

Research Article

Genome Assembly and Genome Annotation of *Leishmania martiniquensis* Isolated from a Leishmaniasis Patient in Thailand

Songtham Anuntakarun ¹, Atchara Phumee ², Vorthon Sawaswong ¹,
Kesmanee Praianantathavorn ³, Witthaya Poomipak ⁴, Rungrat Jitvaropas ⁵,
Padet Siriyasatien ⁶ and Sunchai Payungporn ^{3,7}

¹Program in Bioinformatics and Computational Biology, Graduate School, Chulalongkorn University, Bangkok, Thailand

²Department of Medical Technology, School of Allied Health Sciences, Walailak University, Nakhon Si Thammarat, Thailand

³Department of Biochemistry, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

⁴Research Affairs, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

⁵Division of Biochemistry, Department of Preclinical Science, Faculty of Medicine, Thammasat University, Pathum Thani, Thailand

⁶Vector Biology and Vector Borne Disease Research Unit, Department of Parasitology, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

⁷Research Unit of Systems Microbiology, Chulalongkorn University, Bangkok, Thailand

Correspondence should be addressed to Sunchai Payungporn; sp.medbiochemcu@gmail.com

Received 19 August 2021; Revised 17 December 2021; Accepted 9 March 2022; Published 21 March 2022

Academic Editor: Kwang Poo Chang

Copyright © 2022 Songtham Anuntakarun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Leishmaniasis is a parasitic disease caused by *Leishmania* spp. with worldwide distribution. Autochthonous leishmaniasis has been reported to result from the infection by *Leishmania martiniquensis* in Thailand. This species was isolated in culture and subjected to high-throughput whole-genome sequencing. A total of 30.8 Mb in 36 chromosomes of the whole genome was assembled, annotated, and characterized. The *L. martiniquensis* under study was shown to segregate into the same clade and thus closely related to the previously identified *L. martiniquensis* (LU_Lmar_1.0), as determined by phylogenetic analysis of their genomic sequences along with those of representative kinetoplastid species. The total number of open reading frames genome-wide predicts 8,209 protein-coding genes, of which 359 are putative virulence factors, including two previously known, e.g., cysteine proteinase C and superoxide dismutase B1. The results obtained from this study will be useful for further annotation and comparison with other *Leishmania martiniquensis* in the future.

1. Introduction

Leishmania species are members of the Class Kinetoplastea, Order Trypanosomatida. They are intracellular protozoa that are transmitted through vertebrate hosts by infected female phlebotomine sandflies. There are three major clinical presentations of the disease including cutaneous leishmaniasis (CL), mucocutaneous leishmaniasis (MCL), and visceral leishmaniasis (VL). Symptoms of CL occur on the skin with wet or dry ulcers that are usually painless and localized

lesions, while MCL produces sores on mucosal surfaces, especially the nose, mouth, or throat. VL is the most severe form which occurs in internal organs including the spleen, liver, lymph nodes, and bone marrow. Recently, a novel subgenus *Mundinia* of *Leishmania* parasites has been described, which consists of *L. martiniquensis*, *L. orientalis* n. sp. (previously called *L. siamensis*), *L. enriettii*, and *L. macropodum* (previously called “*Leishmania* sp. AM-2004”) [1–4]. Only *L. martiniquensis* and *L. orientalis* n. sp. have been described as causative agents of leishmaniasis in humans [1, 5, 6].

In Thailand, autochthonous leishmaniasis is caused by *L. martiniquensis* and *L. orientalis* n. sp. (*L. siamensis*). The prevalence of *L. martiniquensis* cases in Thailand have dramatically increased in recent years [7, 8]. Indigenous leishmaniasis cases in Thailand caused by *L. martiniquensis* and *L. orientalis* n. sp. were presented with CL, VL, and overlapping CL and VL. Most of the cases were found in immunocompromised patients especially those with AIDS, and these patients also present a poor response to medical treatment. Amphotericin B is the only antileishmanial agent available for the treatment of indigenous leishmaniasis in Thailand. Cases of relapsed leishmaniasis caused by *L. martiniquensis* were found after receiving amphotericin B treatment [9]. Therefore, the whole-genome sequencing of *L. martiniquensis* would be useful for the understanding of virulence factor genes and interpretation of clinical severity and manifestations.

There have been many studies of the *Leishmania* genome in various species based on next-generation sequencing during the past few years [10, 11]. Currently, it is known that there are virulence factor genes in protozoans including *Leishmania* species. These genes are related to parasite survival and infection in the host cells. For example, proteins such as chaperones and endoribonuclease L-PSP can improve the survival rate of the parasite. In addition, some enzymes are related to migration to facilitate infection of host cells [12]. Particular proteinases, e.g., cysteine-proteinases, metalloproteinases, and serine-proteinases, are also known as virulence factors in *Leishmania* spp. Moreover, they have a wide range of biological roles, including the mechanism of infection [13].

In this study, the *L. martiniquensis* genome was sequenced, assembled, and explored, providing a genetic resource for future exploration of the biology of a clinical isolate of under study species. Subsequently, virulence factor genes in this genome were predicted and analyzed. The candidate virulence factor genes will be validated in further studies.

2. Materials and Methods

2.1. Ethics Statement. This study was approved by the Institutional Review Board of the Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand (COA No. 768/2012). Patients were not involved in this study.

2.2. Promastigote of *L. martiniquensis* Culture. The promastigotes of *L. martiniquensis* (MHOM/TH/2011/CU1) were isolated from the bone marrow of an overlapping CL and VL patient in Songkhla province, Southern Thailand, in 2011 [14–17]. The promastigotes were cultured in Schneider's Insect Medium (Sigma-Aldrich, Missouri, USA) at a pH of 6.7 supplemented with 10% heat-inactivated fetal bovine serum, 100 U/ml penicillin, and 100 µg/ml streptomycin. The promastigotes were incubated at 25 ± 2°C in an incubator and inspected for parasite viability every day under an inverted microscope (Olympus, Tokyo, Japan).

2.3. DNA Extraction. The *L. martiniquensis* promastigotes (10⁶ parasites/ml) were washed with 1x phosphate-buffered saline (PBS) three times (Sigma-Aldrich, Missouri, USA) and centrifuged at 11,000 × g for 10 min. The sample was ground in lysis buffer and used for DNA extraction by using an Invisorb Spin Tissue Mini Kit (STRATEC Molecular, Berlin, Germany), following the manufacturer's instructions. The DNA concentration and purity were quantified by a Qubit 2.0 Fluorometer (Invitrogen, Massachusetts, USA). The extracted DNA samples were used for sequencing immediately, and the rest of the samples were stored at -80°C.

2.4. Library Preparation and High-Throughput Sequencing. DNA (1 µg) was fragmented by using a Covaris M220 focused-ultrasonicator (Covaris, Brighton, UK) with a 20% duty factor, 50 units of peak incident power (W), and 200 cycles per burst for 150 seconds. Then, the fragmented DNA was used for DNA library preparation based on the TruSeq DNA LT Sample Prep Kit (Illumina, California, USA) following the manufacturer's instructions. The DNA library was cleaned up, and the size was selected by AMPure XP beads (Beckman Coulter, USA). The concentration of library DNA was measured by using the KAPA Library Quantification Kit (Kapa Biosystems, Massachusetts, USA). The DNA library was diluted to 6 pM and then paired-end sequenced (2 × 150 bp) based on the MiSeq platform (Illumina, California, USA) by using MiSeq Reagent Kits V2 (300 cycles) according to the standard protocol.

2.5. Quality Filter and Genome Assembly. FASTQ files with 150 bp paired-end reads were checked for the quality of sequences by FastQC [18]. Trimmomatic version 0.39 [19] was used to trim and remove low-quality reads using default parameters. *De novo* assembly was performed using SPAdes version 3.12.0 [20]. The scaffold sequences from the previous step were used to align with the *Leishmania martiniquensis* genome from the NCBI database (accession number CM030396.1–CM030431.1 for chromosome 1–36) [21] using the Artemis comparison tool (ACT) [22]. In addition, the Illumina reads were mapped into the assembled scaffolds using BWA version 0.7.16a [23].

2.6. Gene Prediction and Functional Annotation. AUGUSTUS (Galaxy version 3.3.3) [24] was used to predict genes in the *L. martiniquensis* genome. In this work, the *Leishmania tarentolae* model organism was used in the species parameter for the prediction of gene locations and protein-coding genes. Putative protein-coding sequences from AUGUSTUS were performed in the functional annotation. The EggNOG-mapper version 2 [25] (default parameters) was used to predict functional annotation against EggNOG 5.0 [26]. This database contains functional information from many sources including a Cluster of Orthologous Groups of proteins (COGs) [27], KEGG pathway [28], and GO annotation [29].

2.7. Prediction of Virulence Factor Genes. Putative protein-coding sequences were analyzed by blastP with the protozoa virulence protein database (ProtVirDB) [30] and pathogen host interaction database (PHI-base) [31] for predicting

candidate virulence factor proteins and interaction between hosts and pathogens, respectively. In this study, the criteria for the determination of candidate virulence sequences were an e -value of $10e^{-5}$ or less. For proteinase gene analysis, proteinase genes of *L. martiniquensis* were predicted using sequences from the previous report [13] as a reference.

2.8. Phylogenetic Tree Analysis. OrthoFinder version 2.5.2 [32] with default parameters was used for finding single-copy orthologous genes and alignment of single-copy orthologous genes. In this study, the protein sequence dataset from various species including *Trypanosoma brucei* TREU927 (GCF_000002445.2), *Trypanosoma vivax* Y486 (CA_000227375.1), *Trypanosoma grayi* (GCF_000691245.1), *Trypanosoma cruzi* strain CL Brener (GCF_000209065.1), *Trypanosoma rangeli* (GCF_003719475.1), *Phytomonas* sp. isolate EM1 (GCA_000582765.1), *Leptomonas seymouri* (GCA_001299535.1), *Leptomonas pyrrocoris* (GCF_001293395.1), *Leishmania enriettii* (GCA_017916305.1), *Leishmania martiniquensis* (GCA_017916325.1), *Leishmania tarentolae* (GCA_009731335.1), *Leishmania mexicana* MHOM/GT/2001/U1103 (GCF_000234665.1), *Leishmania major* strain Friedlin (GCF_000002725.1), *Leishmania donovani* (GCF_000227135.1), *Leishmania infantum* JPCM5 (GCF_000002875.1), *Leishmania panamensis* (GCF_000755165.1), *Leishmania braziliensis* MHOM/BR/75/M2904 (GCF_000002845.1), and our *Leishmania martiniquensis* (MHOM/TH/2011/CU1) were used as input of the OrthoFinder tool. The Newick format of phylogenetic tree from OrthoFinder was visualized using Interactive Tree Of Life (iTOL) (<https://itol.embl.de/>) [33].

2.9. Comparison of *L. martiniquensis* Genome with Other *Leishmania* Species. The analysis of the percentage identity of *Leishmania* chromosomes was performed on representative *Leishmania* species including *Leishmania major* strain Friedlin (GCF_000002725.1), *Leishmania infantum* JPCM5 (GCF_000002875.1), *Leishmania donovani* (GCF_000227135.1), *Leishmania mexicana* MHOM/GT/2001/U1103 (GCF_000234665.1), and *Leishmania martiniquensis* MHOM/TH/2012/LSCM1 (LU_Lmar_1.0) (GCA_017916325.1) using Clustal Omega version 1.2.4 with default parameters [34].

3. Results

3.1. Genome Characteristics of *L. martiniquensis* Genome. Paired-end FASTQ files were used for *de novo* assembly using SPAdes. After assembly, there were 6,939 scaffolds with N50 63,362 bp. The percentage of mapped reads to assembled scaffolds between *L. martiniquensis* in this study and *L. martiniquensis* (LU_Lmar_1.0) was 82.36%.

The statistics of *L. martiniquensis* data are shown in Table 1. After the gene prediction step, there were 8,209 protein-coding genes in the final assembly of chromosome 1 to chromosome 36. The chromosome size ranges from 0.24 to 2.8 Mb. The existence of regions in the genome with large variations in the CG content may be caused by over- or underfragmentation during the library con-

TABLE 1: Statistics of *L. martiniquensis* data and *de novo* assembly.

Genome features of <i>L. martiniquensis</i>	
Read length (bp)	150
Raw reads	26,205,720
Q30 reads	23,836,943
Number of scaffolds	6,939
N50 (bp)	63,362
Number of protein-coding genes	8,209
Illumina read mapping to the assembled scaffolds (%)	82.36

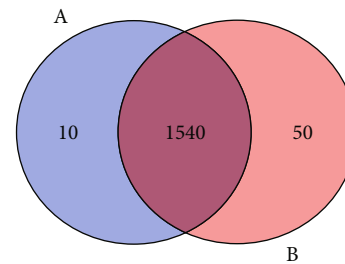


FIGURE 1: Venn diagram analysis of genes between *Leishmania martiniquensis* (LU_Lmar_1.0) genome from the NCBI database (a) and *Leishmania martiniquensis* from our study (b).

struction. The *L. martiniquensis* genome had an average GC content of 59.77%.

3.2. Comparison of *L. martiniquensis* with Other *Leishmania* Species. The genome (36 chromosomes) of *L. martiniquensis* was compared with other *Leishmania* species including *L. infantum*, *L. donovani*, *L. braziliensis*, *L. major* strain Friedlin, *L. mexicana*, and *L. martiniquensis*. In addition, the genome of *L. martiniquensis* in this study was closely related to *L. martiniquensis* (LU_Lmar_1.0). Genome coverage between *L. martiniquensis* in this study and *L. martiniquensis* (LU_Lmar_1.0) was 95.18%. The Venn diagram analysis in Figure 1 shows that 10 genes and 50 genes were found in only *L. martiniquensis* (LU_Lmar_1.0) genome and our *L. martiniquensis*, respectively. In addition, the descriptions of genes are shown in Supplementary Material 3. The result of nucleotide identity is summarized in Supplementary Table 1. The COG functional category in *L. martiniquensis* was compared with other *Leishmania* spp. including *L. infantum*, *L. donovani*, *L. braziliensis*, *L. major*, and *L. mexicana*. Our results showed that *L. martiniquensis* genes occupied similar functional roles to those in other *Leishmania* spp. (Supplementary Table 2). The KEGG pathway analysis and GO annotation are represented in Supplementary Figure 1. In the KEGG pathway analysis (Supplementary Figure 1A), the top three pathways include ribosome, metabolic pathways, and RNA polymerase. Functional annotation is the process of collecting information about the function of genes. Gene Ontology (GO) is the most widespread and extensive functional annotation for gene and protein sequences. There are three categories of terms in Gene Ontology. First, the molecular function comprises the molecular activities of individual

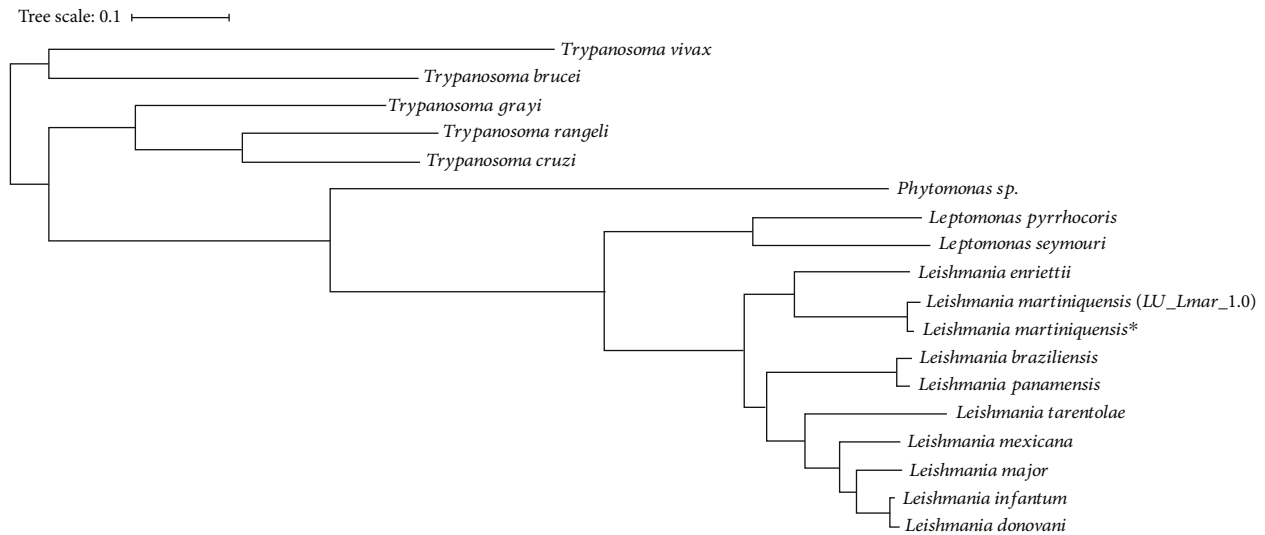


FIGURE 2: The phylogenetic analysis based on 17 single-copy shared orthologous genes obtained from OrthoFinder tool. The star symbol represents our *Leishmania martiniquensis* (MHOM/TH/2011/CU1).

gene products. Second, the cellular component comprises the region of active gene products. Third, the biological process comprises the processes and the pathways in which the activity of gene products is involved. The result of GO analysis in Supplementary Figures 1B-1D shows that the top three molecular functions were structural constituent of ribosome, poly (A) RNA-binding, and DNA-directed RNA polymerase activity. The top three cellular component functions were cytosolic large ribosomal subunit, motile cilium, and intraciliary transport particle B. The top three biological process functions were translation, rRNA processing, and ribosomal large subunit assembly.

3.3. Virulence Factor Gene Analysis. Predicted proteins for the protein-coding genes predicted by AUGUSTUS were analyzed by blastP with the protozoa virulence protein database (ProtVirDB) using the criteria of e -value $< 10e^{-5}$. A total of 359 genes were found as candidate virulence factor genes. These genes were then analyzed for COG functional annotation. The top 3 COG functions were signal transduction mechanism, carbohydrate transport and metabolism, and intracellular trafficking, secretion, and vesicular transport, while the remaining COG functions are shown in Supplementary Figure 2. The annotation lists of the 359 genes from ProtVirDB are shown in Supplementary Material 1. Moreover, forty-three predicted protein sequences that passed the criteria from blastP with PHI-base were related with *Homo sapiens* organisms. The annotation lists of the 43 genes from the PHI-base are shown in Supplementary Material 2. However, the predicted virulence factor gene should be validated in further study.

3.4. Phylogenetic Tree Analysis. The concatenated protein sequences of 17 single-copy orthologous genes were used to create a phylogenetic tree. In Figure 2, the phylogenetic tree indicates that *L. martiniquensis* is related to *Leishmania* spp. Moreover, the outgroup including *Trypanosoma brucei*

TREU927, *Trypanosoma vivax* Y486, *Trypanosoma grayi*, *Trypanosoma cruzi* strain CL Brener, and *Trypanosoma rangeli* is a more distinctly related group of the *Leishmania* species. The result suggests that *L. martiniquensis* in this study is closely related to *L. martiniquensis* (LU_Lmar_1.0) published in April 2021 on the NCBI website.

4. Discussion

In this study, genome assembly and gene prediction of *Leishmania martiniquensis* were performed. The results showed that the COG functional category of *L. martiniquensis* was similar to other *Leishmania* species. However, there was a slight difference in the number of genes in each functional group. The importance of parasite virulence factors has become apparent in recent years [35]. The variability of virulence factor genes within the *Leishmania* species is largely unknown. In our virulence factor gene prediction of *L. martiniquensis*, the result showed that 359 candidate virulence factor genes were found in *L. martiniquensis*. Some of these genes are discussed below.

Heat shock proteins are intracellular molecules of varying molecular weights. They are a large family of molecular chaperones. The role of this protein is maturation, degradation, and refolding [36]. They also play an important role in immune biological functions, especially in *hsp70*. There was a report which showed that *hsp70* induces dendritic cells to generate proinflammatory cytokines [37] and is related to the enhancement of adaptive immunity [38]. In our results, the *hsp70* protein-coding gene in 359 candidate virulence factor genes was found. This gene might be related to the infection of host cells.

Proteinases are enzymes that hydrolyze peptide bonds in proteins and participate in a wide range of biological functions, including the process of infection [13]. There are many classes of proteinase based on catalytic domains [39]. There are only three classes, including aspartyl-, metallo-,

and cysteine-proteinase, which have been extensively studied in *Leishmania* organisms [40, 41]. In a previous review, cysteine proteases were considered to play a crucial role in the pathogenesis of other parasitic protozoan infections [42]. *CPA*, *CPB*, and *CPC* genes in a group of cysteine proteases have been widely studied in *Leishmania* species. In our analysis result, *CPC* genes were found in *L. martiniquensis*. *CPC* plays a potential role in resisting parasite killing by macrophages [43].

In addition, superoxide dismutase gene (*SODB1*) was found only in our genome when compared with *L. martiniquensis* (LU_Lmar_1.0). There is a report that *SODB1* is required for *Leishmania major* pathogenicity in mice and persistence in macrophages [44]. Superoxide dismutases (SODs) are metalloenzymes that convert superoxide to oxygen and hydrogen peroxide in the antioxidant defense system. Nickel (Ni), manganese (Mn), iron (Fe), or copper and zinc (Cu/Zn) are cofactors of superoxide dismutases [45, 46].

The phylogenetic analysis showed that our *L. martiniquensis* is closely related to the previously reported *L. martiniquensis* (LU_Lmar_1.0) reference genome in the NCBI database. Comparative genomics of *L. (Mundinia) martiniquensis* was reported in 2019 [47]. This research undertaken here reported genomes of *L. (Mundinia) martiniquensis* with a protein dataset.

5. Conclusions

In this study, *L. martiniquensis* genomic DNA was successfully sequenced and was assembled into 6939 scaffolds. A total of 30,784,469 bp in 36 chromosomes of the *L. martiniquensis* genome were analyzed. The analysis results showed that the general features of *L. martiniquensis* were similar to other *Leishmania* species, including chromosome sizes, the number of protein-coding genes, and the GC contents. In addition, the results of COG functional annotation were shown to be similar to other *Leishmania* species. In the virulence factor gene prediction result, 359 potential candidate virulence factor genes were found in this study. Most predicted virulence factor genes were related to RNA processing and modification function. However, candidate potential virulence factor genes should be validated in a further study using experimental study.

Data Availability

All data analyzed during this study are included in this article. In addition, the DNA sequences were deposited in Sequence Read Archive (SRA) data of the NCBI server (BioProject ID PRJNA674467).

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgments

The authors would like to acknowledge the supports from Graduate School, Faculty of Science and Faculty of Medicine, Chulalongkorn University (the 100th Anniversary Chulalongkorn University Fund for Doctoral Scholarship; the 90th Anniversary of Chulalongkorn University Ratchadapisek Somphot Endowment Fund). This work was supported by the Thailand Research Fund (TRF) (grant number RSA6180035); National Science and Technology Development Agency (NSTDA) (grant number P-17-51377); Program Management Unit for Human Resources & Institutional Development, Research and Innovation (grant number B16F630071); and Thailand Science Research Innovation (TSRI) (grant number FRB640001).

Supplementary Materials

Supplementary 1. Supplementary Material 1: the annotation lists of the candidate predicted virulence factor genes from ProtVirDB.

Supplementary 2. Supplementary Material 2: the annotation lists of the candidate predicted virulence factor genes from PHI-base.

Supplementary 3. Supplementary Material 3: the annotation lists of the unique genes between *L. martiniquensis* (LU_Lmar_1.0) and our *L. martiniquensis* genome.

Supplementary 4. Supplementary Table 1: comparison of percent identity in *Leishmania* spp. including *L. infantum*, *L. donovani*, *L. major* Friedlin, *L. Mexicana*, *L. braziliensis*, *L. martiniquensis*, and our *L. martiniquensis*. Supplementary Table 2: comparison of functional category of putative protein-coding genes in *Leishmania* spp. genome. The alphabet A-G represent names of *Leishmania* species including *L. martiniquensis* (A), *L. infantum* (B), *L. donovani* (C), *L. braziliensis* (D), *L. major* strain Friedlin (E), *L. mexicana* (F), and *L. martiniquensis* (LU_Lmar_1.0) (G).

Supplementary 5. Supplementary Figure 1: the functional annotation of *L. martiniquensis*. Supplementary Figure 2: the COG functional analysis of candidate virulence factor protein-coding genes.

References

- [1] N. Jariyapan, T. Daroontum, K. Jaiwong et al., “*Leishmania* (*Mundinia*) *orientalis* n. sp. (Trypanosomatidae), a parasite from Thailand responsible for localised cutaneous leishmaniasis,” *Parasites and Vectors*, vol. 11, pp. 3–11, 2018.
- [2] O. A. Espinosa, M. G. Serrano, E. P. Camargo, M. M. G. Teixeira, and J. J. Shaw, “An appraisal of the taxonomy and nomenclature of trypanosomatids presently classified as *Leishmania* and *Endotrypanum*,” *Parasitology*, vol. 145, no. 4, pp. 430–442, 2018.
- [3] J. Barratt, A. Kaufer, B. Peters et al., “Isolation of novel trypanosomatid, *Zelonia australiensis* sp. nov. (Kinetoplastida: Trypanosomatidae) provides support for a Gondwanan origin of dioxenous parasitism in the Leishmaniinae,” *PLoS Neglected Tropical Diseases*, vol. 11, no. 1, pp. e0005215–e0005226, 2017.

- [4] T. Sukmee, S. Siripattanapipong, M. Mungthin et al., "A suspected new species of *Leishmania*, the causative agent of visceral leishmaniasis in a Thai patient," *International Journal for Parasitology*, vol. 38, no. 6, pp. 617–622, 2008.
- [5] S. Chiewchanvit, N. Tovananubutra, N. Jariyapan et al., "Chronic generalized fibrotic skin lesions from disseminated leishmaniasis caused by *Leishmania martiniquensis* in two patients from northern Thailand infected with HIV," *The British Journal of Dermatology*, vol. 173, no. 3, pp. 663–670, 2015.
- [6] T. Pothirat, A. Tantiworawit, R. Chaiwarith et al., "First isolation of *Leishmania* from northern Thailand: case report, identification as *Leishmania martiniquensis* and phylogenetic position within the *Leishmania enriettii* complex," *PLoS neglected tropical diseases*, vol. 8, 2014.
- [7] P. Siriyasatien, S. Chusri, K. Kraivichian et al., "Early detection of novel *Leishmania* species DNA in the saliva of two HIV-infected patients," *BMC Infectious Diseases*, vol. 16, no. 1, p. 89, 2016.
- [8] A. Phumee, S. Chusri, K. Kraivichian et al., "Multiple cutaneous nodules in an HIV-infected patient," *PLoS Neglected Tropical Diseases*, vol. 8, no. 12, pp. e3291–e3298, 2014.
- [9] A. Phumee, N. Jariyapan, S. Chusri et al., "Determination of anti-leishmanial drugs efficacy against *Leishmania martiniquensis* using a colorimetric assay," *Parasite Epidemiology and Control*, vol. 9, article e00143, 2020.
- [10] P. Lypaczewski, J. Hoshizaki, W. W. Zhang et al., "A complete *Leishmania donovani* reference genome identifies novel genetic variations associated with virulence," *Scientific Reports*, vol. 8, no. 1, p. 16549, 2018.
- [11] S. Coughlan, A. S. Taylor, E. Feane et al., "Leishmania naiffian-dLeishmania guyanensis reference genomes highlight genome structure and gene evolution in the Vianniasubgenus," *Royal Society open science*, vol. 5, no. 4, 2018.
- [12] S. Da Fonseca Pires, L. C. Fialho, S. O. Silva et al., "Identification of virulence factors in *Leishmania infantum* strains by a proteomic approach," *Journal of Proteome Research*, vol. 13, no. 4, pp. 1860–1872, 2014.
- [13] M. Silva-Almeida, B. A. S. Pereira, M. L. Ribeiro-Guimarães, and C. R. Alves, "Proteinases as virulence factors in *Leishmania* spp. infection in mammals," *Parasites and Vectors*, vol. 5, pp. 1–10, 2012.
- [14] S. Chusri, T. Hortiwakul, K. Silpapojakul, and P. Siriyasatien, "Consecutive cutaneous and visceral leishmaniasis manifestations involving a novel *Leishmania* species in two HIV patients in Thailand," *The American Journal of Tropical Medicine and Hygiene*, vol. 87, no. 1, pp. 76–80, 2012.
- [15] A. Phumee, K. Kraivichian, S. Chusri et al., "Detection of *Leishmania siamensis* DNA in saliva by polymerase chain reaction," *The American Journal of Tropical Medicine and Hygiene*, vol. 89, no. 5, pp. 899–905, 2013.
- [16] T. Becvar, B. Vojtkova, P. Siriyasatien et al., "Experimental transmission of *Leishmania* (*Mundinia*) parasites by biting midges (Diptera: Ceratopogonidae)," *PLoS Pathogens*, vol. 17, no. 6, p. e1009654, 2021.
- [17] S. Leelayoova, S. Siripattanapipong, A. Hitakarun et al., "Multilocus characterization and phylogenetic analysis of *Leishmania siamensis* isolated from autochthonous visceral leishmaniasis cases, southern Thailand," *BMC Microbiology*, vol. 13, no. 1, p. 60, 2013.
- [18] S. Andrews, *Fast QC: a quality control tool for high throughput sequence data*, 2010.
- [19] A. M. Bolger, M. Lohse, and B. Usadel, "Trimmomatic: a flexible trimmer for Illumina sequence data," *Bioinformatics*, vol. 30, no. 15, pp. 2114–2120, 2014.
- [20] A. Bankevich, S. Nurk, D. Antipov et al., "SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing," *Journal of Computational Biology*, vol. 19, no. 5, pp. 455–477, 2012.
- [21] H. Almutairi, M. D. Urbaniak, M. D. Bates et al., "Chromosome-scale assembly of the complete genome sequence of *Leishmania* (*Mundinia*) *martiniquensis*, isolate LSCM1, strain LV760," *Microbiology Resource Announcements*, vol. 10, article e00058, 2021.
- [22] T. J. Carver, K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrell, and J. Parkhill, "ACT: the Artemis comparison tool," *Bioinformatics*, vol. 21, no. 16, pp. 3422–3423, 2005.
- [23] H. Li and R. Durbin, "Fast and accurate short read alignment with Burrows-Wheeler transform," *Bioinformatics*, vol. 25, no. 14, pp. 1754–1760, 2009.
- [24] M. Stanke and B. Morgenstern, "AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints," *Nucleic Acids Research*, vol. 33, no. Web Server, pp. W465–W467, 2005.
- [25] C. P. Cantalapiedra, A. Hernández-Plaza, I. Letunic, and J. Huerta-Cepas, "eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale," *BioRxiv*, vol. 5, 2021.
- [26] J. Huerta-Cepas, D. Szklarczyk, D. Heller et al., "EggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses," *Nucleic acids research*, vol. 47, pp. D309–D314, 2019.
- [27] M. Y. Galperin, K. S. Makarova, Y. I. Wolf, and E. V. Koonin, "Expanded microbial genome coverage and improved protein family annotation in the COG database," *Nucleic Acids Research*, vol. 43, no. D1, pp. D261–D269, 2015.
- [28] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima, "KEGG: new perspectives on genomes, pathways, diseases and drugs," *Nucleic Acids Research*, vol. 45, no. D1, pp. D353–D361, 2017.
- [29] M. A. Harris, J. Clark, A. Ireland et al., "The Gene Ontology (GO) database and informatics resource," *Nucleic acids research*, vol. 32, pp. 258–261, 2004.
- [30] J. Ramana and D. Gupta, "ProtVirDB: a database of protozoan virulent proteins," *Bioinformatics*, vol. 25, no. 12, pp. 1568–1569, 2009.
- [31] M. Urban, A. Cuzick, K. Rutherford et al., "PHI-base: a new interface and further additions for the multi-species pathogen-host interactions database," *Nucleic Acids Research*, vol. 45, no. D1, pp. D604–D610, 2017.
- [32] D. M. Emms and S. Kelly, "OrthoFinder: phylogenetic orthology inference for comparative genomics," *Genome Biology*, vol. 20, no. 1, p. 238, 2019.
- [33] I. Letunic and P. Bork, "Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation," *Nucleic Acids Research*, vol. 49, no. W1, pp. W293–W296, 2021.
- [34] F. Sievers, A. Wilm, D. Dineen et al., "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega," *Molecular systems biology*, vol. 7, 2011.
- [35] D. A. Urrea, J. Duitama, H. Imamura et al., "Genomic analysis of Colombian *Leishmania panamensis* strains with different

- level of virulence,” *Scientific Reports*, vol. 8, no. 1, pp. 1–16, 2018.
- [36] D. J. Miller and P. E. Fort, “Heat shock proteins regulatory role in neurodevelopment,” *Frontiers in Neuroscience*, vol. 12, pp. 1–15, 2018.
- [37] M. C. Kuppner, R. Gastpar, S. Gelwer et al., “The role of heat shock protein (hsp70) in dendritic cell maturation: Hsp70 induces the maturation of immature dendritic cells but reduces DC differentiation from monocyte precursors,” *European Journal of Immunology*, vol. 31, no. 5, pp. 1602–1609, 2001.
- [38] P. A. MacAry, B. Javid, R. A. Floto et al., “HSP70 peptide binding mutants separate antigen delivery from dendritic cell stimulation,” *Immunity*, vol. 20, no. 1, pp. 95–106, 2004.
- [39] N. D. Rawlings, A. J. Barrett, and A. Bateman, “MEROPS: the peptidase database,” *Nucleic Acids Research*, vol. 38, pp. 227–233, 2010.
- [40] E. Valdivieso, F. Dagger, and A. Rascón, “_Leishmania mexicana_ : identification and characterization of an aspartyl proteinase activity,” *Experimental Parasitology*, vol. 116, no. 1, pp. 77–82, 2007.
- [41] M. Sajid and J. H. McKerrow, “Cysteine proteases of parasitic organisms,” *Molecular and Biochemical Parasitology*, vol. 120, no. 1, pp. 1–21, 2002.
- [42] A. B. Vermelho, M. H. Branquinha, C. M. D’Ávila-Levy, A. L. Souza dos Santos, E. P. de Souza Dias, and A. C. Nogueira de Melo, “Biological roles of peptidases in trypanosomatids~!2009-11-26~!2010-02-15~!2010-03-18~!,” *Open Parasitology Journal*, vol. 4, no. 1, pp. 5–23, 2010.
- [43] M. J. Frame, J. C. Mottram, and G. H. Coombs, “Analysis of the roles of cysteine proteinases of *Leishmania mexicana* in the host–parasite interaction,” *Parasitology*, vol. 121, no. 4, pp. 367–377, 2000.
- [44] B. J. Davenport, C. G. Martin, S. M. Beverley, D. J. Orlicky, A. Vazquez-Torres, and T. E. Morrison, “SODB1 is essential for *Leishmania major* infection of macrophages and pathogenesis in mice,” *PLoS Neglected Tropical Diseases*, vol. 12, no. 10, p. e0006921, 2018.
- [45] C. N. Broxton and V. C. Culotta, “SOD enzymes and microbial pathogens: surviving the oxidative storm of infection,” *PLoS Pathogens*, vol. 12, no. 1, p. e1005295, 2016.
- [46] I. A. Abreu and D. E. Cabelli, “Superoxide dismutases—a review of the metal-associated mechanistic variations,” *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, vol. 1804, no. 2, pp. 263–274, 2010.
- [47] A. Butenko, A. Y. Kostygov, J. Sádlová et al., “Comparative genomics of *Leishmania* (*Mundinia*),” *BMC Genomics*, vol. 20, no. 1, p. 726, 2019.