FULL LENGTH	Iranian Biomedical Journal 29 (1 & 2): 68-81 January & March 2025

Unraveling *Leishmania major* Metacyclogenesis: A Comprehensive Analysis of Transcriptomic and Metabolomic Profiles

Mansour Aminzadeh[1,2], Fariborz Bahrami[2], Zeynab Piryaei[3], Mahdi Vasighi[4], Zahra Kalantari[5], Mohammad Arjmand[1*], and Soheila Ajdary[2*]

[1]Metabolomics Lab, Department of Biochemistry, Pasteur Institute of Iran, Tehran, Iran; [2]Department of Immunology, Pasteur Institute of Iran, Tehran, Iran; [3]Department of Bioinformatics, Kish International Campus, University of Tehran, Kish, Iran; [4]Department of Computer Sciences, Institute for Advanced Studies in Basic Sciences, Zanjan, Iran; [5]Iran Polymer and Petrochemical Institute. Department of Polymerization Engineering, Tehran, Iran

OPEN ACCESS

Article type: Research article
Received: November 13, 2024
Revised: December 14, 2024
Accepted: December 25, 2024
Published online: December 28, 2024

How to cite:
Aminzadeh M, Bahrami F, Piryaei Z, Vasighi M, Kalantari Z, Arjmand M, Ajdary S. Unraveling *Leishmania major* Metacyclogenesis: A Comprehensive Analysis of Transcriptomic and Metabolomic Profiles. *Iran. Biomed. J.* 2024; 29(1&2): 68-81.

This article is licensed under a Creative Commons Attribution-NonDerivatives 4.0 International License.

ABSTRACT

Background: Metacyclogenesis is a critical developmental process in the life cycle of *Leishmania* parasites, particularly in their transition from non-infective procyclic to infective metacyclic promastigotes. This transformation is closely linked to the metabolic adaptation of the parasite, optimizing its survival and study, we integrated metabolomics and transcriptomics data to gain deeper molecular mechanisms of *L. major* metacyclogenesis.

Methods: The metabolic profiles of procyclic and metacyclic promastigotes were first identified using ^{1}H-NMR spectroscopy. Multivariate statistical analysis conducted to distinguish different metabolites between the two forms. Metabolic pathway analysis was performed using the KEGG database to identify the metabolic pathways that significantly altered and overrepresented in the metabolomic profile. Finally, the differential gene expression and pathway enrichment analyses were conducted on transcriptomic data retrieved from public repositories.

Results: Multivariate statistical analysis revealed that 44 metabolites and ten pathways were significantly different between the two forms. Transcriptome genes during metacyclogenesis. These genes underwent GO and KEGG pathway analyses, revealing upregulated GO categories in the metacyclic phase, including protein phosphorylation, ion transport, signal transduction, and phosphorylation reactions, as well as several downregulated GO categories. Integrating metabolomic and transcriptomic data demonstrated seven significantly different KEGG pathways between procyclic and metacyclic forms, including fructose and mannose, galactose, ascorbate and aldarate, arginine and proline, histidine, inositol phosphate, and pyruvate metabolism.

Conclusion: Our findings suggest distinct metabolic profiles and changes in gene expression associated with the transition from procyclic to metacyclic promastigotes. By integrating diverse omics data, we could identify more reliable altered pathways and biomarkers. ***DOI: 10.61882/ibj.4899***

Keywords: Leishmania, Metabolomics, Metacyclogenesis

Corresponding Authors:
Soheila Ajdary
Department of Immunology, Pasteur Institute of Iran, Tehran, Iran; E-mail: sohary@yahoo.com, ajdsoh@pasteur.ac.ir; ORCID ID: 0000-0002-5052-2523
Mohammad Arjmand
Metabolomics Lab., Department of Biochemistry, Pasteur Institute of Iran, Tehran, Iran; E-mail: arjmand2@gmail.com;
ORCID ID : 0000-0001-7825-7883

List of Abbreviations:
AcCoA: acetyl CoA; **BMRB:** Biological Magnetic Resonance Data Bank; **DAM:** differential abundance metabolite; **DEG:** differentially expressed gene; **FDR:** false discovery rate; **GO:** Gene Ontology; **HMDB:** Human Metabolome Database; **KEGG:** Kyoto Encyclopedia of Genes and Genomes; ***L. major*:** *Leishmania major*; **LmAPX:** Ascorbate peroxidase from *L. major;* **OPLS-DA:** orthogonal partial least squares discriminant analysis; **PCA:** principal component analysis; **PLS-DA:** partial least squares discriminant analysis; **RF:** Random Forest; **SRA:** Sequence Read Archive; **VIP:** variable importance in projection

INTRODUCTION

Leishmaniases comprise a spectrum of diseases caused by over 20 species of the *Leishmania* parasite. The clinical manifestations vary widely, from self-healing skin lesions characteristic of cutaneous leishmaniasis to the more severe mucocutaneous leishmaniasis, which affects both the skin and mucous membranes. The most severe form, visceral leishmaniasis, can be fatal if left untreated. Annual reports from the World Health Organization estimate 1.3 million new cases each year, resulting in approximately 70,000 deaths per year[1-3]. *Leishmania* parasites, members of the Trypanosomatidae family, have a digenetic life cycle, alternating between mammalian hosts and the *Phlebotomine* sandfly vector. Each stage is marked by distinct morphological and functional adaptations.

In the sandfly midgut, the parasites exist as promastigotes—slender, motile, and flagellated forms that divide rapidly. These promastigotes consist of two main developmental stages: procyclic and metacyclic parasites. Both stages can also be observed in culturing media during their logarithmic and stationary phases of growth, respectively. Procyclic parasites attach to the epithelial cells of the midgut and exhibit low virulence. In contrast, metacyclic parasites represent the virulent disease-inducing form of *Leishmania*, which is transmitted to mammalian hosts during blood feeding[4]. Notably, morphological and chemical changes occur during differentiation from procyclic to metacyclic forms in the insect gut or culturing media. These changes include a reduction in cell size, a more slender shape, elongation of the flagellum, and increased expression of surface protease gp63 and lipophosphoglycan. Moreover, significant changes occur in the structure and composition of lipophosphoglycan carbohydrates[3,4].

Metabolic pathways play a crucial role in the growth, proliferation, and survival of *Leishmania* spp. during various stages of their life cycle. Numerous studies have used single-omics approaches, including metabolomics, proteomics, and transcriptomics to compare different stages of *Leishmania* differentiation. An earlier investigation employed metabolomics to explore DAMs and metabolic pathways involving in the transition of *L. major* from the procyclic to the metacyclic stage under in vitro conditions[5]. Another research compared the metabolic profiles of the metacyclic forms of *L. major* and *L. tropica* to identify significant DAM and metabolic pathways. The metabolome of *L. infantum* has also been studied using a genome-scale metabolic model, which indicated a unique metabolic organization for the parasite[6].

Proteomic analysis has revealed differentially expressed proteins during the metacyclogenesis of *L. major*[7]. Using RNA sequence data from *L. major*, Farias Amorim and colleagues identified DEGs during the in vitro development of promastigotes[8]. DEGs have also been recognized between stationary phase of promastigotes and amastigotes in *L. infantum*, demonstrating significant transcriptional changes associated with these developmental stages[9]. The transcriptome of different stages of *L. major* during its cyclical development has been analyzed in sandflies, disclosing extensive genetic reprogrammin[8]. However, the integration of omics data in the study of metacyclogenesis in *Leishmania* remains limited. Recent advancements in multi-omics approaches are initiating to illuminate the complex biological processes involving in metacyclogenesis. In this study, we integrated metabolomics and transcriptomics data to identify DAMs and DEGs, with the goal of elucidating the underlying molecular mechanisms of *Leishmania* metacyclogenesis.

MATERIALS AND METHODS

Leishmania promastigote culture

L. major (Friedline strain) was cultured as promastigotes at 23-25 °C in complete media consisting of RPMI 1640 medium (Biowest, France), supplemented with 10% fetal bovine serum (Biosera, France), penicillin (100 U/ml), and streptomycin (100 µg/ml). To maintain virulence, the *L. major* promastigotes were passaged no more than three times, after isolation from BALB/c mice. The promastigotes were cultured in T-75 flasks at a density of 2×10^7 cells/ml. Procyclic forms were obtained after 4 to 6 days of culturing, while metacyclic forms were collected after 8 to 10 days without the addition of fresh medium. Metacyclic forms were isolated from stationary phase cultures using peanut agglutinin. Parasites were washed in PBS by centrifugation at 4,000 ×g at 4 °C for 20 min. The cell pellet was washed twice with ice-cold PBS. Chilled perchloric acid (1.8 M) was added to the cell suspension, which was then vortexed and sonicated at 4 °C for 5 min. The mixture was subsequently centrifuged at 48,000 ×g at 4 °C for 10 min. After adjusting the pH of the supernatant to 6.8, the mixture was placed on ice for 1 h to facilitate the precipitation of potassium perchlorate. Following centrifugation, the supernatant was lyophilized and subjected to ^{1}H-NMR[5].

^{1}H-NMR spectroscopic analysis

Lyophilized samples were reconstituted in deuterated phosphate buffer (pH 7.4), which contained 1 mM of sodium 3-(trimethylsilyl)propionate as a chemical shift reference, 2 mM of imidazole as a pH indicator, and

D_2O water as a blocking agent. After centrifugation to remove any particulate matter, the supernatant was transferred to NMR probe for analysis. ^{1}H-NMR spectra were acquired using a Bruker 400 MHz spectrometer (400.13 MHz, 298 K) with a total of 540 scans. A standard water pre-saturation pulse sequence (NOESY) was employed to suppress the residual water signal.

Metabolomics data analysis

NMR spectral preprocessing

Spectral preprocessing was conducted using the ProMetab function in MATLAB (v. 22) to prepare the data for multivariate analysis. This process involved correcting baseline shifts, reducing noise, and scaling the data. To simplify the analysis, the spectra were aligned and grouped into bins of 0.005 ppm, which facilitated the identification of similar chemical shifts. Additionally, peaks corresponding to water and imidazole were removed from the spectra to enhance data quality and concentrate on metabolites of interest, as these peaks can obscure other signals and hinder accurate analysis. The workflow of integrated metabolomics and transcriptomics analysis of *L. major* metacyclogenesis is summarized in Figure 1.

Univariate and multivariate data analysis

To compare the metabolic profiles of procyclic and metacyclic groups, univariate analyses were conducted on preprocessed and normalized metabolite data utilizing student's t-test. Significant DAMs were identified based on *p* values and FDRs, both set at 0.05. To further explore metabolic alterations occurring during metacyclogenesis, multivariate and pathway analyses were performed. The dataset was analyzed using PCA, PLS-DA, sparse PLS-DA, OPLS-DA, and Random Forest, all of which were implemented in MetaboAnalyst (v. 6.0)[10]. The dataset was processed using Pareto scaling and subjected to multivariate analysis. PCA was initially conducted to assess overall variations, similarities, and natural clusters within the metabolic data of the two groups. PLS-DA was then utilized to identify key metabolites responsible for the discrimination of these clusters, using PLS regression coefficients to highlight the most influential features. The metabolic datasets were analyzed using sparse PLS-DA and OPLS-DA to further refine the data classification. Finally, the RF method was implemented to improve data analysis.

Validation of predictive models

Variable selection for the model was performed using VIP plots with a 95% confidence level. The predictive accuracy of multivariate models was evaluated using various combinations of variables, selected based on VIP > 2 score. The predictive accuracy of the multivariate models was assessed through cross- validation techniques. Model performance was evaluated using metrics such as R^2 and Q^2, which measure the fitness of model to the original data,

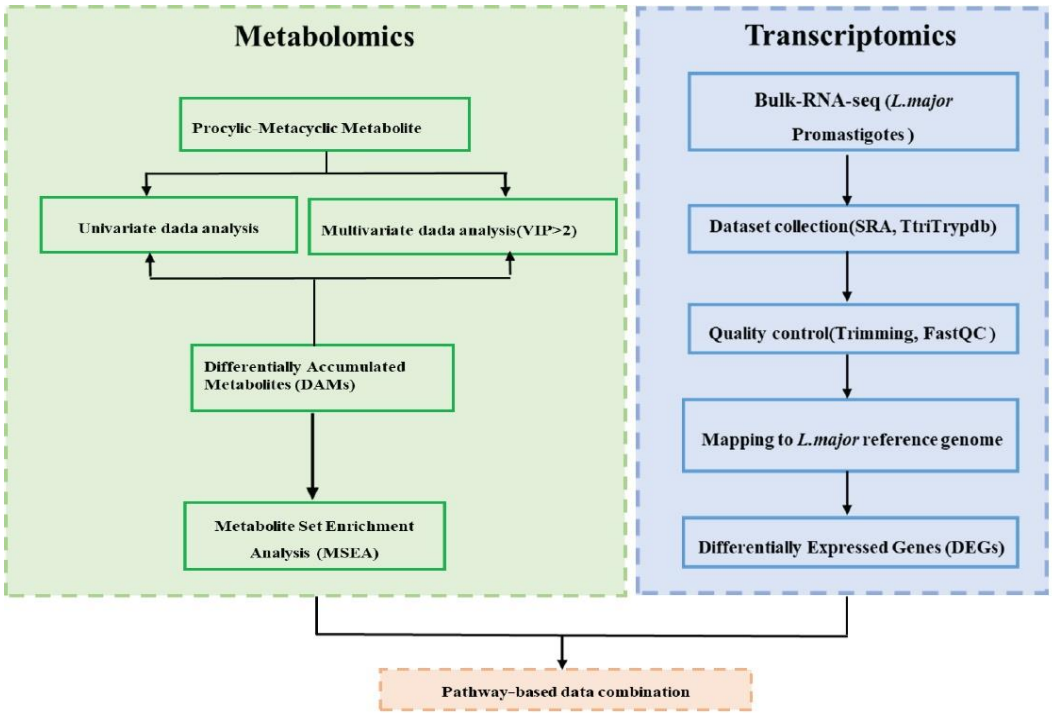

Fig. 1. An overview of the workflow of comprehensive analysis of metabolomics and transcriptomics during *Leishmania* metacyclogenesis.

provide an internal measure of consistency between the original and cross-validation predicted data, respectively. Models with higher R^2, Q^2, and accuracy values indicate desirable performance, which Q^2 is particularly valuable for model selection due to its lower susceptibility to overfitting.

Metabolite annotation

Identification of compounds is crucial for interpreting biological findings in the metabolome. To achieve this, significant chemical shifts were recorded in the previous step and subsequently compared against NMR databases, including the HMDB[11], and BMRB[12].Additionally, pathway analysis was conducted using the KEGG database through the MetaboAnalyst.

Metabolite pathway enrichment analysis

Metabolite pathway enrichment analysis was performed using the KEGG database through MetaboAnalyst to identify the metabolic pathways, significantly altered and overrepresented in the metabolomic profile. Pathways were considered significant if they exhibited an impact > 0.05 and $p < 0.05$.

Transcriptome data analysis

RNA seq dataset selection

We examined samples from the PRJNA252769 project, as reported previously[13]. The goal of the PRJNA252769 study was to obtain comprehensive gene expression profiling during the *L. major* developmental stages in sand fly. The raw sequencing data were obtained from the SRA database using the SRA Toolkit. All 19 SRAs were downloaded using the prefetch command and converted to FASTQ format using fastq-dump **(Table S1)**.

Transcriptome data preprocessing

After data retrieval, the raw sequencing reads were subjected to quality control using Trimmomatic (v. 0.39)[14]. This process involved removing any residual Illumina adapter sequences and trimming bases from the beginning or end of reads with quality scores less than 20. Sequence quality was assessed using the FastQC program provided by the Babraham Institute (www.bioinformatics.babraham.ac.uk/projects/fastqc/). The outcomes of FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) were summarized and visualized using the MultiQC (https://seqera.io/multiqc/) reporting tool[15]. The high-quality, trimmed reads were then aligned to the *L. major* Friedline reference genome (v. 66), which was obtained from the TriTrypDB database[16].

DEG analysis

The trimmed reads were aligned to the *L. major* reference genome (v. 66), using Bowtie2 (v. 2.7.1a) with default settings. Gene expression analysis was performed using edgeR package (v.1.40.2) in R. A |log2 fold-change > 1, and FDR < 0.05 was set as cut-off criteria for identifying DEG[17,18].

GO and gene sets pathway analysis

GO including biological processes, molecular functions, and cellular components, was identified using gProfiler (https://biit.cs.ut.ee/gprofiler/gost) and TriTrypDB (www.tritrypdb.org) databases. Pathway analysis was conducted using the KEGG database (www.genome.jp/kegg) to identify overrepresented metabolic pathways in the lists of DEG[19,20].

Integration analysis of transcriptomic and metabolomic data

To provide a more comprehensive understanding of promastigote development, we conducted integration analysis for DEG and DAM metabolites at the pathway level using Venn diagrams[21].

RESULTS

Multivariate analysis

This study employed various multivariate statistical analysis techniques to distinguish the metabolic profiles between the procyclic and metacyclic phases. The PCA score plot (Fig. 2A) demonstrates the discrimination between the groups. To identify the metabolites responsible for this discrimination, loading plots were generated from the PCA (Fig. 2B), which clearly shows different key metabolites. Additionally, a scree plot was created (Fig. 2C) to analyze the contribution of the first five principal components to the overall variation in the dataset. The PCA biplot (Fig. 2D) for the procyclic and metacyclic groups provided an overview of the relationship between samples and metabolites where samples are represented as points and metabolites as vectors. The clustering pattern was determined using PLS-DA, which identified the most discriminative variables (Fig. 3A). PLS-DA models that demonstrate R^2 and Q^2 values greater than 0.5 suggest a strong fit to the data and robust predictive capabilities. Cross-validation analysis using the first five components yielded a robust model with strong predictive power ($Q^2 = 0.96$), excellent fit to the data ($R^2 = 0.99$), and accuracy of 1 (Fig. 3B). PLS-DA detected 44 significantly different metabolites, based on VIP score > 2 (Fig. 3C). The pairwise scores plots (Fig. 3D).

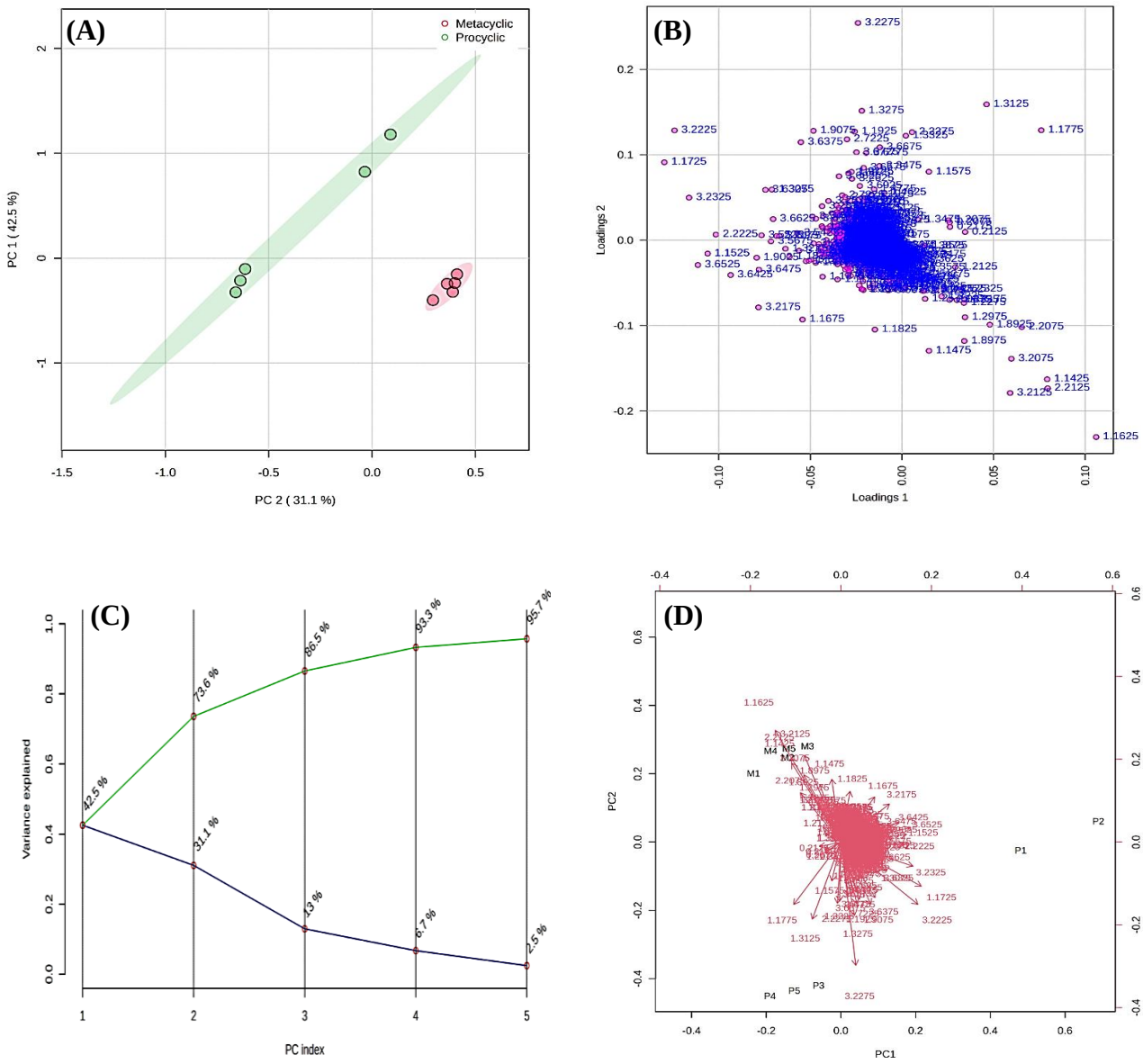

Fig. 2. (A**)** PCA score plot comparing PC1 and PC2 for 1H-NMR *L. major* procyclic and metacyclic stages. Each circle represents one sample; (B) PCA loading plot showing the metabolites that were significantly distant from the origin and played a key role in the discrimination of each group; (C) the scree plot displaying the variance explained by the first five PCs to identify the contribution of each component to the total variation in the dataset. The green line represents the cumulative variance explained by the first fivecomponents, accounting for 95.7%. The blue line indicates the individual variance for each PC. (D) PCA biplot between the selected PCs. This biplot for logarithmic and stationary phases illustrates the relationship between samples and metabolites. Samples are represented as points, and metabolites are represented as vectors.

display the scores for the first five PC of the procyclic and metacyclic groups. At first glance, components 1 and 2 tend to discriminate between the procyclic and metacyclic samples. Furthermore, the data were analyzed using OPLS-DA, which displayed the score plot of all metabolite features between procyclic and metacyclic groups in OPLS-DA (Fig. 4A), based onthe VIP score > 2 (Fig. 4B). The cross-validated cumulative modeled variation of predictive loading and orthogonal components are shown in Figure 4C. The cumulative modeled variation is represented by the quality of the models in the X matrix (R^2X), the cumulative modeled variation in the Y matrix (R^2Y), and the cross-validated predictive ability (Q^2). As shown in Figure 4C, the model demonstrates robustness, with Q^2 > 90% and R^2Y > 90%. The sPLS-DA score plot displayed supervised clustering and discrimination between the procyclic and metacyclic groups **(Fig. S1)**. The model effectively separates the groups of metabolites, as indicated by the clustering. Finally, we employed

Random Forest analysis to enhance our data analysis and gain deeper insights into the metabolic changes occurring during the conversion of the promastigote phases. Figure 5A displays the classification error associated with the RF method, while Figure 5B shows the significant features identified by this method. Table 1 lists the significant metabolites identified in each multivariate analysis.

Metabolic pathway analysis for differentially expressed metabolites

Among 25 enriched metabolic pathways observed in this study (Fig. 6A), significant pathways with $p < 0.05$ and impact > 0.05 were included in the final analysis. KEGG pathway analysis revealed 10 significant pathways including galactose, histidine, starch and sucrose, ascorbate and aldarate, arginine and proline, arginine biosynthesis, beta-alanine and pyruvate metabolism, fructose and mannose, and inositol phosphate metabolism, that can differentiate the procyclic from metacyclic groups (Fig. 6B).

Heat map analysis

Heat map visualization was performed using the normalized concentrations of procyclic and metacyclic groups **(Fig. S2)**. Hierarchical clustering generated a dendrogram demonstrating distinct clustering patterns, which reveal significant differences in metabolite profiles between the procyclic and metacyclic groups.

Identification of DEGs between the developmental stages

DEG analysis of data from the PRJNA252769 project **(Table S1)**, identified 287 upregulated and 49 downregulated genes at significantly different levels in pairwise comparisons between the procyclic and the metacyclic promastigotes. A summary of 10 DEGs is presented in **Table S2**. The DEGs were subsequently analyzed using GO and KEGG pathway analyses to identify cellular functions and processes related to *L. major* metacyclogenesis. The results of the pairwise comparisons revealed the upregulated and downregulated GO categories between the developmental stages **(Tables S3A and S3B)**. Several GO categories were upregulated in the metacyclic phase compared to the procyclic phase, including protein phosphorylation, ion transport, signal transduction, and phosphorylation reactions. Conversely, several GO categories including those associated with ATP biosynthesis, nucleosome and histone modification, cell cycle regulation (particularly cyclin-dependent protein kinases), mitochondrial proton-transporting activities, protein heterodimerization, and ion transporter activity were found to be downregulated during metacyclogenesis. In addition, changes in the expression of genes related to surface macromolecules were observed during promastigote differentiation, including an increase in the expression of mRNA encoding arabinosyltransferase-1. This enzyme modifies surface oligosaccharides and facilitates the detachment of the parasite from the sandfly midgut. The results of KEGG analysis were presented in **Table S3C**.

Integration of metabolomic and transcriptomic data

Among the total KEGG pathways identified from both metabolomics and transcriptomics, seven showed significant enrichment differences between procyclic and metacyclic forms. These pathways include fructose and mannose, galactose, ascorbate and aldarate, arginine and proline, histidine, inositol phosphate, and pyruvate metabolism (Fig. 7).

DISCUSSION

Transition of *L. major* from procyclic to metacyclic promastigotes is critical for the infectivity and survival of a parasite. The development of *Leishmania* promastigote within sandflies and under in vitro conditions involves metabolic changes that prepare the parasite for transmission to its mammalian hosts. The parasites optimize their metabolism to utilize limited resources and adapt to their environment[6,8,22]. This process, known as metacyclogenesis, is essential for the lifecycle of the parasite and its ability to infect mammalian hosts. Different omics approaches have been used to investigate the changes occur during metacyclogenesis.

Transcriptomic analysis of developmental stages in sandflies and in in vitro conditions demonstrated that differentiation from procyclic to metacyclic forms occurs gradually, accompanied by alterations in the metabolism-related genes. Interestingly, a comparison between metacyclics derived from sand flies and those cultured in vitro showed similarities in their transcriptome profiles[8]. ^{1}H NMR spectroscopy has been used as a valuable tool for characterizing metabolites. A preliminary study employing ^{1}H NMR spectroscopy analyzed metabolite outliers from the procyclic and metacyclic phases of *L. major* promastigotes. The results highlighted differences in biochemical profiles and metabolite pathways during Metacyclogenesis[5]. In the present study, ^{1}H NMR spectroscopy was employed to characterize and compare the metabolites and metabolic pathways of the procyclic and metacyclic phases of *L. major*

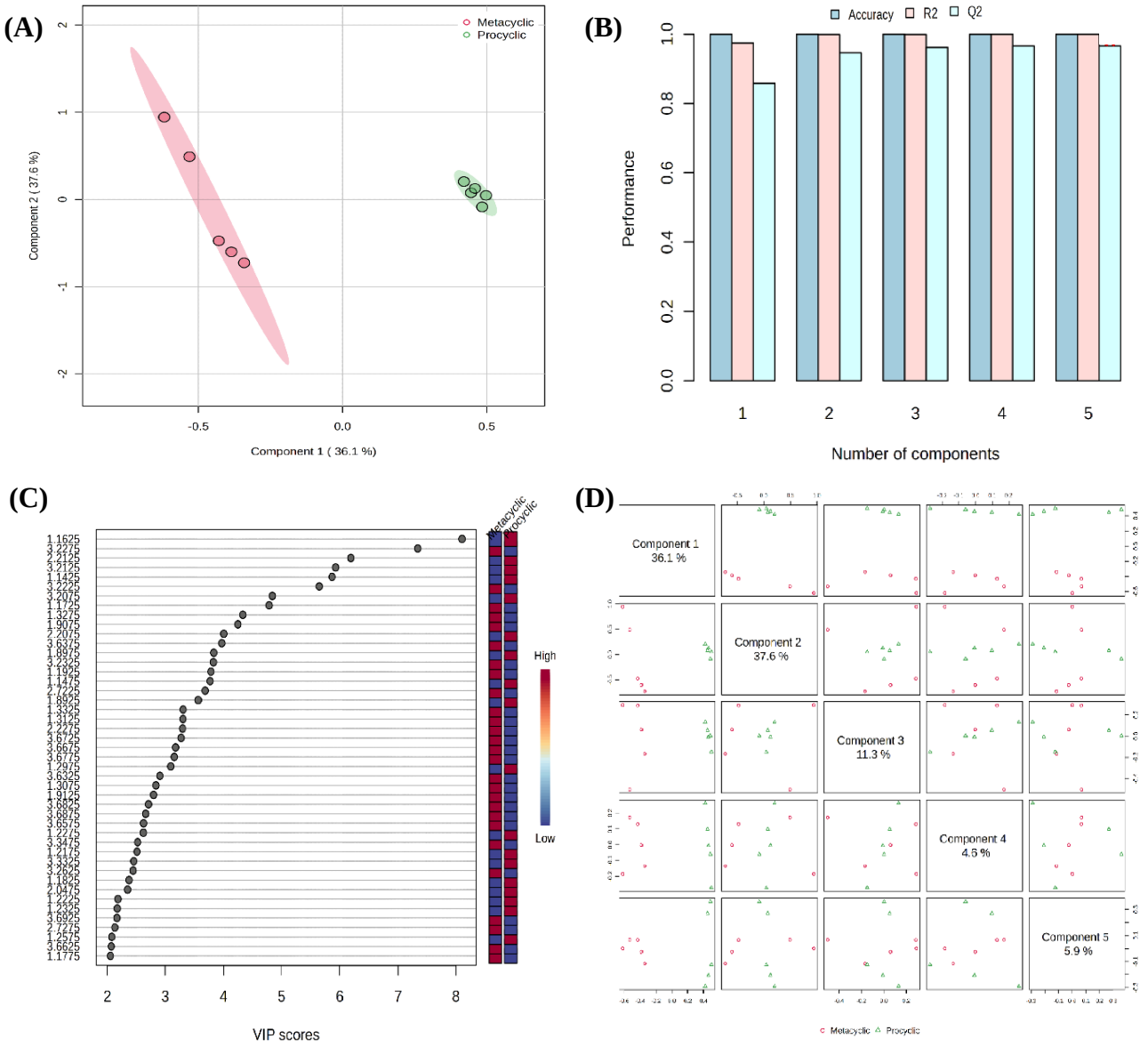

Fig. 3. (A) PLS scores plot between the selected components for ^{1}H-NMR promastigotes in *L. major* spectra obtained from the procyclic and metacyclic phases. (B) PLS-DA classification plot obtained using different numbers of components and following performing a 1000 random permutation test on PLS-DA model. R^2 and Q2 > 0.5 suggest that the model is an acceptable fit and has good predictive power. The cross-validation analysis produced a strong model with high predictability (Q^2 = 0.96), goodness of fit value (R^2 = 1), and accuracy = 1. (C) The VIP scores plot indicating important metabolites VIP > 2, which contribute to the discrimination of groups. The colored boxes on the right indicate the relative concentrations of the corresponding metabolite in each group. (D) Pairwise score plots for the first five components of the PLS-DA analysis. The first component explains 35.9% , second and third components explain 37.6% and 11.1% variability among the groups, respectively. The explained variance of each component is shown in the corresponding diagonal cell.

promastigotes. The findings revealed 44 significantly different metabolites and 10 distinct pathways between the two forms. DEG analysis of data retrieved from public repositories, followed by GO and KEGG pathway analyses, identified several altered gene expressions between procyclic and metacyclic phases. The metabolome analysis results were combined with those of transcriptomics, which identified seven significantly different KEGG pathways between procyclic and metacyclic forms. A pathway-based combination of metabolomics and transcriptomics provides a holistic view of metacyclogenesis. Transcriptomics reveals gene expression patterns, and metabolomics reflects the functional state of the cell by measuring metabolites, which are the end products of gene expression and metabolic pathways. By analyzing

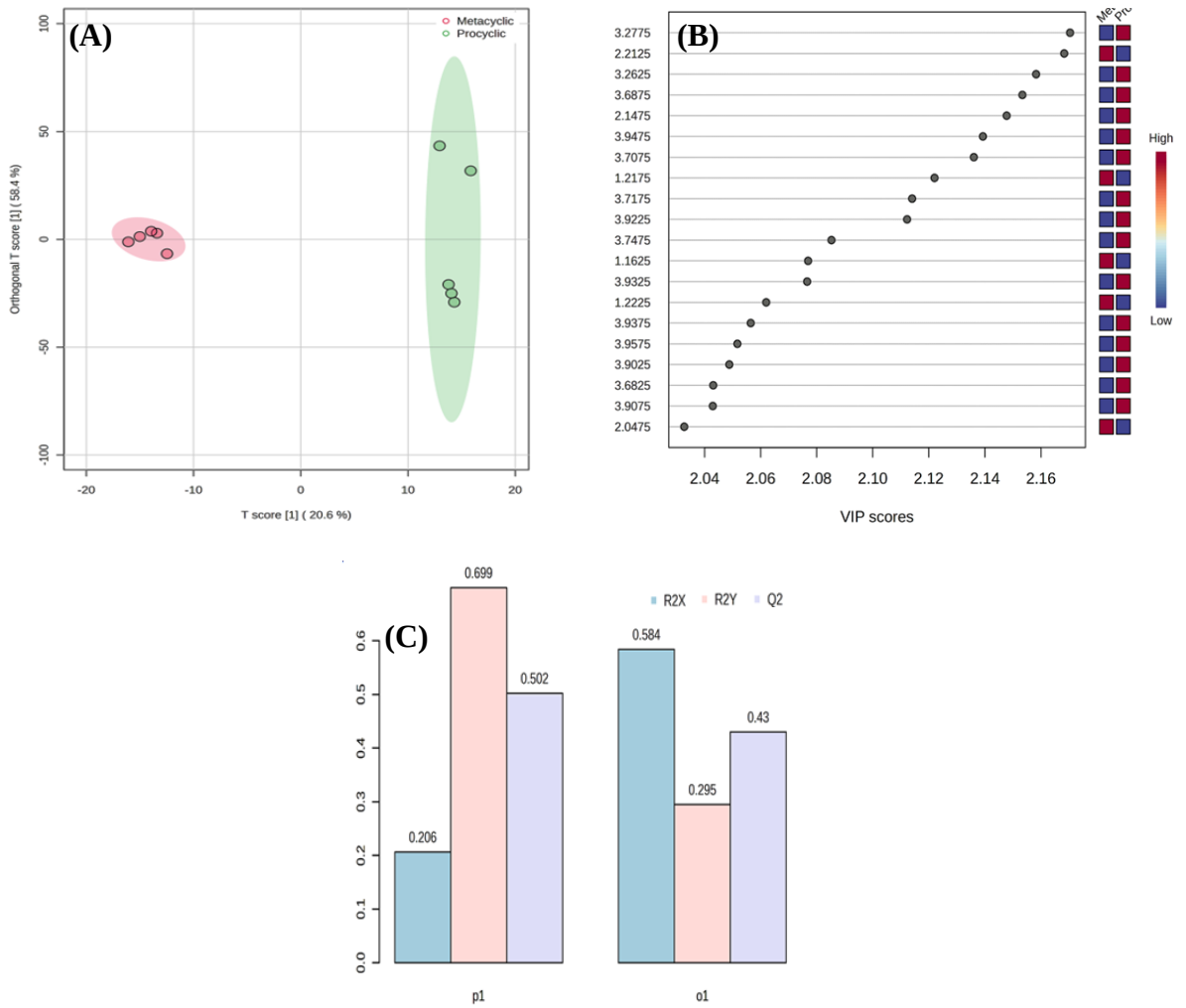

Fig. 4. (A) OPLS-DA score plot of all metabolite features between the logarithmic and stationary phases of promastigotes in *L. major*. (B) VIP scores of OPLS-DA. On the right, the colored boxes indicate the relative abundance of the groups. (C) Cross-validated cumulative modelled variation R2X, R2Y, and Q2 coefficients of predictive loading (p1) and orthogonal (o1) components.

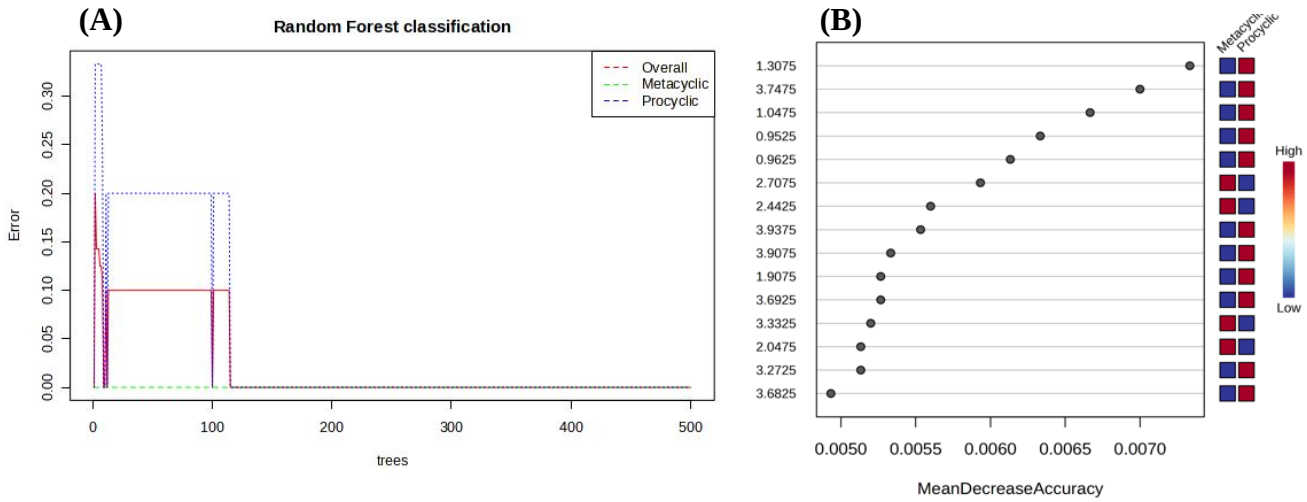

Fig. 5. (A) Cumulative error rates by Random Forest classification. (B) Important features identified using the RF algorithm. These features were subsequently ranked according to the average reduction in classification accuracy caused by permutation.

Table 1. A list of 44 key metabolites obtained from the t-test between procyclic and metacyclic groups

Metabolites name	Univariate analysis				Multivariate analysis		
	t-value	*p* value	− log10(p)	FDR	PLS-DA	OPLS-DA	Random Forest
HMDB0002098	17.987	9.36E-08	7.0287	5.19E-05	VIP scores		MDA
HMDB0002094	-17.043	1.43E-07	6.8457	5.19E-05	1.1625	3.2775	0.007333
HMDB0000660	-12.933	1.21E-06	5.9173	0.00033	3.2275	3.2625	0.005267
HMDB0000122	11.479	3.01E-06	5.5221	0.000656	2.2125	2.2125	0.005267
HMDB0001185	8.7108	2.35E-05	4.6281	0.002143	3.2125	3.6875	0.0052
HMDB0000186	7.0172	0.000111	3.9557	0.006046	1.1425	1.2175	0.005133
HMDB0003466	-6.7603	0.000144	3.8431	0.007463	3.2225	3.9225	0.004933
HMDB0004437	6.6713	0.000157	3.8033	0.007802	3.2075	3.9475	0.0048
HMDB0000148	6.6291	0.000164	3.7843	0.007802	1.1725	3.7075	0.004733
HMDB0000640	-6.1123	0.000286	3.5443	0.01155	1.3275	2.1475	0.004667
HMDB0015472	-5.8888	0.000366	3.4361	0.01258	1.9075	1.1625	0.004333
HMDB0001266	5.8087	0.000401	3.3967	0.01258	2.2075	3.7175	0.003933
HMDB0000508	5.7245	0.000442	3.3548	0.01258	3.6375	3.7475	0.0038
HMDB0001847	-5.6232	0.000497	3.3039	0.013256	1.8975	3.9325	0.003333
HMDB0062590	-5.4666	0.000597	3.2242	0.014533	3.2325	3.2575	0.003333
HMDB0000001	-5.3923	0.000652	3.1858	0.015147	1.1925	3.9075	0.003133
HMDB0031314	5.1017	0.000928	3.0325	0.019119	1.1475	1.2525	0.003133
HMDB0006567	-4.9784	0.001082	2.9659	0.021092	2.7225	3.9375	0.003133
HMDB0001432	-4.9546	0.001114	2.9529	0.02135	1.8925	1.2975	0.003
HMDB0000258	4.7539	0.001438	2.8422	0.025732	1.3325	2.0475	0.003
HMDB0001065	4.6665	0.00161	2.7932	0.027045	1.3125	3.9025	0.002933
HMDB0003559	4.5308	0.001922	2.7162	0.029987	2.2275		
HMDB0001151	-4.3161	0.00256	2.5918	0.035833	3.6725		
HMDB0000108	-17.345	1.24E-07	6.9053	5.19E-05	3.6675		
HMDB0003070	-10.425	6.22E-06	5.2064	0.001103	3.6775		
HMDB0000576	-10.197	7.34E-06	5.1345	0.001103	1.2975		
HMDB0003911	-9.9858	8.58E-06	5.0666	0.001103	3.6325		
HMDB0032055	-9.9089	9.09E-06	5.0415	0.001103	1.3075		
HMDB0000177	-8.9435	1.94E-05	4.712	0.002119	1.9125		
HMDB0001252	-8.7486	2.28E-05	4.6419	0.002143	3.6825		
HMDB0000211	-8.3203	3.29E-05	4.483	0.002762	3.6875		
HMDB0000634	-7.9113	4.73E-05	4.325	0.00369	3.6575		
HMDB0000306	7.6011	6.30E-05	4.2008	0.004585			
HMDB0000044	-7.3977	7.64E-05	4.1171	0.004941			
HMDB0061736	-7.3901	7.69E-05	4.1139	0.004941			
HMDB0000684	-7.3174	8.25E-05	4.0836	0.005005			
HMDB0000042	-7.0843	0.000104	3.9847	0.005954			
HMDB0000062	-6.404	0.000208	3.6815	0.009475			
HMDB0000355	-6.3616	0.000218	3.6618	0.009517			
HMDB0000472	-6.1959	0.000261	3.5841	0.010944			
HMDB0000517	-6.0311	0.000312	3.5053	0.011975			
HMDB0000163	6.0151	0.000318	3.4976	0.011975			
HMDB0000172	-5.826	0.000393	3.4052	0.01258			
HMDB0006577	-5.7778	0.000416	3.3813	0.01258			

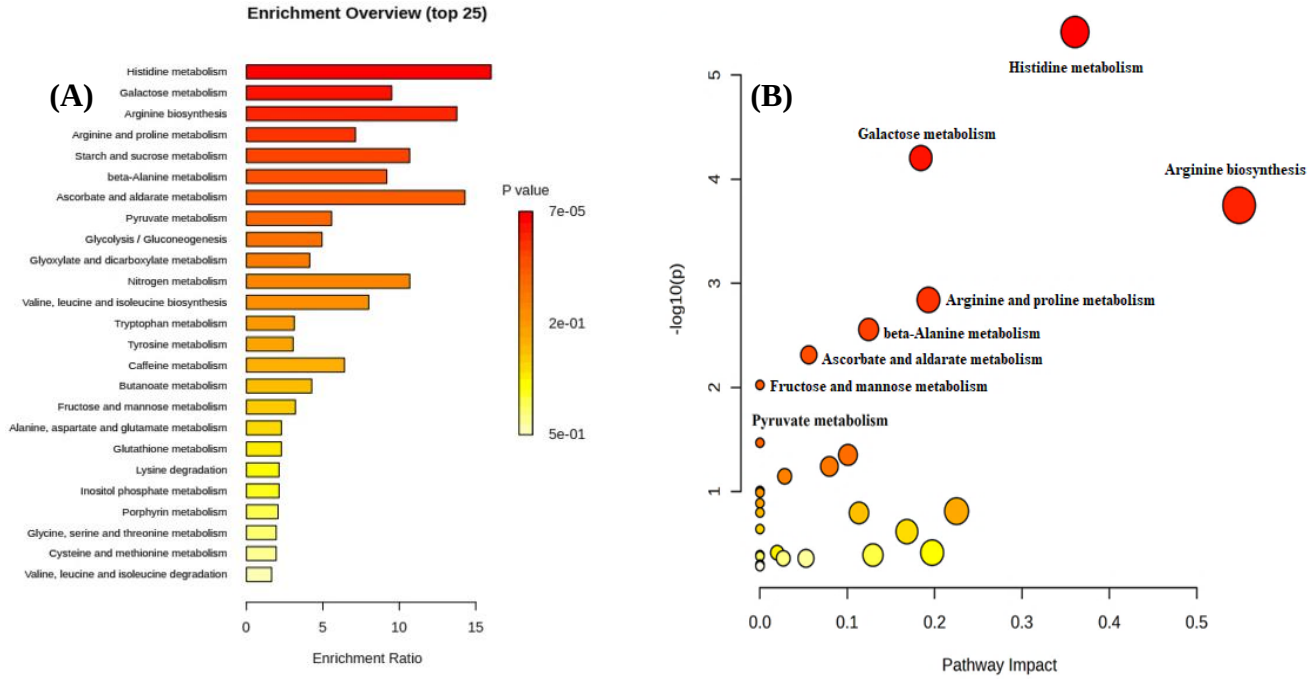

Fig. 6. The affected metabolite pathways during metacyclogenesis (A) for metabolite set enrichment analysis. Rankings are based on the Holm-Bonferroni test. (B) Important nodes in metabolic pathways, which show alterations during promastigote development. The x-axis represents the increasing impact of metabolic pathways based on betweenness centrality from pathway topology analysis.

these two types of data, we can correlate specific metabolic changes with the regulation of genes involved in key pathways. This approach will help elucidate the molecular mechanisms driving metacyclogenesis.

Our integrated metabolomic and transcriptomic analyses indicated significant changes in the fructose and mannose metabolism pathways between procyclic and metacyclic forms of *Leishmania*. In this parasite, the metabolism of fructose and mannose is interconnected with central carbon metabolism. Our transcriptomic analysis indicated a higher expression of the gene LmjF.36.0150 (fructose-6-phosphate 2-kinase/fructose-2,6-bisphosphate-like protein) in metacyclic compared to the procyclic forms. Fructose 2,6-bisphosphate is an important regulator of energy metabolism and a key regulator of glycolysis, which aligns with our metabolomics findings that exhibited fructose among the top 10 DAMs. Fructose plays a role in the production of alanine[23], and as fructose 6-phosphate, contributes to the synthesis of mannogen, a storage carbohydrate, as well as glycoconjugates[24]. Furthermore, our results indicated that the GO category GO:0033925 (mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase) is upregulated in metacyclic compared to procyclic forms. Mannose is essential for *Leishmania* parasites, primarily because they uniquely utilize this monosaccharide to build mannogen[24]. GO:0033925, which involves in the processing of free oligosaccharides within the cytosol, may act a role in mannogen regulation. Mannogen is stored in the cytosol and accumulated in non-dividing metacyclic promastigotes and amastigote stages, serving as a hexose source during glucose starvation conditions inside host cells[25,26]. In addition to mannogen, mannose is a major sugar component of various glycoconjugates on the surface of the parasite, which are essential for virulence and pathogenesis. Research has shown that mutations in mannose pathway enzymes in *L. mexicana* promastigotes disrupt the synthesis of mannose-containing glycoproteins and glycolipids, leading to attenuated infection establishment and reduced viability of the parasites under culture conditions[27].

As observed through the integration of metabolomic and transcriptomic analyses, the ascorbate and aldarate metabolism pathways are among the other altered pathways identified between procyclic and metacyclic forms. These pathways are crucial in protecting cells from oxidative stress and preventing damage to parasite cells. Our transcriptomic data analysis revealed an increased expression of ascorbate peroxidase mRNA (LmjF.34.0070) in the metacyclic form, which supports the findings of Inbar et al.[8]. Additionally, GO analysis indicated an upregulation of ascorbate-related GO categories (GO:0008447, GO:0016688) in metacyclic parasites compared to procyclic forms. LmAPX is central to the redox defense system of the parasite[28,29].

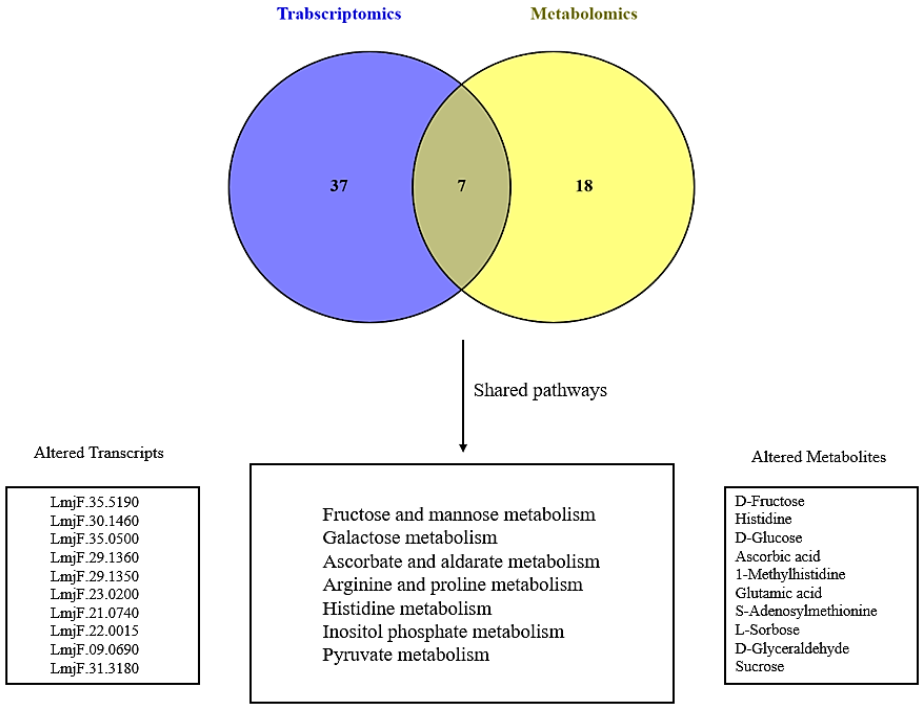

Fig. 7. Venn diagram depicting the integration of transcriptomic and metabolomic enriched pathways.

It is also an important factor to control metacyclogenesis and apoptosis in *L. major*. The loss of APX activity may have secondary effects on *Leishmania* gene expression due to the alterations in the redox balance within the protozoa, potentially resulting in the modulation of lipophosphoglycan formation and the differentiation of promastigotes into infective metacyclic forms[30,31].

Our findings indicated that the transformation of promastigotes influenced the galactose metabolism pathway. Galactose and its related glycoconjugates involve in the metacyclogenesis of *Leishmania* parasites. Furthermore, galactose in the form of galactofuranose, is crucial for specific host recognition and plays a significant role in the growth and virulence of the parasite[32-34]. The fourth different pathways between procyclic and metacyclic forms is the arginine and proline metabolism pathway. This pathway encompasses processes related to arginine and proline, including their synthesis, degradation, and interconversion with other metabolites. Arginine and proline are essential amino acids for the metabolic adaptations and survival of *Leishmania* throughout its lifecycle[35,36]. Pomastigotes use L-arginine to produce nitric oxide, which is important for metacyclogenesis and amastigote differentiation[37]. Additionally, L-arginine can be converted into the amino acids, proline and glutamate, facilitating a multi-metabolic fate during parasite growth and differentiation[38]. Proline is a significant source of energy for *Leishmania* and provides substrates for gluconeogenesis in the absence of glucose[39-41]. This amino acid, along with glutamate, aspartate, and glucose, promote metacyclogenesis[42].

Our results indicated alterations in the pyruvate metabolic pathway during metacyclogenesis, which is consistent with the findings of Arjmand et al. who also reported altered pyruvate metabolism during the metacyclogenesis of *L. major*[5]. Pyruvate metabolism is a central component of carbon metabolism in *Leishmania* parasites and is particularly important for ATP production. Given the increased glucose consumption in the procyclic phase, we can conclude that the pyruvate metabolism pathway is more active in this phase than in the metacyclic phase[43]. Pyruvate, generated from glycolysis, can be further catabolized in the mitochondria to form AcCoA. A significant portion of the AcCoA produced from pyruvate is converted to acetate, which serves as a substrate for fatty acid synthesis, the primary source of energy in amastigotes[22,44,45]. There is an increase in lipid abundance from procyclic to metacyclic forms, reflecting the preparation of parasite to infect mammalian host.

This study found alterations in the inositol phosphate metabolism pathway between metacyclic and procyclic forms, with an increase in myo-inositol metabolites in

the metacyclic stage. Transcriptomic data revealed the upregulation of the GO category (GO:0048017), which is related to inositol lipid-mediated signaling. Moreover, there was an elevation in the mRNA expression of LmjF.30.1850 and LmjF.29.1450 genes associated with the biosynthesis of inositol derivatives. Inositol lipids and their derivatives are essential for energy metabolism, serve as structural components of cell membranes, and involve in signal transduction. *Leishmania* parasites possess a variety of inositol lipids, particularly sphingolipids, which are vital for their survival and virulence. During metacyclogenesis, alterations in inositol phosphate metabolism represent a form of metabolic adaptation. Inositol lipids-null mutants of *L. major* lack inositol phosphoryl ceramides and other sphingolipids, which prevents them from differentiating into infective metacyclic parasites, ultimately leading to their death. This finding underscores the critical role of inositol phosphate metabolism in metacyclogenesis[46,47]. Alteration in the histidine metabolism pathway was also observed in our results. Histidine undergoes various metabolic pathways, leading to the production of urocanate, which is further processed into other metabolites, such as glutamate. Increased urocanate levels have been reported in stationary phase epimastigotes in *T. cruzi*[48]. Our findings showed a change in the level of carnosine metabolites (histidine derivatives), which play a significant role in protecting against oxidative stress a crucial trait for the survival and proliferation of *Leishmania* parasite.

CONCLUSION

The present study identified key metabolic pathways and DEGs involved in metacyclogenesis by integrating metabolomics and transcriptomics data. This integration enhances our understanding of the mechanisms driving this process and informed potential therapeutic strategies for the treatment of leishmaniasis. The combined application of metabolomics and transcriptomics has proved to be highly effective in uncovering altered pathways and identifying biomarkers associated with the infection process.

DECLARATIONS

Acknowledgments

No artificial intelligence services were used for the preparation of this manuscript.

Ethics approval

All experiments were conducted following the ethical code IR.PII.REC.1400.0.56 as a part of Ph.D. thesis.

Consent to participate

Not applicable.

Consent for publication

All authors reviewed the results and approved the final version of the manuscript.

Authors' contributions

MA: data curation, formal analysis, methodology, investigation, and writing original draft; FB: writing review and editing; ZP: methodology, formal analysis, and validation; MV: methodology and formal analysis; ZK: NMR data collection, validation; MA: conceptualization, funding acquisition, project administration, and supervision writing original draft; SA: resources, project administration, supervision, writing review, and editing.

Data availability

All relevant data can be found within the manuscript and its supplementary files.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was financially supported by the Pasteur Institute of IRAN (Grant ZS-9501 to MA).

Supplementary information

The online version contains supplementary material.

REFERENCES

1. Mann S, Frasca K, Scherrer S, Henao-Martínez AF, Newman S, Ramanan P, et al. A Review of Leishmaniasis: Current Knowledge and Future Directions. Curr Trop Med Rep. 2021;8(2):121-32.
2. WHO. Leishmaniasis: WHO 2023 [Available from: https://www.who.int/news-room/fact-sheets/detail/leishmaniasis
3. Torres-Guerrero E, Quintanilla-Cedillo MR, Ruiz-Esmenjaud J, Arenas R. Leishmaniasis: a review. F1000Res. 2017;6:750.
4. Sunter J, Gull K. Shape, form, function and Leishmania pathogenicity: from textbook descriptions to biological understanding. Open Biol. 2017;9(9):170165.
5. Arjmand M, Madrakian A, Khalili G, Najafi Dastnaee A, Zamani Z, Akbari Z. Metabolomics-based study of logarithmic and stationary phases of promastigotes in Leishmania major by 1H NMR spectroscopy. IBJ. 2016;20(2):77-83.
6. Subramanian A, Sarkar RR. Revealing the mystery of metabolic adaptations using a genome scale model of Leishmania infantum. Sci Rep. 2017;7(1):10262.

7. Amiri-Dashatan N, Rezaei-Tavirani M, Zali H, Koushki M, Ahmadi N. Quantitative proteomic analysis reveals differentially expressed proteins in Leishmania major metacyclogenesis. Microb Pathog. 2020;149:104557.
8. Inbur E, Hughitt VK, Dillon LA, Ghosh K, El-Sayed NM, L Sacks D. The transcriptome of Leishmania major developmental stages in their natural sand fly vector. mBio. 2017;4(2):e00029-17.
9. Alcolea PJ, Alonso A, Molina R, Jiménez M, Myler PJ, Larraga V. Functional genomics in sand fly-derived Leishmania promastigotes. Plos Negl Trop Dis. 2019;13(5):e0007288.
10. Xia J, Wishart DS. Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. Curr Protoc Bioinformatics. 2016;55(1):14.10.1-14.10.91.
11. Wishart DS, Guo A, Oler E, Wang F, Anjum A, Peters H, et al. HMDB 5.0: The human metabolome database for 2022. Nucleic Acids Res. 2022;50(D1):622-31.
12. Hoch JC, Baskaran K, Burr H, Chin J, Eghbalnia HR, Fujiwara T, et al. Biological magnetic resonance data bank. Nucleic Acids Res. 2023;51(D1):368-76.
13. Fernandes MC, Dillon LA, Belew AT, Bravo HC, Mosser DM, El-Sayed NM. Dual transcriptome profiling of Leishmania-infected human macrophages reveals distinct reprogramming signatures. mBio. 2016;7(3):e00027-16.
14. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.
15. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32(19):3047-8.
16. Shanmugasundram A, Starns D, Böhme U, Amos B, Wilkinson PA, Harb OS, et al. TriTrypDB: An integrated functional genomics resource for kinetoplastida. Plos Negl Trop Dis. 2023;17(1):e0011058.
17. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357-9.
18. 18. Robinson MD, McCarthy DJ, Smyth GK. Edger: A bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-40.
19. 19. Reimand J, Kull M, Peterson H, Hansen J, Vilo J. G:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic Acids Res. 2007;35(Web Server issue):193-200.
20. 20. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017;45(D1):D353-61
21. Jia A, Xu L, Wang Y. Venn diagrams in bioinformatics. Brief Bioinform. 2021;22(5):bbab108.
22. Marta Silva A, Cordeiro-de-Silva A, Coombs GH. Metabolic variation during development in culture of Leishmania donovani promastigotes. Plos Negl Trop Dis. 2011;5(12):e1451.
23. Yuan M, Vásquez-Valdivieso MG, McNae IW, Michels PAM, Fothergill-Gilmore LA, Walkinshaw MD. Structures of Leishmania fructose-1,6-bisphosphatase reveal species-specific differences in the mechanism of allosteric inhibition. J Mol Biol. 2017;429(20):3075-89.
24. Sernee MF, Ralton JE, Dinev Z, Khairallah GN, O'Hair RA, Williams SJ, et al. Leishmania β-1,2-mannan is assembled on a mannose-cyclic phosphate primer. Proc Natl Acad Sci U S A. 2006;103(25):9458-63.
25. Garami A, Ilg T. Disruption of mannose activation in Leishmania mexicana: GDP-mannose pyrophosphorylase is required for virulence, but not for viability. EMBO J. 2001;20(14):3657-66.
26. Akilov OE, Kasuboski RE, Carter CR, McDowell MA. The role of mannose receptor during experimental leishmaniasis. J Leukoc Biol. 2007;81(5):1188-96.
27. Kumar A, Das S, Purkait B, Sardar AH, Ghosh AK, Dikhit MR, et al. Ascorbate peroxidase, a key molecule regulating amphotericin B resistance in clinical isolates of Leishmania donovani. Antimicrob Agents Chemother. 2014;58(10):6172-84.
28. Xiang L, Laranjeira-Silva MF, Maeda FY, Hauzel J, Andrews NW, Mittra B. Ascorbate-dependent peroxidase (APX) from Leishmania amazonensis is a reactive oxygen species-induced essential enzyme that regulates virulence. Infect immun. 2019;87(12):e00193-19.
29. Santos IFM, Moreira DS, Costa KF, Ribeiro JM, Murta SMF, Santi AMM. Ascorbate peroxidase modulation confirms key role in Leishmania infantum oxidative defence. Parasit Vectors. 2024;17(1):472.
30. Elmahallawy EK, Alkhaldi AAM. Insights into Leishmania molecules and their potential contribution to the virulence of the parasite. Vet Sci. 2021;8(2):33.
31. Damerow S, Hoppe C, Bandini G, Zarnovican P, Buettner FF, Ferguson MA, et al. Leishmania major UDP-sugar pyrophosphorylase salvages galactose for glycoconjugate biosynthesis. Int J Parasitol. 2015;45(12):783-90.
32. Jain S, Sahu U, Kumar A, Khare P. Metabolic pathways of Leishmania parasite: Source of pertinent drug targets and potent drug candidates. Pharmaceutics. 2022;14(8):1590.
33. Rodriguez-Contreras D, Feng X, Keeney KM, Bouwer HG, Landfear SM. Phenotypic characterization of a glucose transporter null mutant in Leishmania mexicana. Mol Biochem Parasitol. 2007;153(1):9-18.
34. Marchese L, Nascimento JF, Damasceno FS, Bringaud F, Michels PAM, Silber AM. The uptake and metabolism of amino acids, and their unique role in the biology of pathogenic trypanosomatids. Pathogens. 2018;7(2):36.
35. Acuña SM, Aoki JI, Laranjeira-Silva MF, Zampieri RA, Fernandes JCR, Muxel SM, et al. Arginase expression modulates nitric oxide production in Leishmania (Leishmania) amazonensis. Plos one. 2017;12(11):e0187186.
36. Muxel SM, Aoki JI, Fernandes JCR, Laranjeira-Silva MF, Zampieri RA, Acuña SM, et al. Arginine and Polyamines Fate in Leishmania Infection. Front Microbiol. 2018;8:2682.
37. Sienkiewicz N, Ong HB, Fairlamb AH. Characterisation of a putative glutamate 5-kinase from Leishmania donovani. FEBS J. 2018;285(14):2662-78.
38. Saunders EC, Naderer T, Chambers J, Landfear SM,

McConville MJ. Leishmania mexicana can utilize amino acids as major carbon sources in macrophages but not in animal models. Mol Microbiol. 2018;108(2):143-58.
39. Burchmore RJS, Rodriguez-Contreras D, McBride K, Merkel P, Barrett MP, Modi G, et al. Genetic characterization of glucose transporter function in Leishmania mexicana. Proc Natl Acad Sci U S A. 2003;100(7):3901-6.
40. Homsy JJ, Granger B, Krassner SM. Some factors inducing formation of metacyclic stages of Trypanosoma cruzi. J Protozool. 1989;36(2):150-3.
41. Blum JJ. Effects of culture age and hexoses on fatty acid oxidation by Leishmania major. J Protozool. 1990;37(6):505-10.
42. Darling TN, Davis DG, London RE, Blum JJ. Metabolic interactions between glucose, glycerol, alanine and acetate in Leishmania braziliensis panamensis promastigotes. J Protozool 1989;36(2):217-25.
43. Lee SH, Stephens JL, Paul KS, T Englund P. Fatty acid synthesis by elongases in trypanosomes . Cell. 2006;126(4):691-9.
44. Van Hellemond JJ, Opperdoes FR, Tielens AG. Trypanosomatidae produce acetate via a mitochondrial acetate:succinate CoA transferase. Proc Natl Acad Sci U S A. 1998;95(6):3036-41.
45. Zhang O, Wilson MC, Xu W, Hsu FF, Turk J, Kuhlmann FM, et al. Degradation of host sphingomyelin is essential for Leishmania Virulence. Plos Pathogens. 2009;5(12):e1000692.
46. Zhang K, Showalter M, Revollo J, Hsu FF, Turk J, Beverley SM. Sphingolipids are essential for differentiation but not growth in Leishmania. EMBO J. 2003;22(22):6016-26.
47. Holeček M. Histidine in health and disease: Metabolism, physiological importance, and use as a supplement. Nutrients. 2020;12(3):848.
48. Mantilla BS, Azevedo C, Denny PW, Saiardi A, Docampo R. The histidine ammonia lyase of trypanosoma cruzi is involved in acidocalcisome alkalinization and is essential for survival under starvation conditions. M Bio. 2021;12(6):e0198121.

.

.