MASS SPECTROMETRY-BASED PROTEOME ANALYSIS OF *Leishmania major* PARASITE IN TWO CLINICAL ISOLATES WHICH EXHIBIT DIFFERENT IMPACT ON VIRULENCE

A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Chemistry

by
Nazlı GÜVENÇ

December 2018
İZMİR

We approve the thesis of **Nazlı GÜVENÇ**

Examining Committee Members:

Prof. Dr. Talat YALÇIN
Department of Chemistry, İzmir Institute of Technology

Prof. Dr. Ahmet ÖZBİLGİN
Department of Parasitology, Celal Bayar University

Doc. Dr. Ferda SOYER DÖNMEZ
Department of Molecular Biology and Genetics, İzmir Institute of Technology

28 December 2018

Prof. Dr. Talat YALÇIN
Supervisor, Department of Chemistry
İzmir Institute of Technology

Prof. Dr. Ahmet EROĞLU
Head of the Department of Chemistry

Prof. Dr. Aysun SOFUOĞLU
Dean of the Graduate School of
Engineering and Sciences

ACKNOWLEDGMENTS

First of all, I would like to thank my thesis supervisor Prof. Dr. Talat YALÇIN for his guidance, suggestions, understanding, encouragement and endless support throughout my thesis study. Also, in addition to his valuable contributions to this work, I would like to thank him for giving us moral support even in most difficult moments with sharing his life experiences whenever it is needed. It was an excellent experience for me to join such a nice and warmth family of Mass Spectrometry team.

I would like to express my special thanks to Prof. Dr. Ahmet ÖZBİLGİN and his team for their collaborated work in the part of isolation of the *L. major* proteins and for providing their support whenever I need protein samples that I use in my experiments. Also, I would like to express my special thanks to İbrahim ÇAVUŞ for his cooperation, patience, helps and understanding in every steps of the isolation of samples and any difficulty about *L. major*.

I am very grateful to my lovely sisters Begüm ÖZYAĞCI to share this adventure with me and she has listened to me without being tired and always be the most supportive person in my life and encouragement me in any difficulty. Besides, I would like to thank Seren ŞEN for her friendship, helps, cheer and, warmth. Also, I would like to thank my laboratory friends Dr. Melda GÜRAY and Dr. Melike DİNÇ for their guidance, contributions and kind helps to this work. All member of the 'Masslab' team has will never be forgotten. We had the most fabulous time together.

I wish to acknowledge my precious teachers in Chemistry Department in İYTE for their guidance and endless support for the times that I spent in this college.

Finally, I would like to express my special thanks to my mother, father, and brothers for their endless love, being in my behind always not matter what, always trusting, supporting and encouraging me in every moment of my life. I am fortunate to have such a great family. I love them so much and dedicate my thesis to my family. Also, I would like to express my thanks to my lovely boyfriend Deniz YAMAÇ for his support, motivation and understanding throughout not only this work but also in my entire life by expressing his kind love that affect all my life.

ABSTRACT

MASS SPECTROMETRY-BASED PROTEOME ANALYSIS OF *Leishmania major* PARASITE IN TWO CLINICAL ISOLATES WHICH EXHIBIT DIFFERENT IMPACT ON VIRULENCE

Leishmaniasis is a disease that covered under the title of neglected tropical diseases caused by protozoan parasites called *Leishmania* which can classify into three groups as visceral, cutaneous and mucocutaneous leishmaniasis. *L. major* is a type of parasite that causes cutaneous leishmaniasis and it is endemic in Iran, and Syria. However, cutaneous type leishmaniasis caused by *L. major* has been begun to detect in Turkey due to its close location to such countries. Moreover, the variety in the infectivity of *L. major* in a different region of Turkey has detected. Therefore, the uncertainty under the virulence effect of *L. major* is aimed to investigate.

Large-scale protein analysis by mass spectrometry based proteomics has cleared up to proteome mapping for different organism recently. Generally, although two methodologies that involve gel-free and gel-based approaches have widely accepted for proteomic analysis, gel-free LC-MS/MS analysis were applied to characterize the proteome analysis of *L. major* parasite in two clinical case exhibiting passive and aggressive virulence effect on leishmaniasis.

Finally, differential and common proteins that can affect the infectivity of *L. major* invastigated by shotgun analysis. As a result, samples showed that there are conflict results with the literature about GP63, secreted acid proteases, cysteine proteases and Peroxiredoxin proteins existence and also in the aggressive *L. major* cystathionine beta-synthase protein which has an role to synthesis of CPs and pyridoxal phosphate binding activity were proposed as a critical protein for *L. major* infectivity due to its association with SAPs, CPs.

ÖZET

Leishmania major PARAZİTİNİN FARKLI VİRÜLENS ETKİ GÖSTEREN İKİ KLİNİK VAKADA KÜTLE SPEKTROMETRESİNE DAYALI PROTEOMİK ANALİZİ

Leşmanyöz, leşmanya parazitinin sebep olduğu ve genelde üç farklı klinik formda bulunan bulaşıcı bir hastalıktır. *L. major* paraziti ise kütanöz leşmanyöz türlerinden olmakla birlikte genelde, Irak ve Suriye gibi bölgelere yayılmıştır. Fakat, ülkemizin jeolojik konumundan dolayı, *L. major* parazitine bağlı deride oluşan leşmanyöz vakaları ülkemizde de saptanmaya başlamıştır. Vakaların genelinde farklı bölgelerde ve farklı şiddette görülen lezyonlar, parazitin virulans etkisinin farklılık gösterebildiğinin kanıtı olmuştur.

Kütle spektrometresine dayalı büyük ölçekli protein incelemeleri, günümüzde birçok proteom analizinde kullanılmaya başlanmıştır. Genelde, jele dayalı ve jelden bağımsız (LC-MS/MS) gibi iki temel prensip dünyaca kabul görmüştür. Bu çalışmada, leşmanyöz üzerinde pasif ve agresif virülans etki gösteren iki klinik vakada *L. major* parazitinin proteome analizini karakterize etmek için jel içermeyen LC-MS / MS analizi uygulanmıştır.

Sonuçta, agresif ve pasif etki gösteren *L. major* parazitinin proteome analizinden elde edilen birbirinden farklı ve ortaklık gösteren proteinler, *L. major*'un virülans etkisinin kaynağının hangi proteinlerden kaynaklandığını görmek için ele alınmıştır. Sonuç olarak, literatürde bulunmuş *L. major* virülansına etki eden GP63, salgılanan asit proteazları, sistein proteazları ve peroksiredoksin proteinleri bu çalışmada da bulunmuştur. Bunun dışında, sadece agresif örnekde gözlenen, sistatiyonin beta-sentaz proteini, sistein sentezinde kullanılması ve ayrıca piridoksal fosfat bağlanması aktivitesinden ötürü *L. major*'ün virülans etkisinde rol oynayacağı öne sürülmüştür.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

Dedicated to;

my supportive and lovely family

for always being my side…

CHAPTER 1

MASS SPECTROMETRY AND PROTEOMICS

1.1. Introduction to Mass Spectrometry

A mass spectrometer can define that is an analytical tool which converts molecules into ions and separates them according to their 'mass to charge' *(m/z)* ratio. It is a powerful technique to identify unknown compounds in a sample and clarify the structure and chemical properties of different molecules.

Mass spectrometers are composed of three main components: an ion source, mass analyzer and detector showed in Figure 1.1. It is important to note that the components of a mass spectrometer must operate under high vacuum conditions (10^{-5} to 10^{-7} Torr), sustained by the vacuum systems such as pumps, in order to eliminate the gas molecules other than analyte gas phase ions. Also, in a vacuum, molecules cannot give a reaction and collide to each other to go back its neutral position.

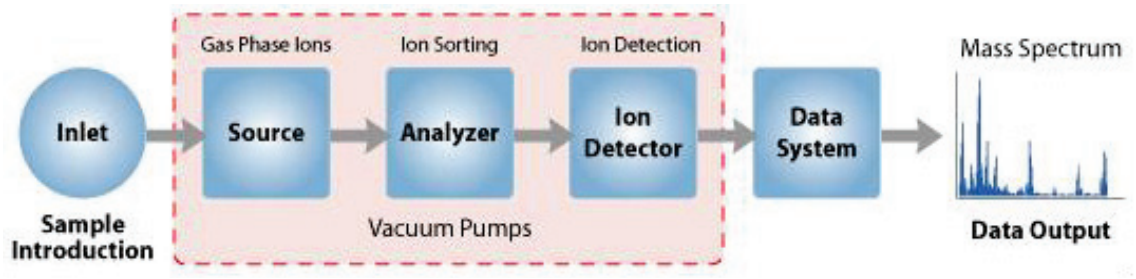


Figure 1.1. Components of Mass Spectrometer

The measurement process depends upon the chemical reactions in the gas phase fragmentation and then characterized by their mass to charge ratios *(m/z)* and relative abundances. This process contains some main steps. Firstly, gas phase ions of the compound (negatively and positively charged) generate by electron ionization, and then these ions undergo fragmentation. These fragmented ions are separated magnetically or electronically based on their mass to charge ratios in the mass analyzer and detected by suitable detector. A mass spectrum of the molecule obtains by displaying the result in the

form of a plot of ion abundance versus mass-to-charge ratio such as shown in Figure 1.2. The figure shows that a mass spectra of the 4-methyl-3-pentene-2-one. It is reflecting the gas phase fragmentation of the molecule given in its *m/z* values in Dalton (Da) that is proportional with their relative intensities (a.u.). Also, The data obtained from a mass spectrum provides information about molecular mass and abundance of the analyte molecules, also their purity, composition and structure.

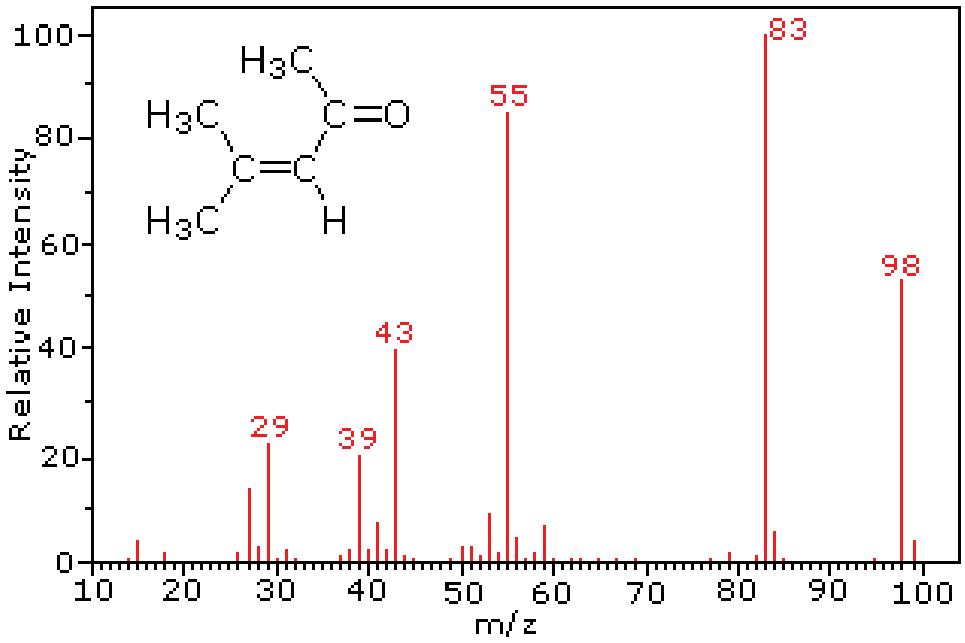


Figure 1.2. An example of mass spectra of the 4-methyl-3-pentene-2-one

The history of mass spectrometry dates back to the 1800s although it showed its first development in the early 20th century. The invention of the mass spectrometry was taken place with the discovery of the electron by J.J. Thomson's cathode ray experiment in 1897s. In the experiment, he built an instrument that measures *e/m* (charge to mass ratio) value by using an electric field inside a cathode ray tube, so detecting the mass of the electron differently.

After this developments, Thomson and with the help of Francis Aston built an instrument, that is recognized later as a mass spectrometer, to measure the masses of the charged atoms. In the 1980s, because studies attempted to the analysis of biomolecules like proteins, carbohydrates and nucleic acids in the field of biochemistry, development of new methods for the mass spectrometry become essential [1,2]. With the invention of 'the soft ionization' techniques , which is matrix-assisted laser desorption ionization

(MALDI) [3,4] and electrospray ionization (ESI) [5], analysis of large, polar and nonvolatile molecules in the gas phase were achieved.

1.2. Proteomics

Recently, the attempts in experimental biology to find out the relation between genes and their proteins has become an important issue. The pathway from genes to proteins includes DNA, transcripts and metabolites but generally proteomics deal with the entire set of proteins expressed by a genome, cell, tissue or organism [6]. Generally, the aim of the proteomic is to complete identification of all proteins since 25 years. It started with Human Genome Project which is an international study to characterize and sequence the entire human genome [7]. The project provided complete and accurate data for the sequence of 3 billion DNA pairs that made finding out all of the estimated nearly 25,000 human genes [8]. Although this genomic study leads to many information about human genome, it changes the conventional considiration as ‘ one gene encodes one protein’ into ‘one gene encodes more than one protein’. Therefore, it can said that if the post-translational modifications at the protein level and DNA splicing considered, the identification might be over a million of human proteins. It shows that there is no strict linear relationship between genes.

In contrast to genome which is static and can encode multiple different proteins, proteins are dynamically adapt the environment and move through the body to interact with other proteins leading to some reaction. In addition to that posttranscriptional modifications such as phosphorylation , glycosylation and some properties that are vital for enzymatic activity, cell signaling, cellular interactions and the immune system cannot be captured by genomics. Since the proteins are the product of an expressed gene, attempt have been focused on the proteomic studies to have an impact on th disease by characterizing protein structure and function.

Features starting with the proteomic studies described by a ‘proteome’ which blended from the terms protein and genome. It was driven in early the 1990’s by Marc Wilkins to bridge over with genomics [9]. While genomics only deal with the total DNA of a cell organism , proteomics studies open the door to analyzing the composition, structure, function and interaction of the proteins directing the activities of each living cell.

For this purposes, proteomic researches that works in the aim of the Human Proteome Organization (HUPO) are;

- Protein expression in different cell types of the organism,
- Protein distribution in subcellular compartments of the organelles,
- Post-translational modifications of the proteins,
- Protein-protein interactions,
- The relation between protein structure and function.

Therefore, it gives the understanding of human biology at the cellular level and leads to recognizing diagnostic and therapeutic approaches to medical applications such as cancer, chronic diseases, and neurodegenerative pathologies.

Generally workflow of proteomic studies can be followed by first protein extraction and their purification from the mixture of complex components such as DNA and RNA in the mixture , separation of proteins by using different methods and digesting with proper enzyme which is generally Tyripsin, then identification and sequencing the protein of interest with different methods by accessing several protein and DNA databases.

The conventional methods for the sequencing of proteins are both Edman sequencing and Western Blot analysis. The basis of the Edman sequencing is Phenyl isothiocyanate (PITC) react with an α-amino group of the N terminal at the peptide chain. Cleavage occurs by Trifluoroacetic acid (TFA), and N-terminal of the chain lose one amino acid and become a part of sequencing. However, this method has some defects to be useful for comprehensive proteomic work. Because it is a very exhaustive method and needs more protein sample, also it is suitable up to 50-60 redues at most [10].

Another method that is used for the sequencing is the Western Bloting that is also known as protein blotting method. It is beneficial for the immunodetection of the proteins, especially the proteins with low aboundance. In this method, proteins are transferred from a sodium dodecyl sulfate (SDS) polyacrylamide gel to an porous membrane. The blotted proteins generate a replica of the SDS-PAGE gel. After this starting step, following experiments is done by an antibody probe that has an interaction between the nitrocellulose proteins [11].

The developments in the area of proteomic studies have accelerated with the new technologies. The real development in this area began with the introduction of the ‘soft ionization’ methods of mass spectrometry joined in 2D electrophoresis separation

techniques. This method is performed by using Matrix Assisted Lazer Desorption/Ionization (MALDI) and Electrospray Ionization (ESI). The key parameters for the mass spectrometry-based proteomics analysis is that the experiments must include high sensitivity, mass accuracy, dynamic range, and resolution. Therefore, the amount of the sample and its qualification is essential to have accurate results for the sequencing and identification.

1.2.1. MS-based Proteomic Studies

As, it is mentioned above, the real development in proteomic studies started with the soft ionization techniquies of mass spectrometery after convensitonal 2D electrophoresis method. Proteomic studies generally performed by using MALDI and ESI sources that are leading to this technique. In both technique, biomolecules like proteins and peptides ionized into gas phase but they have different principles. Also, the mass analyzer used for the proteomic studies is crucial. For that reason, the most commonly used mass analyzers can list as following; quadruopole (Q), iontrap (LTQ), time of flight (TOF) and Fourier-transform ion cyclotron resonance (FTICR) mass analyzer. In addition to that sometimes, this analyzers can be used as a 'hybrid' instrumentation which designed with the combination of Q-q-Q, Q-TOF, TOF-TOF or LTQ-FTICR [12].

In MALDI, the analyte is mixed with a matrix which is an organic compound soluble with the analyte and has an ability to absorp laser energy (generally 337 nm) that comes from the source. Then, it is spotted to the metal target plate and waited for the air-drying. After evaporation on target plate, it placed into MS and laser is addressed different spots on the target. Molecular ions generate by irradiation of laser. During analysis, the analyte is co-crystallized together with a matrix. When the laser hits the spot, matrix absorb the energy and both analyte and matrix generates the gas phase ions. Then, energy release and it leads to desorption of analyte and matrix with formation of negatively and positively charged ions by protonation and deprotonation between analyte and matrix. The principle of the MALDI source was illustrated in Figure 1.3 [14]. Generally, the matrix that is used for analysis is α-cynano-4-hydroxycinnamic acid (CHCA) or it can be 2,5-dihydroxybenzoic acid (DHB). Also, MALDI source is generally combined with TOF-TOF analyzer system to be useful in proteomic studies [13].

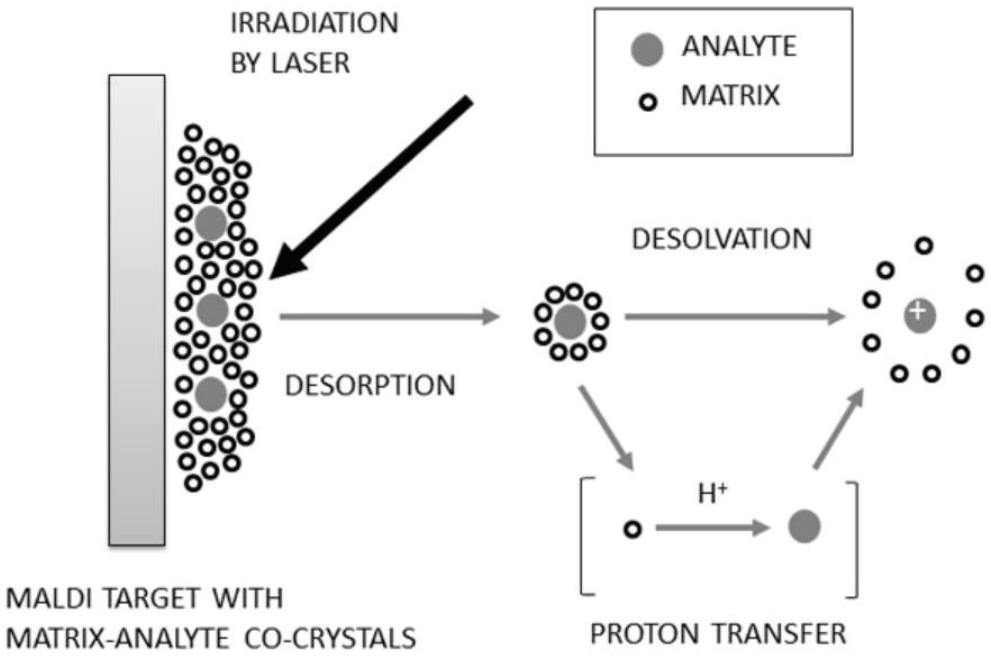


Figure 1.3. A scheme of the principle of MALDI
(Source: Stroobant, 2007)

On the other hand, ESI generally used for non-volatile liquid samples. In principle, the liquid sample flows through a microcapillary tube which has a potential difference, nearly ±2,5V, between the inlet of the mass spectrometry. This process ends with formation of aerosol droplets while heated Nitrogen gas is used to evaporate the solvent in a droplet and causes the droplet getting smaller and smaller and begin negative and positive charges get closer and repel each other. After a critical point where coulombic, repulsive, forces exceed the surface tension of the solvent, droplet explodes and break apart into small particles and then, di-solvated and charged analyte transferred to the analyzer from ESI source [15]. Figure 1.4 [17] shows the illustration of the general principle of the ESI source. Generally, ionic species and neutral compounds, converted to an ionic form, can be utilized for the ESI-MS systems. The precursor ions of the interest could select for further fragmentation with using the collision cell. Mostly, quadruopole and the tandem mass analyzer used with ESI-MS systems. It is based on the principle that ions move through a magnetic and electric field and resulting a movement in the area that is affected by m/z ratio of ions.

This MS/MS system can continue from MS to MS^n. Because of these properties, ESI-MS provides sensitive, robust, and reliable tool for many different study area such as proteomic analysis. When the system combined with HPLC systems for molecule fractionation based on MS analysis, it ensures a powerful technique for the capability of the analyzing both small and large biological molecules of the different protein mixtures [16].

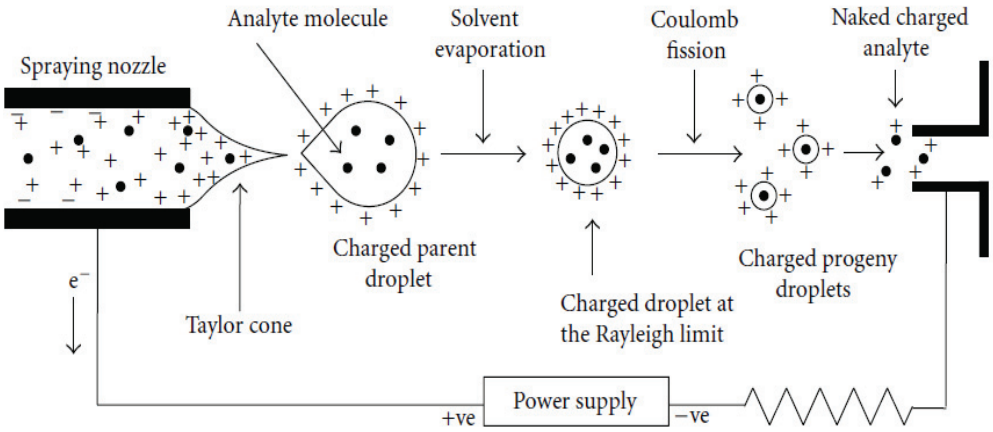


Figure 1.4. General principle of the electrospray ionization process (Source: Banerjee and Mazumdar, 2012)

In both method, sample preparation before analyte introduced into mass spectrometry, is crucial. That is why, several separation methods were developed and combined with MS system. In Figure 1.5 [9], strategies for MS-based proteomic studies listed as a summary. After extraction and purification of the proteins from sample mixture, strategies can be divided into two approaches as bottom-up and top-down to identification and characterization with mass spectrometry. In top down proteomic approaches, intact protein involves the gas phase ionization without any pre-fragmentation and follow with high-resolution mass spectrometry for the direct fragmentation of the intact protein. In other words, the protein is fragmented into the mass spectrometer to generate peptide ions that is highlighting its sequence. Generally, ESI-FT/MS used for the top-down analysis with its high-resolution property. Although, top- down approaches is a robust process, there are some limitations that decrease the efficiency of large scale protein identification and characterization. Therefore,

the top-down analysis of the abundant proteins (>50 kDa, approximately) cannot be efficient because of their comlex tertiary structure in the gas phase.

On the other hand, in bottom-up proteomic approaches, the analyte is introduced into the mass spectrometer (also called as tandem mass spectrometry) as a peptide mixture created from enzymatic digestion of the intact protein. Tandem mass spectrometry provides for selecting every single peptide ion from peptide ion mixture for analysis, and then each peptide ion gives the fragmentation pattern for the amino acid sequence of the peptide. These amino acid sequences used for further analysis by searching sequence databases for the protein identification. Also, before peptide mixture introduces into mass spectrometry, they separated and fractionated chromatically or with a conventional method such as 2D-PAGE. This fractionation provides decreasing the complexity of the mixture before mass spectrometry measurement.

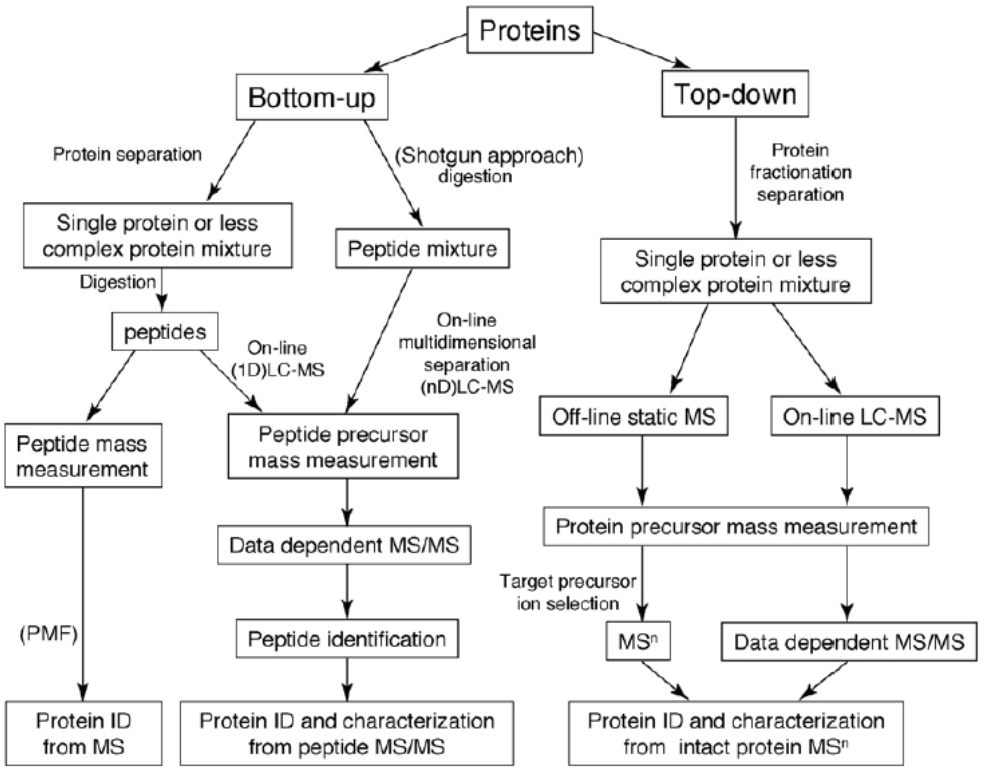


Figure 1.5. Strategies for MS-based protein identification and characterization in proteomic studies (Source: Graves and Haystead, 2002)

Two primary methods used in proteomic studies of bottom-up approaches for the separation method are two-dimentional polyacrylamide gel electrophoresis (2D-PAGE) and liquid chromatography systems such as high pressure liquid

chromatography (HPLC). These techniques are called as gel-based and gel-free methods associated with their techniques used for separation. In 2D-PAGE, large number of protein can be separated. Proteins mixture separates using two-dimensional Electrophoresis approach. The first one base on isoelectric focusing, and then followed molecular size separation. In the end, all proteins spots observed on the SDS-polyacrylamide gel by using different stains such as Coomassie, silver staining or fluorophores. After staining, protein spots are excised from the gel and involved in enzymatic digestion. Then generally, MALDI-TOF mass spectrometry is used for further analysis for identification and characterization of peptide mixtures. The quality of the results coming from database search after mass spectrometry analysis depends on purity, the number of identified peptide mass and efficiency of trypsin digestion after excising spots from the gel. Therefore, 2D-PAGE is a powerful method and it can separate up to thousands of proteins. Also, separation can involve different pH gradients such as pH5-6, pH4-7 or pH3-10 during isoelectric focusing to visualized specific regions on the polyacrylamide gel during analysis. On the other hand, the intensity of spots on the gel can provide relative quantitation by observing how much protein present on the gel with comparison one spot to another spot.

Although 2D-PAGE is a powerful separation method for mass spectrometry based protein analysis, some drawbacks can limit its effectiveness. For example, excising protein spots from the gel is very time-consuming because of the high amount of spots present in the gel. Also, destaining of every excised spots is necessary so it is another time-consuming process. Also, this method can only separate the proteins that their pI values are between 3 and 10. This causes that molecular weight of protein below 15kDa and above 200 kDa cannot detect on the gel. Apart from that, integral membrane proteins and low abundance proteins are not observed on gel sometimes. Recently, attempts have focused on optimizing these drawbacks to 2D-PAGE systems [18]. That is why, several proteomic methods have been investigated recently to overcome the difficulties coming from gel-based proteomic analysis technique. Gel-free proteomic analysis is one of them.

One type of gel-free approach is the mass spectrometry combined with liquid chromatography system for peptide separation. It is also generally called ‘shotgun proteomics’. In this method, complex protein mixture digest with trypsin first and the resulting peptide mixture was fractionated with on-line and off-line HPLC system before MS/MS analysis.

1.2.2. Shotgun Proteomic Approaches

As mentioned above sections, mass spectrometry becomes a central analytical tool for the identification and characterization of the biomolecules, especially in proteomic studies. One of the effective way to use mass spectrometry in proteomic studies is shotgun proteomic approaches that are consist of a multidimensional separation called as MudPIT. In the past decade, shotgun proteomic approaches have been widely used by researches for different proteomic purposes such as proteome profiling, protein quantification and modification, an interaction between proteins and identification, and characterization of the proteins. In this method, isolated protein complex samples are digested with trypsin enzyme generally and, resulted in peptide mixture are fractionated with the assist of the liquid chromatography system, then analyzed by automated MS/MS system. Since the automated system used for MS/MS scan, shotgun proteomic requires the data-dependent acquisition operation mode of mass spectrometry for both MS and MS/MS analysis performed in order.

As in all proteomics studies, sample preparation before peptide fragmentation is an essential step in shotgun studies. Since the complexity of the peptide mixture is too high after enzymatic digestion, the liquid chromatography system decreases complication between peptide in the mixture and provides peptides introduce mass spectrometry in order. However, this fragmentation of peptide mixtures causes loss of the connectivity between the peptides and associated proteins and obtained fragmental spectra can be used for evaluation of each peptide from highly complex peptide pool to identify the correct protein of them. Thereby, to provide the link between peptides and a related protein, a search engine is required to piece together of the high throughput data acquisition coming from MS/MS system. Several algorithms have been developed to automate the identification of the peptide pieces to correct protein. One of the most used database search engines is called MASCOT. Also, the most popular tandem mass spectrum (MS/MS) database matching tools summarized in Table 1.1.

Because of the method described above, experimental simplicity, increased proteomic coverage compared with the gel-based method and obtaining high-throughput data in a short time without protein separation techniques such as 2D-PAGE is the main strength of the shotgun approach.

Table 1.1. Tandem mass spectrum (MS/MS) database matching tools

Search Engine Name	Website
Mascot	www.matrixscience.com
MS-Tag	www.prospector.ucsf.edu
Pepsea	www.pepsea.protana.com
SEQUEST	www.fields.scripps.edu/sequest
Sonar	www.proteometrics.com

However, this method also suffers from the challenges about limited dynamic range, some informatic challenges associated with protein sequence identification from a large number of acquired mass spectral data, protein interference problems, high redundancy, and also complications due to highly mixed peptide samples. These challenges in the shotgun method have accomplished by the use of fractionation that helps to decrease the complexity of the peptide mixtures [19]. In addition to that, generally, since membrane proteins are not sufficiently express themselves in crude protein extracts, they do not analyze efficiently with the shotgun method.

Furthermore, in shotgun proteomics, the first step is tryptic digestion to change a large number of proteins into thousands of peptides that should reduce from complex state to fractionated state. It can be achieved by fractionation method which is generally direct loading into an on-line strong-cation exchange (SCX)/ reverse phase (RP) liquid chromatography or fractionation by collecting peptides in a certain pattern using off-line high pH reversed phased liquid chromatography coupled to LC-MS/MS system after concentration of the peptide fragments [20].

In on-line LC method, a digested complex peptide mixture is loaded directly into fused- silica microcapillary column packed with both reversed phase and strong cation exchanged stationary phases that are the multiple dimensions of solid phase packing. Then, additional sample handling does not require. The process goes after elution of the peptides from the column based on their hydrophobicity and charge. Finally, the peptides which is eluate ionized and isolated by their m/z ratio and fragmented selectively by ESI-MS/MS system. This connectivity between SCX column and RP column provides that salts and contaminations are sent away from the RP column while SCX cloumn holds peptide mixture, then peptides go through RP column. Thereby, this multidimensional

protein identification (MudPIT) provides a solution to overcome the challenges in 2D-PAGE separation [21].

On the other hand, in off-line LC system, digested peptides mixture are first loaded into high pH reversed phase column and fractioned with the help of 96 well plates. After elution, peptides are collected with some particular pattern to reduce down the number of 96 fractions that are collected separately and combine them according to the pattern to concentrate and balance peptide level in every fraction. Then collected fractions are evaporated and resuspended with the mobile phase of the second dimension reversed phase HPLC column. Fractioned peptides are intoduced into RP-LC orderly, and then eluates are directly transferred into the ESI-MS/MS system to analyze. This method developed by Yang F. et al., provides more efficient peptide analysis than conventional SCX chromatography applied for 2D proteomic studies since RP-LC resolves peptides better and get higher peak capacities than SCX-LC system. Another advantage of the reversed phase column is that the use of the RP-LC system with salt-free buffer and this ensures cleaner samples for the downstream LC-MS/MS system. 2D RP-RP-LC systems provides orthogonality to the separation of the peptides mixtures comparing to SCX-RP-LC systems. In addition to that, differences between high pH and low pH RP-RP columns generate differences of separation selectivity with the changes in the charge distribution of the peptide fragments upon changing pH of the eluent. Also, combining peptide fragments and concentrate them, provides a further advantage for the proteome coverage in MS/MS analysis [22].

1.3. Peptide MS and MS/MS Spectrum Acquision

Mass spectrometry and tandem mass spectrometry (MS/MS) are the primary tools for the protein identification by producing gas-phase fragmented ions of the digested peptide for the protein of interest. Protein identification is achieved by the interaction between mass spectrometry and ‘gas-phase peptide chemistry’ that involves which bonds are broken, at what rate, cleavages that effect peptide/protein charge state, size, composition, and sequence. In tandem mass spectrometry, the fragmentation of the protein of interest that is induced into gas-phase the peptide ions, is achieved by Collision-induced dissociation (CID), also known as collisionally activated dissociation (CAD), mostly at positive mode. Generally, tandem mass spectrometry generates partial N- and

C- terminal peptides which are related to where the proton localized. For instance, those peptides in which their all protons can be bound to basic residues such as Arg, the cleavage selectively occurs at the C-terminus of Asp or Glu or in some cases that the number of protons exceed the number of Arg residues, then, cleavages occur at the other sites, especially strong cleavage at N-terminus of Pro. Therefore, different ion types can observed in the interpretation of the ion spectra. For that reason, specific nomenclature generates for the different fragment ion types that can be occur with the cleavages of the different bond along the peptide backbone or side chains and interpret in the spectra [23]. The nomenclature shown in Figure 1.6 [24].

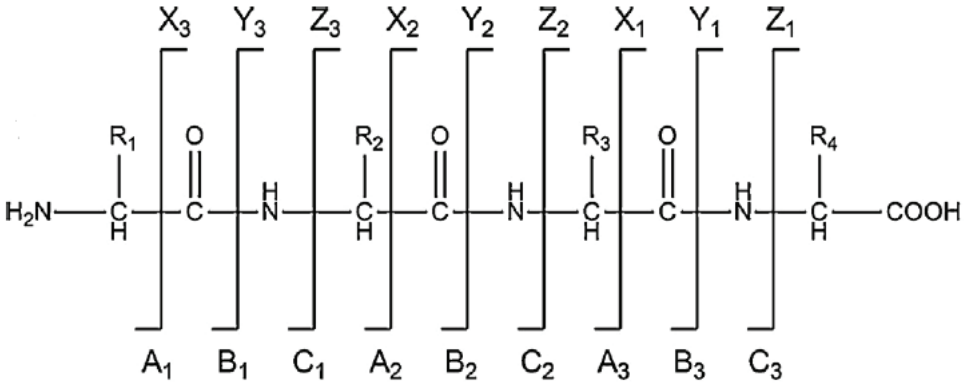


Figure 1.6. Nomenclature for peptide fragment ions in gas-phase
(Source: Chu et al., 2015)

According to this nomenclature illustrated above, N terminus cleavages of the fragmented ions represented by a-, b-, and c- letters and following C-terminus represented by the letters; x-,y-, and z-. Generally, cleavage of backbone occurs at the peptide amide bond and generate b ions if the positive charge remains in the amino terminal or N-terminus. If the charge remains in the carboxy terminal of the peptide or C-terminus, y ions generated. According to that, the most common cleavages occurred in the peptide fragments in CID is a,b, and y ions as it described in the figure above. However, a-ion generated at the lower frequency and they are related to b-ion in such a way that a-b ion pairs can separate from each other by 28u, the mass of the carbonyl group as C=O [25].

Apart from that, since shotgun proteomic is an automated method that combines HPLC and mass spectrometry systems, interpretation of the fragmented peptide ions in

the mass spectrometry system is automated.To do that, for the acquisition of the data in LC-MS/MS studies, data-dependent acquisition (DDA) is the standard method. In DDA method, in each duty cycle, approximately 1s, the instrument cycles through first a short MS survey scan of currently eluting peptides that provides to observe peptide intensity and to identify potential targets to fragment. After that, a series of n, nearly 10-15, MS/MS scan events achieved during a precursor isolated, fragmented and its daughter ions detected. Precursors fragmented in order of decreasing intensity. According to peptide intensities, signal threshold can be programmed to decrease the time of analysis. Finally, the data generated from DDA analysis can be interpreted by specific search engines such as MASCOT.

1.4. Database Search Engine; Mascot

Tandem spectrometers have developed day by day to moving forward about shotgun proteomics to getting closer to the identification of complete proteome. Another significant component to identify proteome is search engines. Interpretation of the data generated from LC-MS/MS system achieved by the one of the most popular search engine, Mascot that is a commercial software package from Matrix Science (www.matrixscience.com). However, this method requires a database that contains the data the associated with amino acid sequences of the proteome of interest. Generally, internet-based sources like NCBI or Uniprot can be utilized for the database requirements. The database of interest can be obtained with genetic code which is translated to the amino acid sequence in FASTA format.

After loading the proper database to the system, there are some particular parameters to be set before the search. All parameters used for the search shown with the interface of the Mascot search engine in Figure 1.7. These parameters for the search involves apart from the selection of the database, enzyme type, the number of allowed miscleavages, fixed and variable modifications, peptide and MS/MS tolerances and peptide charges. Also, data format and ınstrument type can be selected additionally. The entry of the fixed and variable modifications improves the result obtained from the search engine. Also, another important parameter for the data refinement is a decoy search. It is a process that present sequences from databases for the seach have been reversed or

randomized. It just resembles the target database in size amino acid distribution and protein length. Therefore, any exact matches expect to observe from decoy seach.

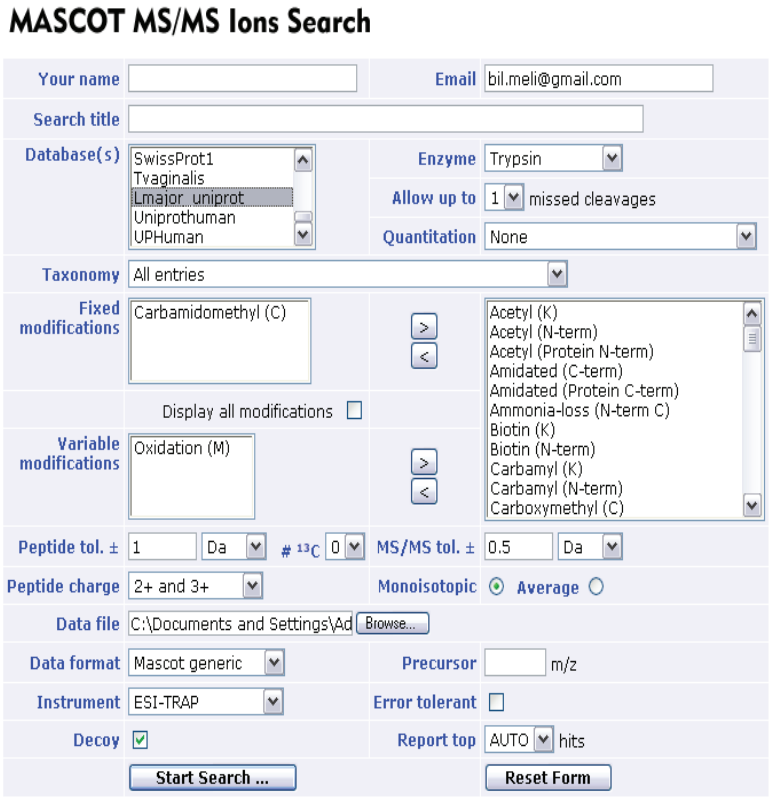


Figure 1.7. Interfaces of Mascot search engine with setting parameters

Using decoy parameter provides self-control of the reliability of the data obtained from Mascot. Generally, 1% decoy match accepted for the proteomic analysis or the lower value can be chosen. On the other hand, Mascot provides a emPAI value that is just a numerical value which is calculated by counting spectrums. However, this value can be utilized in relative quantification with using high-resolution mass spectrometry systems, and also many technical repeats to allow accurate relative quantification.

As similar in the all database search engines, Mascot relies on the principle that calculation of the overall protein score for each protein match. Protein score generally reflects the combination of the entire observed mass spectra matched to any amino acid sequences with the protein of interest. Higher scores defines a more true match between the database of interest and the data of mass spectra. However, the experiments rely on mass spectrometry never produce 'perfect data' for the searching. That is why the analysis of software can generate some random and false matches, or some matches can originate

by chance. For that reason, Mascot always calculates a ‘threshold score’ showing the 5% confidence level in a histogram. Generally, the non-significant area represented in green color.

In a general Mascot search result, possible proteins are ordered with their score and also table shows us to accession number of proteins, their mass, and matches. Matches indicate that the number of MS/MS spectra which were matched with pointed protein. Mostly, a confident detection should be involved two or more matching MS/MS spectra for the protein. Also, matched value can be misleading if the same peptide fragment was searched multiple times. As a result, sofwares such as Mascot aim in recognizing correct matches for the correct peptide but using variable parameters, evoluation of the results depend on personal criticism [26].

CHAPTER 2

Leismania major PARASITE AND LEISHMANIASIS

2.1. A Disease as Leishmaniasis and The genus *Leishmania*

Leishmania spp. is a kind of protozoa parasites which cause a disease called leishmaniasis. This zoonatic infection of the *leishmania* species leads to many human and animals get infected by the bite of a sandfly which belongs to genus *Phlebotomus* and *Lutzomyia*. Therefore, the biting of the sandfly is responsible for the transmisson of the parasite to the living host.

Leishmaniasis is the second most common tropical infections that cause morbidity and mortality and especially in humans, and although this disease shows a significant public health problem with simple tissue scaring cutaneous leishmaniasis (CL) to the lethal visceral leishmaniasis (VL), it does not pay significant attention in many country comparing to other serious diseases such as cancer, AIDS, malaria and tuberculosis. Therefore, it falls under the title as a 'neglected tropical disease'. Because, flies can live under poor conditions such as the places with dust materials and where health conditions are not considered. Generally, third-world countries suffer from such living conditions and cause leishmaniasis in their countries because their economic problems. So, clinically, leishmania disease can present as a three main form: cutaneous, mucocutaneous and visceral. Clinic examples of this there form of leishmaniasis is shown in Figure 2.1.

Mostly, although visceral leishmaniasis causes lethal infections, cutaneous leishmaniasis is not lethal but end up with some severe scar for specific part of the skin and it constitutes the majority of leishmanial infections worldwide [27]. That is why, it can be said that Leishmaniasis is endemic in 98 countries across five continents, Asia, Africa, Europe, North America and South America. Over 350 million people are at risk, with an estimated 12 million infected, and 0.9–1.6 million new cases emerging per year. More than 90% of global visceral leishmaniasis (approximately between 0.2 and 0.4 million cases) occur in six countries: Bangladesh, Brazil, Ethiopia, India, South Sudan, and Sudan.

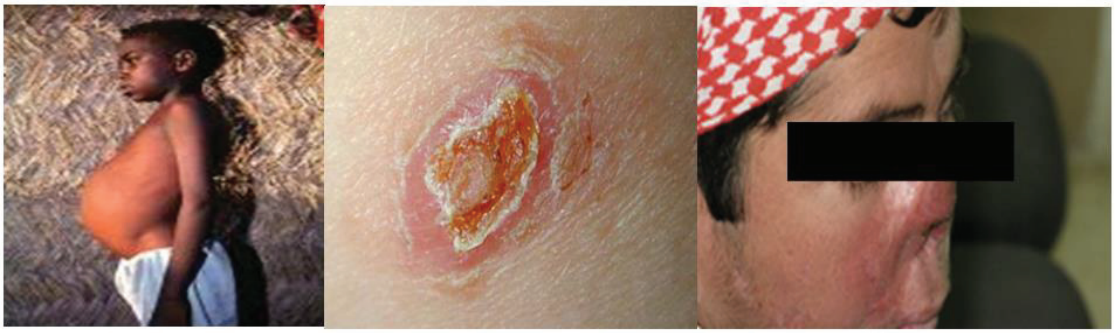


Figure 2.1. Types of the Leishmaniasis in an order of visceral, cutaneous and mucocutaneous

In addition to that over 90% of the cutaneous leishmaniasis (approximately between 0.7 and 1.2 million cases) occur in Afghanistan, Iran, Saudi Arabia, Syria, Brazil and Peru. Within this context, the expressions as a new world and old world type leishmaniasis are used to point out differences in between continents by separating Asia and Africa from Europe and America. Therefore, old-world type leishmaniasis is endemic in Mediterranean and Asia region and is caused by the subgenus leishmania species of *L. tropica, L. major, L. aethiopica, L. mexicana*. In Table 2.1, a subgenus of *leishmania* species which cause a different type of leishmaniasis is illustrated as a summary. This clinical forms of the disease is named according to the body location in where the parasite is proliferated. As it seen in Table 2.1, a different form of leishmaniasis can be related with more than one specie of *leishmania* spp. While *Leishmania tropica, Leishmania major, Leismania aethiopica, Leismania mexicana* are the causative agent for cutaneous leishmaniasis, *Leishmania donovani, Leishmania infantum, Leishmania chagasi* are for visceral leishmaniasis and *Leishmania braziliensis* is for mucocutanaous leishmaniasis.

Several species of leishmania cause cutaneous leishmaniasis and common one of them is *Leishmania major*. It is a common type of leishmaniasis among them and has a short incubation period that is nearly 1 to 4 week, often longer than two months. The lesions appear on the extremities and face as red furunculoid nodule. Also, it can be described as looking somewhat like a volcano with a raised edge and central crater. This lesions may present more than 30 sometimes. Healing can take months or even years but lesions can leave significant scars, especially, it can be disfiguring if they occur on the face. However, visceral leishmaniasis is the most dangerous and lethal form of them. The

incubation period is way more than cutaneous leishmaniasis, approximately 1 to 36 months and usually associated with fever, weight loss, diarrhea, abdominal tenderless and etc. Because it occurs on visceral organs, the most common symptom is enlarging spleen and liver.

Table 2.1. Illustration of the subgenus of *Leishmania* species and their infections caused different Leishmaniasis

***Leishmania* spp.**	**Type of disease**
Leishmania tropica *Leishmania major* *Leismania aethiopica* *Leismania mexicana*	Cutaneous Leishmaniasis
Leishmania braziliensis	Mucocutaneous Leishmaniasis
Leishmania donovani *Leishmania infantum* *Leishmania chagasi*	Visceral Leishmaniasis

This progress can cause other complications such as pneumonia, nephritis, and tuberculosis that can lead to lethal results. On the other hand, in mucocutaneous leishmaniasis, lesions generally spread on a face, especially the noise or month. Also, it may occur months to years after original skin lesion and hard to confirm the diagnosis. As it seen, a different subgenus of *leishmania* can affect alternatively in the living body and it can be originating from immune response of the host. Because leishmaniasis is complicated disease, attemps have still been investigated for treatment. It also has been reported that *leishmania* has resistance against some conventional compounds. It can be said that, the same type of the disease can indicate alternative clinical feature because of possible genetic alternation between *leishmania* species. For the cutaneous and mucocutaneous leishmaniasis, examination of the lesions is easier than visceral leishmaniasis. The diagnosis is based on the detection of the parasite in the stained smear

that is obtained from cutaneous lesions, or spleen by using biopsy with using Giemsa, Wright or Feulgen stain for demonstration of parasite on the tissues. In the treatment of the cutaneous and mucocutaneous type leishmaniasis, non-parenteral therapy or a short course of antimony is used with the combination of antimony allopurinol. If the oral therapy is preffered, then fluconazole is used for 6 weeks. In addition to that, Paromomycin shows significant activity against leishmania species. For the treatment of visceral leishmaniasis, generally, Liposomal amphotericin B shows powerful treatment in many regions of the world [28].

2.2. Morphology and Life Cycle of *Leishmania* spp.

Leishmania is a kind of species that is called a flagellated protozoan parasite with a complex life cycle and involved to a class of kinetoplastida which is a significant group of obligatory parasitic protists belonging to phylum euglenozoa. The organisms in the class of kinetoplastida have a unique organelle that is called as kinetoplast which locate in their mitochondrion. Kinetoplastids are involved in the class of eukaryotes and they are the earliest eukaryotes that those of animals, plants, and even fungi. Therefore, it can be considered that their kinetoplasts are the only distinctive property of them all. In addition to that, kinetoplastids have been divided into two monophyletic groups as the biflagellate bodonids and uniflagellate trypanosomatids according to a number of fragella. This parasite species that involved in the class of kinetoplastida, changes their morphologies during their life cycle adapt the new surroundings in the living host [29]. According to these knowledges, *Leishmania* species exist in two primary body forms as the amastigote and the promastigote. In the promastigote from, the parasite is extracellular and motile. They grow and divides by longitudinal binary fission at 27 ºC in the sandfly with an anterior flagellum. However, in the amastigote form, the parasite is intracellular, non-motile form in the mammalian host without its flagellum. They grows and divides by longitudinal binary fission at 37 ºC. Amastigote form of the parasite is not devoid of a flagellum. It is merely that the flagellum does not protrude beyond the body surface of the parasite.

On the other hand, the life cycle of *leishmania* is summarized in the given Figure 1.7. Firstly, leishmaniasis is transmitted by the bite of the sandfly that is called as phlebotomine. Then, during blood meal of sandfly, infective stage, promastigote form of

the parasite, is injected into the host body. Generally, the parasite gets macrophages under control in host body or also, they can attack different tissues such as neutrophil, dendritic and fibroblasts. The promastigote form are phagocytized by macrophages and alteration of the amastigote form begins. During this stage, the flagellum of the parasite encloses and a vacuole surround the parasite. Then, After that, amastigote form multiplies itself in the infected cells by using vacuole and spread out different tissues in the cytoplasm till it reaches the extracellular space and blood to infect related organs by depending on the subgenus of leishmania.

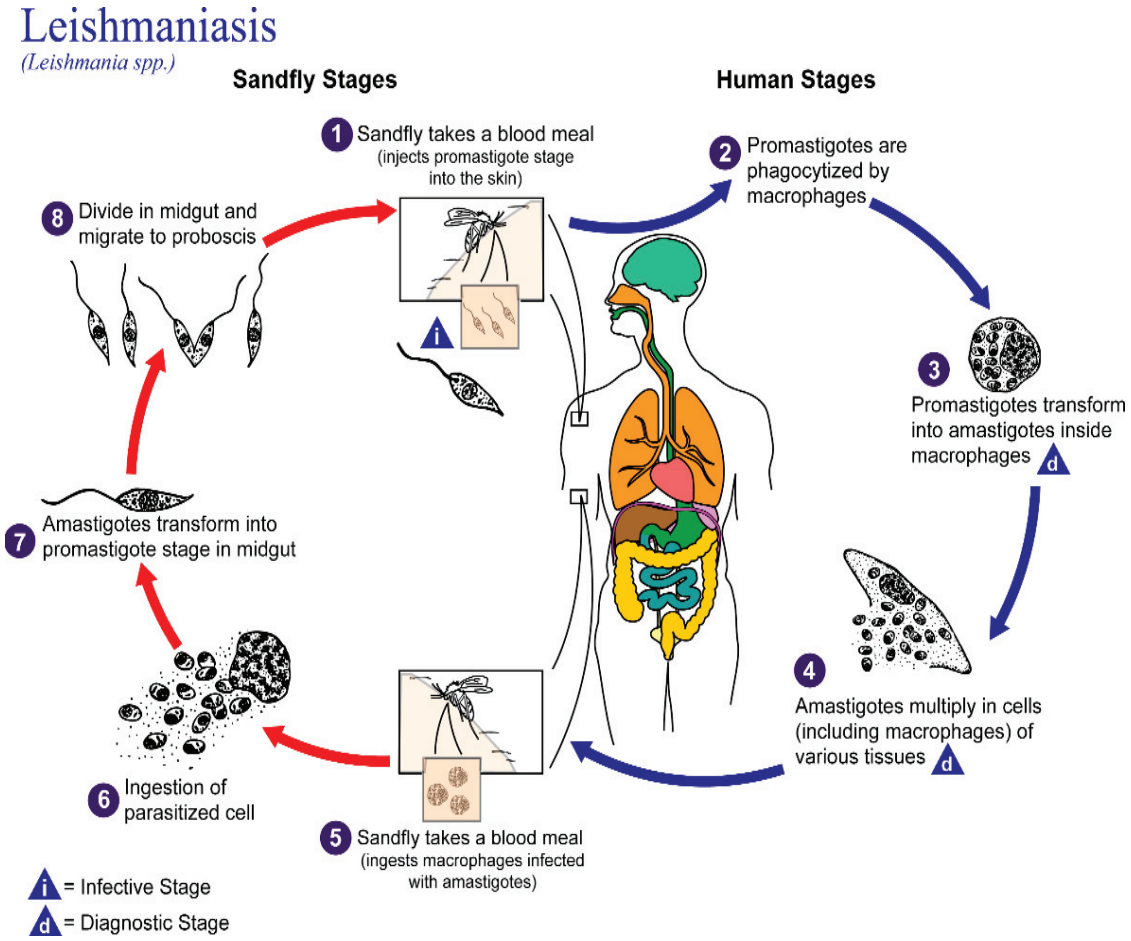


Figure 2.2. Life cycle of *Leishmania* spp.
(Source: CDC/Alexander and Moser, 2005)

As the amastigote form is transmitted into the bloodstream in the host body, it can be transmitted to another sandfly that bites the host. Thereby, the amastigote form of the parasite infects the sandfly and goes through its midgut.

Parasite transforms itself to promastigote form in there to adapt new environment and then, they adhere in midgut epithelium by their flagellum. Finally, this cycle circulates with the moment that sandfly bites another host and infect them [30].

2.3. The Virulence Effect of *Leishmania major* and Aim of the Study

The protozoan parasite *Leishmania major* that is belonging to the subgenus of leishmania, is a very vital causative agent for the cutaneous leishmaniasis. *L. major* has a wide distribution from Africa to Asia. However, not only these reagons infected, but also all old-world type countries infected from *L. major* by transmission of the parasite with female sandflies. Therefore, by looking at this enlargement of the *L. major* infection area, it can be considered that the parasite can find a way to escape from immune respond of the host and leading to surviving its existence in the mammalian host long times. This situation causes that involving possible genetic variances of *L. major* and gets drug resistance for the treatment by upregulating its virulence effect in the mammalian host. Virulence effect of the parasite can be defined as an ability to avoid host defense system and easily damage host by the parasite. For that reason, many attemps for the investigation of virulence effect of *L. major* have been studied. For example, Bifeld et al., have studied geographical sequence variation in the *L. major* and its effects on its virulence factor on the genetic basis. Their results showed that P46 gene that belongs to *L. major* contributes an effect on the virulence factor of *L. major* and infectivity of this parasite both in vivo and in vitro. Thereby, they conclude that this result raises the possibility of P46 gene may be involved in host adaption of *L. major* [31].

Recently, The cases for the cutaneous leishmaniasis by *L. major* have been reported to begin increasing in Iraq, Iran, and Syria. Because of the location of Turkey, cutaneous leishmaniasis has become a significant health problem in the Southeastern Anatolia and Mediterranean regions and also several cases have been reported in the Central Anatolia and Aegean regions. Thereby, A. Özbilgin et at., have investigated the cases that is collected from a total of 18 autochthonous due to the cutaneous leishmaniasis by *L. major* and this was the first study about *L. major* isolating from clinical samples. 11 of 18 patients had only one lesion, mostly on the face but rest of them had at least 3 lesions on the different part of their body.

As a result, findings showed that collected samples from a different patient and region of Turkey indicate the possibility of alternative virulence effect of *L. major* [32].

According to all reason mentioned above, in this study, we aimed to characterize the proteome analysis of *Leishmania major* parasite in two clinical case exhibiting different virulence effect on leishmaniasis disease by using mass spectrometric shotgun method. In addition to that, our results would make a contribution to clarifying differential proteins that cause a change in virulence, later can be improved a proper treatment for *L. major* parasite form our findings and it can help to get a piece of information about proteome profiling of *L. major*.

CHAPTER 3

MATERIALS AND METHODS

3.1. Total Protein Extraction of *Leishmania Major* and Parasite Growth

In this study, All *Leishmania Major* parasite samples were cultivated and extracted by Prof. Dr. Ahmet Özbilgin's group in Celal Bayar University, Faculty of Medicine, Department of Parasitology.

Protein extracts from *L. major* samples were categorized as the one has aggressive virulence effect, passive virulence effect and a reference sample from strain, and they were named as 470, 575 and RF2 respectively.

In the growth and extraction method, firstly, frozen promastigote state *L. major* samples stored in liquid nitrogen were dissolved in a water bath in 37 ºC and allowed to cultivate in NNN (Novy-MCNeal-Nicolle) medium in 26 ºC and then transferred to 5ml RPMI 1640 that also includes 15% FCS for incubation in 26 ºC. When promastigote state reaches up to logarithmic phase, 200 µl of promastigote sample was used to the isolation of DNA by using ROCHE High Pure PCR Template Preparation Kit. After using leishmania ITS1 gene region to subtype separation between *L. tropica*, *L. major* and *L. donovani* complex with RT-PCR, the samples that belong to *L. major* species was prepared to protein isolation in 50 ml falcon tubes. Samples were centrifuged at 4400 rpm for 10 min. at 4 ºC. After centrifugation, 10 ml PBS (Phospahate-buffered saline) was added to rest of pellet and homogenized with mixing. Then again, they were centrifuged at 4400 rpm for 10 min at 4 ºC for 3 times. Finally, 1 ml of Mammalian Protein Extraction Reagent kit (M-PER®, Thermo, USA) was added and incubated in 10 min at 25 ºC. Then, proteins isolation were completed with collecting upper phase after centrifugation at 15000 rpm in 15 min at 4 ºC.

3.2. Total Protein Quantification and Purification

Firstly, acetone precipitation process was applied to all samples in order to concentrate total protein and exchange cell lysis reagent with resuspension buffer which include 7M Urea, 2M Thiourea and 0.1M Tris-HCl at Ph 7-8. This procedure provides samples to clean up from the supernatant and concentrate protein level, also buffer exchange prepare the media for the HPLC separation.

In acetone precipitation method, 4 volume of the cold-iced acetone was added to 1 volume all samples and incubated in –20 ˚C during 4 hours. After incubation, samples were centrifuged at 14000 rpm at 4 ˚C for during 15 min and precipitated proteins at the bottom of the tube was pulled out from the supernatant and dissolved with resuspension buffer. Bradford protein assay was performed to determine the concentration of the total protein level for each sample. In this method, protein molecules binds to Coomassie dye under acidic conditions results in a color change from brown to blue. It actually measures the presence of the basic amino acid residues, arginine, lysine and histidine, which contributes to formation of the protein-dye complex.

BSA standard with range of 25 to 1000 µg protein was prepared to apply Bradford assay. After that 240 µl Bradford reagent was added to first column of the 96 well plate (Granier bio-one flat bottom) and then 5µl BSA standard and 5µl suspension buffer were added to make calibration curve for the quantification for unknown protein samples. Then, the second column was filled with 5µl unknown, 240 µl Bradford reagent and 5µl ultrapure water to balance with standards. The change in blue color was measured with the spectrophotometer. At the end, sample concentration was determined and set with 400µg for each sample.

3.3. Protein In-Solution Digestion Protocol

This protocol provides proteolytic digestion of the total protein extracted from samples to peptides. Generally, in-solution digestion process include trypsin as immobilized protease and digest the proteins into peptides. For digestion, followed protocol shown in Table 3.1 was applied. In addition to that, filter-aided sample preparation method was used to remove excess urea, salts and some small contaminants to the medium with a 10K molecular weight cut-off (MWCO) filters [33].

Firstly, 400 µg total protein in 250 µl resuspension buffer was placed in 10K molecular weight cut-off filter and all digestion process was occurred in there.

Table 3.1. In-solution Digestion protocol and enzymatic digestion step

5 µl of 0.2M Dithiothreitol (DTT) as a reducing agent added and waited 1h.	prepared in 50mM Tris-HCl; pH 8,5
20 µl of 0.2M 2-Iodoacetamide (IAA) as an alkylating agent added and waited 1h.	prepared in 50mM Tris-HCl; pH 8,5 (sensitive to light)
20 µl of 0.2M DTT added to reduce excess IAA and waited 1h.	End of the alkylation and reducing process.
Samples were centrifuged and washed 2 times with 200µl with 50mM Tris-HCl ; pH 8,5	at 14000 rpm and room temperature for 10 min. (Last wash step is centrifuged until desired volume was obtained.)
Trypsin was added as 1:80 ratio, 5µg enzyme and 400 µg sample. (trypsin concentration: 0,2µg/µl)	25µl trypsin +200µl Tris-HCl+ 20µl acetonitrile was added to sample as total and kept at 37 °C overnight.

3.4. Reversed-Phase Chromatography For Separation and Off-line HPLC Fractionation

After digestion, all proteins turned into very complex peptide mixture that was hard to separate. At this point, the off-line HPLC fractionation method provided pre-separation process to reduce the complexity that could be transmitted to LC-MS/MS system due to the high amount of peptide mixture for each sample. For fractionation, HPLC system, SHIMADZU Prominence UFLC (LC version 1.25 SP3) was used with high pH reversed phase C18 column, Teknokroma Mediterraneasea18 (25 cm x0.46 cm x5 µm) and separation were performed by offline-line mode combine with collector, (Sunchrom Micro Fraction Collector).

This approach improves proteome coverage and provides efficient salt-free or low salt separation when it is compared to the low pH columns. This superiority coming from using high pH column is related to the change in charge distribution in peptide

chains over the change in pH of the eluent. Also, using a reversed phase (RP) column in both off-line HPLC and LC-MS/MS systems provides a good strategy for better separation orthogonality comparing to SCX-RP systems [34].

The method that was used for pre-separation is shown in the Table 3.2. For separation, mobile phases were used ;

- Mobile phase A as ultrapure water in 10mM ammonium formate.
- Mobile phase B as 90% acetonitrile in 10mM ammonium formate.

Table 3.2. The method for Off-line HPLC separation and parameters

<table>
<tr><th colspan="2">LC program</th><th>Parameters</th></tr>
<tr><td>Time (min.)</td><td>Mobile Phase B</td><td rowspan="2">Ph of the both mobile phase was adjusted 10.</td></tr>
<tr><td>5.00</td><td>0%</td></tr>
<tr><td>5.10</td><td>5%</td><td rowspan="2">Injection volume was 100 µl.</td></tr>
<tr><td>40.00</td><td>25%</td></tr>
<tr><td>47.00</td><td>30%</td><td rowspan="2">Column temperature was adjusted to 40 C.</td></tr>
<tr><td>54.00</td><td>40%</td></tr>
<tr><td>61.00</td><td>80%</td><td rowspan="2">Flow was set to 0.5 ml/min.</td></tr>
<tr><td>70.00</td><td>80%</td></tr>
<tr><td>85.00</td><td>0%</td><td>Duration time for fractionation was 50min.</td></tr>
</table>

During analysis, the collector was used to collecting each fraction in 96 well plates. Because of handling with that much of fraction, they were decreased to 12 fraction by placing into a new 2ml tube. To do that, fractions were collected by skipping 11 of them each time. The scheme for the collection process was shown in Figure 3.1. Then, all 12 fractions was evaporated. After the evaporation process was done, µ-C18 ZipTip

protocol (Millipore) was used to peptide desalting. However, it must be known that samples must be in an aqueous buffer before this step.

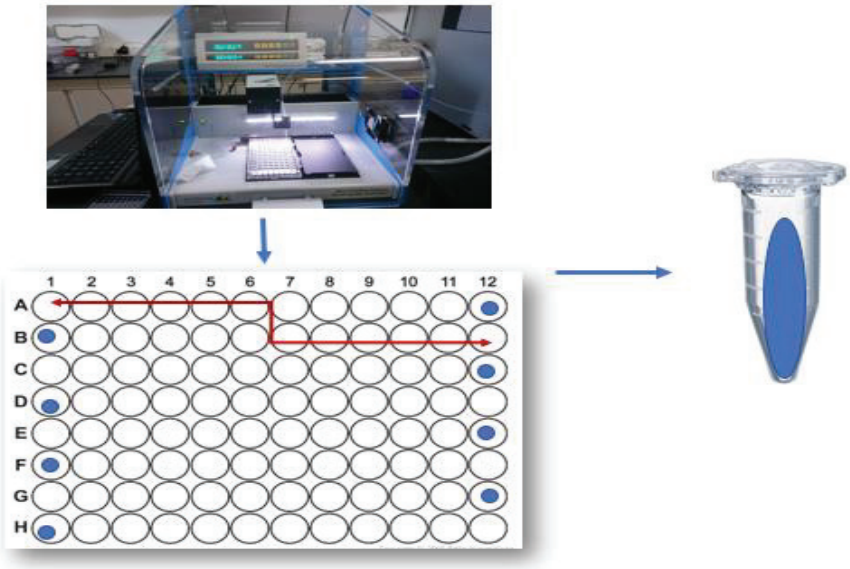


Figure 3.1. Collection and fractionation process

Therefore, 2% acetonitrile, 0.1% formic acid buffer was prepared in ultrapure water and added 20 µl to the all evaporated samples. This step also provided to keep the column of LC-MS/MS system from any clogging that could originate from salts. In addtion to that all 12 fractions for each sample was intoduced to LC-MS/MS system freshly desalted. The method that was used in both LC and MS system was shown in Table 3.3. Other than that, for the second separation associated with MS system, the column, Sigma Supelco Ascentis (15cm x 500 µm ; 2.7 µm) of the HPLC, DIONEX Ultimate 3000 (Chromeleon Version 6.80) was used as reversed phased that was mentioned above at low Ph. Also, mobile phases that used for LC analysis was ;

- Mobile phase A: Ultrapure water in 0.1% Formic Acid
- Mobile phase B: Acetonitrile in 0.1% Formic Acid

Fractionated peptides were separated with gradient elution in 50 min with HPLC system and MS profiling spectra were obtained on LTQ XL ESI Ion Trap Mass Spectrometer, Thermo Scientific. To obtain data from 12 fractions of each L. *major* sample data-depending settings were adjusted before run and every fraction was analyzed two times to get a technical repeat. From the data obtained from MS for each 12 fraction

for one sample was converted to Mascot Genetic Format (mgf) by using Proteome Discoverer 1.4 and then merged to get single data that was uploaded to Mascot server, Version 2.3.

Table 3.3. LC-MS/MS methods and Data-depending parameters

LC Gradient Elution Method		**MS Data-Dependent Setting Parameters**	
Time(min)	***Mobile Phase B***	**Scan events (SE)**	15
0.00	2%	**MS Scan Range**	400-800
4.00	2%	**Activation Tipe**	CID
5.00	5%	**Default Charge State**	2
6.00	8%	**Isolaton width**	2.0
36.00	22%	**Activation Time Q (ms)**	30.00
39.00	35%	**Repeat Count and Duration**	1-30
42.00	60%	**Exclusion List and Duration**	500-30
45.00	90%	**Signal Treshold (counts)**	5000
50.00	90%	**Normalized Collison Energy**	35
55.00	2%	**Data Type**	Centroid

3.5. Mascot Search and Parameters

Database search for the all merged data coming from LC-MS/MS analysis were performed by Mascot server (version 2.3) that is installed and used in laboratory conditions. The seach parameters for each seach were set as shown in the following Table 3.4. Apart from these parameters, missed miscleavages that originate from enzymatic digestion were set as one miss cleavages and experimental mass values were selected as monoisotopic. In addition to that, decoy database was selected for all search to ensure the data that is finding out is reliable. It basically is a process that present sequences from databases for the seach have been reversed or randomized. Therefore, any true matches must be expected from decoy seach.

On the other hand, a database for the seached data was obtained from Uniprot website (www.uniprot.org) by using taxonomy section associated to *L. major*. It was installed in FASTA format and 32.129 entry was found that was belonged to *L. major*.

Table 3.4. Search Parameters for Mascot Server

Database	L. *major* Uniprot
Enzyme used in digestion	Tyrpsin
Taxonomy	All entries
Fixed Modifications	Carbamidomethyl (C)
Variable Modifications	Oxidation (M)
Peptide Tolerance	1 Da
MS/MS Tolerance	0.5 Da
Peptide Charge	2+ and 3+
Data Format	mgf
Instrument	ESI-TRAP

Finally, common and differential proteins between there *L. major* sample which is aggressive, passive and reference was discussed by transferring all data to excel sheet to see wide differences or association between them. Then, Search result were illustrated by using a Venn diagram program that is generated by VIB from Gent University.

CHAPTER 4

RESULTS AND DISCUSSION

In this study, our purpose was to characterize the proteome analysis of *Leishmania major* parasite in two clinical cases exhibiting different virulence effect on leishmaniasis disease by using mass spectrometric shotgun method. For obtaining reliable results, LC-MS/MS system collaborated with Mascot Server. At this point, the parameters that were preferred for searches was crucial. Therefore, the importance of the parameters was evaluated on selected sample. In addition to that, for the evaluation of the results, false discovery rate value (FDR) of both identity threshold and homology threshold was kept under 1% for each sample. FDR values are a really important detail for target-decoy searches on Mascot because it is a vital parameter to see the ratio of the false positive matches to the total of both false positive and true positive matches in the target database. Therefore it must be kept as low as possible to obtain a reliable scoring scheme. Also, emPAI values given in the mascot results aimed to provide getting relative quantitation between the common proteins that was found in all samples. However, results from this emPAI values did not give a valuable result to discuss.

On the other hand, after results evaluated, some parameters were changed and filtered to get FPR value under 1%. Therefore, p-value for the confidence level was set to 0.02 to deal with lower FDR value and kept list more accurate and cut off filter was used to eliminate the matches that are under 20.

As a result, all aggressive, passive and reference *L. major* samples were analyzed twice as being in two group of set and consequential results were determined. The findings were listed as a comparison of differential proteins between aggressive and passive samples associated with reference and common proteins also listed in Table 4.1 and 4.2 respectively. Also, Venn diagram results were given in the following Figure 4.1 for both set of *L. major* samples. According to results, in the first batch of the *L. major* group, 382 protein for aggressive, 736 protein for passive and 703 protein for reference samples were identified. Also, Venn results in the given figure showed that aggressive *L. major* sample has only 27 different protein than other passive and reference samples and passive *L .major* has 155 different protein from all.

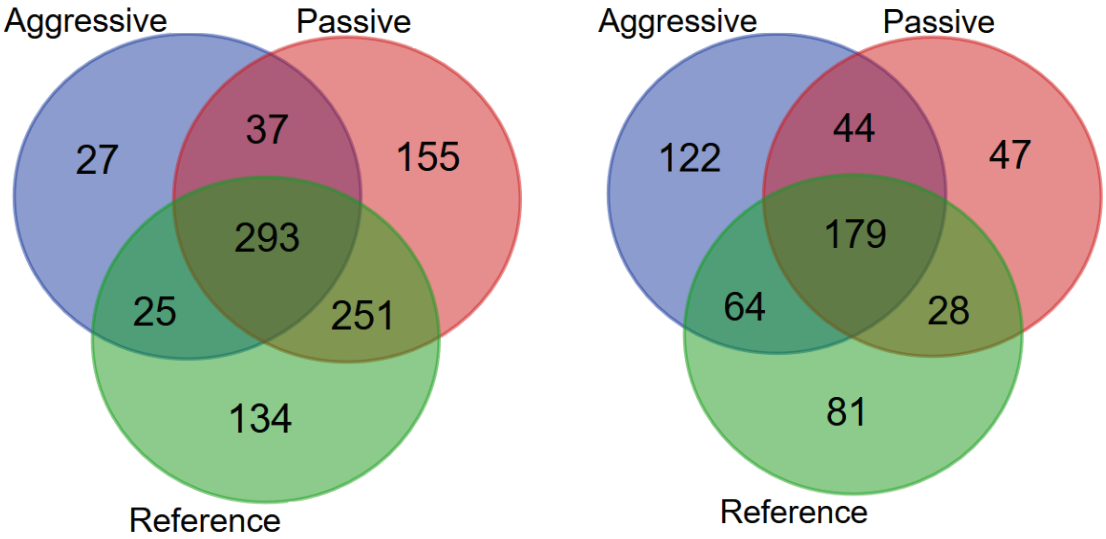


Figure 4.1. Venn results showed in both batch 1 and batch 2

Table 4.1. Differential proteins aggressive between passive samples in batch 1

Protein Name for Aggressive *L. major*	**Uniprot code**	**Score**	**Matches**
Putative phosphoglucomutase	Q4QCF1	109	3
Nucleosome assembly protein-like protein	Q4Q687	79	2
UDP-galactopyranose mutase	Q5EEK0	60	2
Putative NADP-dependent alcohol dehydrogenase	Q4QBD8	59	1
Putative small GTP-binding protein Rab7	Q4QH97	58	1
Receptor-type adenylate cyclase b	Q4QEH9	48	3
Adenosine monophosphate deaminase-like protein	Q4Q520	47	1
Hydrophilic acylated surface protein b	Q4QB56	47	3
Putative pumilio protein 6	Q4Q475	44	1
Putative CAS/CSE/importin domain protein	Q4Q6V0	43	2
Putative glycoprotein 96-92	Q4Q843	43	1
Obg-like ATPase 1	E9ACV1	39	1
Putative rrp44p homologue	Q4Q8N7	35	2
Coatomer subunit epsilon	Q4Q5A7	33	1
Ribonuclease mar1	Q4QGT7	33	1
Putative phosphomevalonate kinase protein OS	Q4QF34	32	1
Putative ATP-binding cassette protein subfamily E,member 1	Q4QCE4	27	1
Topoisomerase-related function protein-like protein	Q4QIM1	27	1
Surface antigen-like protein	Q9XZY1	26	2
Putative uracil phosphoribosyltransferase	Q4Q3A1	25	1
Palmitoyltransferase	Q4QG89	22	1

(cont. on next page)

Table 4.1. (cont.)

Protein Name for Passive *L. major*	**Uniprot code**	**Score**	**Matches**
Putative aspartate--ammonia ligase	Q4Q9B1	171	6
Putative surface antigen protein	Q4QGL8	99	7
Putative rab1 small GTP-binding protein	Q4QH98	57	4
Aromatic amino acid hydroxylase-like	Q6WRI4	111	5
Putative T-complex protein 1, beta subunit	E9ADA4	93	5
Excreted/secreted protein 27	Q1X7K5	85	1
Putative short chain 3-hydroxyacyl-CoA dehydrogenase	Q4Q1W2	82	4
Pyridoxal phosphate homeostasis protein	Q4QAZ0	79	3
Amastin-like surface protein-like protein	Q4QAK4	79	2
Putative proteasome beta 2 subunit	Q4Q246	73	2
Putative N-ethylmaleimide-sensitive factor	Q4QCW5	38	3
Putative ras-related protein rab-5	Q4QDS9	71	2
Putative RNA binding protein rbp16	Q4Q8I6	71	2
3-hydroxyisobutyryl-CoA hydrolase, mitochondrial	Q4Q4Q5	69	1
Putative folate/biopterin transporter	Q4QHI1	69	4
Putative paraflagellar rod component par4	Q4QJJ9	68	2
Flavoprotein subunit-like protein	Q4QIK9	66	4
Eukaryotic translation initiation factor 5A	Q4QA21	66	4
Putative short chain dehydrogenase	E9AEU6	66	1
Putative cytochrome c oxidase VII	Q4Q9Y0	64	2
Vacuolar protein sorting-associated protein 4	E9AEB2	64	1
Putative methionine-S-sulfoxide reductase	Q4QIH2	64	1
Putative argininosuccinate synthase	Q4QBE8	59	2
Oligosaccharyl transferase-like protein	E9AET7	58	3
Hs1vu complex proteolytic subunit-like	Q4Q116	55	3
Putative ribosomal protein L38	Q4Q8X0	55	2
Putative hydroxyacylglutathione hydrolase	Q4QGS1	55	1
Putative acidocalcisomal exopolyphosphatase	E9AC20	54	1
Putative tryptophanyl-tRNA synthetase	E9ADK8	54	1
Isopentenyl-pyrophosphate isomerase	Q5QQ43	54	1
Putative 3'-nucleotidase/nuclease	Q4QGQ3	54	2
Putative aspartyl-tRNA synthetase	Q4Q7R2	52	3
Prefoldin subunit 3	Q4Q956	52	2
Stomatin-like protein	Q4QJ98	52	1
Phosphatase-like protein	Q4QCN8	51	2
Putative proteasome regulatory non-ATP-ase subunit 3	E9ADC7	49	2
Putative ubiquitin carrier protein	Q4QIK2	49	1
Putative carboxylase	E9ABZ4	49	1
2-methoxy-6-polyprenyl-1,4-benzoquinol methylase	Q4Q0M6	49	1
Glucose transporter, lmgt2	Q4Q0D1	49	2
Adenine phosphoribosyltransferase	Q4Q9H8	48	1
Putative paraflagellar rod component	Q4QHP3	47	2
Farnesyl pyrophosphate synthase	Q4QBL1	46	2
Putative cytochrome b-domain protein	E9AC79	44	2
Diphosphomevalonate decarboxylase	Q4QE40	43	1
Pteridine reductase 1	Q01782	43	1
Putative QA-SNARE protein	E9AF95	43	1
Putative glycogen synthase kinase	Q4QE15	42	2
Putative cyclophilin 12	E9AC11	42	1
Putative nucleoside transporter 1	Q4QF58	42	1
Putative iron-sulfur cluster assembly protein	E9AFI6	42	1
Putative elongation initiation factor 2 alpha subunit	E9ACP3	41	2
Inosine-guanosine transporter	Q4Q1M9	41	5

(cont. on next page)

Table 4.1. (cont.)

Protein Name for Passive *L. major*	**Uniprot code**	**Score**	**Matches**
Putative ATP-binding cassette protein subfamily C,member 6	Q4Q6D4	41	1
Elongation factor Ts, mitochondrial	Q4FXY8	40	1
Putative phosphoglycerate mutase family member 5	Q4Q108	40	1
Putative eukaryotic translation initiation factor 5	Q4Q3H3	40	1
Putative X-pro, dipeptidyl-peptidase,serine peptidase, Clan SC, family S15	Q4Q871	39	1
Methionine aminopeptidase	Q4QDG7	38	1
Profilin	Q4Q5N1	38	3
Cation-transporting ATPase	Q4QII2	38	1
Sphingosine-1-phosphate lyase	Q4Q758	36	1
Putative 3-hydroxyacyl-ACP dehydratase	Q4QIQ0	36	1
Glyceraldehyde-3-phosphate dehydrogenase-like protein	E9AFV0	36	3
Nucleobase transporter	Q4QG33	36	1
Putative signal recognition particle protein	Q4QDE2	35	1
Putative electron transfer flavoprotein-ubiquinone oxidoreductase	Q4QIN2	35	1
V-type proton ATPase subunit F	Q4QGP0	35	1
Alkyl dihydroxyacetonephosphate synthase	Q7YWB6	35	3
Putative glycosomal membrane protein	Q4Q838	35	2
Coatomer subunit beta	Q4QCR1	35	3
Sucrose-phosphate synthase-like protein	Q4QES5	34	1
Transcription factor-like protein	E9ADK4	34	1
Oxidoreductase-like protein	Q4QG82	34	1
Acyl carrier protein	E9AD06	34	1
Polyprenol reductase	Q4Q9R2	33	4
2-oxoglutarate dehydrogenase, e3 component,lipoamidedehydrogenase-like protein	Q4Q5Z7	33	1
Receptor-type adenylate cyclase a	Q4QEI3	33	2
Clathrin light chain	Q4QDL0	33	1
Protein kinase A catalytic subunit	Q27687	32	1
Putative threonine synthase	Q4QFV0	32	1
Putative branched-chain amino acid aminotransferase	E9ADI3	32	2
Putative leucine-rich repeat protein	Q4QHX1	30	1
Phosphomannomutase-like protein	Q4Q2G5	30	1
Protein disulfide isomerase	Q4QIX1	30	1
Methylthioribose-1-phosphate isomerase	Q4Q0R9	30	1
Aldose 1-epimerase-like protein	Q4QBD1	30	1
Amastin-like protein	Q4Q3F7	30	1
Dynein intermediate-chain-like protein	Q4QAV3	29	1
Cytochrome c oxidase assembly factor-like protein	Q4QCW2	29	1
Putative DNAJ domain protein	Q4QAS9	29	1
Diphthine synthase-like protein	Q4Q6A1	29	1
Cystathionine beta-lyase-like protein	Q4QFT8	29	1
Putative intraflagellar transport (IFT) protein	Q4Q794	29	1
Elongation factor G, mitochondrial	Q4Q219	28	1
Putative DNA-directed rna polymerase I largest subunit	Q4QEN2	28	1
Putative kinesin K39	Q4QFM4	28	2
V-type proton ATPase subunit a	Q4QAY7	28	2
Putative Ran-binding protein	Q4QB12	27	1
Putative centrin	Q4Q2W1	27	1

(cont. on next page)

Table 4.1. (cont.)

Protein Name for Passive *L. major*	**Uniprot code**	**Score**	**Matches**
Phosphatidic acid phosphatase protein-like protein	Q4QD76	26	1
Putative ATP-binding cassette protein subfamily A,member 8	E9AD74	25	1
Protein SEY1 homolog	Q4Q5P8	25	1
Acidocalcisomal pyrophosphatase	Q4QH59	25	1
Putative ATP-binding cassette protein subfamily	Q4QBD6	24	1
CTP synthase	Q4QCY4	24	1
Putative mitochondrial oligo_U binding protein TBRGG1	Q4Q7M7	23	1
Glucokinase	Q4Q1I9	22	1
Putative prostaglandin f synthase	Q4Q5N9	22	1
Putative major surface protease gp63	Q4Q8L3	22	2
Putative AMP deaminase	E9AFV5	21	1

On the contrary, in the second batch of the *L. major* group, 409 protein for aggressive, 298 protein for passive and 352 protein for reference sample were identified. Also, Venn results showed in Figure 3.1 indicate that aggressive *L. major* sample has only 122 different protein than other passive and reference samples and passive *L. major* has 47 different protein from all. Tables for the illustration of the differential proteins for the first batch of aggressive and passive samples are shown in Table 4.2. According to the second batch results, the number of total protein for each sample was decreased compared to the first batch. Apart from that, the number of specific proteins that were found in both aggressive and passive are increased proportionally. This result can originates from any process before the samples introduce into LC-MS/MS system.

In the first batch, for the aggressive sample, 27 protein of the total 382 protein were found different than both passive sample and reference. Also, 100 uncharacterized proteins were detected for the aggressive sample (not listed). Among these 27 identified proteins, majors ones are predominantly involved in the catalytic activity as a molecular function and they are localized in the cell and membrane part of the parasite. For the passive sample, 155 protein of the total 736 protein were found as unique for the sample. Also, like the aggressive sample, 220 uncharacterized proteins were found in the mascot list and they are not included in the assessment since their functionalities have not been investigated clearly, yet. Conversely, this 155 protein that were found as differential, mostly, are involved in the catalytic activity of the sample and unlike aggressive sample, they are distributed in organelles, protein-containing complexes and mitochondrial matrixes with cell and membrane part of the parasite. These findings were repeated with

the new batch of all samples to get a reliability to the results. According to that, the results for the differential proteins of the second batch is shown in Table 4.2.

Table 4.2. Differential proteins aggressive between passive samples in batch 2

Protein Name for Aggressive *L. major*	**Uniprot code**	**Score**	**Matches**
Putative 40S ribosomal protein S21	Q4QH01	213	8
Probable citrate synthase, mitochondrial	Q4QDX3	139	9
Tubulin-specific chaperone A	Q9U1D9	126	4
Macrophage migration inhibitory factor-like protein	Q4Q412	113	2
Putative aminopeptidase P	E9AF59	97	6
Putative cytochrome c oxidase subunit V	Q4Q922	96	5
Putative Pyridoxal kinase	Q4Q7H9	82	6
Putative glutamine synthetase	Q4QJ42	82	3
Cystathionine beta-synthase	Q4QEG9	78	2
Inosine-5'-monophosphate dehydrogenase	Q4QD53	78	2
Putative 40S ribosomal protein S15A	Q4QGW3	76	2
Aminopeptidase	E9AE86	72	2
Putative calmodulin-related protein	Q4Q6V3	66	3
Putative proteasome regulatory ATPase subunit 5	Q4QBT5	64	1
GMP reductase	Q4QEB3	63	1
Putative ATP synthase, epsilon chain	Q4Q6S8	63	1
Putative nucleolar RNA binding protein	Q4QF42	61	2
Kinesin-like protein	Q9NF78	60	1
Putative paraflagellar rod component par4	Q4QJJ9	60	2
Putative cytochrome c oxidase VII	Q4Q9Y0	59	3
Chimeric folate transporter	B0FRA4	57	4
Putative high mobility group proteinhomolog tdp-1	E9ADT8	55	1
Putative aspartate aminotransferase	Q4QAU4	55	1
Putative chaperone protein DNAj	Q4Q0D3	54	1
Putative small myristoylated protein 4	Q5SDH3	54	1
Clathrin light chain	Q4QDL0	51	2
Glucose-6-phosphate 1-dehydrogenase	Q4Q3K1	49	1
Putative 60S ribosomal protein L27A/L29	E9AFK0	49	2
Putative 60S ribosomal protein L35	Q4Q8V7	47	1
14-3-3 protein II	Q4QH45	46	2
Putative proteasome regulatory non-ATP-ase subunit 3	E9ADC7	45	1
ALP5	Q4QF49	44	1
V-type proton ATPase subunit a	Q4QAY7	44	1
Obg-like ATPase 1	E9ACV1	43	2
60S ribosomal protein L30	E9AEK1	43	1
Putative cytochrome c oxidase VIII (COX VIII)	Q4Q6A5	42	3
Putative 60S ribosomal protein L23	E9AFK3	42	1
Putative dynein	Q4Q5H6	42	1
Putative peptidase M20/M25/M40	Q4Q426	41	2
ADP-ribosylation factor-like protein 1	Q4QEJ1	40	1
40S ribosomal protein S24	Q4Q1D2	40	1

(cont. on next page)

Table 4.2. (cont.)

Protein Name for Aggressive *L. major*	**Uniprot code**	**Score**	**Matches**
Receptor-type adenylate cyclase b	Q4QEH9	40	1
Putative ribosomal protein L24	Q4Q1W6	39	1
Putative translation initiation factor IF-2	Q4Q3R1	39	1
Acidocalcisomal pyrophosphatase	Q4QH59	39	1
Dihydrolipoamide acetyltransferase OS	Q4QCG0	38	1
Nucleosome assembly protein-like protein	Q4Q687	38	1
Putative rrp44p homologue	Q4Q8N7	37	2
Putative small ubiquitin protein	Q4QIC2	37	1
Putative NADH-dependent fumarate reductase	E9AEU1	37	1
Putative intraflagellar transport protein component	Q4QDI9	36	1
Putative 60S ribosomal protein L10a	Q4QDX9	36	2
V-type proton ATPase subunit F	Q4QGP0	36	1
Putative mitochondrial intermediate peptidase	Q4Q0W9	35	2
Putative asparaginyl-tRNA synthetase	Q4Q2W6	35	2
3-hydroxy-3-methylglutaryl coenzyme A reductase	Q9Y0F3	35	1
Mannose-6-phosphate isomerase	Q4Q5C3	35	1
Putative seryl-tRNA synthetase	Q4QH70	35	1
Putative ER--golgi transport protein p24	Q4Q574	35	2
Translation initiation factor-like protein	Q4QIB4	34	1
T-complex protein 1 subunit delta	Q4QC87	34	1
Alanine aminotransferase	Q4QGN0	34	2
Eukaryotic translation initiation factor 3 subunit I	Q4Q127	34	1
Farnesyltransferase alpha subunit	Q8WR01	33	1
Putative proteasome regulatory non-ATPase subunit 6	E9ACB2	31	1
Putative 4-methyl-5(Beta-hydroxyethyl)-thiazole monophosphate synthesis protein	Q4Q7N0	30	1
Putative signal recognition particle protein	Q4QDE2	30	1
Putative proteasome regulatory ATPase subunit 2	Q4QG45	30	1
Putative chaperonin TCP20	Q4QFY8	30	1
Putative intraflagellar transport (IFT) protein	Q4Q794	29	1
Putative eukaryotic translation initiation factor 2 subunit	Q4QHR7	29	1
Cytochrome c	Q4QEN5	28	2
Putative prolyl-tRNA synthetase	Q4QDS0	25	1
Protein phosphatase 2C-like protein	Q4Q2U5	23	1
Putative ATP-dependent zinc metallopeptidase	Q4Q5D1	23	1
Putative nucleoside diphosphate kinase	E9AFL1	22	1
Tetratricopeptide repeat (TPR) protein	Q4QH31	21	1
Probable tRNA N6-adenosine threonylcarbamoyltransferase	Q4Q6Q4	21	1

(cont. on next page)

According to results showed in Table 3.3, the number of different proteins that were found for the aggressive sample was increased and also, the number of the differential proteins for passive sample showed a dramatic decrease compared to the first batch of their samples.

Table 4.2. (cont.)

Protein name for passive *L. major*	Uniprot code	Scores	Matches
Putative argininosuccinate synthase	Q4QBE8	200	6
Acetylornithine deacetylase-like protein	Q4QIR7	100	4
Thiol-dependent reductase 1	Q70GE8	90	3
Putative serine peptidase family S51, peptidase E	Q4QHW4	82	4
Putative carboxypeptidase	Q4QFW7	74	2
V-type proton ATPase subunit C	Q4QDY6	68	2
Small nuclear ribonucleoprotein SmD2	Q4Q3L7	65	2
1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	Q4QCU9	57	1
Prostaglandin f2-alpha synthase (Fragment)	A1Y2D8	53	3
Putative ATPase	Q4QJ96	50	2
V-type proton ATPase subunit	Q4QJ88	49	1
Adenylyl cyclase-associated protein	Q4Q0K2	48	2
Riboflavin kinase/fmn adenylyltransferase-like protein	E9AFE0	46	1
Putative proteasome regulatory non-ATPase subunit	E9ADL4	45	1
Putative nucleoside transporter 1	Q4QF58	45	1
Aspartyl putative aminopeptidase	E9AE98	43	2
Putative tyrosyl-tRNA synthetase	Q4QFJ7	42	2
Putative dihydrolipoamide dehydrogenase OS	E9AE44	42	1
Putative folate/biopterin transporter OS	Q4QHI0	41	1
Putative U3 small nucleolar ribonucleoprotein protein MPP10	E9ADS7	40	1
Putative dynein light chain	Q4QES6	38	1
Cyclophilin 40	E9AFV2	38	1
Mitochondrial RNA binding protein 2	Q4QHR3	38	1
Arginine N-methyltransferase-like protein	Q4QGG2	35	1
Putative short chain 3-hydroxyacyl-CoA dehydrogenase	Q4Q1W2	35	1
Putative heat shock protein-like protein	Q4Q584	35	1
Putative 40S ribosomal protein S3	Q4Q4A0	34	1
C2 domain protein	Q9BHF8	34	2
Possible ABC transporter	Q9BHG2	33	1
Putative 60S ribosomal protein L6	Q4Q4D3	31	1
Putative metallopeptidase	Q4Q8U8	29	1
Ribonucleoside-diphosphate reductase	Q4Q8H8	28	1
Putative paraflagellar rod protein	Q4QJB0	26	1
Putative QA-SNARE protein	E9AF95	26	1
Choline dehydrogenase, like protein	Q4QC34	24	1
Putative adaptor complex subunit medium chain 3	Q4Q2T6	21	1
DNA repair protein RAD51 homolog	O61127	20	1
Methylmalonyl-coa epimerase-like protein	Q4Q9I9	20	1
Putative udp-glc 4'-epimerase	Q4Q3V7	20	1

As a result, 122 proteins of the total 409 protein for the aggressive sample were found as differential and 47 proteins of the total 298 protein were found unlike than both aggressive and reference sample. Unlike the first batch of the aggressive sample, differential protein distribution in the cellular component of the sample was showed a

distinction by localizing microtubules and intracellular organelle lumen for 5 protein out of 78. However, they are mostly involved in catalytic activity as similar to the previous batch.

On the other hand, as it seen the table above, 47 differential proteins were detected for the passive sample. Parallel to results that were found before, they are predominantly involved in catalytic activity and shows the same distribution in the cell of the samples. Furthermore, molecular functions for all proteins for aggressive, passive and L. *major* samples were shown in Figure 4.2.

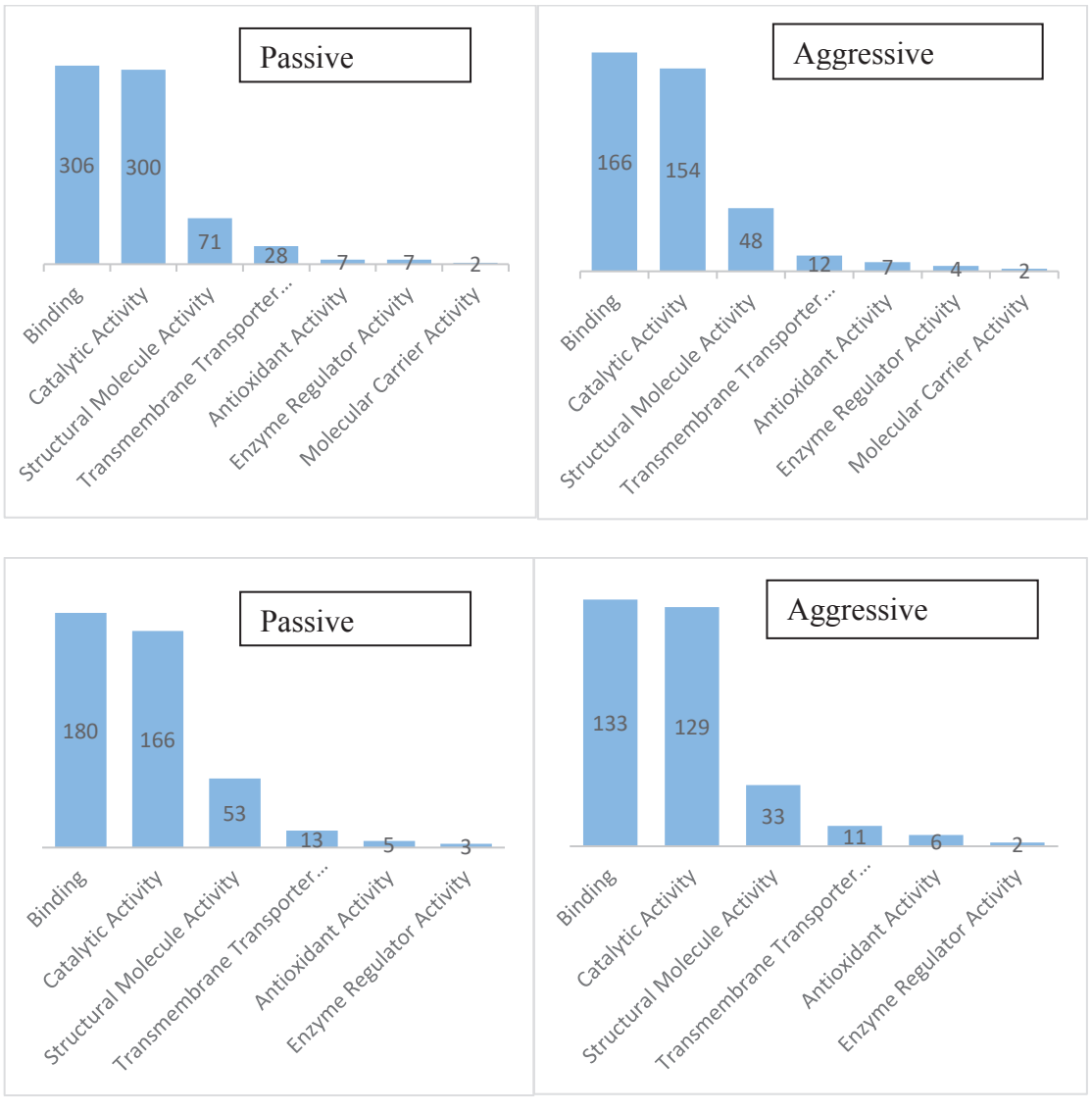
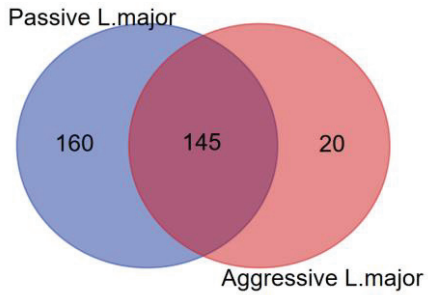
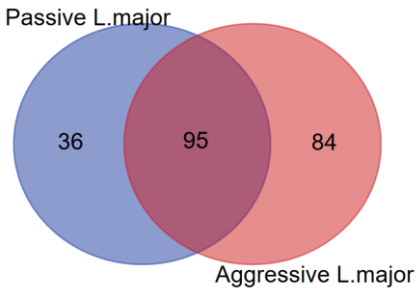


Figure 4.2. Molecular functions of passive (left) and aggressive (right) *L. major* samples. First batch of the passive and aggressive samples (up) and second batch of the passive and aggressive samples (bottom) are shown.

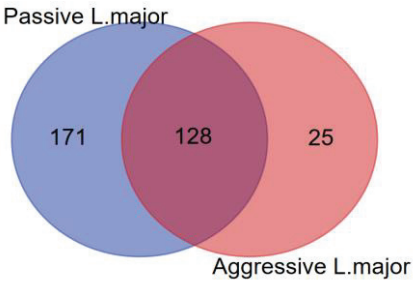
Batch 1 **Batch 2**

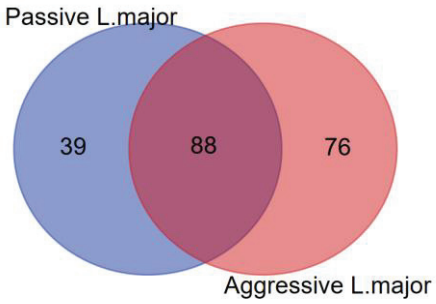
Binding Proteins



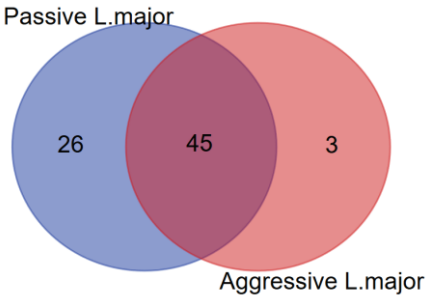


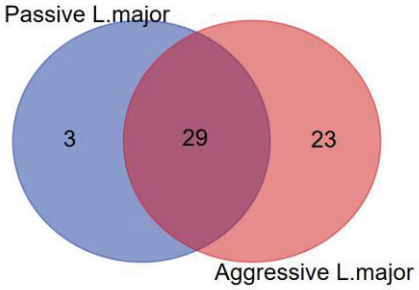
Catalytic Activity Proteins



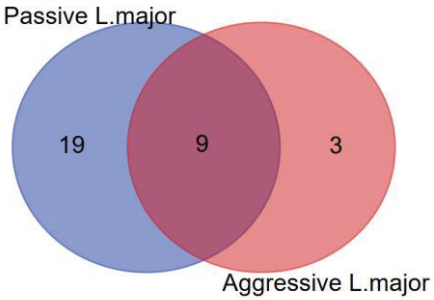


Structural Molecule Activity Proteins





Transmembrane Transporter Activity Proteins



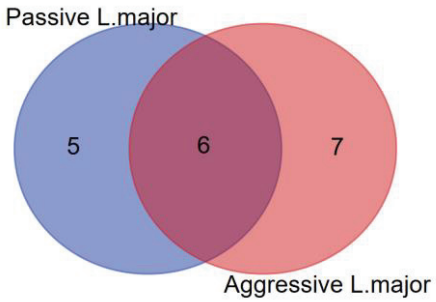


Figure 4.3. Binary Venn diagrams distributions of *L. major* samples associated to different molecular functions for both first and second batches

As it shown in the figure above, distribution of the proteins in both batches according to their molecular functions have the same functionality pattern. Also, it is observed that the passive *L. major* samples in both batches have increased the number of protein, mostly observable in binding and catalytic activities, compared to aggressive *L. major* sample. This decrease in aggressive *L. major* sample might originate from the virulence factor differences between aggressive and passive samples. In addition to that, binary comparison between passive and aggressive *L. major* samples according to their molecular function. Binary Venn diagram distributions of *L. major* samples associated with different molecular functions illustrated as in Figure 4.3 for first batch.

All Venn diagrams above were investigated in detail to recognize differential proteins coming from the different molecular function. These differences can be originated from the virulence factor of the *L. major* samples both passive and aggressive and play important role in the differencing mechanism of the parasite. Identification of these proteins might open a door for the characterization of the proteins that cause virulence effect in *L. major* infectivity. Therefore, the lists below in Table 4.3, differential proteins associated with the binding activity are illustrated for both the first and second batches. Comparison between both batches provides to detect reliable and repetitive proteins that are found in both batches.

Table 4.3. Differential proteins that is found according to their binding activities in both first and second batch for passive *L. major*

Protein Name for Passive *L. major*	**Batch 1**	**Batch 2**
Putative 40S ribosomal protein S3	√	√
T-complex protein 1 subunit epsilon	√	√
Putative short chain 3-hydroxyacyl-CoA dehydrogenase	√	√
Thiol-dependent reductase 1	√	√
Putative N-ethylmaleimide-sensitive factor	√	√
Putative 2-oxoglutarate dehydrogenase subunit	√	√
Pyruvate, phosphate dikinase	√	√
Putative RNA binding protein rbp16	√	√
Dynein light chain	√	√
Putative cysteinyl-tRNA synthetase	√	√
Cyclophilin 40	√	√
1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	√	√
Putative argininosuccinate synthase	√	√
Putative mitotubule-associated protein Gb4	√	√
Putative QA-SNARE protein	√	√
Adenylyl cyclase-associated protein	√	√

(cont. on next page)

Table 4.3. (cont.)

Protein Name for Passive *L. major*	**Batch 1**	**Batch 2**
Putative 3'-nucleotidase/nuclease	√	x
Cation-transporting ATPase	√	x
Ubiquitin-conjugating enzyme-like protein	√	x
Serine/threonine-protein phosphatase	√	x
Putative mitochondrial oligo_U binding protein TBRGG1	√	x
Cystathionine beta-lyase-like protein	√	x
Cytochrome c	√	x
Elongation factor Ts, mitochondrial	√	x
Acidocalcisomal pyrophosphatase	√	x
Putative ATP-binding cassette protein subfamilyA,member8	√	x
RuvB-like helicase	√	x
Eukaryotic translation initiation factor 3 subunit L	√	x
Inosine-5'-monophosphate dehydrogenase	√	x
Farnesyl pyrophosphate synthase	√	x
Putative cystathione gamma lyase	√	x
Putative chaperonin TCP20	√	x
Putative aspartyl-tRNA synthetase	√	x
Putative prolyl-tRNA synthetase	√	x
Putative thimet oligopeptidase	√	x
Replication protein A subunit	√	x
Putative mitogen-activated protein kinase	√	x
Putative peptidase t	√	x
Alkyl dihydroxyacetonephosphate synthase	√	x
Putative T-complex protein 1, beta subunit	√	x
Putative fumarate hydratase	√	x
Putative CCR4 associated factor	√	x
Putative reiske iron-sulfur protein	√	x
Putative aspartate aminotransferase	√	x
Putative 40S ribosomal protein S16	√	x
CTP synthase	√	x
Putative heat shock protein HslVU, ATPase,subunit HslU	√	x
Eukaryotic translation initiation factor 5A	√	x
Polyadenylate-binding protein	√	x
Elongation factor G, mitochondrial	√	x
Cysteine synthase	√	x
Receptor-type adenylate cyclase a	√	x
Acyl carrier protein	√	x
Putative methionyl-tRNA synthetase	√	x
Putative proteasome regulatory ATPase subunittcc1l8.3	√	x
Prefoldin 5-like protein	√	x
Eukaryotic translation initiation factor 3 subunit 7-like protein	√	x
GrpE protein homolog	√	x
Putative leucyl-tRNA synthetase	√	x
Putative ribosomal protein L38	√	x
Sphingosine-1-phosphate lyase	√	x
Putative calmodulin	√	x
Putative carboxylase	√	x
Tyrosine aminotransferase	√	x
Ribokinase	√	x
Putative acetyl-CoA carboxylase	√	x
Putative proteasome regulatory ATPase subunit 5	√	x
Putative cytochrome b-domain protein	√	x
Aldose 1-epimerase-like protein	√	x

(cont. on next page)

Table 4.3. (cont.)

Protein Name for Passive *L. major*	**Batch 1**	**Batch 2**
I/6 autoantigen-like protein	√	x
ADF/Cofilin	√	x
Putative 60S Ribosomal protein L36	√	x
ATP-dependent 6-phosphofructokinase	√	x
Phosphomannomutase-like protein	√	x
Alanine aminotransferase	√	x
Putative 60S ribosomal protein L23a	√	x
Putative rab1 small GTP-binding protein	√	x
Putative signal recognition particle protein	√	x
Formate--tetrahydrofolate ligase	√	x
Elongation factor 1B alpha	√	x
Prefoldin subunit 3	√	x
Kinesin-like protein	√	x
Aromatic amino acid hydroxylase-like	√	x
Putative ribosomal protein L24	√	x
Putative dipeptylcarboxypeptidase	√	x
5-methyltetrahydropteroyltriglutamate-homocystein e S-methyltransferase	√	x
Putative RNA-binding protein	√	x
Tyrosyl or methionyl-tRNA synthetase-like protein	√	x
Putative 40S ribosomal protein S15	√	x
Putative isovaleryl-coA dehydrogenase	√	x
Putative eukaryotic translation initiation factor 5	√	x
Putative 60S ribosomal protein L2	√	x
Protein SEY1 homolog	√	x
Putative ribosomal protein L1a	√	x
Glucokinase	√	x
Putative glycogen synthase kinase	√	x
Putative prefoldin subunit 2	√	x
Putative carbamoyl-phosphate synthase	√	x
Putative calmodulin-related protein	√	x
Putative lanosterol 14-alpha-demethylase	√	x
Putative iron-sulfur cluster assembly protein	√	x
Putative high mobility group protein homolog tdp-1	√	x
Putative elongation initiation factor 2 alpha subunit	√	x
Putative dynein	√	x
Pyridoxal phosphate homeostasis protein	√	x
ATPase ASNA1 homolog	√	x
Putative ATP-binding cassette protein subfamily C	√	x
Serine hydroxymethyltransferase	√	x
Eukaryotic translation initiation factor 6	√	x
Dynein light chain roadblock	√	x
Putative ATP-binding cassette protein subfamily F	√	x
Mitogen-activated protein kinase	√	x
Histone H4	√	x
Putative kinesin K39	√	x
Putative 60S ribosomal protein L23	√	x
Eukaryotic translation initiation factor 3 subunit E	√	x
Diphosphomevalonate decarboxylase	√	x
Mevalonate kinase	√	x
60S ribosomal protein L30	√	x
Putative 60S ribosomal protein L7	√	x
Putative la RNA binding protein	√	x

(cont. on next page)

Table 4.3. (cont.)

Protein Name for Passive *L. major*	**Batch 1**	**Batch 2**
Putative histone H3	√	x
Flavoprotein subunit-like protein	√	x
Putative serine/threonine-protein kinase	√	x
Intraflagellar transport protein-like protein	√	x
V-type proton ATPase subunit a	√	x
Methionine aminopeptidase	√	x
Putative valyl-tRNA synthetase	√	x
Isopentenyl-pyrophosphate isomerase	√	x
Putative GTP-binding protein	√	x
Putative 60S ribosomal protein L26	√	x
Transcription factor-like protein	√	x
2-oxoglutarate dehydrogenase, e3 component,lipoamidedehydrogenase-like protein	√	x
Putative tryptophanyl-tRNA synthetase	√	x
Putative NADH:flavin oxidoreductase/NADH oxidase	√	x
Proteasome regulatory ATPase subunit	√	x
Putative 40S ribosomal protein S10	√	x
Putative small GTP-binding protein Rab1	√	x
Prefoldin subunit 4	√	x
Putative DNA-directed rna polymerase I largest subunit	√	x
GMP reductase	√	x
Vacuolar protein sorting-associated protein 4	√	x
Putative centromere/microtubule binding protein cbf5	√	x
Putative paraflagellar rod component	√	x
Putative ubiquitin carrier protein	√	x
Putative centrin	√	x
Putative 26S protease regulatory subunit	√	x
Putative ATP-binding cassette protein subfamily	√	x
Putative electron transfer flavoprotein-ubiquinone oxidoreductase	√	x
Profilin	√	x
Transcription elongation factor-like protein	√	x
Nucleolar protein family a member-like protein	√	x
Succinate dehydrogenase [ubiquinone] flavoprotein subunit	√	x
Dynein intermediate-chain-like protein	√	x
Putative ras-related protein rab-5	√	x
Putative 60S ribosomal protein L19	√	x
Phosphotransferase	√	x
LmRab7 GTP-binding protein	√	x
T-complex protein 1 subunit delta	√	x
Protein kinase A catalytic subunit	√	x
Choline dehydrogenase, like protein	√	x
Putative developmentally regulated GTP-binding protein 1	x	√
Ribonucleoside-diphosphate reductase	x	√
Putative small nuclear ribonucleoprotein	x	√
Putative dihydrolipoamide dehydrogenase	x	√
Alanine--tRNA ligase	x	√
Putative tyrosyl-tRNA synthetase	x	√
Putative ATPase	x	√
DNA repair protein RAD51 homolog	x	√
Putative 40S ribosomal protein S18	x	√
Mitochondrial RNA binding protein 2	x	√
Putative heat shock protein-like protein	x	√
Possible ABC transporter	x	√

(cont. on next page)

Table 4.3. (cont.)

Protein Name for Passive *L. major*	**Batch 1**	**Batch 2**
Aspartyl putative aminopeptidase	x	√
Putative 60S ribosomal protein L6	x	√
Putative udp-glc 4'-epimerase	x	√
Metallo-peptidase, Clan MA	x	√
Putative paraflagellar rod protein	x	√
Putative 2-hydroxy-3-oxopropionate reductase	x	√
Putative acyl-CoA dehydrogenase	x	√

Table 4.4. Differential proteins that is found according to their binding activities in both first and second batch for aggressive *L. major*

Protein Names for Aggressive *L. major*	**Batch 1**	**Batch 2**
Obg-like ATPase 1	√	√
Tubulin-specific chaperone A	√	√
Cystathionine beta-synthase	√	√
Receptor-type adenylate cyclase b	√	√
Putative cytochrome c oxidase copper chaperone	√	x
Putative CAS/CSE/importin domain protein	√	x
40S ribosomal protein S5B	√	x
Casein kinase 1 isoform 2	√	x
Putative NADP-dependent alcohol dehydrogenase	√	x
Mitochondrial RNA binding protein 1	√	x
Putative phosphomevalonate kinase protein	√	x
Putative ATP-binding cassette protein subfamily E,member 1	√	x
Putative phosphoglucomutase	√	x
ADP-ribosylation factor 1	√	x
Putative pumilio protein 6	√	x
Putative small GTP-binding protein Rab11	√	x
Putative small GTP-binding protein Rab7	√	x
Putative ras-like small GTPases	√	x
Ca2+-binding EF-hand protein	√	x
Galactokinase-like protein	√	x
Putative asparaginyl-tRNA synthetase	x	√
Mannose-6-phosphate isomerase	x	√
Ubiquitin-conjugating enzyme-like protein	x	√
Putative 60S ribosomal protein L27A/L29	x	√
Uncharacterized protein L7836.08	x	√
Cytochrome c	x	√
40S ribosomal protein S14	x	√
Acidocalcisomal pyrophosphatase	x	√
Putative calmodulin-related protein	x	√
Putative chaperone protein DNAj	x	√
Putative 60S ribosomal protein L10a	x	√
Putative high mobility group protein homolog tdp-1	x	√
Putative dynein	x	√
Inosine-5'-monophosphate dehydrogenase	x	√
Putative 40S ribosomal protein L14	x	√
Eukaryotic translation initiation factor 3 subunit I	x	√
Histone H4	x	√
14-3-3 protein II	x	√
Putative chaperonin TCP20	x	√

(cont. on next page)

Table 4.4. (cont.)

Protein Names for Aggressive *L. major*	**Batch 1**	**Batch 2**
Putative nucleolar RNA binding protein	x	√
Putative aminopeptidase P	x	√
Putative prolyl-tRNA synthetase	x	√
60S ribosomal protein L30	x	√
3-hydroxy-3-methylglutaryl coenzyme A reductase	x	√
Putative thimet oligopeptidase	x	√
60S ribosomal protein L11	x	√
Putative RNA binding protein	x	√
Putative fumarate hydratase	x	√
Putative ADP-ribosylation factor	x	√
V-type proton ATPase subunit a	x	√
60S ribosomal protein L18a	x	√
Putative 60S ribosomal protein L35	x	√
Putative aspartate aminotransferase	x	√
Putative ribosomal protein L3	x	√
40S ribosomal protein S2	x	√
Putative GTP-binding protein	x	√
Putative eukaryotic translation initiation factor 1A	x	√
Transketolase	x	√
9,11-endoperoxide prostaglandin H2 reductase	x	√
Putative translation initiation factor IF-2	x	√
Putative heat shock protein DNAJ	x	√
Putative ADP ribosylation factor 3	x	√
Probable tRNA N6-adenosine threonylcarbamoyltransferase	x	√
Putative ATP-dependent DEAD/H RNA helicase	x	√
Eukaryotic translation initiation factor 3 subunit 7-like protein	x	√
GTP-binding nuclear protein	x	√
Putative nucleoside diphosphate kinase	x	√
Histone H2A	x	√
GMP reductase	x	√
Putative leucyl-tRNA synthetase	x	√
Myosin XXI	x	√
Putative proteasome regulatory ATPase subunit 2	x	√
Glucose-6-phosphate 1-dehydrogenase	x	√
Putative isoleucyl-tRNA synthetase	x	√
Putative mitochondrial intermediate peptidase	x	√
Succinate--CoA ligase [ADP-forming] subunit alpha	x	√
Putative proteasome regulatory ATPase subunit 5	x	√
ADP-ribosylation factor-like protein 1	x	√
ADF/Cofilin	x	√
Putative ATP-dependent zinc metallopeptidase	x	√
ATP-dependent 6-phosphofructokinase	x	√
Aminopeptidase	x	√
Putative serine peptidase	x	√
Putative eukaryotic translation initiation factor 2 subunit	x	√
Putative seryl-tRNA synthetase	x	√
Alanine aminotransferase	x	√
Translation initiation factor-like protein	x	√
6-phosphogluconate dehydrogenase, decarboxylating	x	√
Putative paraflagellar rod protein 1D	x	√
Protein phosphatase 2C-like protein	x	√
Putative signal recognition particle protein	x	√
Phosphotransferase	x	√

(cont. on next page)

Table 4.4. (cont.)

Protein Names for Aggressive *L. major*	**Batch 1**	**Batch 2**
Putative ribosomal protein L24	x	√
T-complex protein 1 subunit delta	x	√
Putative vacuolar ATP synthase subunit b	x	√
Putative kinesin	x	√
Putative ribosomal protein L1a	x	√
Putative Transitional endoplasmic reticulum ATPase	x	√
Putative 60S ribosomal protein L23	x	√
Kinesin-like protein	x	√

As it seen in both tables above, the binding proteins that are overlapped with both batches for the passive and aggressive samples ensure that these protein may have a vital role for infectivity of *L. major* associated with its binding activities. Same investigation have been applied for rest of the molecular function of obtained *L. major* proteins in fallowing tables.

Table 4.5. Differential proteins that is found according to their catalytic activities in both first and second batch for passive *L. major*

Protein Names for Passive *L. major*	**Batch 1**	**Batch 2**
Putative 40S ribosomal protein S3	√	√
Putative cytochrome c oxidase subunit VI	√	√
Putative short chain 3-hydroxyacyl-CoA dehydrogenase	√	√
Thiol-dependent reductase 1	√	√
Putative N-ethylmaleimide-sensitive factor	√	√
Putative 2-oxoglutarate dehydrogenase subunit	√	√
Pyruvate, phosphate dikinase	√	√
Putative phosphatase 2C	√	√
Dynein light chain	√	√
Putative cysteinyl-tRNA synthetase	√	√
Cyclophilin 40	√	√
Elongation of fatty acids protein	√	√
1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	√	√
Putative argininosuccinate synthase	√	√
Riboflavin kinase/fmn adenylyltransferase-like protein	√	√
Putative mitotubule-associated protein Gb4	√	√
Putative X-pro, dipeptidyl-peptidase,serine peptidase, Clan SC	√	√
Putative 3'-nucleotidase/nuclease	√	x
Putative ubiquitin hydrolase	√	x
Cation-transporting ATPase	√	x
Ubiquitin-conjugating enzyme-like protein	√	x
Serine/threonine-protein phosphatase	√	x
Cystathionine beta-lyase-like protein	√	x
Cytochrome c	√	x
Putative cytochrome c oxidase subunit V	√	x
Acidocalcisomal pyrophosphatase	√	x
Putative ATP-binding cassette protein subfamily A,member 8	√	x
RuvB-like helicase	√	x

(cont. on next page)

Table 4.5. (cont.)

Protein Names for Passive *L. major*	**Batch 1**	**Batch 2**
Adenine aminohydrolase	√	x
Putative cystathione gamma lyase	√	x
Adenine phosphoribosyltransferase	√	x
Putative aspartyl-tRNA synthetase	√	x
Putative prolyl-tRNA synthetase	√	x
Putative threonine synthase	√	x
Putative thimet oligopeptidase	√	x
Putative mitogen-activated protein kinase	√	x
Putative peptidase t	√	x
Methylthioribose-1-phosphate isomerase	√	x
Alkyl dihydroxyacetonephosphate synthase	√	x
Putative deoxyribose-phosphate aldolase	√	x
Putative fumarate hydratase	√	x
Putative CCR4 associated factor	√	x
Putative reiske iron-sulfur protein	√	x
Putative short chain dehydrogenase	√	x
Putative aspartate aminotransferase	√	x
Putative AMP deaminase	√	x
CTP synthase	√	x
Putative heat shock protein HslVU, ATPase,subunit HslU	√	x
Putative branched-chain amino acid aminotransferase	√	x
Elongation factor G, mitochondrial	√	x
Cathepsin L-like protease	√	x
Cysteine synthase	√	x
Receptor-type adenylate cyclase a	√	x
Putative methionyl-tRNA synthetase	√	x
Trifunctional enzyme alpha subunit, mitochondrial-like protein	√	x
Putative vacuolar-type proton translocating pyrophosphatase 1	√	x
Putative proteasome regulatory ATPase subunittcc1l8.3	√	x
Putative U-box domain protein	√	x
V-type proton ATPase subunit H	√	x
Putative hydroxyacylglutathione hydrolase	√	x
Phosphatidic acid phosphatase protein-like protein	√	x
Oligosaccharyl transferase-like protein	√	x
Proteasome endopeptidase complex	√	x
Putative leucyl-tRNA synthetase	√	x
Proteasome subunit beta	√	x
Sphingosine-1-phosphate lyase	√	x
Putative carboxylase	√	x
Proteasome subunit beta type	√	x
Tyrosine aminotransferase	√	x
2-methoxy-6-polyprenyl-1,4-benzoquinol methylase, mitochondrial	√	x
Ribokinase	√	x
Putative acetyl-CoA carboxylase	√	x
Putative proteasome regulatory ATPase subunit 5	√	x
Oxidoreductase-like protein	√	x
Aldose 1-epimerase-like protein	√	x
Polyprenol reductase	√	x
Putative major surface protease gp63	√	x
Formate--tetrahydrofolate ligase	√	x
Hs1vu complex proteolytic subunit-like	√	x
Kinesin-like protein	√	x
Aromatic amino acid hydroxylase-like	√	x
Putative cytochrome c oxidase VII	√	x

(cont. on next page)

Table 4.5. (cont.)

Protein Names for Passive *L. major*	**Batch 1**	**Batch 2**
Putative dipeptylcarboxypeptidase	√	x
5-methyltetrahydropteroyltriglutamate-homocystein e S-methyltransferase	√	x
Tyrosyl or methionyl-tRNA synthetase-like protein	√	x
Putative isovaleryl-coA dehydrogenase	√	x
Protein SEY1 homolog	√	x
Thiolase protein-like protein	√	x
Protein disulfide isomerase	√	x
Glucokinase	√	x
Putative glycogen synthase kinase	√	x
Pteridine reductase 1	√	x
Cysteine protease	√	x
Putative carbamoyl-phosphate synthase	√	x
Putative lanosterol 14-alpha-demethylase	√	x
Putative iron-sulfur cluster assembly protein	√	x
Putative dynein	√	x
Putative 3-hydroxyacyl-ACP dehydratase	√	x
Pyridoxal phosphate homeostasis protein	√	x
ATPase ASNA1 homolog	√	x
Putative ATP-binding cassette protein subfamily C,member 6	√	x
Serine hydroxymethyltransferase	√	x
Dynein light chain roadblock	√	x
Putative ATP-binding cassette protein subfamily F,member 2	√	x
Mitogen-activated protein kinase	√	x
Cytochrome c oxidase assembly factor-like protein	√	x
Putative kinesin K39	√	x
Putative NADH-dependent fumarate reductase	√	x
Diphosphomevalonate decarboxylase	√	x
Mevalonate kinase	√	x
Ubiquitin carboxyl-terminal hydrolase	√	x
Putative prostaglandin f synthase	√	x
Flavoprotein subunit-like protein	√	x
Putative serine/threonine-protein kinase	√	x
Putative N-acyl-L-amino acid amidohydrolase	√	x
V-type proton ATPase subunit a	√	x
Methionine aminopeptidase	√	x
Oligopeptidase b	√	x
Putative valyl-tRNA synthetase	√	x
Isopentenyl-pyrophosphate isomerase	√	x
Putative GTP-binding protein	√	x
Glyceraldehyde-3-phosphate dehydrogenase-like protein	√	x
Putative glutamine synthetase	√	x
Proteasome regulatory ATPase subunit	√	x
Putative acidocalcisomal exopolyphosphatase	√	x
Putative small GTP-binding protein Rab1	√	x
Putative glutamamyl carboxypeptidase	√	x
3-hydroxyisobutyryl-CoA hydrolase, mitochondrial	√	x
Putative DNA-directed rna polymerase I largest subunit	√	x
GMP reductase	√	x
Vacuolar protein sorting-associated protein 4	√	x
Putative centromere/microtubule binding protein cbf5	√	x
Putative ubiquitin carrier protein	√	x
Putative 26S protease regulatory subunit	√	x
Putative beta eliminating lyase	√	x
Putative ATP-binding cassette protein subfamily	√	x

(cont. on next page)

Table 4.5. (cont.)

Protein Names for Passive *L. major*	**Batch 1**	**Batch 2**
NADH-cytochrome b5 reductase	√	x
Sucrose-phosphate synthase-like protein	√	x
Putative Pyridoxal kinase	√	x
Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	√	x
Dynein intermediate-chain-like protein	√	x
Putative ras-related protein rab-5	√	x
Phosphotransferase	√	x
LmRab7 GTP-binding protein	√	x
Protein kinase A catalytic subunit	√	x
Putative proteasome beta 2 subunit	√	x
Putative cytoskeleton-associated protein CAP5.5	√	x
Choline dehydrogenase, like protein	x	√
Ribonucleoside-diphosphate reductase	x	√
Citrate synthase	x	√
Putative dihydrolipoamide dehydrogenase	x	√
Alanine--tRNA ligase	x	√
Putative tyrosyl-tRNA synthetase	x	√
Putative metallopeptidase	x	√
Arginine N-methyltransferase-like protein	x	√
Inositol-3-phosphate synthase	x	√
DNA repair protein RAD51 homolog	x	√
Putative serine peptidase family S51, peptidase E	x	√
V-type proton ATPase subunit	x	√
Acetylornithine deacetylase-like protein	x	√
Putative carboxypeptidase	x	√
Possible ABC transporter	x	√
Aspartyl putative aminopeptidase	x	√
V-type proton ATPase subunit C	x	√
Prostaglandin f2-alpha synthase	x	√
Putative udp-glc 4'-epimerase	x	√
Methylmalonyl-coa epimerase-like protein	x	√
Putative 2-hydroxy-3-oxopropionate reductase	x	√
Putative acyl-CoA dehydrogenase	x	√
2-oxoglutarate dehydrogenase, e3 component,lipoamidedehydrogenase-	√	x
V-type proton ATPase subunit F	√	x
Putative ubiquitin-activating enzyme e1	√	x
Phosphomannomutase	√	x
Putative ATP synthase, epsilon chain	√	x
Succinyl-diaminopimelate desuccinylase-like protein	√	x
Putative tryptophanyl-tRNA synthetase	√	x
Diphthine synthase-like protein	√	x
Putative NADH:flavin oxidoreductase/NADH oxidase	√	x
ATP-dependent 6-phosphofructokinase	√	x
Putative carnitine/choline acetyltransferase	√	x
Similarity to endo-1-like protein	√	x
Phosphomannomutase-like protein	√	x
Alanine aminotransferase	√	x
Leucine carboxyl methyltransferase 1	√	x
Putative rab1 small GTP-binding protein	√	x
Putative cyclophilin 12	√	x
Putative methionine-S-sulfoxide reductase	√	x
Inosine-5'-monophosphate dehydrogenase	√	x
Mannose-1-phosphate guanyltransferase	√	x
Farnesyl pyrophosphate synthase	√	x

(con. on next page)

Table 4.5. (cont.)

Putative cytochrome c oxidase subunit 10	√	x
Putative aspartate--ammonia ligase	√	x
Putative electron transfer flavoprotein-ubiquinone oxidoreductase	√	x

Table 4.6. Differential proteins that is found according to their catalytic activities in both first and second batch for aggressive *L. major*

Protein Names for Aggressive *L. major*	**Batch 1**	**Batch 2**
Dihydrolipoamide acetyltransferaselike protein	√	√
Putative ATP synthase F1 subunit gamma protein	√	√
UDP-galactopyranose mutase	√	√
Obg-like ATPase 1	√	√
Putative rrp44p homologue	√	√
Cystathionine beta-synthase	√	√
Receptor-type adenylate cyclase b	√	√
Putative 3-methylcrotonoyl-CoA carboxylase beta subunit	√	x
Ribonuclease mar1	√	x
Putative electron transfer flavoprotein	√	x
Thioredoxin-like protein	√	x
Casein kinase 1 isoform 2	√	x
Putative uracil phosphoribosyltransferase	√	x
Putative NADP-dependent alcohol dehydrogenase	√	x
Putative phosphomevalonate kinase protein	√	x
Putative ATP-binding cassette protein subfamily E,member 1	√	x
Phosphatidate cytidylyltransferase	√	x
Putative phosphoglucomutase	√	x
Putative ATP synthase	√	x
Putative small GTP-binding protein Rab11	√	x
Putative small GTP-binding protein Rab7	√	x
Adenosine monophosphate deaminase-like protein	√	x
Topoisomerase-related function protein-like protein	√	x
Palmitoyltransferase	√	x
Galactokinase-like protein	√	x
Putative asparaginyl-tRNA synthetase	x	√
Mannose-6-phosphate isomerase	x	√
Putative ubiquitin hydrolase	x	√
Ubiquitin-conjugating enzyme-like protein	x	√
Cytochrome c	x	√
Biotin/lipoate protein ligase-like protein	x	√
Regulatory subunit of protein kinase a-like protein	x	√
Putative cytochrome c oxidase subunit V	x	√
Acidocalcisomal pyrophosphatase	x	√
Putative dynein	x	√
Methyltransferase	x	√
Inosine-5'-monophosphate dehydrogenase	x	√
Putative cytochrome c oxidase VIII	x	√
Putative peptidase M20/M25/M40	x	√
Putative NADH-dependent fumarate reductase	x	√
Putative aminopeptidase P	x	√
Putative prolyl-tRNA synthetase	x	√
Glutamate dehydrogenase	x	√
3-hydroxy-3-methylglutaryl coenzyme A reductase	x	√
Putative thimet oligopeptidase	x	√

(cont. on next page)

Table 4.6. (cont.)

Protein Names for Aggressive *L. major*	**Batch 1**	**Batch 2**
Putative GTP-binding protein	x	√
Transketolase	x	√
9,11-endoperoxide prostaglandin H2 reductase	x	√
Putative translation initiation factor IF-2	x	√
V-type proton ATPase subunit F	x	√
Probable tRNA N6-adenosine threonylcarbamoyltransferase	x	√
Phosphomannomutase	x	√
Putative ATP synthase, epsilon chain	x	√
Putative ATP-dependent DEAD/H RNA helicase	x	√
Malate dehydrogenase	x	√
Putative glutamine synthetase	x	√
GP63, leishmanolysin	x	√
GTP-binding nuclear protein	x	√
Adenylosuccinate lyase	x	√
Probable citrate synthase, mitochondrial	x	√
Putative nucleoside diphosphate kinase	x	√
GMP reductase	x	√
Putative leucyl-tRNA synthetase	x	√
Myosin XXI	x	√
Putative proteasome regulatory ATPase subunit 2	x	√
Glucose-6-phosphate 1-dehydrogenase	x	√
Putative isoleucyl-tRNA synthetase	x	√
Proteasome subunit beta type	x	√
Putative mitochondrial intermediate peptidase	x	√
Succinate--CoA ligase [ADP-forming] subunit alpha, mitochondrial	x	√
Putative proteasome regulatory ATPase subunit 5	x	√
4-coumarate:coa ligase-like protein	x	√
Putative ATP-dependent zinc metallopeptidase	x	√
Putative Pyridoxal kinase	x	√
Aminopeptidase	x	√
ATP-dependent 6-phosphofructokinase	x	√
Putative eukaryotic translation initiation factor 2 subunit	x	√
Putative seryl-tRNA synthetase	x	√
Farnesyltransferase alpha subunit	x	√
Alanine aminotransferase	x	√
6-phosphogluconate dehydrogenase, decarboxylating	x	√
Protein phosphatase 2C-like protein	x	√
Phosphotransferase	x	√
Kinesin-like protein	x	√
Putative Transitional endoplasmic reticulum ATPase	x	√
Putative cytochrome c oxidase VII	x	√
Uncharacterized protein	x	√
Putative vacuolar ATP synthase subunit b	x	√
Putative kinesin	x	√
Putative 2,4-dihydroxyhept-2-ene-1,7-dioic acid aldolase	x	√
Putative aspartate aminotransferase	x	√
V-type proton ATPase subunit a	x	√
Putative fumarate hydratase	x	√
Aldehyde dehydrogenase, mitochondrial	x	√

For the passive and aggressive *L. major* samples, alternative Venn diagrams have been generated based on the molecular function distribution, that is included as

binding, catalytic activity, structural molecular activity, and transmembrane transporter activity, in both first and second batch. The differences between both passive and aggressive samples in each molecular function group were illustrated as a table list showed in Figure 4.3 to Figure 4.10.

Table 4.7. Differential proteins that is found according to their structural molecular activities in both first and second batch for passive *L. major*

Protein Names Passive *L. major*	**Batch 1**	**Batch 2**
Putative 40S ribosomal protein S15A	√	x
60S ribosomal protein L32	√	x
Putative 40S ribosomal protein S10	√	x
Putative ribosomal protein l35a	√	x
Putative ribosomal protein L38	√	x
Putative 40S ribosomal protein S19 protein	√	x
Putative 40S ribosomal protein S3	√	√
Putative 60S ribosomal protein L23	√	x
60S ribosomal protein L30	√	x
Putative 60S Ribosomal protein L36	√	x
Putative 60S ribosomal protein L7	√	x
Putative 60S ribosomal protein L23a	√	x
Coatomer subunit beta	√	x
Putative 60S ribosomal protein L28	√	x
Putative 60S ribosomal protein L19	√	x
Putative 60S ribosomal protein L37a	√	x
40S ribosomal protein SA	√	x
Putative ribosomal protein L24	√	x
Putative 60S ribosomal protein L26	√	x
Putative 40S ribosomal protein S16	√	x
Putative 40S ribosomal protein S15	√	x
Clathrin light chain	√	x
Putative 40S ribosomal protein S23	√	x
Putative 60S ribosomal protein L2	√	x
Putative proteasome regulatory non-ATPase subunit 6	√	x
Putative ribosomal protein L1a	√	x
Putative 60S ribosomal protein L6	x	√
Putative 40S ribosomal protein S18	x	√

Differential proteins coming from each molecular functions actually can show a different behaviour mechanism of the passive and aggressive *L. major* species that is represents themselves in the mammalian host. This condition might originate from variances in virulence factors of both of them. Additionally, in the illustrated tables, common proteins that were found in both first and second batch were shaded in red to represent repeatability of the experiment. According to these results, for binding proteins, 16 proteins were found in common for both batches of the passive sample and 4 common

proteins were detected for the aggressive sample. For the catalytic activities, 17 common proteins were found for a passive sample while 7 proteins were recognized as common between 2 batches. However, for structural molecular activity and transmembrane transporter activity, the number of obtained common proteins are lower than first two molecular functions mentioned above.

Table 4.8. Differential proteins that is found according to their structural molecular activities in both first and second batch for aggressive *L. major*

Protein Names for Aggressive *L. major*	**Batch 1**	**Batch 2**
Coatomer subunit epsilon	√	x
40S ribosomal protein S5B	√	x
Putative ubiquitin/ribosomal protein S27a	√	x
Putative 60S ribosomal protein L23	x	√
60S ribosomal protein L30	x	√
Putative 60S ribosomal protein L27A/L29	x	√
Putative 40S ribosomal protein S15A	x	√
40S ribosomal protein S24	x	√
Putative 60S ribosomal protein L17	x	√
Putative ribosomal protein S20	x	√
60S ribosomal protein L11	x	√
40S ribosomal protein S14	x	√
Putative 60S ribosomal protein L28	x	√
40S ribosomal protein S12	x	√
60S ribosomal protein L18a	x	√
Putative 60S ribosomal protein L10a	x	√
Putative 60S ribosomal protein L35	x	√
Putative ribosomal protein L3	x	√
Putative ribosomal protein L24	x	√
40S ribosomal protein S2	x	√
Putative 40S ribosomal protein L14	x	√
Clathrin light chain	x	√
Putative 60S ribosomal protein L21	x	√
Putative 40S ribosomal protein S21	x	√
Putative proteasome regulatory non-ATPase subunit 6	x	√
Putative ribosomal protein L1a	x	√

Therefore, only 1 protein was found as a common protein in structural molecular activity for the passive sample, but for the aggressive sample there is no any protein were detected for the activity. Lastly, 2 proteins were detected as common protein for the passive sample in transmembrane transporter activity while only 1 protein was found for the aggressive sample for the same activity.

On the other hand, other than differential proteins, common proteins between the samples have been investigated in this study. In the first batch 293 protein and in the second batch 179 protein was detected for all 3 sample of *L. major*. In addition to that,

similarities have been analyzed as a binary groups of samples between each other. These implications have been also compared with the literature to find out the reliability of our results. In literature, researches showed that GP63 protein which is also called as *Leishmania* Surface Protease or Leishmanolysin, is the main virulence factor for this parasite since the gene encoding GP63 is highly conserved among other genes.

Table 4.9. Differential proteins that is found according to their transmembrane transporter activities in both first and second batch for passive *L. major*

Protein Names for Passive *L. major*	**Batch 1**	**Batch 2**
Putative vacuolar-type proton translocating pyrophosphatase 1	√	x
Cation-transporting ATPase	√	x
Putative nucleoside transporter 1	√	√
Putative cytochrome c oxidase VII	√	x
Nucleobase transporter	√	x
Putative ATP-binding cassette protein subfamily	√	x
Inosine-guanosine transporter	√	x
Glucose transporter, lmgt2	√	x
Putative mitochondrial phosphate transporter	√	x
Putative ATP-binding cassette protein subfamily C,member 6	√	x
Putative cytochrome c oxidase subunit V	√	x
V-type proton ATPase subunit H	√	x
V-type proton ATPase subunit F	√	x
Cytochrome c oxidase assembly factor-like protein	√	x
Putative ATP-binding cassette protein subfamily A,member 8	√	x
Putative cytochrome c oxidase subunit 10	√	x
Putative ATP synthase, epsilon chain	√	x
Putative cytochrome c oxidase subunit VI	√	√
V-type proton ATPase subunit a	√	x
V-type proton ATPase subunit C	x	√
Possible ABC transporter	x	√
V-type proton ATPase subunit	x	√

Table 4.10. Differential proteins that is found according to their transmembrane transporter activities in both first and second batch for aggressive *L. major*

Protein Names for Aggressive *L. major*	**Batch 1**	**Batch 2**
Putative ATP-binding cassette protein subfamily E,member 1	√	x
Putative ATP synthase	√	x
Putative ATP synthase F1 subunit gamma protein	√	√
Putative cytochrome c oxidase VII	x	√
Putative cytochrome c oxidase VIII	x	√
Putative cytochrome c oxidase subunit V	x	√
V-type proton ATPase subunit F	x	√
Putative ATP synthase, epsilon chain	x	√
V-type proton ATPase subunit a	x	√

GP63 protein is a kind of zinc-dependent metalloprotease which localized at the surface of the parasite though glycophosphatidylinositol (GPI) anchor. The role of the GP63 as a promiscuous enzyme is to provide direct interaction between promastigotes form and the host macrophages receptors since it affect virulence of the parasite. Also, it can degrade the components of extra-cellular matrix (ECM). Moreover, its abundance in the promastigote form of the *Leishmania* is nearly 1% which makes this protein is the most abundant surface protein. Furthermore, in amastigote form, GP63 protein level lower itself but it still remains detectable. Especially, *L. major* and *L. donovani* parasites that have similar pathogenicity, show 80% identity for GP63 sequence. Recent works for GP63 also shows that the enzyme can get access to the macrophage cytoplasm and nucleus while parasite infect the host cell. This might open a door for differences in virulence effect of the *L. major* parasite by understanding the mechanism of GP63 in the host macrophages. In our study, GP63 proteins have been detected in the Mascot results. In the first bath of the *L. major* passive, aggressive and reference samples involve GP63 protein. In addition to that, the one with the different Uniprot code for GP63 was observed only in passive *L. major* sample but its function and the way express itself in the parasite was the same as others. Surprisingly, in the second batch results, while aggressive and reference *L. major* samples involve this GP63 protein, there was no sign for GP63 in passive sample. Another important protein for virulence factor of *L. major* is secreted acid phosphatases (SAPs) which is exist on the surface of the *Leishmania* species since derivatives of this protein have large group of substrates such as phosphorylated sugars and inositol phosphates to phosphorylated proteins. Researches have been showed that this kind of proteins can play an important role for the preparation of the nutrients for *Leishmania* species and in this way, they can modulate the host environment by dephosphorylation. Also, it is shown that SAPs have a strong resistance to proteolytic degradation which makes the parasite is protected from digestive enzymes either in the sand fly midgut or macrophages in the mammalian host and longer its infection in the host. On the other hand in this study, SAPs have been observed parallel to literature in both batches. They have been showed up for all passive, aggressive and reference samples. However, this conclusion could not associated with virulence factor of all samples since SAPs was found common for all samples [35].

Furthermore, studies have established that cysteine proteases (CPs) which are coming from the family of calpain-like proteases, also involve the virulence factor of all kinetoplastids. According to that, Mottram J.C. et al., have focused on the role of CPs in

virulence effect of *L. major*. They have investigated single-copy cysteine protease A (CPA), multi-copy cysteine protease B (CPB) and single-copy cysteine protease C (CPC) which is all belong to Clan CA family C1 of CPs. Most of the CPs have an important role between interactions in the parasite-host group. From this point of view, *L. major* has CPs of eight families within Clan CA. In addition, Clan CA involves the family of C1 which contains CPA, CPB and CPC genes. Other than that, there are many genes encoding a total of 65 CPs for *L. major* parasite. Surprisingly, they have found 27 genes with calpain-like domains of *L. major* apart from C1 family. Calpains have a role in transduction pathways of the mammals and alteration in the cytoskeletal and or membrane attachments. As a result, this investigation in the genomes of the *L. major* provides new signs into the diversity of CPs in the *Leishmania* species since some of them involved in several biological processes in the parasite. Also, D-alanyl-glycyl endopeptidase like protein, metacaspase, pyroglutamyl-peptidase I and PfPI have been found that they have more specific roles in CPs family since their homologs do not exist in mammalian hosts [36]. However, apart from Calpain-like peptidase in second batch and Cysteine proteases in the first batch, there has been no sign for the other CPs in our study. Also, surprisingly, Ubiquitin hydrolase protein which involved in Clan CA of the Family C19 was found in all aggressive, passive and reference samples of *L. major*.

Another protein group that can affect the virulence effect of the *L. major* are peroxiredoxins. They used to reduce peroxide molecules to balance the parasite metabolism since they lack of enzyme catalase. Peroxiredoxins are very important for the parasite's defense in the host and absence of this protein may cause damage in the parasite metabolism. For that reason, redundancy of this protein can affect the level of infectivity and virulence factor of the parasite. The findings in this study also approve the existence of peroxiredoxin by Ascorbate peroxidase protein that was found for both aggressive and passive *L. major* samples.

Besides from all important proteins that were found in literature, cystathionine beta-synthase protein (CBS) was found for only aggressive *L. major* with a property of cysteine synthase activity and pyridoxal phosphate binding. To our best knowledge, these two property may be considered crucial due to their linkage between proteins which were mentioned above; cysteine proteases (CPs) and secreted acid phosphatases (SAPs). Cysteine synthase activity of CBS may strengthen interaction between host and parasite by triggering formation of cysteine amino acids to cysteine proteases. Also, phosphate binding activity of this protein may provide good environment for nutrient supply of

parasite in the host. Also, in another study, Romerio et al., have focused on the upregulation of cysteine synthase and cystathionine beta-synthase proteins effect on *L. brazilliensis* which is similar infectivity with *L. major*. As a conclusion, it was observed that overexpressing of CPs and CBSs provides organism to get resistance for oxidative stress [37]. These decrease in oxidative stress of the parasite also can affect parasite infectivity strongly and at the end, parasite can show its virulence effect better on the host.

Finally, if these findings are taken into account, it can be proposed that cystathionine beta-synthase protein may play important role in the virulence effect of *L. major* in the host environment.

CHAPTER 5

CONCLUSION

Several proteins and enzyme mechanism have a crucial role in the infection of *Leishmania* species. This virulence effect can be based on all biological processes that continue to form in the body of the parasite. As the proteins are originated from different gene encoding, any alteration in the genomic level can affect the virulence factor of the parasite as well. Therefore, in this study, a comprehensive proteomic analysis was investigated among *L. major* samples. Comparisons have been achieved among passive *L. major* isolate, aggressive *L. major* isolate and as a reference *L. major* strain. Also, to our knowledges, this study is the first proteomic study that was achieved with clinical isolates of *L. major* species.

All experiments have been done with 2 technical repeat of all passive, aggressive and reference samples. LC-MS/MS provides sensitive and reliable results for the proteome analysis among all three samples. According to obtained results, for the first batch, 293 common proteins among all was detected while 27 protein for aggressive, 155 protein for passive and 134 protein for reference sample have been detected as the differential. On the contrary, for the second batch, 179 proteins were found as common among all and 122 protein for aggressive, 47 protein for passive and 81 protein for reference sample were found as differential proteins. These differences in protein level between two batches might originate from either sample preparation parts such as in solution digestion or zip tip process or the process in parasite growth and protein extraction due to the fact that both batches have been handled at different times. Also, obtained results from differences have been detailed according to their molecular function distribution. Our findings showed that there are repetitive proteins between two batches and conclude that these proteins might have a more significant role in the virulence factor of *L. major* parasite.

As a result, preliminary comparisons among all samples showed that there is conflict results with the literature about GP63, SAPs, CPs and peroxiredoxin proteins existence. Putative major surface protease gp63, inositol-3-phosphate synthase, putative

calpain-like cysteine peptidase, cysteine peptidase C, ascorbate peroxidase was found in all *L. major* samples. However, their enzymatic activities among *L. major* cannot be quantified in our study due to quantitation problems with shotgun analysis enzymatic activity changes associated with these proteins could not be detected but quantitation which requires at least 5 batch of protein samples can be taken into account as a future aspect of the work.

Also, associated to this work, in the aggressive *L. major* batches cystathionine beta-synthase protein which has a role to the synthesis of the Cysteine amino acid was found in only aggressive *L. major* different that passive *L. major* samples. It is proposed that this protein may have an important role on the virulence effect of *L. major* due to its cysteine synthase activity and pyridoxal phosphate binding properties. On the other hand, also in the literature search, it was observed that overexpressing of CPs and CBSs provides organism to get resistance for oxidative stress.

Finally, our results have consistency for all literature studies mentioned above at the proteomic level. At this point of view, it can be said that expected proteins that is associated with the virulence factor of the *L. major*, have been detected and also some new proteins were proposed for virulence effect of the parasite. However, the mechanism of the virulence effect of this parasite still remains unclear.

REFERENCES

1. III, J. R. Y., A century of mass spectrometry: from atoms to proteomes. Natural Method 2011, 8, 633-637.

2. Griffiths, J., A brief history of mass spectrometry. Anal Chem 2008, 80 (15), 5678-83.

3. Tanaka, K. W., H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. , Protein and polymer analyses up to m/z 100,000 by laser ionization time-of-flight mass spectrometry. . Rapid Commun. Mass Spectrom. 1988, 2, 151-153.

4. Karas, M.; Hillenkamp, F., Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Analytical Chemistry 1988, 60 (20), 2299-2301.

5. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray ionization for mass spectrometry of large biomolecules. Science (New York, N.Y.) 1989, 246 (4926), 64-71.

6. Sabido, E.; Selevsek, N.; Aebersold, R., Mass spectrometry-based proteomics for systems biology. Current opinion in biotechnology 2012, 23 (4), 591-7.

7. Sawicki, M. P.; Samara, G.; Hurwitz, M.; Passaro, E., Jr., Human Genome Project. The American Journal of Surgery 1993, 165 (2), 258-264.

8. Sharma, G., The Human Genome Project and its promise. Journal of Indian College of Cardiology 2012, 2 (1), 1-3.

9. Graves, P. R.; Haystead, T. A. J., Molecular biologist's guide to proteomics. Microbiology and molecular biology reviews : MMBR 2002, 66 (1), 39-63.

10. Smith, J. B., Peptide Sequencing by Edman Degradation. ENCYCLOPEDIA OF LIFE SCIENCES 2001.

11. Kurien, B. T.; Scofield, R. H., Western blotting. Methods (San Diego, Calif.) 2006, 38 (4), 283-93.

12. Han, X.; Aslanian, A.; Yates, J. R., 3rd, Mass spectrometry for proteomics. Current opinion in chemical biology 2008, 12 (5), 483-490.

13. Jonsson, A. P., Mass spectrometry for protein and peptide characterisation. Cellular and molecular life sciences : CMLS 2001, 58 (7), 868-84.

14. Vincent Stroobant, E. d. H., Mass Spectrometry Principles and Applications, Ion sources. John Wiley & Sons: 2007.

15. El-Aneed, A.; Cohen, A.; Banoub, J., Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. Applied Spectroscopy Reviews 2009, 44 (3), 210-230.

16. Ho, C. S.; Lam, C. W. K.; Chan, M. H. M.; Cheung, R. C. K.; Law, L. K.; Lit, L. C. W.; Ng, K. F.; Suen, M. W. M.; Tai, H. L., Electrospray ionisation mass spectrometry: principles and clinical applications. The Clinical biochemist. Reviews 2003, 24 (1), 3-12.

17. Banerjee, S.; Mazumdar, S., Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte %J International Journal of Analytical Chemistry. 2012, 2012, 40.

18. Lin, D.; Tabb, D. L.; Yates, J. R., 3rd, Large-scale protein identification using mass spectrometry. Biochimica et biophysica acta 2003, 1646 (1-2), 1-10.

19. Domon, B.; Aebersold, R., Mass Spectrometry and Protein Analysis. Science (New York, N.Y.) 2006, 312 (5771), 212.

20. Matallana-Surget, S.; Leroy, B.; Wattiez, R., Shotgun proteomics: concept, key points and data mining. Expert review of proteomics 2010, 7 (1), 5-7.

21. Wu, C. C.; MacCoss, M. J., Shotgun proteomics: tools for the analysis of complex biological systems. Current opinion in molecular therapeutics 2002, 4 (3), 242-50.

22. Yang, F.; Shen, Y.; Camp, D. G., 2nd; Smith, R. D., High-pH reversed-phase chromatography with fraction concatenation for 2D proteomic analysis. Expert review of proteomics 2012, 9 (2), 129-134.

23. Wysocki, V. H.; Resing, K. A.; Zhang, Q.; Cheng, G., Mass spectrometry of peptides and proteins. Methods (San Diego, Calif.) 2005, 35 (3), 211-22.

24. Chu, I. K.; Siu, C.-K.; Lau, J. K.-C.; Tang, W. K.; Mu, X.; Lai, C. K.; Guo, X.; Wang, X.; Li, N.; Xia, Y.; Kong, X.; Oh, H. B.; Ryzhov, V.; Tureček, F.; Hopkinson, A. C.; Siu, K. W. M., Proposed nomenclature for peptide ion fragmentation. International Journal of Mass Spectrometry 2015, 390, 24-27.

25. Peptide Fragmentation Nomenclature b, y and a ions. In De Novo Peptide Sequencing Tutorial, 2016.

26. In Understanding Mascot Reports, MS/MS ion search results, matrixscience.

27. Pawar, H.; Pai, K.; Patole, M. S., A novel protein coding potential of long intergenic non-coding RNAs (lincRNAs) in the kinetoplastid protozoan parasite Leishmania major. Acta tropica 2017, 167, 21-25.

28. Machado-Pinto, J.; Laborne, L., 5 - Leishmaniasis. In Tropical Dermatology (Second Edition), Tyring, S. K.; Lupi, O.; Hengge, U. R., Eds. Elsevier: 2017; pp 42-49.

29. Donelson, J. E.; Gardner, M. J.; El-Sayed, N. M., More surprises from Kinetoplastida. Proceedings of the National Academy of Sciences of the United States of America 1999, 96 (6), 2579-2581.

30. Hailu, A.; Musa, A. M.; Royce, C.; Wasunna, M., Visceral leishmaniasis: new health tools are needed. PLoS medicine 2005, 2 (7), e211.

31. Bifeld, E.; Chrobak, M.; Zander, D.; Schleicher, U.; Schonian, G.; Clos, J., Geographical sequence variation in the Leishmania major virulence factor P46. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases 2015, 30, 195-205.

32. Özbilgin, A.; Çulha, G.; Uzun, S.; Harman, M.; Topal, S. G.; Okudan, F.; Zeyrek, F.; Gündüz, C.; Östan, İ.; Karakuş, M.; Töz, S.; Kurt, Ö.; Akyar, I.; Erat, A.; Güngör, D.; Kayabaşı, Ç.; Çavuş, İ.; Bastien, P.; Pratlong, F.; Kocagöz, T.; Özbel, Y., Leishmaniasis in Turkey: first clinical isolation of Leishmania major from 18 autochthonous cases of cutaneous leishmaniasis in four geographical regions. Tropical Medicine & International Health 2016, 21 (6), 783-791.

33. Manza, L. L.; Stamer, S. L.; Ham, A. J.; Codreanu, S. G.; Liebler, D. C., Sample preparation and digestion for proteomic analyses using spin filters. Proteomics 2005, 5 (7), 1742-5.

34. Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C., Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. Journal of separation science 2005, 28 (14), 1694-703.

35. Atayde, V. D.; Hassani, K.; da Silva Lira Filho, A.; Borges, A. R.; Adhikari, A.; Martel, C.; Olivier, M., Leishmania exosomes and other virulence factors: Impact on innate immune response and macrophage functions. Cellular immunology 2016, 309, 7-18.

36. Joshi, P. B.; Kelly, B. L.; Kamhawi, S.; Sacks, D. L.; McMaster, W. R., Targeted gene deletion in Leishmania major identifies leishmanolysin (GP63) as a virulence factor. Molecular and biochemical parasitology 2002, 120 (1), 33-40.

37. Romero, I.; Téllez, J.; Romanha, A. J.; Steindel, M.; Grisard, E. C., Upregulation of Cysteine Synthase and Cystathionine β-Synthase Contributes to Leishmania braziliensis Survival under Oxidative Stress. Antimicrobial agents and chemotherapy 2015, 59 (8), 4770-4781.

APPENDIX A

COMMON PROTEINS THAT IS FOUND IN PASSIVE, AGGRESSIVE AND REFERENCE *L. major* GROUPS IN BOTH BATCHES

Table A.1. Common Proteins that is found in all *L. major* samples in both batches

Protein Names	**Batch 1**	**Batch 2**
14-3-3 protein I	√	√
2,3-bisphosphoglycerate-independent phosphoglycerate mutase	√	√
40S ribosomal protein S3a	√	√
40S ribosomal protein S4	√	√
40S ribosomal protein S5	√	√
40S ribosomal protein S6	√	√
40S ribosomal protein S7	√	√
40S ribosomal protein S8	√	√
60S acidic ribosomal protein P0	√	√
60S acidic ribosomal protein P2	√	√
Acetyl-coenzyme A synthetase	√	√
Aconitate hydratase	√	√
Actin	√	√
Activated protein kinase c receptor (Fragment)	√	√
Adenosylhomocysteinase	√	√
ADP-ribosylation factor 1	√	√
ATP synthase subunit beta	√	√
ATPase alpha subunit	√	√
Chaperonin HSP60, mitochondrial	√	√
Clathrin heavy chain	√	√
Cysteine peptidase C (CPC)	√	√
Cytochrome c oxidase subunit IV	√	√
Dihydrolipoamide acetyltransferase component of pyruvate	√	√
Dipeptidyl peptidase 3	√	√
D-isomer specific 2-hydroxyacid dehydrogenase-like protein	√	√
Elongation factor 1-alpha	√	√
Elongation factor 1-beta	√	√
Elongation factor 2	√	√
Enolase	√	√
Fructose-bisphosphate aldolase	√	√
Glucose-6-phosphate isomerase	√	√
Glyceraldehyde-3-phosphate dehydrogenase	√	√
Glycosomal malate dehydrogenase	√	√
Heat shock protein 70-related protein	√	√
Heat shock protein 83-1	√	√
Isocitrate dehydrogenase [NADP]	√	√
META domain containing protein	√	√

(cont. on next page)

Table A.1. (cont.)

Protein Names	Batch 1	Batch 2
Metallo-peptidase, Clan ME, Family M16	√	√
Mitochondrial RNA binding protein 1	√	√
NAD-specific glutamate dehydrogenase	√	√
Nucleoside diphosphate kinase	√	√
Paraflagellar rod protein 2C	√	√
Peptidyl-prolyl cis-trans isomerase	√	√
Peptidylprolyl isomerase	√	√
Peroxidoxin	√	√
Phosphoglycerate kinase	√	√
Plasma membrane ATPase	√	√
Proteasome subunit alpha type	√	√
Protein disulfide-isomerase	√	√
Putative 10 kDa heat shock protein	√	√
Putative 3,2-trans-enoyl-CoA isomerase mitochondrial	√	√
Putative 40S ribosomal protein S33	√	√
Putative 60S acidic ribosomal protein P2	√	√
Putative 60S ribosomal protein L10	√	√
Putative 60S ribosomal protein L12	√	√
Putative 60S ribosomal protein L13	√	√
Putative 60S ribosomal protein L13a	√	√
Putative 60S ribosomal protein L5	√	√
Putative 60S ribosomal protein L7a	√	√
Putative 60S ribosomal protein L9	√	√
Putative aldose 1-epimerase	√	√
Putative aminopeptidase	√	√
Putative ATP-dependent Clp protease subunit,heat shock protein 78	√	√
Putative ATP-dependent RNA helicase	√	√
Putative calpain-like cysteine peptidase	√	√
Putative delta-1-pyrroline-5-carboxylate dehydrogenase	√	√
Putative dynein heavy chain	√	√
Putative endoribonuclease L-PSP (Pb5)	√	√
Putative eukaryotic initiation factor 4a	√	√
Putative glucose-regulated protein 78	√	√
Putative glycosomal phosphoenolpyruvate carboxykinase	√	√
Putative heat shock 70-related protein 1,mitochondrial	√	√
Putative heat shock protein	√	√
Putative heat-shock protein hsp70	√	√
Putative lipophosphoglycan biosynthetic protein	√	√
Putative microtubule-associated protein	√	√
Putative nucleolar protein	√	√
Putative proteasome activator protein pa26	√	√
Putative pyruvate dehydrogenase E1 component alpha subunit	√	√
Putative ribonucleoprotein p18, mitochondrial	√	√
Putative ribosomal protein L27	√	√
Putative RNA helicase	√	√
Putative small myristoylated protein-3	√	√
Putative T-complex protein 1, theta subunit	√	√
Putative trypanothione synthetase	√	√
Putative ubiquitin/ribosomal protein S27a	√	√
Putative vacuolar ATP synthase catalytic subunit A	√	√
Pyruvate dehydrogenase E1 component subunit beta	√	√
Pyruvate kinase	√	√

(cont. on next page)

Table A.1. (cont.)

Protein Names	Batch 1	Batch 2
Succinate--CoA ligase [ADP-forming] subunit beta, mitochondrial	√	√
Metallo-peptidase, Clan MA(E), Family M32	√	√
Succinyl-CoA:3-ketoacid-coenzyme A transferase	√	√
Superoxide dismutase	√	√
Thiol-specific antioxidant (Fragment)	√	√
Trypanothione reductase	√	√
Tryparedoxin	√	√
Tubulin alpha chain	√	√
Tubulin beta chain	√	√
Ubiquitin-60S ribosomal protein L40	√	√
SHERP	√	√
4-coumarate:coa ligase-like protein	√	x
Putative 60S ribosomal protein L21	√	x
Putative 40S ribosomal protein S11	√	x
Glucose-6-phosphate 1-dehydrogenase	√	x
Inositol-3-phosphate synthase	√	x
Putative proteasome regulatory non-ATPase subunit	√	x
Putative phenylalanyl-tRNA synthetase	√	x
Putative threonyl-tRNA synthetase	√	x
Phosphodiesterase	√	x
Chaperonin subunit alpha	√	x
Proliferating cell nuclear antigen	√	x
Putative acyl-CoA dehydrogenase	√	x
Adenylosuccinate synthetase	√	x
Myosin XXI	√	x
Putative serine peptidase	√	x
Putative proteasome regulatory non-ATP-ase subunit	√	x
Putative heat shock protein 20	√	x
Methyltransferase	√	x
Putative nucleolar RNA binding protein	√	x
Putative IgE-dependent histamine-releasing factor	√	x
Triosephosphate isomerase	√	x
Putative 60S ribosomal subunit protein L31	√	x
Putative 2,4-dihydroxyhept-2-ene-1,7-dioic acid aldolase	√	x
Putative heat shock protein DNAJ	√	x
Putative adenylate kinase	√	x
Ribosomal protein L15	√	x
Calcium-transporting ATPase	√	x
Aspartyl putative aminopeptidase	√	x
Putative arginyl-tRNA synthetase	√	x
Calmodulin-like protein	√	x
Universal minicircle sequence binding protein	√	x
Putative paraflagellar rod protein 1D	√	x
Putative ATP synthase	√	x
Putative prolyl oligopeptidase	√	x
Putative 60S ribosomal protein L35	√	x
Transmembrane 9 superfamily member	√	x
Splicing factor ptsr1-like protein	√	x
Putative proteasome regulatory ATPase subunit 1	√	x
Putative 40S ribosomal protein S27-1	√	x
40S ribosomal protein S24	√	x
Putative ribosomal protein L3	√	x

(cont. on next page)

Table A.1. (cont.)

Protein Names	Batch 1	Batch 2
60S ribosomal protein L18a	√	x
Putative 2-oxoglutarate dehydrogenase E1 component	√	x
Putative aminopeptidase P	√	x
Putative glutamyl-tRNA synthetase	√	x
Putative 40S ribosomal protein L14	√	x
60S ribosomal protein L11	√	x
Aminopeptidase	√	x
Aldehyde dehydrogenase, mitochondrial	√	x
Rab GDP dissociation inhibitor	√	x
Putative chaperone protein DNAj	√	x
Putative adenosine kinase	√	x
Eukaryotic translation initiation factor 3 subunit I	√	x
Putative 60S ribosomal protein L10a	√	x
Putative ribosomal protein S20	√	x
Putative sm-f snRNP core complex protein	√	x
ATP/ADP translocase	√	x
40S ribosomal protein S12	√	x
Transketolase	√	x
Putative 40S ribosomal protein S21	√	x
Adenylosuccinate lyase	√	x
Elongation factor 2 (Fragment)	√	x
Elongation factor 1B gamma	√	x
Glutamate dehydrogenase	√	x
14-3-3 protein II	√	x
GTP-binding nuclear protein	√	x
Lysine--tRNA ligase	√	x
Histone H2A	√	x
Putative 60S ribosomal protein L17	√	x
Mannose-6-phosphate isomerase	√	x
Putative 40S ribosomal protein S17	√	x
Kinetoplast membrane protein 11	√	x
Putative RNA binding protein	√	x
Succinate--CoA ligase [ADP-forming] subunit alpha, mitochondrial	√	x
HASPA1	√	x
GP63, leishmanolysin	√	x
Putative glycerol kinase, glycosoma	√	x
Nascent polypeptide associated complex subunit-like protein, copy	√	x
19S proteasome regulatory subunit	√	x
Putative ATP-dependent DEAD-box RNA helicase	√	x
Putative isoleucyl-tRNA synthetase	√	x
Putative vacuolar ATP synthase subunit b	√	x
Cytochrome c oxidase subunit I	√	x
Probable citrate synthase, mitochondrial	√	x
Macrophage migration inhibitory factor-like protein	√	x
Possible 3-ketoacyl-CoA thiolase	√	x
Dihydrolipoyl dehydrogenase	√	x
Probable ubiquitin-conjugating enzyme e2-17 kDa	√	x
Putative eukaryotic translation initiation factor 3 subunit 8	√	x
Putative calreticulin	√	x
Aspartate aminotransferase	√	x
Putative Transitional endoplasmic reticulum ATPase	√	x
Acetylornithine deacetylase-like protein	√	x

(cont. on next page)

Table A.1. (cont.)

Protein Names	Batch 1	Batch 2
Putative kinesin	√	x
Putative cytochrome c1, heme protein,mitochondrial	√	x
Putative seryl-tRNA synthetase	√	x
Histone H2B	√	x
Biotin/lipoate protein ligase-like protein	√	x
Regulatory subunit of protein kinase a-like protein	√	x
Uncharacterized protein L7836.08	√	x
Putative surface antigen protein 2	√	x
Putative ADP ribosylation factor 3	√	x
Metallo-peptidase, Clan MA(E), family 32	√	x
Putative eukaryotic translation initiation factor 1A	√	x
Putative glycyl tRNA synthetase	√	x
Alanine--tRNA ligase	√	x
Putative 40S ribosomal protein S9	√	x
Ribosomal protein S25	√	x
T-complex protein 1 subunit gamma	√	x
Putative 60S ribosomal protein L27A/L29	√	x
Putative 40S ribosomal protein S18	√	x
Putative ADP-ribosylation factor	√	x
Putative 60S ribosomal protein L34	√	x
Putative T-complex protein 1, eta subunit	√	x
Malate dehydrogenase	√	x
40S ribosomal protein S2	√	x
Putative nima-related protein kinase	√	x
Putative udp-glc 4'-epimerase	√	x
6-phosphogluconate dehydrogenase, decarboxylating (Fragment)	√	x
Coatomer subunit gamma	√	x
Ribonucleoside-diphosphate reductase	√	x
Prostaglandin F synthase	√	x
Infective insect stage-specific protein	√	x
Putative 1,2-Dihydroxy-3-keto-5-methylthiopentene dioxygenase	√	x
Putative eukaryotic translation initiation factor 2 subunit	√	x
Putative p450 reductase	√	x
Putative intraflagellar transport protein IFT88	√	x
Coatomer subunit alpha	√	x
Putative long-chain-fatty-acid-CoA ligase	√	x
Putative SNAP protein	√	x
UV excision repair RAD23-like protein	√	x
S-adenosylmethionine synthase	√	x
Putative long chain fatty Acyl CoA synthetase	√	x
Putative ATP-dependent DEAD/H RNA helicase	√	x
Putative sodium stibogluconate resistance protein	√	x
Putative RNA-binding protein, UPB2	√	x
Putative aldehyde dehydrogenase	√	x
Serine palmitoyltransferase 1	√	x
Uncharacterized protein L8530.05	√	x
Ascorbate peroxidase	√	x
Stress-inducible protein STI1 homolog	√	x
UTP--glucose-1-phosphate uridylyltransferase	√	x
Putative 3-methylcrotonoyl-CoA carboxylase beta subunit	√	x
Phosphatidate cytidylyltransferase	√	x

(cont. on next page)

Table A.1. (cont.)

Protein Names	Batch 1	Batch 2
Galactokinase-like protein	√	x
Ca2+-binding EF-hand protein	√	x
Dihydrolipoamide acetyltransferase	√	x
Putative cytochrome c oxidase copper chaperone	√	x
Casein kinase 1 isoform 2	√	x
Putative ATP synthase F1 subunit gamma protein	√	x
Putative electron transfer flavoprotein	√	x
Cystathionine beta-synthase	√	x
Putative ras-like small GTPases	√	x
Putative small GTP-binding protein Rab11	√	x
Thiolase protein-like protein	x	√
Diphosphomevalonate decarboxylase	x	√
Putative mitochondrial RNA binding protein	x	√
Putative calmodulin	x	√
Putative vacuolar-type proton translocating pyrophosphatase 1	x	√
Putative 40S ribosomal protein S19 protein	x	√
Putative 40S ribosomal protein S23	x	√
GrpE protein homolog	x	√
Putative short chain dehydrogenase	x	√
Uncharacterized protein	x	√
5-methyltetrahydropteroyltriglutamate-homocystein e S-methyltransferase	x	√
Putative cystathione gamma lyase	x	√
Putative cytoskeleton-associated protein CAP5.5	x	√
Putative 60S ribosomal protein L2	x	√
Putative 60S ribosomal protein L23a	x	√
Proteasome endopeptidase complex	x	√
Putative 40S ribosomal protein S13	x	√
Putative T-complex protein 1, beta subunit	x	√
Cathepsin L-like protease	x	√
Putative small GTP-binding protein Rab1	x	√
Polyadenylate-binding protein	x	√
Putative nucleosome assembly protein	x	√
Tubulin-specific chaperone A	√	x
Thioredoxin-like protein	√	x
Putative p22 protein	√	x
60S ribosomal protein L18	√	x
Putative Qc-SNARE protein	√	x
Putative asparaginyl-tRNA synthetase	√	x
Transaldolase	√	x

APPENDIX B

REFERENCE *L. major* COMPARISON WITH PASSIVE AND AGGRESSIVE *L. major* FOR BOTH BATCHES

Table B.1. Common and Differential Proteins between Reference, Passive and Aggressive *L. major* for Batch 1

Protein Names	**Aggressive *L. major***	**Passive *L. major***
Putative 3-methylcrotonoyl-CoA carboxylase beta subunit	√	x
Phosphatidate cytidylyltransferase	√	x
Tubulin-specific chaperone A	√	x
Thioredoxin-like protein	√	x
Putative p22 protein	√	x
Galactokinase-like protein	√	x
Ca2+-binding EF-hand protein	√	x
Dihydrolipoamide acetyltransferase	√	x
Putative cytochrome c oxidase copper chaperone	√	x
Casein kinase 1 isoform 2	√	x
Putative ATP synthase F1 subunit gamma protein	√	x
Putative electron transfer flavoprotein	√	x
Cystathionine beta-synthase	√	x
Mitochondrial RNA binding protein 1	√	x
Putative ras-like small GTPases	√	x
ADP-ribosylation factor 1	√	x
Putative small GTP-binding protein Rab11	√	x
40S ribosomal protein S5	√	x
Putative ubiquitin/ribosomal protein S27a	√	x
Phosphotransferase	x	√
Putative carnitine/choline acetyltransferase	x	√
Putative ER--golgi transport protein gp25L	x	√
Putative 60S ribosomal protein L23	x	√
Putative heat shock protein HslVU, ATPase,subunit HslU	x	√
Putative small myristoylated protein 4	x	√
Elongation factor 1B alpha	x	√
Putative cytochrome c oxidase subunit V	x	√
Serine/threonine-protein phosphatase	x	√
Replication protein A subunit	x	√
Putative 60S Ribosomal protein L36	x	√
Putative chaperonin containing t-complex protein	x	√
Putative mitochondrial RNA binding protein	x	√
Coatomer subunit delta	x	√
Putative ribosomal protein l35a	x	√
Mitogen-activated protein kinase	x	√
Putative acetyl-CoA carboxylase	x	√
Putative N-acyl-L-amino acid amidohydrolase	x	√
Mevalonate kinase	x	√
Putative serine/threonine-protein kinase	x	√

(cont. on next page)

Table B.1. (cont.)

Protein Names	Aggressive *L. major*	Passive *L. major*
Cytochrome c	x	√
Putative vacuolar-type proton translocating pyrophosphatase 1	x	√
Putative ATP-binding cassette protein subfamily F,member 2	x	√
Putative beta eliminating lyase	x	√
Putative NADH-dependent fumarate reductase	x	√
Putative proteasome regulatory ATPase subunit 5	x	√
Putative nuclear transport factor 2	x	√
Putative histone H3	x	√
Putative leucyl-tRNA synthetase	x	√
Putative small ubiquitin protein	x	√
Proteasome regulatory ATPase subunit	x	√
Elongation of fatty acids protein	x	√
Succinyl-diaminopimelate desuccinylase-like protein	x	√
Putative dynein	x	√
Putative leucine carboxyl methyltransferase	x	√
Putative 40S ribosomal protein S3	x	√
Proteasome subunit beta	x	√
Phosphomannomutase	x	√
Putative peptidase t	x	√
Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	x	√
Putative calmodulin-related protein	x	√
Putative chaperonin TCP20	x	√
Stress-induced protein sti1	x	√
Putative methionyl-tRNA synthetase	x	√
Putative glutamine synthetase	x	√
Putative ATP synthase, epsilon chain	x	√
Putative cystathione gamma lyase	x	√
Putative GTP-binding protein	x	√
Putative 60S ribosomal protein L23a	x	√
Putative reiske iron-sulfur protein	x	√
Oligopeptidase b	x	√
Putative cytochrome c oxidase subunit VI	x	√
Ubiquitin carboxyl-terminal hydrolase	x	√
T-complex protein 1 subunit delta	x	√
Putative NADH:flavin oxidoreductase/NADH oxidase	x	√
Formate--tetrahydrofolate ligase	x	√
Putative 40S ribosomal protein S15A	x	√
Inosine-5'-monophosphate dehydrogenase	x	√
Putative carbamoyl-phosphate synthase	x	√
Putative 60S ribosomal protein L26	x	√
Cathepsin L-like protease	x	√
Ribokinase	x	√
Intraflagellar transport protein-like protein	x	√
Putative 40S ribosomal protein S15	x	√
Adenine aminohydrolase	x	√
Putative lanosterol 14-alpha-demethylase	x	√
Small nuclear ribonucleoprotein SmD2	x	√
Putative small GTP-binding protein Rab1	x	√
Putative mitogen-activated protein kinase	x	√

(cont. on next page)

Table B.1. (cont.)

Protein Names	Aggressive *L. major*	Passive *L. major*
60S ribosomal protein L30	x	√
Putative fumarate hydratase	x	√
Putative phosphatase 2C	x	√
Eukaryotic translation initiation factor 3 subunit L	x	√
Tyrosyl or methionyl-tRNA synthetase-like protein	x	√
Putative nucleosome assembly protein	x	√
Putative glutamamyl carboxypeptidase	x	√
60S ribosomal protein L37a	x	√
Putative valyl-tRNA synthetase	x	√
Putative proteasome regulatory ATPase subunittcc1l8.3	x	√
Thiolase protein-like protein	x	√
Putative 60S ribosomal protein L7	x	√
ATP-dependent 6-phosphofructokinase	x	√
Proteasome subunit beta type	x	√
Putative high mobility group protein homolog tdp-1	x	√
Putative 40S ribosomal protein S10	x	√
GMP reductase	x	√
Putative calmodulin	x	√
Kinesin-like protein	x	√
Putative proteasome regulatory non-ATPase subunit 6	x	√
Putative ubiquitin hydrolase	x	√
Putative RNA-binding protein	x	√
60S ribosomal protein L32	x	√
1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	x	√
Dynein light chain roadblock	x	√
40S ribosomal protein SA	x	√
Similarity to endo-1-like protein	x	√
Trifunctional enzyme alpha subunit, mitochondrial-like protein	x	√
Alanine aminotransferase	x	√
Nascent polypeptide-associated complex subunit beta	x	√
Putative aspartate aminotransferase	x	√
Putative ribosomal protein L24	x	√
Pyruvate, phosphate dikinase	x	√
Putative 2-oxoglutarate dehydrogenase subunit	x	√
Histone H4	x	√
Transcription elongation factor-like protein	x	√
Putative 40S ribosomal protein S19 protein	x	√
Putative ribosomal protein L1a	x	√
Putative proteasome regulatory non-ATP-ase subunit 2	x	√
Putative 40S ribosomal protein S23	x	√
Serine hydroxymethyltransferase	x	√
Putative dynein light chain, flagellar outer arm	x	√
GrpE protein homolog OS	x	√
ATPase ASNA1 homolog	x	√
Putative cytochrome b5-like protein	x	√
Eukaryotic translation initiation factor 6	x	√
Putative vacuolar protein sorting-associated protein 45	x	√
Uncharacterized protein	x	√
Putative 60S ribosomal protein L28	x	√

(cont. on next page)

Table B.1. (cont.)

Protein Names	**Aggressive** ***L. major***	**Passive** ***L. major***
Putative cysteinyl-tRNA synthetase	x	√
Nucleolar protein family a member-like protein	x	√
Putative 40S ribosomal protein S16	x	√
Cysteine synthase	x	√
NADH-cytochrome b5 reductase	x	√
Prefoldin 5-like protein	x	√
5-methyltetrahydropteroyltriglutamate-homocystein e S-methyltransferase	x	√
Putative U-box domain protein	x	√
Putative cytoskeleton-associatedprotein CAP5.5	x	√
Putative 60S ribosomal protein L2	x	√
Putative dipeptylcarboxypeptidase	x	√
Putative prolyl-tRNA synthetase	x	√
ADF/Cofilin	x	√
Putative deoxyribose-phosphate aldolase	x	√
Putative thimet oligopeptidase	x	√
Putative cleavage and polyadenylation specificity factor	x	√
Proteasome endopeptidase complex	x	√
Putative cytochrome c oxidase subunit 10	x	√
Cysteine protease (Fragment)	x	√
Putative 26S protease regulatory subunit	x	√
Cyclophilin 40	x	√
Eukaryotic translation initiation factor 3 subunit 7-like protein	x	√
Putative mitochondrial phosphate transporter	x	√
Putative prefoldin subunit 2	x	√
Putative 60S ribosomal protein L19	x	√
LmRab7 GTP-binding protein	x	√
I/6 autoantigen-like protein	x	√
Thiol-dependent reductase 1	x	√
Adenylyl cyclase-associated protein	x	√
Tyrosine aminotransferase	x	√
Polyadenylate-binding protein	x	√
Putative isovaleryl-coA dehydrogenase	x	√
Putative mitotubule-associated protein Gb4	x	√
Putative Pyridoxal kinase	x	√
Eukaryotic translation initiation factor 3 subunit E	x	√
Prefoldin subunit 4	x	√
Putative small glutamine-rich tetratricopeptiderepeat protein	x	√
RuvB-like helicase	x	√
Riboflavin kinase/fmn adenylyltransferase-like protein	x	√
Mannose-1-phosphate guanyltransferase	x	√
Putative ubiquitin-activating enzyme e1	x	√
Putative CCR4 associated factor	x	√
Ubiquitin-conjugating enzyme-like protein	x	√
Putative centromere/microtubule binding protein cbf5	x	√

Table B.2. Common and Differential Proteins between Reference Passive and Aggressive *L. major* for Batch 2

Protein Names	**Aggressive *L. major***	**Passive *L. major***
Phosphotransferase	√	x
Glutamate dehydrogenase	√	x
4-coumarate:coa ligase-like protein	√	x
Putative 60S ribosomal protein L21	√	x
GTP-binding nuclear protein OS	√	x
Putative 60S ribosomal protein L17	√	x
Histone H2A	√	x
Amastin-like surface protein-like protein	√	x
ATP-dependent 6-phosphofructokinase	√	x
Kinetoplast membrane protein 11	√	x
Succinate--CoA ligase [ADP-forming] subunit alpha, mitochondrial	√	x
Putative RNA binding protein	√	x
Proteasome subunit beta type	√	x
HASPA1	√	x
GP63, leishmanolysin	√	x
Myosin XXI	√	x
Putative serine peptidase	√	x
Prostaglandin F synthase	√	x
Putative ubiquitin hydrolase	√	x
Putative heat shock protein 20	√	x
Methyltransferase	√	x
Infective insect stage-specific protein	√	x
Putative 2,4-dihydroxyhept-2-ene-1,7-dioic acid aldolase	√	x
Putative heat shock protein DNAJ	√	x
Putative isoleucyl-tRNA synthetase	√	x
Putative vacuolar ATP synthase subunit b	√	x
Ubiquitin-conjugating enzyme-like protein	√	x
Histone H4	√	x
Putative fumarate hydratase	√	x
Putative paraflagellar rod protein 1D	√	x
Putative ribosomal protein L1a	√	x
Putative intraflagellar transport protein IFT88	√	x
Putative leucyl-tRNA synthetase	√	x
Hydrophilic acylated surface protein b	√	x
Putative Transitional endoplasmic reticulum ATPase	√	x
Putative kinesin	√	x
Putative ribosomal protein L3	√	x
Putative 60S ribosomal protein L28	√	x
Phosphomannomutase	√	x
Putative ATP-dependent DEAD/H RNA helicase	√	x
60S ribosomal protein L18a	√	x
40S ribosomal protein S14	√	x
Putative 40S ribosomal protein L14	√	x
60S ribosomal protein L11	√	x

(cont. on next page)

Table B.2. (cont.)

Protein Names	**Aggressive *L. major***	**Passive *L. major***
Aldehyde dehydrogenase, mitochondrial	√	x
Putative GTP-binding protein	√	x
Biotin/lipoate protein ligase-like protein	√	x
Regulatory subunit of protein kinase a-like protein	√	x
Uncharacterized protein L7836.08	√	x
ADF/Cofilin	√	x
Putative ADP ribosylation factor 3	√	x
Metallo-peptidase, Clan MA(E), family 32	√	x
Putative thimet oligopeptidase	√	x
Putative eukaryotic translation initiation factor 1A	√	x
Putative ribosomal protein S20	√	x
Eukaryotic translation initiation factor 3 subunit 7-like protein	√	x
Putative ATP synthase F1 subunit gamma protein	√	x
40S ribosomal protein S12	√	x
UDP-galactopyranose mutase	√	x
Stress-inducible protein STI1 homolog	√	x
Transketolase	√	x
Putative ADP-ribosylation factor	√	x
Adenylosuccinate lyase	√	x
Malate dehydrogenase	√	x
40S ribosomal protein S2	√	x
6-phosphogluconate dehydrogenase, decarboxylating (Fragment)	√	x
Putative 2-hydroxy-3-oxopropionate reductase	x	√
Inositol-3-phosphate synthase	x	√
Putative epsin	x	√
Metallo-peptidase, Clan MA(E), Family M3	x	√
Putative chaperonin containing t-complex protein	x	√
Citrate synthase	x	√
Putative acyl-CoA dehydrogenase	x	√
Putative RNA binding protein rbp16	x	√
Pyruvate, phosphate dikinase	x	√
Putative 2-oxoglutarate dehydrogenase subunit	x	√
Putative small nuclear ribonucleoprotein	x	√
Elongation of fatty acids protein	x	√
Putative X-pro, dipeptidyl-peptidase,serine peptidase, Clan SC, family S15	x	√
Putative cysteinyl-tRNA synthetase	x	√
Putative cytochrome c oxidase subunit VI	x	√
Alanine--tRNA ligase	x	√
Putative N-ethylmaleimide-sensitive factor	x	√
Putative 40S ribosomal protein S18	x	√
Putative developmentally regulated GTP-binding protein 1	x	√
Elongation factor 2 (Fragment)	x	√
Putative phosphatase 2C	x	√
Putative mitotubule-associated protein Gb4	x	√