Data in Brief

Whole transcriptome expression analysis and comparison of two different strains of *Plasmodium falciparum* using RNA-Seq

Hiasindh Ashmi Antony [a], Vrushali Pathak [b], Subhash Chandra Parija [a,*], Kanjaksha Ghosh [c], Amrita Bhattacherjee [d]

[a] *Department of Microbiology, Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Puducherry 605006, India*
[b] *Department of Haematogenetics, National Institute of Immunohaematology (NII), Mumbai, Maharashtra 400012, India*
[c] *Surat Raktadan Kendra & Research Centre, Surat, Gujarat 395002, India*
[d] *SciGenom Labs Pvt. Ltd., Cochin, Kerala 682037, India*

ARTICLE INFO

ABSTRACT

The emergence and distribution of drug resistance in malaria are serious public health concerns in tropical and subtropical regions of the world. However, the molecular mechanism of drug resistance remains unclear. In the present study, we performed a high-throughput RNA-Seq to identify and characterize the differentially expressed genes between the chloroquine (CQ) sensitive (3D7) and resistant (Dd2) strains of *Plasmodium falciparum*. The parasite cells were cultured in the presence and absence of CQ by in vitro method. Total RNA was isolated from the harvested parasite cells using TRIzol, and RNA-Seq was conducted using an Illumina HiSeq 2500 sequencing platform with paired-end reads and annotated using Tophat. The transcriptome analysis of *P. falciparum* revealed the expression of ~5000 genes, in which ~60% of the genes have unknown function. Cuffdiff program was used to identify the differentially expressed genes between the CQ-sensitive and resistant strains. Here, we furnish a detailed description of the experimental design, procedure, and analysis of the transcriptome sequencing data, that have been deposited in the National Center for Biotechnology Information (accession nos. *PRJNA308455 and GSE77499*).

Specifications

Organism/cell line/tissue	*Plasmodium falciparum* 3D7 (chloroquine sensitive) and Dd2 (chloroquine resistant) strains
Sex	N/A
Sequencer or array type	Illumina HiSeq 2500
Data format	Raw and analyzed
Experimental factors	Laboratory strains. Dd2 strain grown in 160 nM chloroquine (MIC)
Experimental features	Gene expression profiling of *P. falciparum* 3D7 and Dd2 strains grown in the presence and absence of chloroquine using RNA-seq
Consent	N/A
Sample source location	Jawaharlal Institute of Postgraduate Medical Education and Medical Education, Puducherry, India

1. Direct link to deposited data

The raw data have been deposited in the BioProject database of NCBI with the accession no PRJNA308455 (http://www.ncbi.nlm.nih.gov/bioproject/308455). The processed data have been submitted in NCBI under the Gene Expression Omnibus (GEO) with the accession no GSE77499 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE77499).

2. Experimental design, materials and methods

2.1. Experimental design

Drug resistance in *Plasmodium falciparum* is a challenging major problem around the world. The first drug resistance was reported in 1960 for the therapeutic drug, chloroquine, which is cost-effective and available easily [1]. Since then many reports have been documented for various antimalarial drugs and reverse the decreased mortality and morbidity rate achieved through malaria control program. To better understand the genes involved in drug resistance and molecular mechanism, we conducted a high-throughput RNA-Seq for the chloroquine (CQ) sensitive and resistant strains of *P. falciparum*. Also, the

* Corresponding author.
E-mail address: subhashparija@yahoo.co.in (S.C. Parija).

CQ-resistant strain was cultured in the presence CQ to understand the changes in gene expression in exposure to the drug.

2.2. Cell culture

The laboratory adapted chloroquine sensitive (3D7), and resistant (Dd2) strains of *P. falciparum* were cultured continuously by in vitro technique as described [2]. The parasite cells were grown in 'O+' red blood cells at 1–5% parasitemia with 5% hematocrit in RPMI 1640 (Gibco) media supplemented with 5% Albumax (Gibco), 25 mM HEPES buffer, 2.1 g/L sodium bicarbonate, 0.05% hypoxanthine (Sigma), and 200 μg/mL gentamycin. Synchronization of the parasite cells was carried out when the parasitemia reaches ~10% using 5% D-Sorbitol (Sigma). The synchronized cells were maintained further and harvested at 48 h with 5% parasitemia containing ~90% tropho-zoites. *P. falciparum* Dd2 strain was cultured and harvested in the presence of minimum inhibitory concentration of CQ (160 nM/L) [3], in addition to the above-described growth condition. The growth of the Dd2 strain was reduced in the presence of CQ when compared to their growth in the absence of CQ.

2.3. RNA isolation and quality control

Total RNA was isolated from the harvested parasite cells with RNAiso Plus reagent (Takara, Japan), based on the manufacturer's instruction. The concentration of the isolated RNA was quantified using Nanodrop (Thermo Scientific, USA) and Qubit (Invitrogen, USA), and integrity of the RNA was measured by a 2100 Bioanalyzer (Agilent Technologies, CA).

2.4. RNA-Seq and assembly

Paired-end sequencing cDNA libraries were constructed from the three samples using a TruSeq RNA Sample Preparation Kit v2 low sample (LS) protocol (Illumina Inc., CA), based on the manufacturer's instructions. RNA-Seq was conducted for the three libraries on an Illumina HiSeq 2500 system (Illumina Inc., CA) with 2× coverage [4]. The data quality of the fastq files was checked with FastQC tool (Babraham Bioinformatics) and the reads with ≥30 Phred quality score were taken into account for analysis. The processed paired-end reads were mapped to the reference *P. falciparum* genome available in the Ensemble Genome database (Release 26) using Tophat program version 2.0.8 [4,5]. The numbers of paired-reads mapped to the reference genome were as follows: 37,469,173 (3D7, CQ-sensitive; 85.4%), 35,650,111 (Dd2, CQ-resistant; 89.1%) and 20,856,492 (Dd2w/CQ, cultured in 160 nM CQ; 84%).

2.5. Annotation and differential expression analysis

The mapped reads were assembled using the cufflinks program (version 2.0.2) with default parameters [5], to determine gene and transcript expression. The differentially expressed genes were determined with the p-value ≤0.05 between the CQ-sensitive (3D7, control) and resistant (Dd2 and Dd2w/CQ, test) strains using the cuffdiff tool in the cufflinks program. The differentially expressed genes were characterized further based on gene ontology enrichment using the strategy tool in the PlasmoDB database [6]. The volcano plot describing the significantly up and down-regulated genes between the 3D7, Dd2 and Dd2w/CQ strains is shown in Fig. 1. The cuffdiff analysis has identified a total of 316 differentially expressed genes of which 89 and 227 genes were up and down-regulated respectively between the 3D7 and Dd2 strains. Similarly, a total of 454 genes were differentially expressed, in which 45 and 409 genes were up down-regulated genes between the 3D7 and Dd2w/CQ strains. Most of the genes encoding for surface antigens involved in invasion, pathogenesis, showed a significant difference in gene expression between the sensitive and resistant strains.

Conflict of interest

The authors declare no conflict of interests.

Transparency document

The Transparency document associated with this article can be found, in online version.
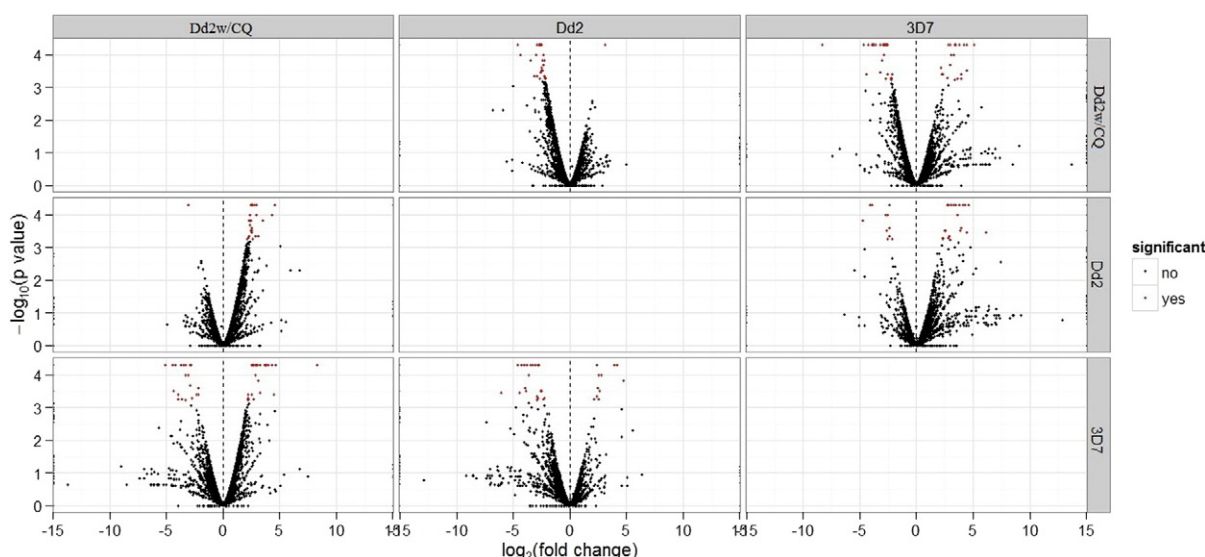
Acknowledgments

Fig. 1. Volcano plot for the significantly expressed genes for the three *P. falciparum* samples.

References

[1] D. Payne, Spread of chloroquine resistance in *Plasmodium falciparum*. Parasitol. Today 3 (1987) 241–246.

[2] W. Trager, J.B. Jensen, Human malaria parasites in continuous culture. Science 193 (1976) 673–675.

[3] World Health Organization, In Vitro Micro-test (Mark III) for the Assessment of the Response of *Plasmodium falciparum* to Chloroquine, Mefloquine, Quinine, Amodiaquine, Sulfadoxine/Pyrimethamine and Artemisinin. WHO, Geneva, 2001.

[4] K.R. Kukurba, S.B. Montgomery, RNA sequencing and analysis. Cold Spring Harb. Protoc. 11 (2015) 951–969.

[5] C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D.R. Kelley, et al., Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7 (2012) 562–578.

[6] C. Aurrecoechea, J. Brestelli, B.P. Brunk, J. Dommer, S. Fischer, B. Gajria, et al., PlasmoDB: a functional genomic database for malaria parasites. Nucleic Acids Res. (2009) D539–D543.