https://doi.org/10.33380/2305-2066-2022-11-1-23-31 UDC 615





Research article

Comparative transcriptomic analysis of the malaria parasites Plasmodium falciparum and Plasmodium vivax sensitive and resistant strains

Shrutika Sakpal^{1,2}*, Sayyed Zara Abdeen², Shanker Lal Kothari¹, Virupaksha Bastikar³

- ¹ Amity Institute of Biotechnology, Amity University Rajasthan, Jaipur, 303002, Rajasthan, India
- ² Department of Biotechnology, Dr. Homi Bhabha State University, The Institute of Science, Mumbai, 400032, India
- ³ Amity Institute of Biotechnology, Amity University, Maharashtra, 410206, India

*Corresponding author: Shrutika Sakpal, E-mail: shrutika.r.sakpal@gmail.com

ORCID: Shutika Sakpal – https://orcid.org/0000-0001-6939-1183; Sayyed Zara Abdeen- https://orcid.org/0000-0002-8077-1117; Shanker Lal Kothari- https://orcid.org/0000-0002-0378-4897; Virupaksha Bastikar- https://orcid.org/0000-0002-5036-8191

Received: 17.06.2021 Revised: 10.01.2022 Published: 25.02.2022

Abstract

Introduction. Malaria is the sixth leading cause of death worldwide. According to a WHO survey (2019-2020), the total number of malaria deaths is estimated to be 409000. *Plasmodium falciparum* and *Plasmodium vivax* are major malaria parasites, particularly in subtropical areas.

Materials and methods. In the present study, we used a transcriptome analysis of raw RNA sequence data to identify and characterize the differentially expressed genes in *Plasmodium falciparum* chloroquine-sensitive and chloroquine-resistant strains, as well as *Plasmodium vivax* primaquine sensitive and primaquine resistant strains. The raw RNA sequence data were obtained from the NCBI SRA database using the Accession IDs PRJNA308455, SRR14191963, and SRR332573.

Results and discussion. The sequence of raw RNA was quantified, mapped, and annotated. The total number of reads mapped to the reference genome for *Plasmodium falciparum* was found to be 45474448 and for *Plasmodium vivax* was 38226870. The Cufflinks-Cuffdiff tool was used to identify differentially expressed genes in *Plasmodium falciparum* and *Plasmodium vivax* sensitive and resistant strains. This differentially expressed gene was further annotated and plotted using the "Limma" package of Bioconductor. The PPI network was constructed in String Database and Cytoscape software. Pathway enrichment analysis of list differentially expressed gene performed using KEGG and GO tool.

Conclusion. In sensitive and resistant strains, comparative transcriptome analysis revealed differentially regulated gene expression patterns.

Keywords: Gene annotation, KEGG, Malaria, Protein-protein interaction, Transcriptome analysis

Conflict of interest. The authors declare that they have no obvious and potential conflicts of interest related to the publication of this article.

Contribution of the authors. Shrutika Sakpal developed the research plan and proposed an experimental strategy. Sayyed Zara Abdeen followed the directions to the letter. All necessary guidance was provided by Shanker Lal Kothari and Virupaksha Bastikar. All authors equally contributed to writing the current manuscript and approval of its text.

For citation: Sakpal S., Abdeen S. Z., Kothari S. L., Bastikar V. Comparative transcriptomic analysis of the malaria parasites *Plasmodium falciparum* and *Plasmodium vivax* sensitive and resistant strains. *Razrabotka i registratsiya lekarstvennykh sredstv* = *Drug development & registration*. 2022;11(1):23–31. https://doi.org/10.33380/2305-2066-2022-11-1-23-31

INTRODUCTION

The malaria parasite cyclically infects humans via female Anopheles mosquitos. Malaria parasites enter humans through the bite of a female Anopheles mosquito. After the invasion, these malarial parasites multiply and grow in the liver's hepatocytes before moving to red blood cells (RBC). Malaria parasite broods grow in the RBC and destroy them, releasing daughter parasites (merozoites) that continue the infection cycle. Malaria symptoms are caused by blood-stage parasites. When blood-stage parasites are consumed by a female Anopheles mosquito, they mate in the mosquito's gut and begin a cycle of growth and multiplication in the mosquito itself. Sporozoites migrate to the mosquito's salivary glands after 10–18 days. When an

Anopheles mosquito sucks blood from another human, the sporozoites are injected with anticoagulant saliva and migrate to the liver, starting a new cycle [1–19]. As a result, the infected mosquito acts as a "vector" for the disease, while infected humans transmit the parasite to the mosquito. The mosquito vector, unlike the human host, is unaffected by the presence of parasites.

The most common parasite causing malaria is *Plasmodium falciparum and Plasmodium vivax*. *Plasmodium falciparum* and *Plasmodium vivax* are the most common parasites that cause malaria. *Plasmodium falciparum* and *Plasmodium vivax* malaria parasites cause chronic blood-stage infections in humans that can last for weeks, months, or even years. The parasite growth processes, in which asexual merozoites invade RBC and replicate every 48 hours. *P. vivax* can survive in humans through

a reservoir of hypnozoites in the liver, in addition to blood-stage parasite replication. After being inoculated with *P. vivax* sporozoites from an infectious mosquito, some sporozoites develop into hypnozoites and remain dormant in the liver for weeks to years, until they activate to cause new blood-stage infections.

Relapses are blood-stage infections caused by the activation of P. vivax hypnozoites, and it is a distinguishing feature [5] between P. vivax and P. falciparum. Relapses can be avoided if an effective Primaquine is used. However, even with directly observed treatment, Primaquine does not guarantee clearance of hypnozoites in human patients with a low CYP2D6 metabolizer phenotype [19]. Chloroquine resistance, on the other hand, is a major cause of global increases in malaria mortality and morbidity in Plasmodium falciparum. Point mutations in the pfcrt gene have been identified as the cause of chloroquine (CQ) resistance in recent clinical studies. According to the WHO (2019), one child dies from malaria every two minutes. It is one of the major issues because of the inability to control disease in endemic and tropical areas, population migration, and serious complications caused by the disease itself. Because of the emergence of CQ resistant strains of Plasmodium falciparum, P. falciparum malaria is more severe than *Plasmodium vivax* malaria [8, 19]. Until the last decade, few studies in Asian countries demonstrated that P. vivax is capable of causing severe disease due to several important biological differences, such as the development of the dormant stage in the liver causing relapse and greater transmission potential of P. vivax at low parasite densities, which are responsible for these observations. P. vivax isolates from Southeast Asia (e.g., Thailand and Myanmar), Pacific Oceania (Papua New Guinea), and South America (Mexico, Peru, and Colombia) have significantly higher nucleotide diversity at the genome level than P. falciparum [11]. This could be due to frequent gene flow via human movement, intense transmission, or host susceptibility variation. In Papua New Guinea, for example, P. vivax infections had a 3.5-fold higher rate of polyclonality and nearly double the multiplicity of infection than P. falciparum infections [11]. However, Plasmodium vivax emerging resistance to CQ endangers the health of hundreds of millions of people who are routinely exposed to the risk of infection with this organism [18]. Since 1946, CQ has been the first-line therapy for *P. vivax* – caused malaria. Plasmodium falciparum developed CQ resistance in the 1950s, and it is now found worldwide. P. vivax resistance was unknown until 1989 when Australians who returned from Papua New Guinea failed routine treatment [11]. CQ, along with the combination of Primaguine, is still used as the first-line treatment for malaria in some countries. Only CQ-sensitive P. vivax was discovered in Thailand, the Philippines, and Vietnam. In Indonesia, surveys revealed a low prevalence of CRPV in the west (10 %) and a higher risk in the east (45 %). [5, 8, 11]. The advancement of NGS techniques, such as transcriptome studies, has provided insights into DNA and RNA, revealing gene expression variation that occurs in *Plasmodium* species during stage transitions and developmental processes. In this study, we characterized the transcriptional profiles of *Plasmodium falciparum* and *Plasmodium vivax* and compared resistant and sensitive strains to investigate the differentially expressed genes and their functions between these two strains.

MATERIALS AND METHODS

Data collection. Raw sequence data from a laboratory-adapted sensitive and resistant strain of *Plasmodium falciparum* were obtained from the NCBI SRA and deposited in the NCBI Bio project database under the accession number PRJNA308455 [12]. The raw sequence data of *Plasmodium vivax* sensitive and resistant strains are stored in the NCBI SRA under the accession numbers SRR14191963 and SRR332573.

Sequence Database curation. The raw sequence data quality in fastq files was checked using the FastQC tool [16]. Ligated adaptors were removed using the Trimmomatic tool (Bolger, A.M. et al., Germany), and all sequence quality reports were aggregated into a single report using the Multiqc tool (Philips Ewels, Sweden).

Mapping or alignment. The paired-end reads from *Plasmodium falciparum* and *Plasmodium vivax* were mapped to their respective reference genomes using the BWA-MEM [14] tool.

Quantification. The BWA-MEM mapping matrices were used for further quantification, using Samtool stats and samtool idx. (Heng Li, U.S.).

Assembling and DEA. Plasmodium falciparum and Plasmodium vivax mapped reads were assembled with their respective reference genome sequences using the Cufflinks program (Lior Patcher et al., U.S.) with default parameters. Cuffmerge was used to combine Plasmodium falciparum results into a single report. For Plasmodium vivax, we used the same method. To determine the differentially expressed genes between CQ sensitive and CQ resistant strains of Plasmodium falciparum and between PQ resistant and PQ sensitive strains of Plasmodium vivax, we used the Cuffdiff program to proceed with the Cuffmerge report of Plasmodium falciparum and Plasmodium vivax. The differentially expressed genes identified by Cuffdiff in Plasmodium falciparum and Plasmodium vivax were further investigated using the Plasmodb database.

Annotations and Identification. Annotations and identifications of differentially expressed genes were performed using the Gene identifier from the Cufflinks-Gene expression output to further classify and characterize gene types and their roles. Annotations were made with the Gene Annotations and Identifiers tool in the "Plasmodb Research resources" section (Beatrice Amos. et al., U.S.).

PPI Network analysis. To find potential interactions between Differentials expressed genes of *P. falciparum* CQ sensitive and CQ resistance strain and *P. vivax* PQ sensitive and PQ resistance strain, the string database tool (Szklarczyk D. et al., German) were used to construct

PPI network. The interaction score was adjusted to >0.4, Disconnected nodes were filtered out from the network and only active interaction sources such as Textmining, Experiment, gene fusion was applied to create a PPI network. The resulted in the PPI network being exported into Cytoscape software version 3.8.2 (Shannon. et al., U.S.) to visualize the PPI network of differential expressed genes. The confidence score of the PPI network was adjusted to >0.40. In the PPI network, the nodes represent frontness and edges represent the interaction between different proteins. The protein with high edges connection is termed as Interactome or nodes and its neighbor protein is called interactors. To evaluate the nodes in the PPI network, we used a topological measure such as Degree(K) in Cytoscape. Degree (K) and Between centrality (BC) often use for detecting the Hub and Bottleneck in the network. The degree (K) of any node is defined as the number of edges linked to it. A node with a high degree denotes a hub, has a connection with many neighbor's proteins. Here we consider the protein nodes with a high degree (K) for functional enrichment analysis.

Functional enrichment Analysis. Functional enrichment analysis of each protein with a high degree (K) along with its neighbor protein was performed to check whether these proteins associate with any biological function and pathways. The functional enrichment and pathway analysis were done using KEGG (Minoru Kanehisa. et al., Japan) and the GO tool (Ashburner et al., U.S.).

RESULTS AND DISCUSSION

The whole-genome transcriptome analysis and comparison of sensitive and resistant *Plasmodium falciparum* and *Plasmodium vivax* strains were completed successfully. Output data revealed a high degree of differentiation between sensitive and resistant strains based on gene expression and gene regulation; many genes and their functions were identified.

Quality check. The raw sequence quality of both *Plasmodium falciparum* and *Plasmodium vivax* strains was

checked using the FastQC tool, and the data were further aggregated into single reports using the Multiqc tool. For further analysis, we used Phred 30 as the quality score for each sequence.

Mapped reads. After each sequence was quality-checked, it was mapped to its reference genome. The total number of *Plasmodium falciparum* paired reads mapped to the reference genome was found to be 40,454,002 (88.96 %) for CQ sensitive and 39,175,728 (93.98 %) for CQ resistant. The total number of *Plasmodium vivax* paired reads mapped to the reference genome was 28,226,870 (67.71 %) for PQ sensitive and 37,300,760 (89.47 %) for PQ resistant.

Assembly and differentially expressed genes. To identify differentially expressed genes and transcripts, total mapped reads from Plasmodium falciparum and Plasmodium vivax were assembled using the Cufflinks program. Using the Cuffdiff program, the differentially expressed genes were identified based on their P-values and Log2 Fold change (Fc). The P-values between CQ sensitive and CQ resistant Plasmodium falciparum strains were discovered to be 0.05. The total number of differentially expressed genes identified by Cuffdiff analysis was 727, with 208 up-regulated and 519 downregulated. Whereas the P-values between *Plasmodium* vivax PQ sensitive and PQ resistant were found to be 0.05. The total number of differentially expressed genes was 2540, with 1284 up-regulated and 1256 down-regulated (Figure 1). Using the Bioconductor Limma package, we plotted P-values against Log2Fc in a volcano plot (Figure 2) to describe significantly expressed genes between CQ, PO sensitive, and resistant strains.

Genes Annotations. The differentially expressed genes were further identified and characterized based on their gene identifier using Plasmodb's tool. The identified genes appeared to be protein-coding genes involved in the production of surface antigens, invasion, pathogenesis, and virulence, revealing a significant difference between sensitive and resistant strains. We began by determining the functions of five highly expressed genes in both species (Table 1 and 2). The five most abundant

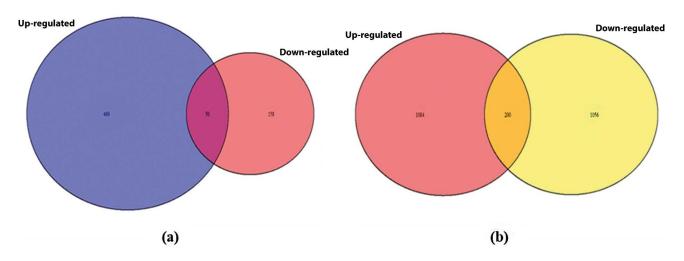


Figure 1. Up-regulated and down-regulated genes of (a) P. falciparum and (b) P. vivax

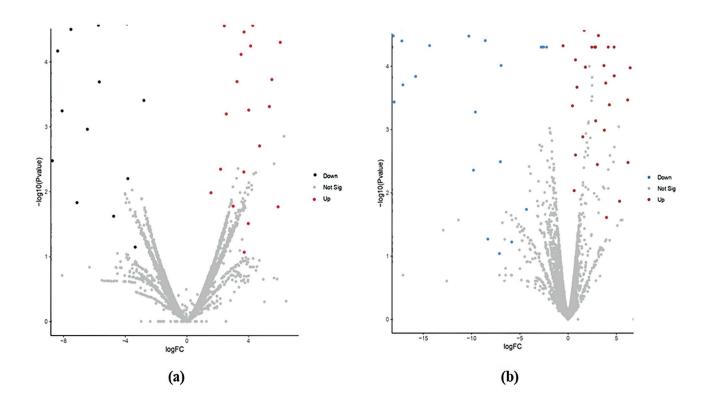


Figure 2. Differentially expressed gene of (a) P. falciparum and (b) P. vivax

genes in *P. vivax* strains were discovered to be genes coding for Reticulocyte binding protein, ADP Ribosylation factor, RNA pseudouridylate synthase, Rhoptry neck protein-5, and E3 ubiquitin-protein ligase. Multidrug resistance protein (PfMRP), Erythrocyte membrane protein (PfEMP), Hsp70, Hsp90, and Glyceraldehyde-3-phosphate dehydrogenase were discovered to be the five most highly expressed genes in *Plasmodium falciparum* strains

PPI Network Analysis. The PPI network was generated with help of a string database tool using DEGs of *P. falciparum* CQ sensitive and CQ resistance strain. The PPI network visualized in Cytoscape consists of 50 nodes and 71 edges (Figure 3). The result of the PPI network in

Cytoscape, show that *Plasmodium falciparum* Nucleoside Diphosphate Kinase(PfNDK) was hub(with largest degree K = 13) following with sub interactome *Plasmodium falciparum* Mitochondrial processing peptidase-α (PfMPP-α), *Plasmodium falciparum* Proliferating Cell Nuclear Antigen (PfPCNA),1-deoxy-D-5-phosphate reductoisomerase (DXR), *Plasmodium falciparum* Ribosomal protein L-35 (PfRPL-35), *Plasmodium falciparum* Multidrug resistance protein-1 (PfMDRP-1), *Plasmodium falciparum* Glycerophosphodiesterase (PfGDPD), *Plasmodium falciparum* TOC-75(PfTOC-75) with many interconnected nodes. PfNDK protein shows direct interaction with neighbor proteins such as GCP-2, RP-L35, PfExP, PCNA, Kinase, DXR, etc. The giant components of the PPI network

Table 1. Gene Annotation of Plasmodium falciparum

Gene_Id	Description	Function	Genomic Location
Gene- PF3D7_1462800	PF3D7_1462800 Glyceraldehyde-3-phosphate dehydrogenase Production of Energy through pathway require for developing		Pf3D7_14_v3:2,558,0552,560,979(+)
Gene- PF3D7_0818900 Hsp-70		Refolding of nascent protein. Assembly-dissemble of a multiprotein complex, Protein translocation, Protein degradation, Prion replication	Pf3D7_08_v3:858,550862,980(+)
1 (3ene- PE31) / () / ()84()() Hsn-9()		Development of malaria parasite and regulation of cell cycle	Pf3D7_07_v3:380,449385,546(+)
PF3D7_0523000 Multidrug Resistance Protein		Energy-dependent efflux pump, responsible for effluxing drug accumulation across the membrane	Pf3D7_05_v3:955, 955.963,095(+)
PF3D7_0600600 Erythrocytes Membrane Protein		Highly variable proteins, mediate cytoadhesion of infected erythrocytes to human endothelium facilitating further binding with uninfected RBC, helping parasitic schizont to invade and infect new RBC	Pf3D7_06_v3:29, 618.31,484(-)

Table 2. Gene Annotation of Plasmodium vivax

Gene_ld	Description	Function	Genomic Location
Gene-PVX_089055	E3 ubiquitin-protein ligase, putative	Protein ubiquitylation. Example: Tagging proteins for proteasomal degradation, or targeting to other distinct Sub-cellular locations	Pv_Sal1_chr05:259,893285,703(+)
Gene-PVX_089530	PVX_089530 Rhoptryneck protein 5, putative P.v RON-5 play a crucial role during host cell invasion, in erythrocyte, hepatocytes, mosquito salivar gland		Pv_Sal1_chr05:634,671642,746(+)
Gene-PVX_080660	7X_080660 RNA pseudouridylate synthase, putative Pseudouridylation, methylation, control multiple stages of gene expression regulations.		Pv_Sal1_chr10:829,185863,703(-)
PF3D7_1252400 Reticulocyte Binding Protein		Antigenic protein highly expresses in schizont responsible for invasion and infection in erythrocyte cells and targeting human immune response	Pf3D7_12_v3:2,136,0532,144,874(+)
PVX_001905	ADP-Ribosylation Factor	ADP-Ribosylation factor is a member of the ARF family of GTP binding protein. It is a major component of the Vesicular trafficking pathway, which plays a crucial role in the pathogenesis and survival of the malarial parasite	Pv_Sal1_chr06:240,794242,171(-)

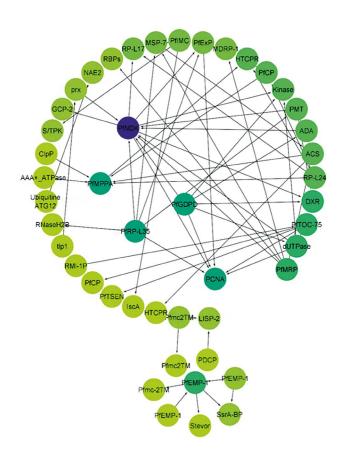


Figure 3. Protein-Protein Interaction (PPI) pathway of DEGs of *P. falciparum* strains

of *P. vivax* using DEGs were also generated in string database and visualized in Cytoscape using Degree(K) parameters (Figure 4). The results of the PPI network showing 70 nodes and 148 edges. Proteins with the highest degree were found to be Helicase (K = 28), DNA polymerase δ -1(POLD1) and Flap endonuclease-1(FEN-1) (K = 16), Reticulocyte binding protein(RBPs) (K = 11), RNA helicase protein(RNAHP) (K = 10) following with ADP Ribosylation factor (ARFs), Replication factor C-1 (RFC1), *Plasmodium falciparum* Dynein light chain-1 (PfDLC-1),

Histone acetyltransferase (HAT), Transporter protein and Splicing factor. Protein with higher degree showing more connection with their neighbor proteins like Exonuclease, Replication factor C-1 (RFC-1), AAA+Adenosine Triphosphatase, Stevor.

Functional enrichment analysis:. Gene ontology using protein with the highest degree was analyzed to check whether this protein associate with any biological function. According to our finding, *Plasmodium falciparum* Nucleoside Diphosphate Kinase (PfNDK) has the highest interactors showing interaction with other protein and protein molecules such as Ribosomal Protein L-35 (RPL-35), Kinase, Mitochondrial processing Peptidase α (MPP- α), Multi-Drug resistance protein (PfMRP). All these proteins sharing multiple interconnected functions in the Drug resistance pathway, Dna-Rna replication, DNA damage response and repair mechanism, Energy production, Cell cycle, and growth regulation inside the host body, which is a significant feature of any drug resistance parasite.

The protein molecules with the highest degree, Helicase, Flap endonuclease-1 (FEN-1), DNA Polymerase δ -1 (POLD1), Histone Acetyltransferase (HAT)in PPI network of *P. vivax* sharing interaction with other protein network and subnetwork, such as ADP ribosylation factors (ARFs), RNA Helicase, Replication Factor-C (RFC), Transporter protein, Splicing factor. The Association of these protein molecules has biological functions in an event such as erythrocyte invasion, Vehicular trafficking pathway, Viral genome replication and transcription in host, Pathogenesis, and drug resistance pathway.

The function of all the interactomes along with its neighbor protein (given below in Table 3 and 4) showing a significant role in immune tolerance, immune response, invasion, and Viral genome replication and translation inside the host body.

Plasmodium falciparum and Plasmodium vivax are the main strains causing malaria. P. falciparum CQ resistance and P. vivax Primaquine resistance are the major cause of increases in malaria mortality and morbidity rate worldwide. The progress in Computational Biology and

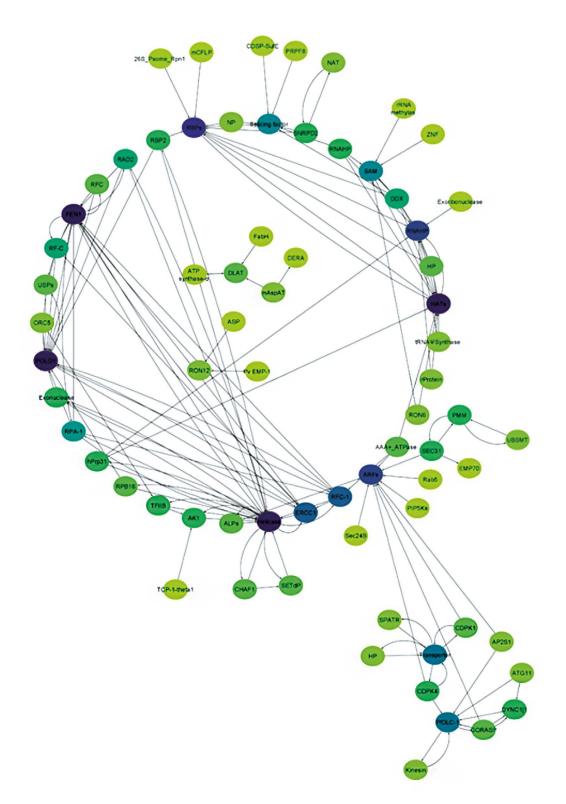


Figure 4. Protein-Protein Interaction (PPI) pathway of DEGs of P. vivax strains

Bioinformatics has simplified most of the research and study of malaria agents, it also has minimized the duration of the research from years to a month. Transcriptome analysis is one of those computational techniques. The transcriptome is a comprehensive analysis of whole sets of RNA of an organism's cell, tissue, or organ at a particular

time and condition. Whole-genome transcriptome analysis of *Plasmodium falciparum* and *Plasmodium vivax* provided the path for gene expression study using the whole genome sequence of malarial parasite. This is the first study to investigate the difference between CQ sensitive and CQ resistance of *P. falciparum* and PQ

Table 3. Function analysis of *P. falciparum* PPI network Protein

Sr. No	Protein name	Description	Function
1	PfNDK	Plasmodium falciparum Nucleoside Diphosphate Kinase	The catalytic enzyme requires for the production of nucleotides. Require for intracellular parasites proliferation, cell secretion, energy production, Apoptosis, DNA binding, and DNA repair mechanism
2	PfMPP α	P. falciparum Mitochondrial Processing Peptidase-α	DNA binding. Regulations of DNA synthesis, Leading strand elongation, Mismatch repair
3	PfRPL-35	P. falciparum Ribosomal Protein L-35	Structure constituents of ribosomes. Require for mRNA binding, Maturation of LSU reran from triastronic rRNA transcript
4	ERCC-1	P. falciparum Excision Repair ERCC1	Regulate drug resistance pathway. Help in nucleotide excision repair pathway
5	PfPCNA	P. falciparum Proliferating Cell Nuclear Antigen	Participate in DNA damage response pathway. Overexpression of PCNA help in the survival of malarial parasites in a genotoxic environment. Require for nucleic acid metabolism
6	PfDXR	1-deoxy-D-xylose 5-phosphate reductoisomerase	Contribute to MEP (2-C-Methyl-D-erythritol-4-phosphate) pathway for Isoprenoid synthesis
7	PfMRP	P. falciparum Multidrug Resistance Protein	Member of ATP binding cassette (ABC) transport superfamily. Energy-driven efflux pump responsible for effluxing drug accumulation in food vacuoles of parasites
8	PfGDPD	P. falciparum Glycerophosphodiesterase	Mg (2+) dependent enzymes catalyze glycerophospholipids during the growth of asexual intraerythrocytic malarial parasites. Hence im- portant in parasite's growth and pathogenicity
9	PfToc-75	P. falciparum Toc-75	Toc-75 play important role in protein secretion, outer membrane protein assembly, and bacterial growth
10	PfRPL-35	P. falciparum Ribosomal protein L-35	Regulate viral genome replication, transcription, and initiation of translation in the host cell
11	PfEMP-1	P. falciparum Erythrocyte Membrane Protein-1	Act as adhesion antigenic protein. Mediate adhesion of infected erythrocytes to human endothelium, facilitating binding with uninfected RBC by sequestration and resetting process. Hence spreading parasites pathogens in blood

Table 4. Function analysis of *P. vivax* PPI network Proteins

Sr. No	Protein name	Description	Function
1	Helicase	Helicase	Essential for regulation of viral genome replication-translation, DNA repair, mRNA splicing, Autophagy, Homeostasis
2	PvFEN	P. vivax Flap endonuclease	Play a crucial role in DNA repair mechanism. Exhibit RNase H activities
3	PvPOLD1	<i>P. vivax</i> DNA polymerase δ-1	Catalytic components of trimeric and tetrameric DNA polymerase require for high fidelity genome replication, lagging strand synthesis, and repair
4	PvHAT	P. vivax Histone Acetyltransferase	Require for parasites cell growth and differentiation. Acetylate conserved lysine residue on histone protein present in euchromatin, hence causing transcriptional activation
5	PvRBP	P. vivax Reticulocyte Binding Protein	In <i>P. vivax</i> , RBP helps parasites schizont in RBC invasion and targeting immune response to cause immune tolerance in parasites
6	PvRNA-H	P. vivax RNA Helicase	Function in RNA metabolism such as unwinding, annealing of RNA molecules, clamping protein complex on RNA remolding ribonucleo-protein complex
7	PvARF	P. vivax ADP Ribosylation Factor	ARF is a member of the ARF family GTP binding protein of the Ras superfamily. It is a major component of the Vesicular trafficking pathway, plays a crucial role in parasite survival and pathogenesis
8	PvRF-C	P. vivax Replication Factor-C	Replication Factor-C acts as a clamp loader in replication, involve in catalyzing to loading Proliferation cell nuclear antigen onto DNA So that it can encircle the DNA
9	PvTransporter protein	Transporter protein	An integral membrane protein facilitates the movement of nutrients, metabolites, inorganic ions into and out of the intraerythrocytic as well as between the subcellular compartment of parasites. Transporter protein is thought to efflux drug compounds out of the cellular compartments
10	PvDLC-1	P. vivax Dynein light chain-1	The dynein light chain is a member of the leucine-rich protein family. Help in controlling movement and accessibility of late trophozoites and schizonts into RBC

sensitive and PQ resistance of *P. vivax* at a genomic level using the whole genome sequence of malarial parasite. Through comparative transcriptome study between resistance and sensitivity of both P. falciparum and P. vivax parasite, differentially expressed genes including protein-coding genes involved in pathogenesis, invasion, surface antigens, cell cycle development, etc. were revealed. PfMRP, PfEMP, Hsp70, Hsp90, PfGDPD are the top 5 highly expressed genes between resistance and sensitive strain of P. falciparum. These genes code proteins, responsible for Multidrug resistance, Erythrocyte cell invasion and adhesion, nascent protein folding, assemble-dissemble of a multiprotein complex, protein translocation, protein degradation, prion replication, cell cycle regulation, and energy metabolism. Whereas highly expressed gene differencing between sensitive and resistant strains of *P. vivax* is gene code for Reticulocyte binding protein, ADP Ribosylation factor, E3 ubiquitinprotein ligase, Rhoptry neck protein 5 (RON-5), RNA pseudouridylate synthase. Reticulocytes binding protein play a crucial role in infecting reticulocyte subpopulation of red blood cells for invasion and targeting human immune response.ADP Ribosylation factor play role in the regulation of the Vesicular trafficking pathway and actin remolding. The vesicular trafficking pathway increases malarial parasite survival rate and pathogenesis inside the host. Ubiquitin ligase is involved in protein ubiquitination. For example, tagging proteins for proteasomal degradation, or targeting other distinct subcellular locations, P. vivax RON-5 helps host cell invasion, such as human erythrocytes, hepatocytes, or mosquito salivary glands. Gene coding for RNA of pseudouridylate synthase responsible for pseudouridylation, methylation, control multiple stages of gene expression regulations. Through PPI network analysis and functional enrichment analysis, we can find protein correlation with other proteins, which helps us to understand DEGs' function in the particular biological pathway and biological function. In the PPI network, most of the protein is protein molecules i. e. enzymes. In the PPI network of Plasmodium falciparum, proteins such as Nucleoside diphosphate kinase, Toc-75,1-deoxy-D-xylulose 5-phosphate reductoisomerase, Mitochondrial processing peptidase, and Proliferation cell nucleus antigen sharing interconnection. All these proteins together work in DNA replication, DNA repair mechanism, RNA transcription and translation, Cell proliferation, Parasite growth. Further interaction of Helicase, Flap endonuclease, Histone acetyltransferase, RNA helicase, Replication factor-c, and Transporter protein with each other share same multifunction such as host cell attachment, parasite genome, replication, RNA transcription, and mRNA translation in the host.

Through the overall study, we find that gene coding for proteins such as Multiple drug resistance protein, Erythrocytes membrane protein, proliferating cell nuclear antigen, Glycerophosphodiesterase, HSP-70/90 may be the reason behind drug resistance and immune tolerant

property of CQ resistance *P. falciparum* strain. In *P. vivax*, the gene coding for protein such as ADP Ribosylation factor, Reticulocyte binding protein, Rhoptry neck protein, Transporter protein may be the reason behind PQ resistance.

A limitation of this study is the lack of wet-lab experiments. All the results are based on the dry lab method and experiments. Hence for confirmation, the need of wet lab experiment is needed.

CONCLUSION

Comparative transcriptome analysis of P. falciparum CQ sensitive and CQ resistant strains, as well as P. vivax PQ sensitive and PQ resistant strains, was performed using whole-genome sequences. Comparative transcriptome analysis revealed differentially regulated gene expression patterns in sensitive and resistant strains. The total number of differentially expressed genes between resistant and sensitive P. falciparum strains was 727, with 208 up-regulated and 519 down-regulated genes, and the total number of differentially expressed genes between resistant and sensitive P. vivax strains was 2540, with 1284 up-regulated and 1256 down-regulated genes. Differentially expressed genes in P. falciparum and P. vivax sensitive and resistant strains code for the surface antigen, proteins involved in invasion and pathogenesis, virulence, and cell cycle development. The most common genes were discovered to code for Multidrug resistance, Reticulocyte binding protein, Erythrocyte membrane protein, Transporter protein, Vir-9, Rifin, Transmembrane, and Conserved Plasmodium proteins, which are involved in transportation, attachment, invasion, and pathogenesis, drug resistance, immune tolerance. The highly expressed genes between the sensitive and resistant strain of P. falciparum aid in signal transduction, nascent protein refolding, enzyme activation-deactivation, drug resistance, and erythrocyte infection. Whereas gene between P. vivax sensitive and resistant strains are in charge of protein ubiquitylation as well as host cell invasion, pathogenesis. When compared to sensitive strains, all of these genes were highly expressed in *P. falciparum* and *P. vivax* resistance strains. As a result, these genes may be responsible for CQ and PQ drug resistance.

REFERENCES

- Aurrecoechea C., Brestelli J., Brunk B.P., Dommer J., Fischer S., Gajria B., Gao X., Gingle A., Grant G., Harb O.S., Heiges M., Innamorato F., Iodice J., Kissinger J.C., Kraemer E., Li W., Miller J. A., Nayak V., Pennington C., Pinney D. F., Roos D. S., Ross C., Stoeckert C. J., Treatman C., Wang H. PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Research*. 2009;37(Database):D539–D543. DOI: 10.1093/nar/gkn814.
- Ford A., Kepple D., Abagero, B. R., Connors J., Pearson R., Auburn S., Getachew, Ford C., Gunalan K., Miller L. H., Janies D. A., Rayner J. C., Yan G., Yewhalaw D., Lo, E. Whole genome sequencing of *Plasmodium vivax* isolates reveals frequent sequence and structural polymorphisms in erythrocyte binding genes. *PLoS Neglected Tropical Diseases*. 2020;14(10):e0008234. DOI: 10.1371/journal.pntd.0008234

- Bolger A. M., Lohse M., Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–2120. DOI: 10.1093/bioinformatics/btu170.
- Andrews S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. Available online at: http://www. bioinformatics.babraham.ac.uk/projects/fastqc/ Accessed: 25.11.2021.
- Bousema T., Drakeley C. Epidemiology and infectivity of Plasmodium falciparum and Plasmodium vivax gametocytes in relation to malaria control and elimination. Clinical microbiology reviews. 2011;24(2):377–410.
- Hartzell J. D., Aronson N. E., Weina P. J., Howard R. S., Yadava A., Wortmann G. W. Positive rK39 serologic assay results in US servicemen with cutaneous leishmaniasis. *The American journal of tropical medicine and hygiene*. 2008;79(6):843–846.
- Bozdech Z., Llinás M., Pulliam B. L., Wong E. D., Zhu J., DeRisi J. L., Ward G. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS biology*. 2003;1(1):e5.
- 8. Cui L., Mascorro C. N., Fan Q., Rzomp K. A., Khuntirat B., Zhou G., Chen H., Yan G., Sattabongkot J. Genetic diversity and multiple infections of *Plasmodium vivax* malaria in Western Thailand. *The American journal of tropical medicine and hygiene*. 2003;68(5):613–619. DOI: 10.4269/ajtmh.2003.68.613.
- Blankenberg D., Von Kuster G., Bouvier E., Baker D., Afgan E., Stoler N., Galaxy Team, Taylor J., Nekrutenko A. Dissemination of scientific software with Galaxy ToolShed. *Genome biology*. 2014;15(2):1–3.
- Afgan E., Baker D., Batut B., van den Beek M., Bouvier D., Cech M., Chilton J., Clements D., Coraor N., Grüning B. A., Guerler A., Hill-man-Jackson J., Hiltemann S., Jalili V., Rasche H., Soranzo N., Goecks J., Taylor J., Nekrutenko A., Blankenberg, D. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic acids research*. 2018;46(W1):W537– W544. DOI: 10.1093/nar/qky379.
- 11. Fola A. A., Harrison G. A., Hazairin M. H., Barnadas C., Hetzel M. W., Iga J., Siba P. M., Mueller I., Barry A. E. Higher complexity of infec-

- tion and genetic diversity of *Plasmodium vivax* than *Plasmodium falciparum* across all malaria transmission zones of Papua New Guinea. *The American journal of tropical medicine and hygiene*. 2017;96(3):630–641. DOI: 10.4269/ajtmh.16-0716.
- Antony H. A., Pathak V., Parija S. C., Ghosh K., Bhattacherjee A. Whole transcriptome expression analysis and comparison of two different strains of *Plasmodium falciparum* using RNA-Seq. *Genome Data*. 2016;8:110–112. DOI: 10.1016/j.gdata.2016.04.004.
- Kukurba K. R., Montgomery S. B. RNA Sequencing and Analysis. Cold Spring Harbor Protocols. 2015;2015(11):pdb.top084970. DOI: 10.1101/pdb.top084970.
- Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013. Available at: https://arxiv.org/ abs/1303.3997. Accessed: 25.11.2021.
- Ritchie M. E., Phipson B., Wu D. I., Hu Y., Law C. W., Shi W., Smyth G. K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*. 2015;43(7):e47–e47.
- PlasmoDB: An integrative database of the Plasmodium falciparum genome. Tools for accessing and analyzing finished and unfinished sequence data. Nucleic Acids Research. 2001;29(1):66– 69. DOI: 10.1093/nar/29.1.66.
- Trapnell C., Roberts A., Goff L., Pertea G., Kim D., Kelley D. R., Pimentel H., Salzberg S. L., Rinn J. L., Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols*. 2012;7(3):562–578.
- World Health Organization. In vitro micro-test (Mark III) for the assessment of the response of *Plasmodium falciparum* to chloroquine, mefloquine, quinine, amodiaquine, sulfadoxine (No. CTD/ MAL/97.20 Rev. 2 2001). World Health Organization. 2001.
- World malaria report 2019. World malaria report 2019. In WHO Regional Office for Africa. Available at: https://www.who.int/ news-room/fact-sheets/detail/malaria. Accessed: 25.11.2021.