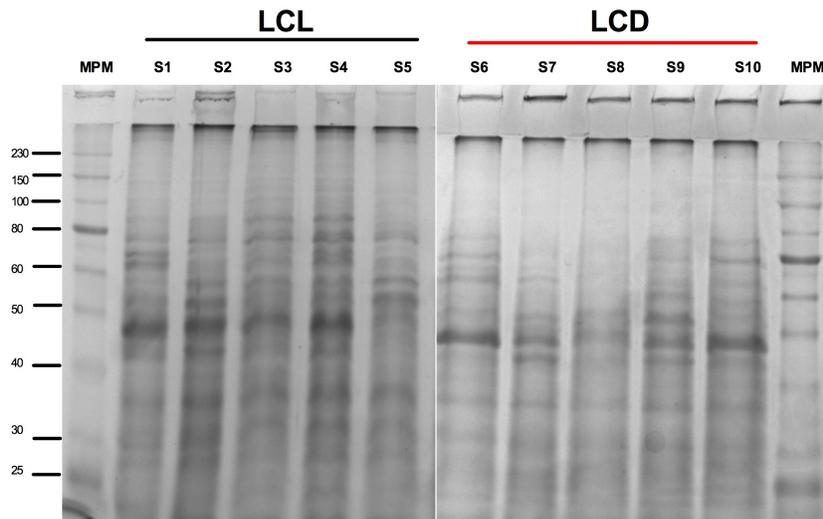


Figura 8. SDS-PAGE de extractos protéicos de *Leishmania mexicana* aislados de diferentes cuadros clínicos. S1 a S5 aislados de pacientes con leishmaniasis cutánea localizada (LCL) y S6 a S10 aislados de lesiones de pacientes con leishmaniasis cutánea difusa (LCD). MPM= marcador de peso molecular.

6.2 Identificación de proteínas a partir de los datos LC-MS/MS

Del total de espectros (177,659) en los datos crudos, 43,370 cubrieron los umbrales establecidos (>95% del umbral de probabilidad para péptidos), los cuales fueron asignados a 779 proteínas identificadas (mínimo 2 péptidos identificados por proteína y 0.43% ProteinProphet FDR). Además, se identificaron 6 contaminantes: 4 queratinas, tripsina y albúmina de suero bovino.

De las 779 proteínas, 10 fueron exclusivas para los aislados que derivan de la forma clínica localizada (LCL) mientras que 3 son exclusivas para LCD. Para seleccionar las proteínas exclusivas en cada grupo se filtraron aquellas que estaban presentes en al menos 4 de las 5 muestras de un grupo y ausentes en el otro grupo (Figura 9).

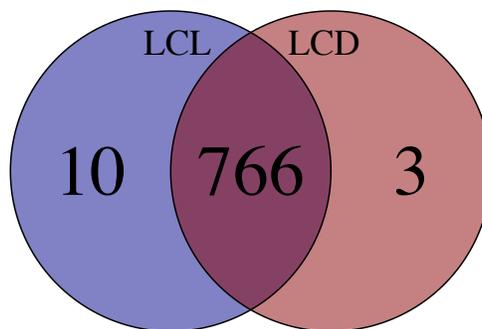


Figura 9. Diagrama de Venn. Proteínas únicas y compartidas para los aislados LCL y LCD. En el diagrama de Venn observamos que hay 766 proteínas encontradas en común, 10 proteínas exclusivas para el grupo LCL y 3 para el grupo LCD.

6.3 Proteínas exclusivas para LCL

Las proteínas exclusivas para el grupo LCL fueron 10 (Tabla 2).

Tabla 2. Proteínas identificadas en los 5 aislados de LCL con ausencia en todos los LCD.

Número de Acceso (UniProt)	Nombre
E9B0Q4	Piridoxal Putativa
LMEX_000153200.1	Peptidil-prolil cis-trans isomerasa Putativa tipo FKBP
E9AU61	Proteína Chaperona Putativa de ADN
E9AWI3	Subunidad A de la ATPasa de protones tipo V
E9ALY7	Flavoproteína succinato dehidrogenasa
E9AQ14	Proteína Fosfatasa 2C Putativa
E9ALJ4	Hidrolasa Ubiquitina Putativa
E9B496	Aminopeptidasa Putativa
E9ANQ5	Proteína Putativa Transportadora Sec31
E9B4S1	p-glicoproteína

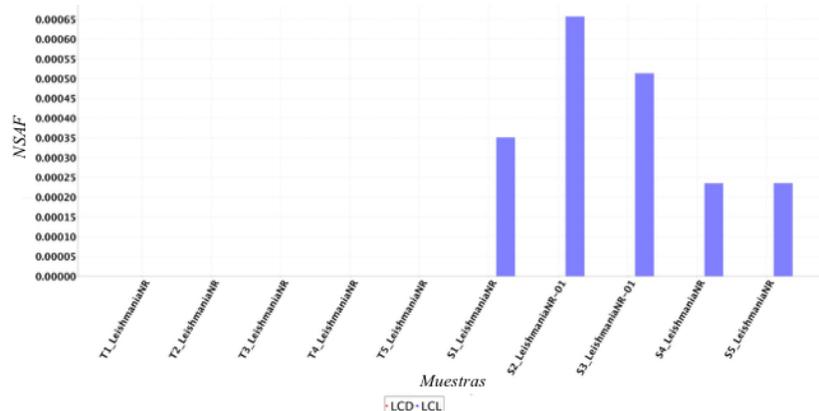


Figura 10. Abundancia de las cuentas espectrales normalizadas de la proteína FKBP en los promastigotes en fase estacionaria de aislados de *L. mexicana*. En barras azules se muestran las abundancias pertenecientes a esta proteína en los aislados del grupo LCL.

6.4 Proteínas exclusivas para LCD

Las proteínas exclusivas para el grupo LCD se describen en la tabla 3 y un ejemplo de la presencia exclusiva en aislados LCD se muestra en la figura 11.

Tabla 3. Proteínas identificadas en los 5 aislados de LCD con ausencia en todos los LCL.

Nombre de Acceso (UniProt)	Nombre
E9APZ1_LEIMU	Proteína Putativa no caracterizada.
E9ALU6_LEIMU	Subunidad Clp de la proteasa Putativa dependiente de ATP, proteína de choque térmico 100 (HSP100)
E9ALC3_LEIMU	ATPasa transportadora de cationes

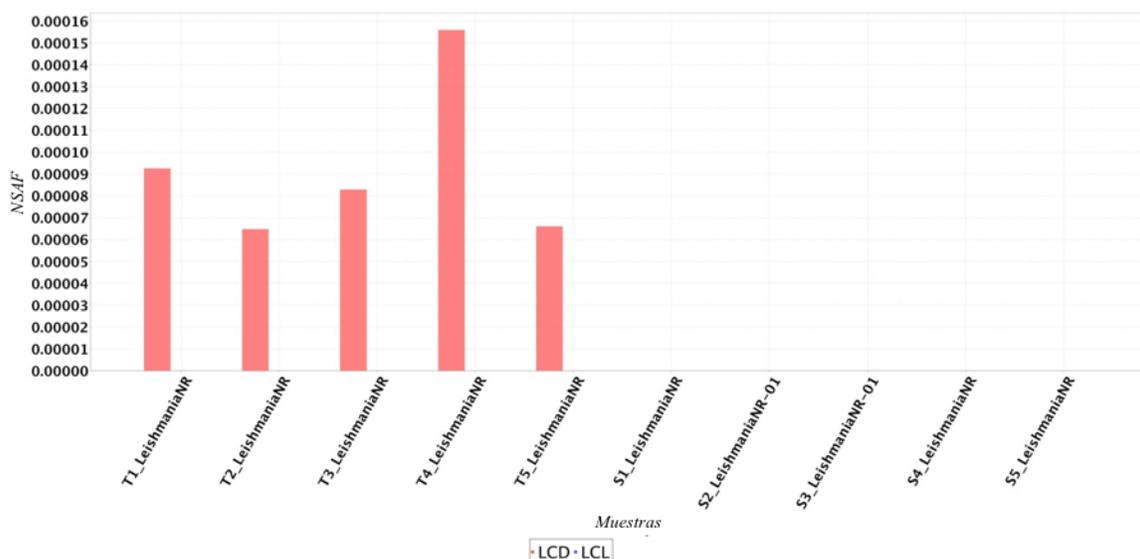


Figura 11. Abundancia normalizada de la proteína de choque térmico 100 (HSP100) en los promastigotes en fase estacionaria de aislados de *L. mexicana*. En barras rojas se muestran las abundancias pertenecientes a esta proteína en los aislados del grupo LCD.

6.5 Expresión diferencial

Para realizar una cuantificación relativa y normalización de los datos se optó por aplicar el algoritmo conocido como NSAF (por sus siglas en inglés Normalized Spectral Abundance Factor), ecuación 1 (ver página 17).

6.5.1 Distribución de la expresión de proteínas por muestra.

Para observar los valores extremos y la homogeneidad de NSAF en las diferentes muestras se realizó una gráfica de cajas (Figura 12). Se puede observar que las muestras normalizadas presentan medianas y distribuciones homogéneas que los hacen comparables para un análisis de expresión diferencial.

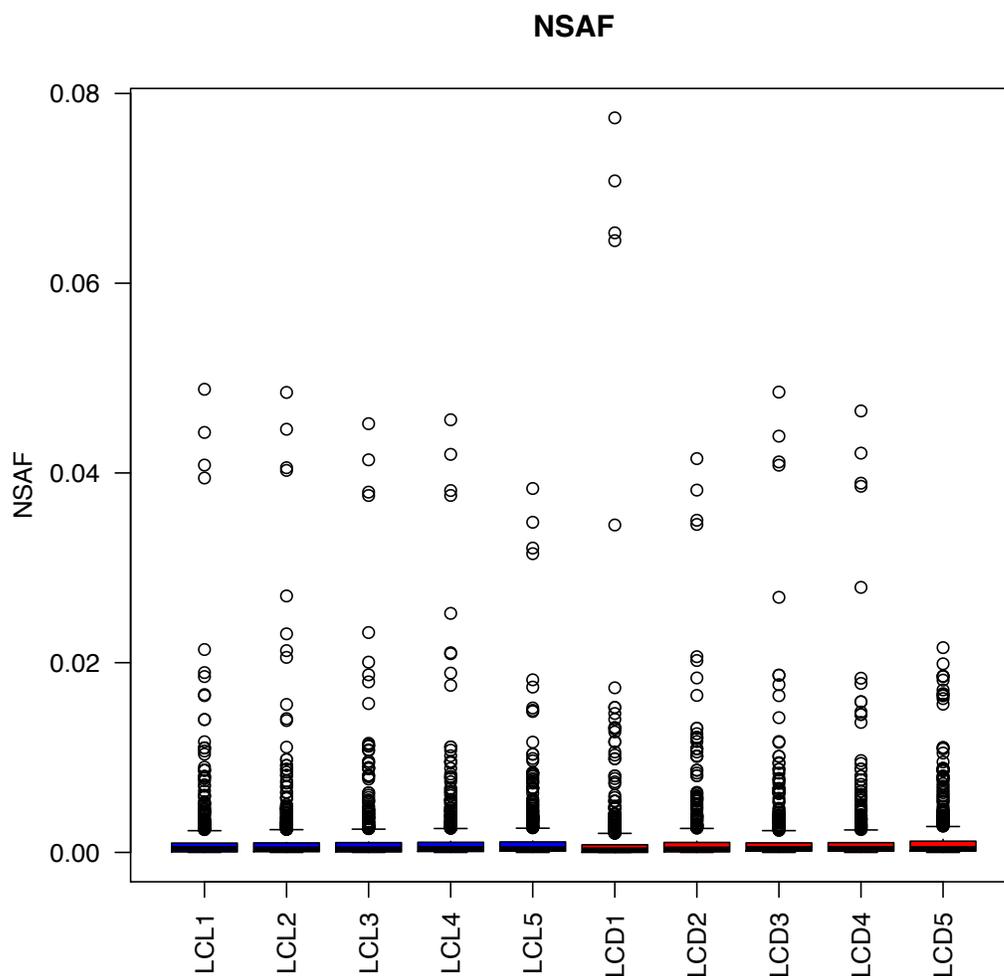


Figura 12. Gráfica de cajas de las muestras normalizadas por NSAF. LCD en rojo y LCL en azul.

6.5.2 Gráfica de densidad

Para observar la distribución de los valores NSAF en las diferentes muestras se realizó un histograma (Figura 13), donde se puede observar una distribución normal y homogénea entre las muestras analizadas que las hace comparables para un análisis de expresión diferencial porque se fija una misma distribución de densidad para todos los valores de abundancia de las muestras.

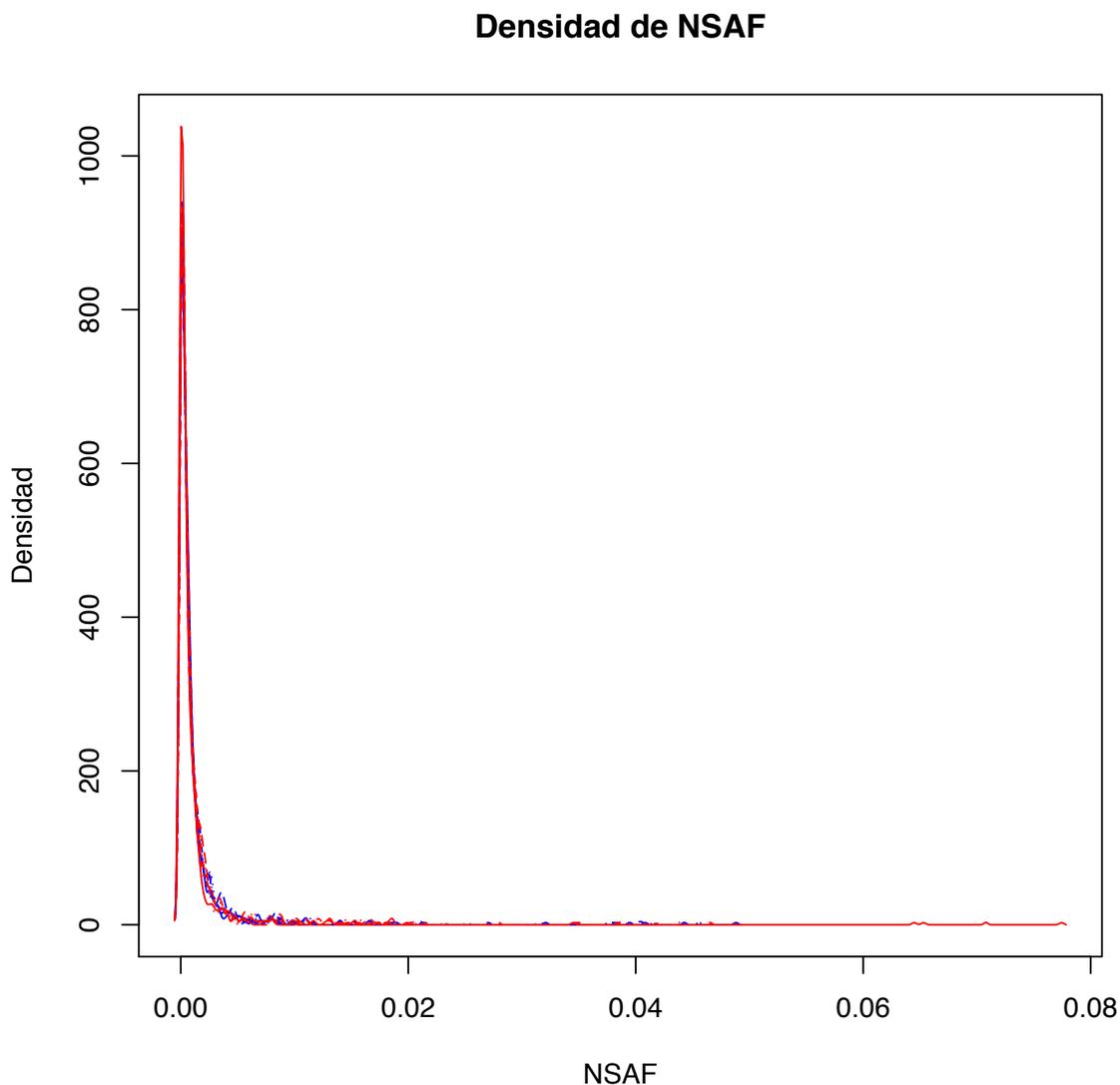


Figura 13. Gráfica de densidad de los valores NSAF. LCD en rojo y LCL en azul.

Según Ning *et al.*, 2012, NSAF (conteos espectrales) permite analizar la proteómica libre de marcaje de una forma robusta, obteniendo una correlación con los datos de transcriptoma. En este trabajo no se consideraron replicas técnicas sino biológicas, por lo que de las corridas de LC-MS/MS se usaron los conteos espectrales.

6.6 Proteínas diferencialmente expresadas

Al graficar el logaritmo (base 2) de la tasa de cambio diferencial (ecuación 3) con respecto al logaritmo del inverso del valor p, mostrado en el eje de las ordenadas de la figura 14, se puede visualizar la distribución de las proteínas diferencialmente expresadas y su nivel de significancia. Se realizó un corte considerando una tasa de cambio superior o igual a 2 ($\log_2 FCh > 1$) y con una confianza del 5% (valor $p \leq 0.05$). La tasa de cambio de cada proteína fue representada utilizando una gráfica de volcán (Figura 14), en el eje de las abscisas se muestra la tasa de cambio en logaritmo base 2 y en el eje de las ordenadas se representa la confianza estadística (valor p) en logaritmo base 10. De tal forma que, aquellos puntos en color rojo son proteínas sobre expresadas en las muestras LCD respecto a las LCL y aquellos en color verde las proteínas diferencialmente sobreexpresadas en el grupo LCD respecto al grupo LCL, (todas ellas con un valor $p \leq 0.05$). Como resultado de este análisis se encontraron 57 proteínas diferencialmente expresadas y que están enlistadas en la Tabla 4.

$$FCh = \frac{\text{media LCD}}{\text{media LCL}} \quad (\text{Ecuación 2})$$

$$\log_2 FCh = \log_2 \left(\frac{\text{media LCD}}{\text{media LCL}} \right) \quad (\text{Ecuación 3})$$

Ecuaciones 2 y 3 utilizadas para calcular la razón de cambio de las proteínas entre las muestras LCL y LCD a partir de los valores de NSAF normalizados.

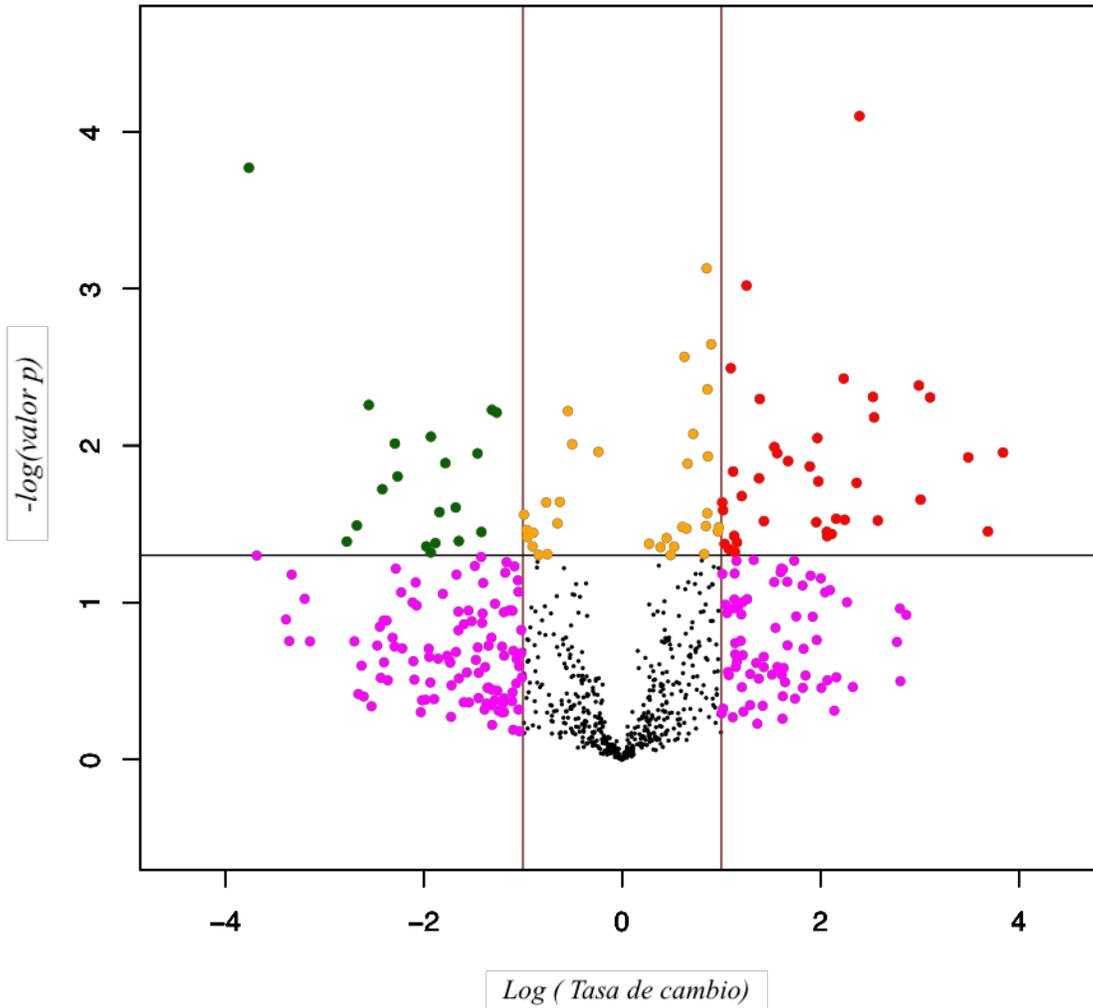


Figura 14. Gráfico de Volcán. Puntos rojos representan las proteínas sobre-expresadas en el grupo LCD vs LCL; verde, proteínas sub-expresadas en el grupo LCD vs LCL. En color amarillo se muestran las proteínas con significancia estadística, pero con una tasa de cambio menor al umbral seleccionado (Logaritmo de la tasa de cambio mayor a 1). En violeta se muestran las proteínas con una tasa de cambio mayor al umbral seleccionado, pero sin significancia estadística (valor p menor a 0.05). En negro son aquellas proteínas que no son significativas ni tienen una tasa de cambio mayor al umbral seleccionado.

Tabla 4. Descripción de proteínas diferencialmente expresadas en las muestras LCD vs LCL.

Nombre de Acceso (UniProt)	Nombre	Valor P	log2(LCD/LCL)	Notas
Traducción				
E9AJC9	Proteína no caracterizada	0.000	2.39	subunidad pequeña proteína ribosomal S7e
E9APE6	Proteína ribosomal 40S S4	0.038	1.13	
E9ASM7	Proteína ribosomal putativa L24	0.030	2.24	
E9AVY4	Proteína ribosomal putativa 40S L14	0.005	1.39	
E9B3G8	Proteína putativa ribosomal 60S L2	0.037	2.11	
E9B423	Proteína nucleolar 1 de unión a GTP	0.012	3.49	Biogénesis del ribosoma
E9B3L1	Proteína Putativa con cassette de union a ATP	0.017	2.37	Factor de Iniciación
E9AMU9	Factor de elongación 1-gamma	0.001	1.25	Factor de Traducción
E9B052	Factor 3 subunidad E de inicio de la traducción	0.026	1.02	eIF-3
Transcripción				
E9ALZ2	Proteína de alta movilidad homóloga a tdp-1	0.025	-1.68	
Metabolismo de aminoácidos				

E9AJD5	Beta Liasa	0.014	1.89	Metabolismo de Glicina, Serina y Treonina
E9APY0	Serina Hidroximetiltransferasa	0.042	1.03	Metabolismo de Glicina y Serina. Metabolismo de folato.
E9AUV6	Aminoacilasa Putativa	0.007	2.54	N-acil-L-amino Amidohidrolasa
E9AUV7	Aminoacilasa Putativa	0.004	2.23	N-acil-L-amino Amidohidrolasa
E9AZA7	Aminotransferasa de aminoácidos de cadena ramificada	0.011	1.56	Metabolismo de Leucina, Isoleucina y Valina
E9B627	Arginasa	0.010	1.54	Biosíntesis de arginina
E9B1E9	5-metil-tetrahidropteroil-triglutamato-homocisteinemetiltransferasa	0.044	-1.97	Metabolismo de Cisteína y Metionina
Metabolismo de Ácidos Grasos				
E9ALJ3	Enoil-CoA hidratasa	0.004	2.99	Degradación de ácidos grasos
E9B6F2	Acil-CoA dehidrogenasa	0.035	2.06	Degradación de ácidos grasos
LMEX_000022600.1	Ligasa putativa de ácido graso de cadena larga-CoA	0.032	-2.67	Síntesis de ácidos grasos
Metabolismo de Carbohidratos				
E9B469	Epimerasa 4' putativa tipo udp-glc	0.035	3.69	Epimerización de UDP-Glucosa a UDP-Galactosa
Metabolismo de Inositol Fosfato				

E9APY4	Sintetasa Mio-inositol-1-fosfato	0.017	1.98	La superficie de <i>Leishmania</i> está cubierta de moléculas ancladas a GPI que contienen mio-inositol como leishmanolisina (gp63), gp46/PSA2 o proteo fosfoglicanos unidos a membrana (mPPG).
Chaperonas HSP60				
LMEX_000200100.1	Chaperonina TCP20 putativa	0.029	2.16	
E9AV93	Proteína 1 Subunidad Delta del complejo T	0.046	1.08	
E9B6R4	Proteína putativa 1 subunidad eta del complejo T	0.048	1.14	
Señalización				
E9B5Z3	Oligosacaril transferasa	0.030	2.58	N-Glicosilación
Defosforilación				
E9AKB4	Proteína fosfatasa de Serina/treonina	0.041	1.16	
E9AT34	Fosfatasa ácida de unión a membrana	0.013	1.67	
Fosforilación				
E9B5Y1	Cinasa de Caseína Putativa	0.023	1.01	Cinasa de Serina/treonina tipo CK1

E9B573	Subunidad regulatoria de cinasa tipo a	0.006	-1.26	Cinasa
Proteina Putativa con dominio de unión a RNA				
LMEX_000041900.1	Proteína putativa de unión a RNA	0.019	-2.42	
LMEX_000558700.1	Proteína putativa de unión a RNA	0.000	-3.76	
Proteasoma				
E9AZZ2	Subunidad 2 del Proteasoma regulatorio no-ATPasa	0.006	-1.31	
E9ANH0	Subunidad alfa del Proteasoma	0.036	-1.42	Proteasoma 20S subunidad alpha 4
Aminopeptidasa				
E9ANK8	Aminopeptidasa putativa (Metalopeptidasa clan mf familia m17)	0.010	-2.29	Aminopeptidasa NPEPL1
Antígeno Inmunodominante				
E9APT8	Antígeno inmunodominante putativo (antígeno tipo Tc40)	0.031	1.96	
Metabolismo energético				
E9AP45	Proteína tipo SURF1	0.021	1.21	Complejo IV
E9APU4	Proteína mitocondrial transportadora de ADP/ATP	0.011	3.84	
Ciclo de Krebs				
E9ATC6	Componente E1 de la 2-oxoglutarato deshidrogenasa	0.011	-1.46	
Citoesqueleto				

E9AYH5	Proteína putativa asociada al mitotúbulo Gb4	0.042	-1.88	
Biosíntesis de aminoacil-tRNA				
E9AZ37	Sintetasa Putativa del arginil-tRNA	0.016	1.38	
Metabolismo de Cofactores y Vitaminas				
E9AP13	Oxidoreductasa de NADH:flavin /NADH oxidasa	0.003	1.10	
Transporte de Calcio				
E9AL78	ATPasa transportadora de Calcio	0.022	3.01	ATPasa vacuolar putativa de Ca ²⁺
No Caracterizadas				
E9B6G9	Proteína no caracterizada putativa	0.005	3.11	
E9ALS6	Proteína no caracterizada	0.005	2.53	
E9AZS9	Proteína no caracterizada putativa	0.038	2.07	
E9B307	Proteína no caracterizada	0.030	1.43	
E9B625	Proteína no caracterizada putativa	0.009	1.97	
E9B660	Proteína no caracterizada putativa	0.015	1.12	
E9AN19	Proteína no caracterizada	0.048	-1.93	Posible Seril-tRNA
E9ANE3	Proteína no caracterizada	0.016	-2.26	
E9ANM6	Proteína no caracterizada putativa	0.027	-1.84	

E9ATG0	Proteína no caracterizada	0.006	-2.55	
E9AW79	Proteína no caracterizada putativa	0.041	-1.65	
E9B0I3	Proteína no caracterizada	0.041	-2.78	
E9B2K8	Proteína no caracterizada	0.009	-1.93	
E9B6C4	Proteína no caracterizada putativa	0.013	-1.78	

6.7 Mapa de calor

Se realizó un mapa de calor (Figura 15) como método de agrupamiento jerárquico no supervisado y visualización de las proteínas diferencialmente expresadas. Cada fila es una proteína diferencialmente expresada, las columnas representan cada aislado. Se observa en la barra superior roja las muestras LCD y en color azul las LCL. Los cuadros en blanco representan una proteína que no fue identificada en esa muestra, los cuadros rojos las proteínas sobre-expresadas, en azul las sub-expresadas. Se puede observar que las muestras se agrupan de acuerdo al grupo al que pertenecen. Adicionalmente, podemos observar que, de acuerdo con el perfil de expresión de proteínas por cada muestra, se agrupan proteínas sobre-expresadas en las muestras LCL y sub-expresadas en las muestras LCD, y también viceversa.

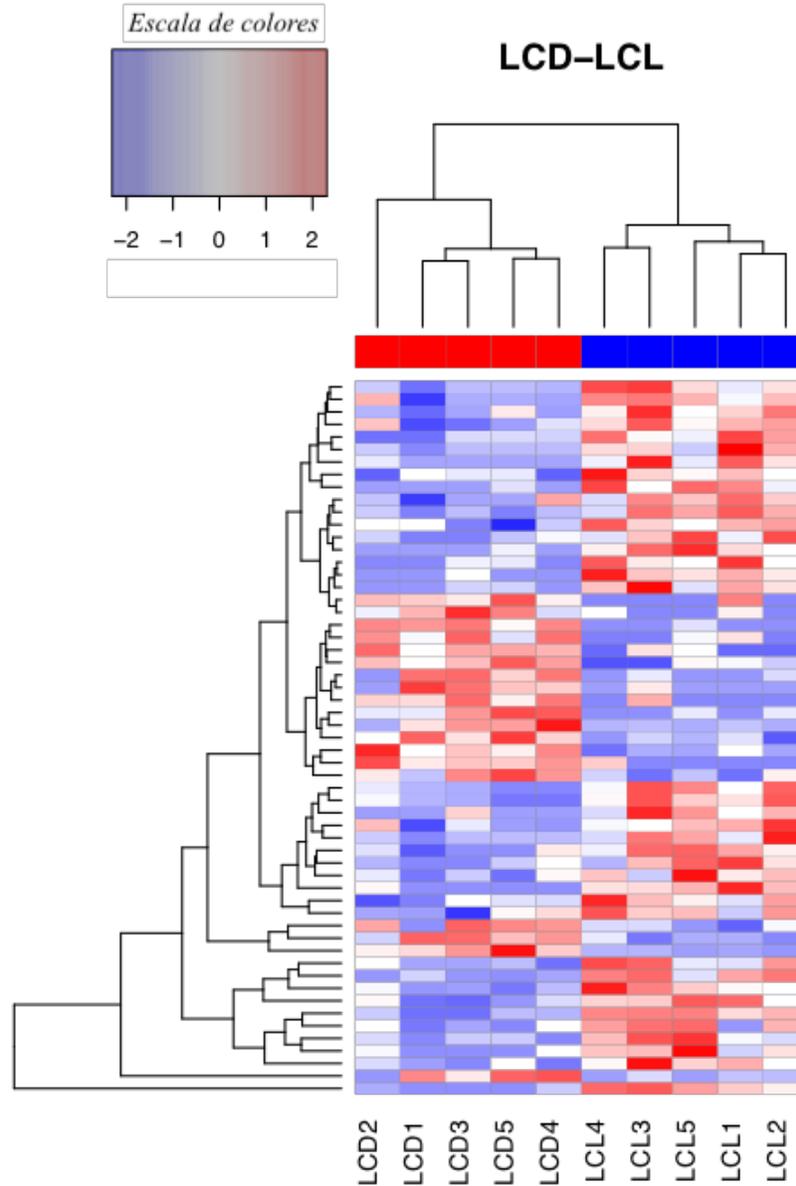


Figura 15. Mapa de calor de las proteínas diferencialmente expresadas. Las columnas representan cada muestra de *L. mexicana*. La barra superior indica el grupo de cada muestra, rojo para LCD y azul para LCL. En color rojo se muestran las proteínas sobre-expresadas y en azul las sub-expresadas. El dendrograma vertical agrupa cada muestra de acuerdo con la similitud en el perfil de expresión de proteínas (filas), mientras que el dendrograma horizontal agrupa el perfil de expresión de las proteínas.

6.8 Análisis de enriquecimiento

Dado que la anotación del genoma de *L. mexicana* no está adecuadamente curado se complementó la anotación de las 779 proteínas identificadas por LC-MS/MS con un identificador de la base de datos de Ontología de Genes (GO) que clasifica cada proteína de acuerdo con su función molecular, componente celular y proceso biológico. Lo anterior se realizó utilizando el programa Blast2GO. Con los datos de las proteínas diferencialmente expresadas se realizó un análisis de enriquecimiento donde se identificaron aquellos procesos más representados en las proteínas diferencialmente expresadas. El proceso involucra una búsqueda en BLAST de las secuencias de proteínas, posteriormente una asociación funcional a la base de datos de GO e InterPro. En la figura 16 se muestran los resultados obtenidos del enriquecimiento de identificadores de Gene Ontology.

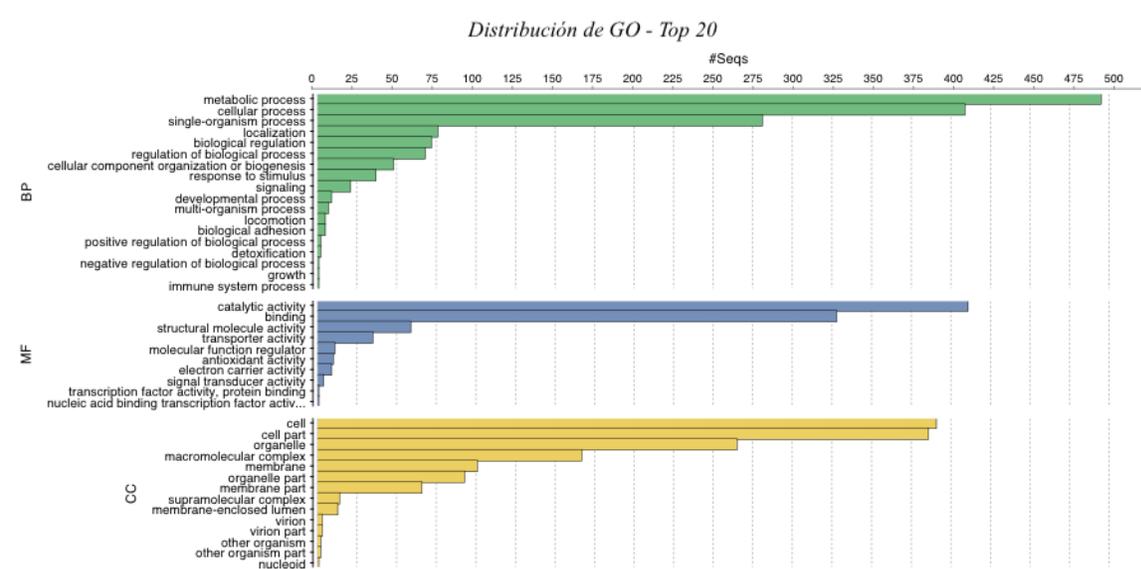


Figura 16. Distribución GO por nivel. Top 20 de las 779 proteínas identificadas por LC-MS/MS. La primera gráfica con barras en color verde clasifica las proteínas de acuerdo con su proceso biológico (BP). La gráfica con barras en color azul clasifica las proteínas de acuerdo con su función molecular (MF). Por último, la gráfica de barras de color amarillo clasifica a las proteínas de acuerdo a su componente celular (CC).

El análisis de enriquecimiento busca aquellas categorías funcionales que están sobre representadas en un conjunto de datos comparándolo con la base de datos completa reportada hasta el momento. En este caso utilizamos las 57 proteínas diferencialmente expresadas como conjunto de datos para el análisis de enriquecimiento.

Aquellos términos GO más representados fueron las vías de la actividad de aminoacilasa, la actividad de hidrolasa, las asociadas a la unión a cofactores, y chaperoninas (Figura 17).

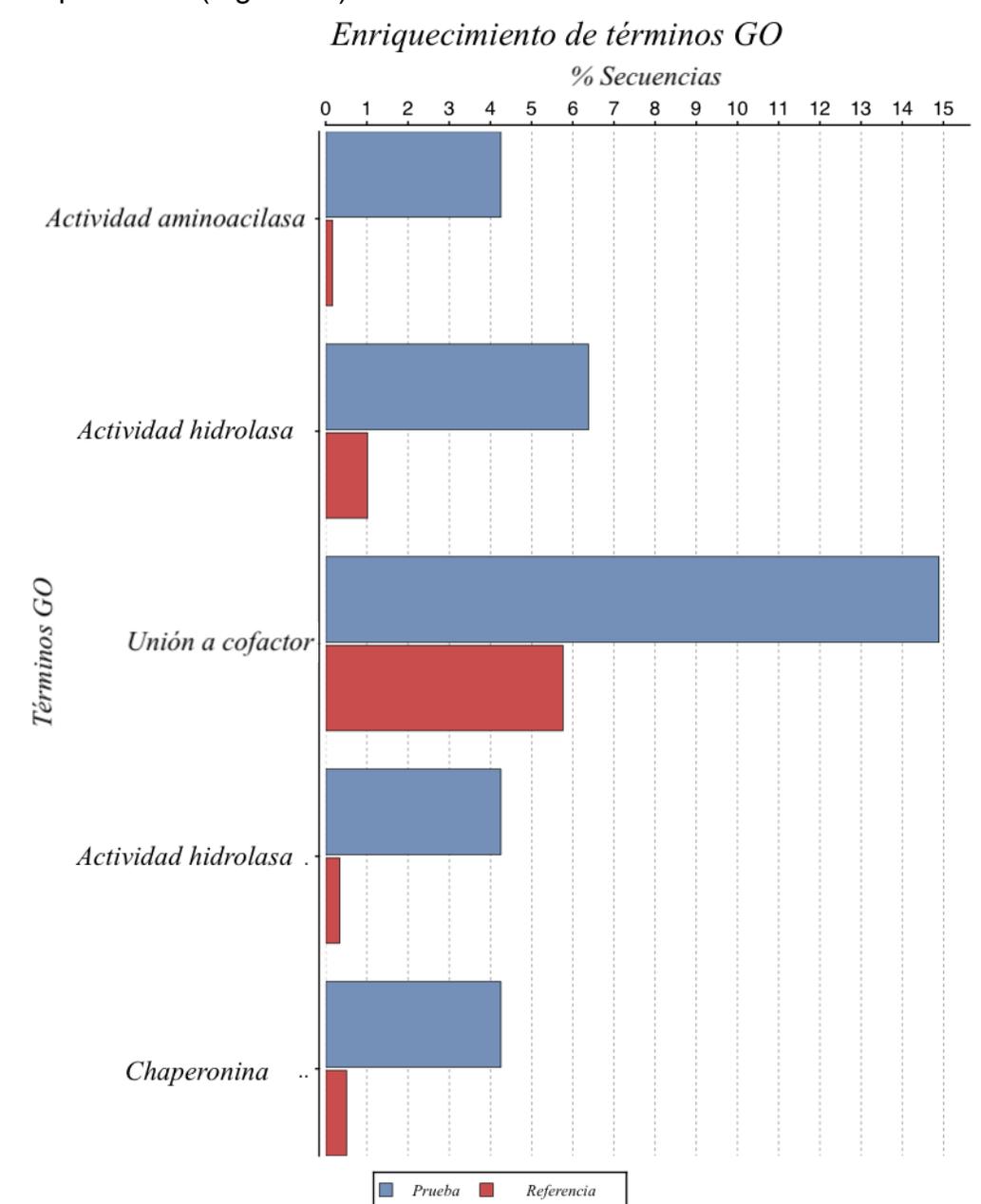


Figura 17. Términos Gene Ontology (GO) sobre representados en el grupo de las proteínas diferencialmente expresadas con respecto a las 779 proteínas identificadas.

6.9 Proteogenómica

Mejoramiento de la anotación de *Leishmania mexicana*

De las 779 proteínas identificadas en este estudio y correlacionando estos datos con los del genoma de *L. mexicana*, 39 proteínas no han sido reportadas y publicadas (NCBI o UniProt). En la siguiente tabla (Tabla 5) se enlistan:

Tabla 5.- Proteínas no anotadas en UniProt/NCBI identificadas por LC-MS/MS.

Proteínas Identificadas	Nombre de Acceso	MW
Aminopeptidasa P1 putativa	LMEX_000013300.1	69 kDa
Delta-1-pirrolina-5-carboxilato dehidrogenasa putativa	LMEX_000022300.1	62 kDa
Ligasa putativa de ácido graso de cadena larga-CoA	LMEX_000022600.1	79 kDa
Helicasa tipo caja DEAD/DEAH	LMEX_000027100.1	247 kDa
Factor 2a de inicio de traducción	LMEX_000030100.1	53 kDa
ATPasa tipo P translocadora de calcio	LMEX_000030200.1	112 kDa
Proteína hipotética conservadora	LMEX_000030800.1	59 kDa
Peptidasa de cisteína Clan CA familia C2	LMEX_000034500.1	98 kDa
Sintetasa de Espermidina	LMEX_000035900.1	33 kDa
Aciltransferasea	LMEX_000040600.1	40 kDa
Fructosa-16-bifosfatasa	LMEX_000041700.1	39 kDa
Proteína putativa de unión a RNA	LMEX_000041900.1	35 kDa
Actina	LMEX_000042500.1	42 kDa
Proteína putativa asociada a microtúbulo	LMEX_000046600.1	94 kDa
Proteína dehidrogenasa de esteroides dependiente de NAD(P)	LMEX_000060100.1	46 kDa
RNA helicasa putative dependiente de ATP	LMEX_000073900.1	61 kDa
Fumarato deshidrogenasa putativa	LMEX_000092800.1	34 kDa
Proteína putativa flagelar radial	LMEX_000106800.1	64 kDa
Factor de transcripción	LMEX_000113600.1	115 kDa
Ribonucleasa mitocondrial putativa	LMEX_000117000.1	25 kDa
Amastina, Glicoproteína de superficie putativa	LMEX_000146300.1	56 kDa
Proteína transportadora SEC23	LMEX_000147200.1	101 kDa
Proteína de dominio GRIP	LMEX_000159200.1	177 kDa
Chaperonina TCP20 putativa	LMEX_000200100.1	66 kDa
Factor de elongación 1-alfa	LMEX_000253300.1	71 kDa

Subunidad C de la ATP sintasa vacuolar	LMEX_000274100.1	47 kDa
Fibrilarina	LMEX_000336100.1	31 kDa
Proteína hipotética conservada	LMEX_000367500.1	277 kDa
Plectina putativa	LMEX_000405400.1	402 kDa
Peptidasa de cisteína tipo Calpaina	LMEX_000541500.1	89 kDa
Peptidasa de cisteína tipo Calpaina	LMEX_000541700.1	558 kDa
Proteína putativa de unión a RNA-	LMEX_000558700.1	50 kDa
Metiltioribulosa-1-fosfato deshidratasa putativa	LMEX_000583000.1	27 kDa
Superóxido dismutasa de hierro putativa	LMEX_000688700.1	23 kDa
Cinasa de nucleósido difosfato b	LMEX_000701100.1	18 kDa
Reductasa 1 dependiente de Tiol	LMEX_000713900.1	50 kDa
Componente paraflagelar putativo	LMEX_000718300.1	22 kDa
Peptidasa tipo m20/m25/m40	LMEX_000729400.1	53 kDa
Ciclofilina 4 putativa	LMEX_000729700.1	32 kDa
Proteína ribosomal 60S L5 putativa	LMEX_000809800.1	35 kDa

Para corroborar que estas 39 proteínas no reportadas probablemente son proteínas funcionales en el parásito se realizaron más búsquedas *in silico* para investigar su posible existencia en otros genomas del género *Leishmania*.

En la siguiente imagen (Figura 18), se muestra un ejemplo de la búsqueda realizada en BLASTp (búsqueda a una base de datos de proteínas) de una de las secuencias anotadas a partir de genoma y los datos de RNASeq de los aislados del parásito. Esta secuencia LMEX_000541500.1 corresponde a una peptidasa de cisteína tipo calpaína. La secuencia más cercana es una proteína de la misma familia reportada en *Leishmania major* cepa Friedlin.

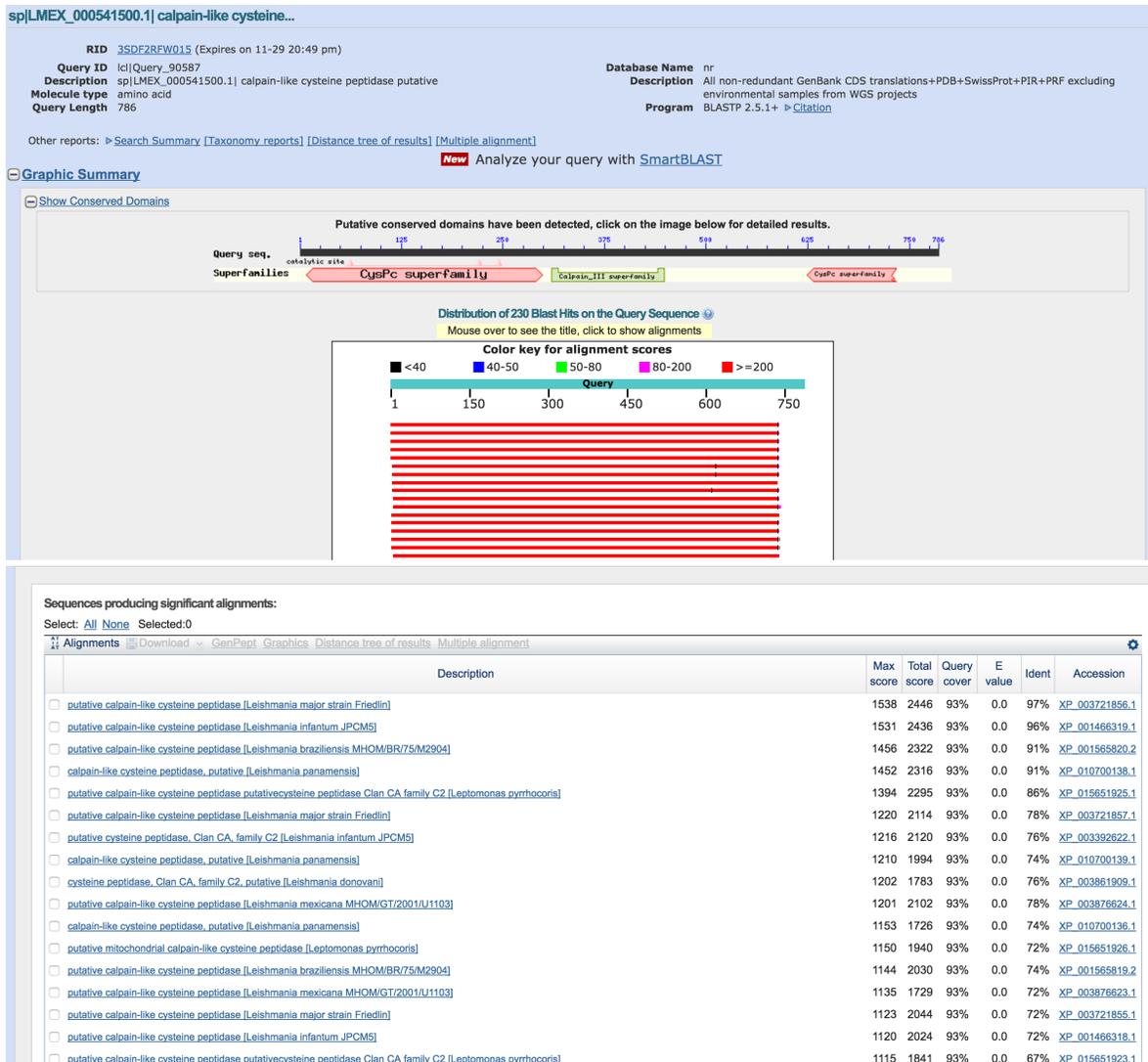


Figura 18. Resultado de la búsqueda de la proteína LMEX_000541500.1 en la base de datos no redundante de proteínas del NCBI usando la herramienta BLASTp.

Al realizar la búsqueda en otra base de datos de nucleótidos del NCBI (Figura 19), tBLASTn, de la peptidasa de cisteína tipo calpaína, se encuentra que la secuencia con mayor similitud es una región del cromosoma 27 de *Leishmania mexicana*, confirmando que se trata de una secuencia proveniente de la especie estudiada.

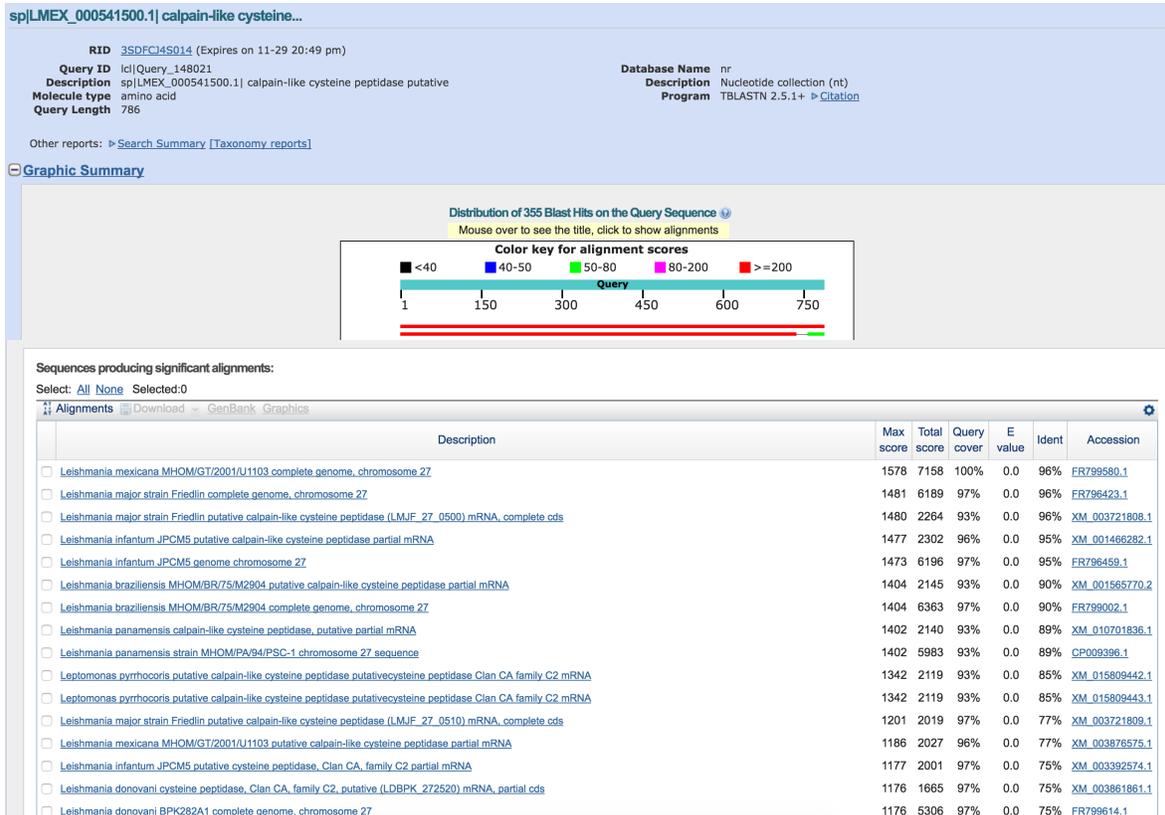


Figura 19. Resultado de la búsqueda de la proteína LMEX_000541500.1 en la base de datos no redundante de nucleótidos del NCBI usando la herramienta tBLASTn.

Analizando la cobertura de los péptidos identificados de la peptidasa de cisteína tipo calpaína podemos observar (Figura 20) una alta cobertura (32%) de péptidos únicos que permitieron identificar esta proteína con una alta certeza.

sp|LMEX_000541500.1| (100%), 89,438.8 Da

Peptidasa de cisteína tipo Calpaína

9 péptidos únicos exclusivos, 11 espectros únicos exclusivos, 89 espectros totales, 253/786 aminoácidos (cobertura de 32%)

```

MSAE E V P D S N K L F S D A V F D R E N A H I A R E W O R I T E V Y P A G V N O P L L P E A F S
R E O F G G O N H Y E C F M I S A L A T L I R F P D V I R N C F V T K K V R O D G R Y T F O F F R G
O E W V R V E I D D R I A M E E G E V L Y V R S P T E H W W P L L L E K A Y A K F Y T A Y D H L E G
C T L O E T F H D L T G N P V L N I P M D A K L A K A A G A E V T E G F Y W L D L A O R I O S G O F
I A S V L T K D I E L E T M G L O R E O O Y G V L E I F S L T G T S S I N D I V I H M H N P F E D E
E Y L Y K G P L N S K D T T W D V K O R A K H D V D D E R S I F L P L N T F L K I M N S L O L C Y I
A P V E A D A T Y F H D E W K G E S A G G N P T S V S W R K N P L Y F V R N S G S T A F E I V V V I
K O E D O R R F T S P D E M T K Y L O C G M V L I H Y S Y A N P I P T F W V T G N N H K P I H K S L
F L N S R E V A N A M T I P P N S L C Y L V P S C M H K G A D G P F T L A L Y R M K D V D Y S D I T
I N K L A I P G M N W S N C A T K N V E L C O K E K D R V D F Y V D E G T D I H I L M H O T K P Y V
S K S G G D A M T E D Y M G M Y L Y D D T D R K I A G V H A A T N F R E I S I I H R L P R S G R Y A
I S I T C P R A K G E V P A Y I T I V G S H A S N V R I V D P P F D A T M F D D F D I I D F G D F A

```

Figura 20. Cobertura de la proteína LMEX_000541500.1. Se identificaron 9 péptidos únicos por LC-MS/MS. Con una cobertura de 32%. En amarillo se muestran los péptidos de esta proteína identificados y en verde los aminoácidos con modificaciones.

6.10 Validación

Una de las proteínas encontradas sobreexpresada en los aislados LCL con respecto a LCD fue la PP2C. Para validar este dato se realizó un WB utilizando extractos totales (20 μ g) de los 10 aislados utilizados en este estudio (5 LCL y 5 LCD), más un aislado de un paciente con leishmaniasis mucocutánea y un aislado de una paciente LCD que generó la forma difusa en un corto periodo de tiempo (1 año). Se utilizó como control positivo una proteína PP2C purificada de *L. mexicana*. La expresión de la proteína PP2C es mayor en los aislados de *Leishmania mexicana* de pacientes diagnosticados con LCL y LMC comparándola con la de aislados LCD (figura 21). El resultado obtenido corrobora los resultados obtenidos por LC-MS/MS. Esta proteína juega un papel importante en la modulación de la respuesta inmune en el hospedero, guiando la infección a una respuesta tipo Th1 en un aislados con LCL según lo descrito por Escalona *et. al*, 2016. Es importante señalar que la PP2C utilizada en este artículo fue purificada de un aislado de la lesión de paciente con LCL. En este trabajo se presenta la expresión entre los aislados localizados y difusos, dando un panorama más amplio de la expresión de proteínas entre aislados con diferente virulencia.



Figura 21. Expresión de la PP2C en 12 diferentes aislados de *Leishmania mexicana*. Western blot de la PP2C en 5 aislados LCL, 6 aislados LCD y 1 aislado mucocutáneo (MCQR). La muestra correspondiente a Isabel es un control para el grupo LCD.

7. DISCUSIÓN

Tomando en cuenta las evidencias en la literatura de la respuesta inmune del hospedero a la infección por *Leishmania mexicana*, se ha observado que una de las diferencias dependiendo el cuadro clínico corresponde al tipo de respuesta inmune que se produce en la infección. Mientras que en la leishmaniasis cutánea localizada se presenta una respuesta inmune celular que favorece la eliminación del parásito (tipo Th1), induciendo la producción de proteínas pro-inflamatorias (TLR-2, TNF- α , IFN- γ), en la leishmaniasis cutánea difusa se desencadena una respuesta inmune celular tipo Th2, que favorece la sobrevivencia del parásito produciendo proteínas antiinflamatorias.

Lo anterior correlaciona con lo observado a nivel proteómico en el parásito, es decir, los aislados que derivan de pacientes con LCD tienen un perfil de expresión diferencial de proteínas relacionadas con una respuesta antiinflamatoria o que han sido reportadas en la literatura como factores de virulencia. Algunas de estas proteínas que resultan importantes para entender la diferencia entre los tipos de aislados son: HSP100, que se menciona en el artículo de Maxwell-Silverman *et al.*, 2010 con un papel determinante para la inducción de la respuesta tipo Th1 cuando está subexpresada. En la tabla 4, se observa una sobreexpresión de los procesos relacionados a síntesis de ácidos grasos y generación de proteínas de membrana como gp63, lo cual nos indica una tendencia a la sobreexpresión de los aislados que derivan a LCD para estas proteínas de virulencia.

7.1 Proteínas presentes exclusivamente en LCL

Escalona et al., 2016 comprobó que la Proteína Fosfatasa 2C (PP2C) induce la producción de $TNF\alpha$, IL-1 β , IL-12p70 and IL-10 en macrófagos de humanos; por lo tanto, al secretar PP2C, promastigotes y amastigotes de *L. mexicana* promueven una respuesta inflamatoria; lo cual correlaciona con el fenotipo observado para LCL, este grupo de investigación utilizó un aislado de *L. mexicana* el cual justamente proviene de un aislado de LCL.

La proteína Peptidil-prolil cis-trans isomerasa Putativa tipo FKBP tiene una actividad de chaperona, cataliza la isomerización cis/trans de los puentes peptídicos seguidos por una prolina. De acuerdo a Kromina et al., 2008, tiene un papel importante para la supervivencia intracelular de *L. major*. Además, se incluye en una lista de potenciales blancos para desarrollo de fármacos (Crowther, 2010. Identification of Attractive Drug Targets in Neglected Disease Pathogens Using an in-silico Approach).

7.2 Proteínas presentes exclusivamente en LCD

Cuando la proteína HSP 100 Putativa dependiente de ATP se elimina en *L. donovani* se observa una respuesta proinflamatoria y un fenotipo Th1 que promueve la eliminación del parásito; el caso contrario tiene una respuesta antiinflamatoria, de acuerdo con Maxwell-Silverman, et al., 2010. En este trabajo encontramos que esta proteína esta sobre-expresada en los aislados LCD, lo que podría sugerir que esto beneficia a los parásitos para inducir una respuesta que le permite sobrevivir dentro de la célula hospedera.

Tanto la presencia de la proteína PP2C, como la ausencia de HSP100 al momento de la infección podrían jugar un papel importante en la respuesta inmune del hospedero.

Por otro lado, la proteína PP2C que induce una respuesta pro-inflamatoria según Escalona et al. en 2016, se encuentra sub expresada en los aislados LCL. Por otro lado, se encontraron proteínas con alto potencial para ser usadas como blancos para el desarrollo de nuevos fármacos; tal es el caso de la FKBP y la CBP.

7.3 Proteogenómica

Las 39 proteínas descritas en la tabla 5 son contribuciones directas a la base de datos del proteoma de *Leishmania mexicana* de UniProt.

Se observa en los resultados una identidad del 97% entre la peptidasa de cisteína tipo calpaína de *L. major* y la secuencia LMEX_000541500.1. Sin embargo, también se observó que el ORF más cercano corresponde a *L. mexicana*, aunque la identidad es apenas del 96%. Las diferencias entre la cepa de referencia y la variabilidad genómica de los aislados que fueron secuenciados pueden explicar esta diferencia; la evidencia proteómica es muy contundente.

8. CONCLUSIONES

Resumiendo todo lo encontrado, se concluye que existe un perfil diferencial de expresión de proteínas entre aislados que derivan la forma difusa y localizada de la leishmaniasis cutánea.

Las proteínas encontradas diferencialmente en el parásito, pueden apuntar a que éstas están modulando la respuesta inmune en el hospedero, incluso de forma diferente para cada una de las formas clínicas. La respuesta tipo Th1 se tiene que activar para microorganismos intracelulares, sin embargo, la Th2 es la que predomina en los pacientes con LCD, lo que se sugiere que los parásitos que derivan a LCD pueden inducir una respuesta tipo Th2.

El estudio proteómico realizado sugiere que existen proteínas que inducen o inhiben una respuesta inmune tipo Th1 o Th2 en el hospedero es clave para el entendimiento del papel del parásito en la infección; se proponen como casos de estudio las proteínas: Proteína Fosfatasa 2C (PP2C), Proteína de choque térmico 100 (HSP100), Peptidasa de Cisteína Clan CA familia C19, metaloproteasa clan MF familia m17 (E9ANK8), p-Glicoproteína (E9B4S1), Ubiquitin hidrolasa (E9ALJ4), Peptidil-prolil cis-trans isomerasa Putativa tipo FKBP, Antígeno inmunodominante putativo tipo Tc40 (E9APT8), HSP60, Beta liasa (E9AJD5) y las proteínas no caracterizadas (E9ALS6, E9B660, E9B0I3).

Las proteínas que pueden ser estudiadas para su potencial clínico son las FKBP, HSP60 por la importancia que tienen estas proteínas para la supervivencia del parásito.

Utilizando los resultados del genoma y transcriptoma de los parásitos se pudo completar la base de datos de genes para *Leishmania mexicana*, necesarios para la identificación proteómica

Desde el punto de vista bioinformático, se creó un flujo de análisis para la integración de diferentes tecnologías ómicas (proteómica, genómica y transcriptómica), en este caso un término recientemente descrito como proteogenómica. Este trabajo contribuye a presentar la evidencia biológica a diferentes niveles genómicos (transcritos o proteínas).

9. REFERENCIAS

- [Abbas, 2004] Abbas A, Litchman A, Pober J. Inmunología celular y molecular. Mc Graw-Hill Interamericana. 2004. 5ta ed. México. p.p.563.
- [Alcolea et al., 2011] Alcolea, P. J., Alonso, A., and Larraga, V. (2011). Proteome profiling of leishmania infantum promastigotes. *Journal of Eukaryotic Microbiology*, 58(4):352–358.
- [Alvar et al., 2012] Alvar, J., Vélez, I. D., Bern, C., Herrero, M., Desjeux, P., Cano, J., Jannin, J., den Boer, M., Team, W. L. C., et al. (2012). Leishmaniasis worldwide and global estimates of its incidence. *PloS one*, 7(5):e35671.
- [Berrueta, 2015] Berrueta, T. U. (2015). Leishmaniosis o leishmaniasis. departamento de microbiología y parasitología, facultad de medicina, UNAM.
- [Biyani et al., 2011] Biyani, N., Singh, A. K., Mandal, S., Chawla, B., and Madhubala, R. (2011). Differential expression of proteins in antimony-susceptible and- resistant isolates of leishmania donovani. *Molecular and biochemical parasitology*, 179(2):91–99.
- [Bjornson et al., 2007] Bjornson, R. D., Carriero, N. J., Colangelo, C., Shifman, M., Cheung, K. H., Miller, P. L., & Williams, K. (2007). X!! Tandem, an improved method for running X! tandem in parallel on collections of commodity computers. *The Journal of Proteome Research*, 7(1), 293-299.
- [Brotherton et al., 2012] Brotherton, M.-C., Racine, G., Ouameur, A. A., Leprohon, P., Papadopoulou, B., and Ouellette, M. (2012). Analysis of membrane-enriched and high molecular weight proteins in leishmania infantum promastigotes and axenic amastigotes. *Journal of proteome research*, 11(8):3974–3985.
- [DNDi, 2010] DNDi, Annual report 2010. Developing, Implementing, Consolidating. Recuperado 22 de agosto de 2017, a partir de <https://www.dndi.org/about-dndi/annual-reports/annual-report-2010/>
- [Fernández-Figueroa, 2013] Fernández-Figueroa, E. A. (2013). Análisis de Genes y Proteínas de Células NK y Monocitos de Pacientes con Leishmaniasis Cutánea Localizada y Difusa. PhD thesis, Universidad Nacional Autónoma de México, Facultad de Medicina, UNAM.
- [Guthals, 2012] Guthals, A., & Bandeira, N. (2012). Peptide identification by tandem

mass spectrometry with alternate fragmentation modes. *Molecular & Cellular Proteomics*, 11(9), 550-557. <https://doi.org/10.1074/mcp.R112.018556>

[Ivens et al., 2005] Ivens, A. C., Peacock, C. S., Worthey, E. A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.-A., Adlem, E., Aert, R., et al. (2005). The genome of the kinetoplastid parasite, *leishmania major*. *Science*, 309(5733):436–442.

[James, 1997] James, P. (1997). Protein identification in the post-genome era: the rapid rise of proteomics. *Quarterly reviews of biophysics*, 30(04):279–331.

[Liu, 2004] Liu, H., Sadygov, R. G., & Yates, J. R. (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Analytical Chemistry*, 76(14), 4193-4201. <https://doi.org/10.1021/ac0498563>

[Lupi, 2009] Lupi O, Bartlett BL, Haugen RN, Dy LC, Sethi A, Klaus SN, et al. Tropical dermatology: Tropical diseases caused by protozoa. *J Am Acad Dermatol*. 2009 Jun;60(6):897-925; quiz 6-8. PubMed PMID: 19467364. Epub 2009/05/27. eng.

[Myler and Fasel, 2008] Myler, P. J. and Fasel, N. (2008). *Leishmania: after the genome*. Horizon Scientific Press.

[Paape and Aebischer, 2011] Paape, D. and Aebischer, T. (2011). Contribution of proteomics of *leishmania* spp. to the understanding of differentiation, drug resistance mechanisms, vaccine and drug development. *Journal of proteomics*, 74(9):1614–1624.

[Pawar et al., 2012] Pawar, H., Sahasrabudhe, N. A., Renuse, S., Keerthikumar, S., Sharma, J., Kumar, G., Sameer, S., Venugopal, A., Sekhar, N. R., Kelkar, D. S., et al. (2012). A proteogenomic approach to map the proteome of an unsequenced pathogen—*leishmania donovani*. *Proteomics*, 12(6):832–844.

[Pawar et al., 2014] Pawar, H., Renuse, S., Khobragade, S. N., Chavan, S., Sathe, G., Kumar, P., Mahale, K. N., Gore, K., Kulkarni, A., Dixit, T., et al. (2014). Neglected tropical diseases and omics science: proteogenomics analysis of the promastigote stage of *leishmania major* parasite. *Omics: a journal of integrative biology*, 18(8):499–512.

[Qian et al., 2005] Qian, W. J., Liu, T., Monroe, M. E., Strittmatter, E. F., Jacobs, J. M., Kangas, L. J., ... & Smith, R. D. (2005). Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: the human proteome. *Journal of proteome research*, 4(1), 53-62.

[Rogers et al., 2011] Rogers, M. B., Hilley, J. D., Dickens, N. J., Wilkes, J., Bates, P. A., Depledge, D. P., Harris, D., Her, Y., Herzyk, P., Imamura, H., et al. (2011). Chromosome and gene copy number variation allow major structural change between species and strains of leishmania. *Genome research*, 21(12):2129–2142.

[RutaN, 2015] RutaN. (2015). UN ENCUENTRO PARA HABLAR DE ENFERMEDADES OLVIDADAS. Recuperado el 30 de agosto de 2017, desde: <https://www.rutanmedellin.org/es/actualidad/noticias/item/red-leish-2015-un-encuentro-para-el-bienestar-de-los-mas-necesitados-7>

[Sardar et al., 2013] Sardar, A. H., Kumar, S., Kumar, A., Purkait, B., Das, S., Sen, A., Kumar, M., Sinha, K. K., Singh, D., Equbal, A., et al. (2013). Proteome changes associated with leishmania donovani promastigote adaptation to oxidative and nitrosative stresses. *Journal of proteomics*, 81:185–199.

[Schneider, 1992] Schneider P, Rosat JP, Bouvier J, Louis J, Bordier C. *Leishmania major*: differential regulation of the surface metalloprotease in amastigote and promastigote stages. *Exp Parasitol*. 1992 Sep;75(2):196-206. PubMed PMID: 1516667. Epub 1992/09/01.eng

[Seidelin, 1912] Seidelin, H. (1912). Leishmaniasis and babesiasis in yucatan. *Annals of Tropical Medicine & Parasitology*, 6(2):295–300.

[Spath, 2000] Spath GF, Epstein L, Leader B, Singer SM, Avila HA, Turco SJ, et al. Lipophosphoglycan is a virulence factor distinct from related glycoconjugates in the protozoan parasite *Leishmania major*. *Proc Natl Acad Sci U S A*. 2000 Aug 1;97(16):9258-63. PubMed PMID: 10908670. Epub 2000/07/26. eng.

[Spath, 2003] Spath GF, Garraway LA, Turco SJ, Beverley SM. The role(s) of lipophosphoglycan (LPG) in the establishment of *Leishmania major* infections in mammalian hosts. *Proc Natl Acad Sci U S A*. 2003 Aug 5;100(16):9536-41. PubMed PMID: 12869694. Epub 2003/07/19. eng.

[Tsigankov et al., 2013] Tsigankov, P., Gherardini, P. F., Helmer-Citterich, M., Späth, G. F., and Zilberstein, D. (2013). Phosphoproteomic analysis of differentiating leishmania parasites reveals a unique stage-specific phosphorylation motif. *Journal of proteome research*, 12(7):3405–3412.

[Villaseñor, 2008] Villaseñor MI, Saliza N, Delgado J, Gutiérrez L, Pérez A, et al. Mast cells are activated by *Leishmania mexicana* LPG and regulate the disease outcome depending on the genetic background of the host. *Parasite Immunol*. 2008;30: 425–434.

[Von Stebut, 2007] Von Stebut E. Immunology of cutaneous leishmaniasis: the role of mast cells, phagocytes and dendritic cells for protective immunity. *Eur J Dermatol.* 2007 Mar- Apr;17(2):115-22. PubMed PMID: 17337393. Epub 2007/03/06. eng.

[WHO, 2013] WHO (2013). Sustaining the drive to overcome the global impact of neglected tropical diseases. *SECOND WHO REPORT ON NEGLECTED TROPICAL DISEASES.* World Health Organization.

[Yates III et al., 1995] Yates III, J. R., Eng, J. K., and McCormack, A. L. (1995). Mining genomes: correlating tandem mass spectra of modified and unmodified peptides to sequences in nucleotide databases. *Analytical chemistry*, 67(18):3202–3210.

[Zhang et al., 2013] Zhang, Y., Fonslow, B. R., Shan, B., Baek, M.-C., and Yates III, J. R. (2013). Protein analysis by shotgun/bottom-up proteomics. *Chemical reviews*, 113(4):2343–2394.