



## Data in Brief

## Dataset of natural antisense transcripts in *P. vivax* clinical isolates derived using custom designed strand-specific microarray



P.A. Boopathi <sup>a,1</sup>, Amit Kumar Subudhi <sup>a,1</sup>, Shilpi Garg <sup>a</sup>, Sheetal Middha <sup>b</sup>, Jyoti Acharya <sup>b</sup>, Deepak Pakalapati <sup>a</sup>, Vishal Saxena <sup>a</sup>, Mohammed Aiyaz <sup>c</sup>, Bipin Chand <sup>c</sup>, Raja C. Mugasimangalam <sup>c</sup>, Sanjay K. Kochar <sup>b</sup>, Parmendra Sirohi <sup>b</sup>, Dhanpat K. Kochar <sup>d</sup>, Ashis Das <sup>a,\*</sup>

<sup>a</sup> Department of Biological Sciences, Birla Institute of Technology and Science (BITS), Pilani, Rajasthan, India

<sup>b</sup> Department of Medicine, Sardar Patel Medical College, Bikaner, Rajasthan, India

<sup>c</sup> Genotypic Technology Pvt. Ltd., Bangalore, India

<sup>d</sup> Rajasthan University of Health Sciences, Jaipur, Rajasthan, India

## ARTICLE INFO

## Article history:

Received 12 May 2014

Received in revised form 18 June 2014

Accepted 19 June 2014

Available online 11 July 2014

## Keywords:

Natural antisense transcripts

Strand specific microarray

*Plasmodium vivax*

Complicated malaria

Uncomplicated malaria

## ABSTRACT

Natural antisense transcripts (NATs) have been detected in many organisms and shown to regulate gene expression. Similarly, NATs have also been observed in malaria parasites with most studies focused on *Plasmodium falciparum*. There were no reports on the presence of NATs in *Plasmodium vivax*, which has also been shown to cause severe malaria like *P. falciparum*, until a recent study published by us. To identify *in vivo* prevalence of antisense transcripts in *P. vivax* clinical isolates, we performed whole genome expression profiling using a custom designed strand-specific microarray that contains probes for both sense and antisense strands. Here we describe the experimental methods and analysis of the microarray data available in Gene Expression Omnibus (GEO) under GSE45165. Our data provides a resource for exploring the presence of antisense transcripts in *P. vivax* isolated from patients showing varying clinical symptoms. Related information about the description and interpretation of the data can be found in a recent publication by Boopathi and colleagues in *Infection, Genetics and Evolution* 2013.

© 2014 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

## Specifications

Organism/cell line/tissue	<i>Plasmodium vivax</i>
Source	<i>P. vivax</i> parasites from patients' blood
Sequencer or array type	Agilent custom <i>P. vivax</i> 244 K array. (AMADID: 019103)
Data format	Raw data: txt files, normalized data: SOFT, MINIMAL, TXT
Experimental factors	Complicated and uncomplicated <i>P. vivax</i> malaria
Experimental features	Identification of sense and antisense transcripts in <i>P. vivax</i> isolated from patients showing differing clinical symptoms.
Consent	The patient samples were collected on informed consent by the team of clinicians at S. P. Medical College, Bikaner, India according to hospital guidelines.
Sample source location	Bikaner, Rajasthan, India.

\* Corresponding author.

E-mail addresses: [boopathiarunachalam@gmail.com](mailto:boopathiarunachalam@gmail.com) (P.A. Boopathi), [amit4help@gmail.com](mailto:amit4help@gmail.com) (A.K. Subudhi), [shilpi\\_garg19@yahoo.co.in](mailto:shilpi_garg19@yahoo.co.in) (S. Garg), [sheetumig21@gmail.com](mailto:sheetumig21@gmail.com) (S. Middha), [jyotiacharya2@gmail.com](mailto:jyotiacharya2@gmail.com) (J. Acharya), [deepakpchowdary@gmail.com](mailto:deepakpchowdary@gmail.com) (D. Pakalapati), [vishalsaxena12@gmail.com](mailto:vishalsaxena12@gmail.com) (V. Saxena), [aiyaz@genotypic.co.in](mailto:aiyaz@genotypic.co.in) (M. Aiyaz), [bipin.c@genotypic.co.in](mailto:bipin.c@genotypic.co.in) (B. Chand), [raja@genotypic.co.in](mailto:raja@genotypic.co.in) (R.C. Mugasimangalam), [drskkochar@rediffmail.com](mailto:drskkochar@rediffmail.com) (S.K. Kochar), [drpsirohi@gmail.com](mailto:drpsirohi@gmail.com) (P. Sirohi), [drdkkochar@yahoo.com](mailto:drdkkochar@yahoo.com) (D.K. Kochar), [ashisd28@gmail.com](mailto:ashisd28@gmail.com), [ashdas28@gmail.com](mailto:ashdas28@gmail.com), [adas@pilani.bits-pilani.ac.in](mailto:adas@pilani.bits-pilani.ac.in) (A. Das).

<sup>1</sup> These authors contributed equally to this work.

## Direct link to deposited data

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE45165>

## Experimental design, materials and methods

## Sample collection and preparation

Venous blood samples (~5 ml) were collected from 8 *Plasmodium vivax* infected adult patients on informed consent at S. P. Medical College, Bikaner, India. The patients were diagnosed as either complicated (n = 7) or uncomplicated malaria (n = 1). The infection with *P. vivax* was confirmed by detailed investigation of peripheral blood films (PBFs) and rapid diagnostic tests (RDTs) (Optimal test; Diamed AG, Cressier sur Morat, Switzerland, Falcivax test; Zephyr Biomedical System, Goa, India). All other laboratory investigations as described in [4] were also performed. Peripheral blood mononuclear cells (PBMCs) were separated from infected and non-infected erythrocytes using density gradient based separation (Histopaque 1077, Sigma Aldrich, USA) according to manufacturer's instructions. The infected and non-infected erythrocytes were washed with phosphate buffered saline (PBS) and lysed using Tri-Reagent (Sigma Aldrich, USA) and preserved immediately

at  $-80^{\circ}\text{C}$ . All the samples were then transported in cold chain to BITS, Pilani for further processing. The infection with only *P. vivax* was confirmed by 18S rRNA based multiplex PCR and 28S rRNA based nested PCR [6,7].

#### Microarray hybridization and scanning

Total RNA and DNA was isolated from complicated ( $n = 7$ ) and uncomplicated ( $n = 1$ ) malaria blood samples, according to manufacturer's protocol (Tri-Reagent, Sigma Aldrich, USA). The quality of the isolated total RNA samples was analyzed by denaturing agarose gel and also by using RNA 6000 Nano Lab Chip on the 2100 Bioanalyzer (Agilent, Palo Alto, CA) following manufacturer's protocol. Quantity and purity of the total RNA was measured by the NanoDrop® ND-1000 UV-vis Spectrophotometer (Nanodrop technologies, Rockland, USA). The total RNA from 7 complicated malaria samples was pooled in an equimolar amount. Total RNA (500 ng) from each of the pooled complicated and uncomplicated malaria samples was amplified and labeled in the presence of Cyanine 5-CTP and Cyanine 3-CTP respectively using Low RNA Input Fluorescent Linear Amplification Kit (Agilent Technologies) following manufacture's protocol. After labeling, the cRNA was cleaned and the quality was assessed for yields and specific activity. Then 1500 ng of each Cy3 and Cy5-labeled samples was mixed, fragmented and hybridized to the array at  $65^{\circ}\text{C}$  for 16 h using Gene Expression Hybridization Kit (Agilent Technologies, Part Number 5188–5242). The hybridized slides were washed using Gene Expression wash buffers (Agilent Technologies, Part No: 5188–5327) and scanned using the Agilent Microarray Scanner (Agilent Technologies, Palo Alto, CA, G Model G2565BA) at  $5\ \mu\text{m}$  resolution. Image analysis was conducted using Agilent Feature extraction software (Agilent Technologies).

#### 244K custom array designing

A custom genome-wide strand specific *P. vivax* 244K microarray was designed on an Agilent platform using the RightDesign (Genotypic Technology, Bangalore, India) probe design workflow to choose the best probe(s) for a transcript by balancing several criteria: GC content, sequence complexity, cross hybridization potential and secondary structure. The array contains 232756 60-mer oligonucleotide probes (sense and antisense probes) representing *P. vivax* Sal-I transcript sequences from PlasmoDBv5.3 [1,3], expressed sequence tags (ESTs) and whole genome shotgun assemblies of *P. vivax* from NCBI (2007) and apicoplast sequences of *P. vivax* [9] and *Plasmodium falciparum* [12]. Annotations of all the probes were again updated according to PlasmoDB v8.2 and NCBI database (2012). Information about the array was described in detail in [2]. Summary of the array details has been submitted and available under GEO accession number GPL16492. Information such as feature number, oligonucleotide probe ID and sequence, target gene identifier ID against which probe has been designed and the gene description is provided in the array details. Here, we also give additional information about the orientation of probe against the target gene in the array (Supplementary Table S1).

#### Data analysis

The raw signal intensities and background intensities were obtained for each channel and analyzed separately for uncomplicated (Green channel) and complicated (Red channel) malaria. Analyses of probe hybridization for the PlasmoDBv8.2 [1,3] transcripts were discussed in the paper. A total of 5317 genes represented by both the sense and antisense probes were only considered for the analysis. Ratio was calculated for each probe by dividing the raw signal intensity and background signal intensity. Probes (sense or antisense) showing  $\geq 2$ -fold the background intensity were filtered. Probe data were converted to gene based data. For genes with multiple probes, median of raw signal intensity and background intensity was considered. Genes with

**Table 1**  
Genes expressing sense and antisense transcripts in complicated and uncomplicated malaria.

	Genes expressing		
	Only sense (S) transcripts	Only antisense (AS) transcripts	Both S and AS transcripts
PVC	959	106	942
PVU	782	117	934

Number of genes expressing AS transcripts in PVC is 1048.

Number of genes expressing AS transcripts in PVU is 1051.

Unique number of genes expressing AS transcripts in PVC & PVU is 1348.

$\geq$ twice the median background signal intensity were included in further analysis. To these genes, we have applied a stringent filtering criterion by considering genes expressing only sense(S), antisense (AS) or both S and AS transcripts represented by at least 3 probes. We categorized genes based on the type of transcripts they expressed (1) genes with only S transcripts (2) genes with only AS transcripts and (3) genes with both S and AS transcripts. Classification of genes with S and AS transcripts in complicated *P. vivax* malaria (PVC) and uncomplicated *P. vivax* malaria (PVU) is shown in Table 1. We detected a total of 1348 Natural Antisense transcripts using strand-specific custom designed microarray. Detailed analysis of this study has been published [2].

#### Discussion

Here we describe information about microarray dataset obtained from our custom designed strand-specific genome-wide *P. vivax* array on an Agilent platform. The dataset comprises whole genome transcriptome profiling of *P. vivax* isolated from patients showing differing clinical symptoms. The dataset was analyzed in recently published study and is the first study to reveal the presence of NATs in *P. vivax* clinical isolates. Discovery of NATs in *P. vivax* and *P. falciparum* [2,5,8,10,11] suggests that they might play an important role in regulating gene expression. Results from this microarray dataset thus would greatly assist investigations of gene regulation in future.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.gdata.2014.06.024>.

#### Conflict of interest

The authors declare that they have no conflicts of interest.

#### Acknowledgments

We thank all the patients and technical workers for their participation in support of this project. A.K.D., S.K.K. and D.K.K. acknowledges Department of Biotechnology (DBT), New Delhi, India for the financial support through the grant BT/PR7520/BRB/10/481/2006 and Birla Institute of Technology and Science, Pilani, India and S.P. Medical college, Bikaner, India for providing the required infrastructural facilities during this study. P.A.B. acknowledges Basic Scientific Research fellowship from University Grant Commission, New Delhi, India and Project Assistantship from Department of Biotechnology (DBT), New Delhi, India. A.K.S. acknowledges Senior Research Fellowship from the Council of Scientific and Industrial Research (CSIR), New Delhi, India and Project Assistantship from Department of Biotechnology (DBT), New Delhi, India. We thank the PlasmoDB team for making available the genome of the Sal I strain. We also thank Genotypic Technology Pvt. Ltd., Bangalore, India for the microarray hybridization and preliminary data analysis service provided by them.

## References

- [1] C. Aurrecochea, J. Brestelli, B.P. Brunk, J. Dommer, S. Fischer, B. Gajria, X. Gao, A. Gingle, G. Grant, O.S. Harb, M. Heiges, F. Innamorato, J. Iodice, J.C. Kissinger, E. Kraemer, W. Li, J.A. Miller, V. Nayak, C. Pennington, D.F. Pinney, D.S. Roos, C. Ross, C.J. Stoeckert Jr., C. Treatman, H. Wang, PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Res.* 37 (2009) D539–D543.
- [2] P.A. Boopathi, A.K. Subudhi, S. Garg, S. Middha, J. Acharya, D. Pakalapati, V. Saxena, M. Aiyaz, B. Chand, R.C. Mugasimangalam, S.K. Kochar, P. Sirohi, D.K. Kochar, A. Das, Revealing natural antisense transcripts from *Plasmodium vivax* isolates: evidence of genome regulation in complicated malaria. *Infect. Genet. Evol.* 20 (2013) 428–443.
- [3] J.M. Carlton, J.H. Adams, J.C. Silva, S.L. Bidwell, H. Lorenzi, E. Caler, J. Crabtree, S.V. Angiuoli, E.F. Merino, P. Amedeo, Q. Cheng, R.M. Coulson, B.S. Crabb, H.A. Del Portillo, K. Essien, T.V. Feldblyum, C. Fernandez-Becerra, P.R. Gilson, A.H. Gueye, X. Guo, S. Kang'a, T.W. Kooij, M. Korsinczky, E.V. Meyer, V. Nene, I. Paulsen, O. White, S.A. Ralph, Q. Ren, T.J. Sargeant, S.L. Salzberg, C.J. Stoeckert, S.A. Sullivan, M.M. Yamamoto, S.L. Hoffman, J.R. Wortman, M.J. Gardner, M.R. Galinski, J.W. Barnwell, C.M. Fraser-Liggett, Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. *Nature* 455 (2008) 757–763.
- [4] D.K. Kochar, A. Das, S.K. Kochar, V. Saxena, P. Sirohi, S. Garg, A. Kochar, M.P. Khatri, V. Gupta, Severe *Plasmodium vivax* malaria: a report on serial cases from Bikaner in northwestern India. *Am.J.Trop. Med. Hyg.* 80 (2009) 194–198.
- [5] M.J. Lopez-Barragan, J. Lemieux, M. Quinones, K.C. Williamson, A. Molina-Cruz, K. Cui, C. Barillas-Mury, K. Zhao, X.Z. Su, Directional gene expression and antisense transcripts in sexual and asexual stages of *Plasmodium falciparum*. *BMC Genomics* 12 (2011) 587.
- [6] D. Pakalapati, S. Garg, S. Middha, J. Acharya, A.K. Subudhi, A.P. Boopathi, V. Saxena, S.K. Kochar, D.K. Kochar, A. Das, Development and evaluation of a 28S rRNA gene-based nested PCR assay for *P. falciparum* and *P. vivax*. *Pathog. Glob. Health* 107 (2013) 180–188.
- [7] D. Pakalapati, S. Garg, S. Middha, A. Kochar, A.K. Subudhi, B.P. Arunachalam, S.K. Kochar, V. Saxena, R.P. Pareek, J. Acharya, D.K. Kochar, A. Das, Comparative evaluation of microscopy, OptiMAL(R) and 18S rRNA gene based multiplex PCR for detection of *Plasmodium falciparum* & *Plasmodium vivax* from field isolates of Bikaner, India. *Asian Pac. J. Trop. Med.* 6 (2013) 346–351.
- [8] C.A. Raabe, C.P. Sanchez, G. Randau, T. Robeck, B.V. Skryabin, S.V. Chinni, M. Kube, R. Reinhardt, G.H. Ng, R. Manickam, V.Y. Kuryshev, M. Lanzer, J. Brosius, T.H. Tang, T.S. Rozhdvestvensky, A global view of the nonprotein-coding transcriptome in *Plasmodium falciparum*. *Nucleic Acids Res.* 38 (2010) 608–617.
- [9] V. Saxena, S. Garg, J. Tripathi, S. Sharma, D. Pakalapati, A.K. Subudhi, P.A. Boopathi, G. S. Saggi, S.K. Kochar, D.K. Kochar, A. Das, *Plasmodium vivax* apicoplast genome: a comparative analysis of major genes from Indian field isolates. *Acta Trop.* 122 (2012) 138–149.
- [10] K. Sorber, M.T. Dimon, J.L. DeRisi, RNA-Seq analysis of splicing in *Plasmodium falciparum* uncovers new splice junctions, alternative splicing and splicing of antisense transcripts. *Nucleic Acids Res.* 39 (2011) 3820–3835.
- [11] A.K. Subudhi, P.A. Boopathi, S. Garg, S. Middha, J. Acharya, D. Pakalapati, V. Saxena, M. Aiyaz, H.B. Orekondy, R.C. Mugasimangalam, P. Sirohi, S.K. Kochar, D.K. Kochar, A. Das, Natural antisense transcripts in *Plasmodium falciparum* isolates from patients with complicated malaria. *Exp. Parasitol.* 141C (2014) 39–54.
- [12] R.J. Wilson, P.W. Denny, P.R. Preiser, K. Rangachari, K. Roberts, A. Roy, A. Whyte, M. Strath, D.J. Moore, P.W. Moore, D.H. Williamson, Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. *J. Mol. Biol.* 261 (1996) 155–172.