ISSN (Print) 0023-4001
ISSN (Online) 1738-0006

Korean J Parasitol Vol. 52, No. 1: 99-103, February 2014
http://dx.doi.org/10.3347/kjp.2014.52.1.99

▣ BRIEF COMMUNICATION

Whole Mitochondrial Genome Sequence of an Indian *Plasmodium falciparum* Field Isolate

Suchi Tyagi[1], Veena Pande[2], Aparup Das[1,*]

[1]*Evolutionary Genomics and Bioinformatics Laboratory, National Institute of Malaria Research, Sector- 8 Dwarka, New Delhi 110077, India;*
[2]*Department of Biotechnology, Kumaun University, Nainital, Uttarakhand 263001, India*

Abstract: Mitochondrial genome sequence of malaria parasites has served as a potential marker for inferring evolutionary history of the *Plasmodium* genus. In *Plasmodium falciparum*, the mitochondrial genome sequences from around the globe have provided important evolutionary understanding, but no Indian sequence has yet been utilized. We have sequenced the whole mitochondrial genome of a single *P. falciparum* field isolate from India using novel primers and compared with the 3D7 reference sequence and 1 previously reported Indian sequence. While the 2 Indian sequences were highly divergent from each other, the presently sequenced isolate was highly similar to the reference 3D7 strain.

Key words: *Plasmodium falciparum*, malaria, whole genome sequencing, mitochondrial genome, India

Malaria is a vector-borne infectious disease, endemic to many of the tropical and subtropical countries of the globe including India. Approximately half (273 million) of the high risk population outside Africa resides in India [1]. *Plasmodium falciparum* malaria is the leading cause of deaths and accounts for about 50% of malaria cases in India [2]. The problem is further compounded by high virulence and emergence of drug resistance in *P. falciparum*, and till now no effective vaccine is available. The primary huddle to design an effective vaccine that would work in all malaria endemic populations is highly observed genetic diversity in *P. falciparum* field isolates, as this parasite uses the genetic diversity to fight against the anti-malarial drugs and host immunity [3]. Therefore, the analysis of within-species genetic diversity is very important for understanding evolutionary processes both at the population and the genetic level which will not only enlightens the origin, historical migration, and demography of different populations, but also inform if new parasite genotypes of high virulence and drug resistance are emerging and spreading to different populations. Such understanding on the long run will definitely be of help in devising effective population-based control measures.

• Received 8 July 2013, revised 25 September 2013, accepted 9 December 2013.
* Corresponding author (aparup@mrcindia.org).

To this respect, mitochondrial (*mt*) genome has served as an ideal marker to understand evolutionary history of many model and non-model organisms and has been a marker of choice for reconstructing historical patterns of population demography and phylogenetic studies [4]. Several characteristics features; such as high mutation rate [5], maternal inheritance [6], and lack of recombination [7] have made the *mt* genome an ideal extra-nuclear genome to reconstruct evolutionary histories of the species. In malaria parasites, *mt* genome is of particular relevance, due to (i) its small size (~6 kb), (ii) haploid, and (iii) contains 3 protein-coding genes, cytochrome c oxidase I (*cox1*), cytochrome c oxidase III (*coxIII*) and cytochrome b (*cytb*) [8].

All these 3 genes are essential for a range for cellular processes; like membrane potential maintenance, heme and coenzyme Q biosynthesis, and oxidative phosphorylation [9]. Most importantly, the *cytb* gene of *P. falciparum* mitochondria is a potential target for an antimalarial drug, atovaquone [10]. Moreover, the *mt* genome of parasites evolves neutrally and shows no signs of recombination or selection [11], hence the whole genome behaves as a single locus and all sites share a common genealogy, which makes it ideal for studying within-species variations and phylogenetic analysis. While *mt* genome sequences have been reported from many malaria endemic countries of the world [11], only 1 complete *mt* genome sequence of *P. falciparum* isolate originating from unknown location in India has so far been reported [12]. Considering *P. falciparum* malaria is widespread in India, lack of *mt* genome se-

quence information from multiple Indian isolates has debarred us in unravelling the net diversity of Indian *P. falciparum*. To fill this gap in knowledge and to initiate population genomic studies of Indian *P. falciparum*, we herewith report the whole *mt* genome sequence of a single Indian field isolate, and the results of preliminary comparative genomic studies between the existing Indian data [12] and the reference sequence (3D7 isolate).

Finger-pricked blood sample of a microscopically diagnosed *P. falciparum* infected malaria patient from Bilaspur (Chhattisgarh state, India) was spotted (2-3 spots) on Whatmann filter paper, dried, and brought to the laboratory in New Delhi. Genomic DNA was extracted from these dried blood spots using QIAamp mini DNA isolation kit (Qiagen, Hilden, Germany). Since both *P. falciparum* and *Plasmodium vivax* are endemic to India, we used the isolated genomic DNA to perform PCR diagnostic assay to identify mixed infection following the nested PCR with genus and species-specific primers based on 18S rRNA gene [13]. The study was approved by the Ethics Committee of the National Institute of Malaria Research (NIMR), India, and written informed consent of the patients have been obtained.

To sequence the whole *mt* genome of the present *P. falciparum* isolate originated from Bilaspur (referred as Blsp1), we checked for PCR amplification using the sets of primers reported to sequence the Indian *P. falciparum* isolate PfPH10 [12]. However, we could not PCR-amplify any of the fragments of the *mt* genome of the Blsp1 isolate with repeated attempts. Therefore, we used the whole *mt* genome sequence information of the published 3D7 isolate, downloaded from the NCBI (www.ncbi.nlm.nih.gov/) database with GenBank accession no. AY282930, and chopped the whole *mt* genome into 19 different DNA fragments. In order to sequence the whole *mt* genome of *P. falciparum* Bslp1 isolate, we designed 19 novel primer-pairs. Two online computer programs, Primer3 and SIGMA primer-calculator were used to design these novel primer-pairs so as to keep the length of each fragment below 600 nucleotide base pairs (bp) with ~150 bp of overlapping sequences between each adjoining fragment. The length of each of the 19 PCR-amplified fragments was deliberately kept below 600 bp, as we have performed DNA sequencing following Sanger technology (see below).

All PCR amplification reactions were carried out in a final volume of 25 µl, which included 1 µl of each primer (10 *p*mol/µl), 0.2 mM of dNTP, 1 unit of Taq DNA polymerase (Merck) with 2.5 µl of 1x Taq DNA polymerase buffer and 1 µl of DNA. Amplifications were performed with the following cycling conditions: 95℃ for 5 min, then 35 cycles of 1 min denaturation at 95℃, 1 min annealing at different temperatures for different DNA fragments (Table 1), 1 min extension at 72℃ followed by 5 min final extension at 72℃. Successfully amplified PCR products were further purified by incubating with

Table 1. Details of primer sequences employed in PCR amplification and sequencing of the whole *mt* genome of *Plasmodium falciparum* along with annealing temperatures of the respective fragments

S. no.	Primer name	Primer sequence (5' to 3')	Annealing temperature (°C)
1	M_1F	TGCTATTGGATTCAACGTCC	63.7
	M_1R	GTCCTGCATGAACGGTGTA	
2	M_2F	TCGTAACCATGCCAACACAT	63.2
	M_2R	GCTGGGCATTTAATCCACTC	
3	M_3F	GGGTATCCAATCCAGTGCTC	63.7
	M_3R	CAAACACTAGCGGTGGAACA	
4	M_4F	AGGGAACAAACTGCCTCAAG	63
	M_4R	GGCATTTTGTTGAAATAGTCTGG	
5	M_5F	ACTTCCTTTCTCGCCATTTG	63.7
	M_5R	GCATCATGTATGAGTGCATGTT	
6	M_6F	TTGTAGAGATGCAAAACATTCTCC	60.4
	M_6R	GCACATCTAGTTTCATATCCTGCA	
7	M_7F	CAGAATAAAAACTTTCTCGAATAGG	61.8
	M_7R	AAGTACGCGATCTCTTGTATGG	
8	M_8F	CGCAGCCTTGCAATAAATAA	61.8
	M_8R	CATGAGGCTCGGATATAAATGA	
9	M_9F	GAACGCTTTTAACGCCTGAC	52.7
	M_9R	AGTCCATCCAGTTCCACCAC	
10	M_10F	CCAGGATTATTCGGAGGATT	52.8
	M_10R	CAGGATGTCCAAAATACCAGA	
11	M_11F	CCGGTTTTAACTGGAGGAGT	62.6
	M_11R	GCTACATCAATGGCAGCAT	
12	M_12F	CCGGTACAAAAGTATTTAACTGGA	62.6
	M_12R	GGTCATTGTTGTCCCAATAGAA	
13	M_13F	GCATTTCAAGATAATTTCTTTGGT	62.6
	M_13R	AAACATCTGGTGTATATCGACTTG	
14	M_14F	CACACTTAATAAATTACCCATGTCCA	62.6
	M_14R	GGATCACTCACAGTATATCCTCCA	
15	M_15F	TTGTCTTACCATGGGGTCAAA	62.6
	M_15R	CCAGCTGGTTTACTTGGAACA	
16	M_16F	TCACATCCTGATAATGCTATCG	62.6
	M_16R	CGAAGCATCCATCTACAGCTA	
17	M_17F	TTACAGCTCCCAAGCAAACA	62.6
	M_17R	GACGGTTTTCTGCGAAATCTA	
18	M_18R	GGGAGTTGGCAAGTTAAAGAAG	62.6
	M_18R	GGAAGTACGAATTGAAGTGTGG	
19	M_19F	CCTGGCTAAACTTCCCAATG	62.6
	M_19R	AGAAACAGTCGGTGCGAAGT	

Exonuclease-I and Shrimp Alkaline Phosphatase (Fermentas, Life Sciences, Berlington, Ontario, Canada) and DNA sequencing was performed in an ABI 3730XL DNA Analyzer (Applied Biosystems. Foster City, California, USA), an in-house sequencing facility of NIMR. All the 19 PCR-amplified fragments were sequenced in both forward and reverse directions (2x coverage) and each DNA sequence was individually edited and assembled using the EditSeq and MegAlign modules of the Lasergene (DNASTAR, Madison, Wisconsin, USA) computer program. All the 19 finally edited sequences were manually assembled to form a single whole *mt* genome sequence and aligned with the 3D7 reference sequence and the published Indian *P. falciparum* sequence (PfPH10) [12] using the MEGA v5.0 computer program [14] to ascertain nucleotide differences, if any, by comparing the *mt* genome sequences of 3 different isolates. Further, to understand the phylogenetic interrelationship among the 2 Indian (Blsp1 andPfPH10), the reference 3D7, and other 21 *Plasmodium* species, Neighbor-joining (NJ) phylogenetic tree was constructed using the MEGA v5.0 computer program [14] with 100 bootstrap replicates.

For phylogenetic analysis, the whole *mt* genome sequences of *Plasmodium* species infecting primates (GenBank no. AB354573, AB434919, AB434920, AB354574, AB354572, AY722799, AB354575, NC_007232, AB434918, NC_002235), rodents (GenBank no. AB379663, AB599931, AB558173), birds (GenBank no. AB599930, AB250415, AB302215), Lizard (GenBank no. NC_009961) and humans (GenBank no. M76611, NC_007243, AB354570, AB354571, AY282930), were downloaded from the NCBI web database (www.ncbi.nlm.nih.gov) and aligned using MEGA v5.0 computer program [14]. The whole *mt* genome sequence of the Blsp1 isolate has been deposited in GenBank public domain sequence database with accession number KJ144901.

Using the 19 novel primer-pairs (Table 1) designed in the present study, we could successfully sequenced the whole *mt* genome of a single *P. falciparum* field isolate (Blsp1) from an endemic locality of India with 2x coverage and compared with the whole *mt* genomes of 2 other isolates (3D7 and the previously reported PfPH10 isolate from India) [12]. A detailed list of novel primers designed in the present study is provided in Table 1, and the approximate locations of each primer-pair in the circular *mt* genome of *P. falciparum* is presented in Fig. 1. While the total length of the whole *mt* genome from 2 Indian *P. falciparum* isolates (PfPH10 and Blsp1) was similar in size of 5,967 bp (Fig. 1), surprisingly, the alignment of these 2 isolates revealed nucleotide differences in 22 positions; 2 of these were in *coxIII* gene, 5 in *cox1* and, 7 in *cytb* genes (Table 2), suggesting very high amount of variation in the *mt* genome in Indian *P. falciparum*. However, when the presently sequenced *mt* genome from Blsp1 isolate was aligned with the 3D7 isolate, only 1 nucleotide difference could be observed between these 2 genomes (Table 2).

This observation was in contrast to an earlier report on high sequence variation in *mt* genome between the worldwide and Indian isolates involving the PfPH10 isolate [12]. In any case,

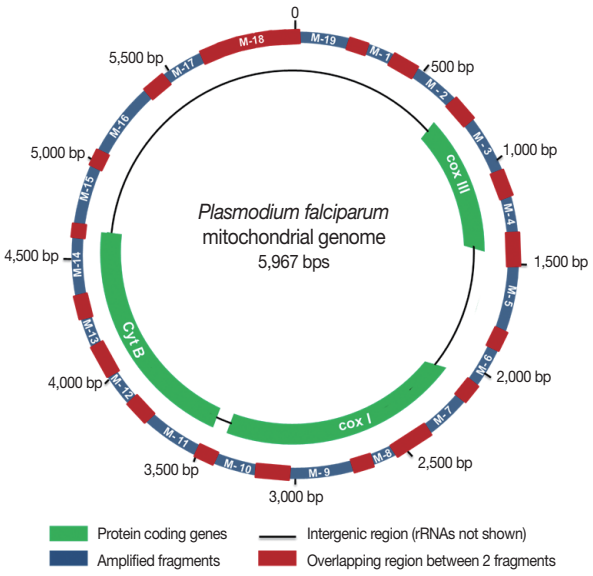

Fig. 1. Schematic overview of the ~6 kb *mt* genome and the approximate locations of primers to amplify the whole *mt* genome of Indian *P. falciparum*.

Table 2. Alignment showing variations in the 3 *mt* [2 Indian (Blsp1 and PfPH10) and 1 reference (3D7)] genomes of *Plasmodium falciparum* isolates

Nt Positions	208	222	230	510	615	1122	1339	2175	2768	3330	3433	3444	3764	3766	3868	3985	4352	4353	4420	4640	4759	4952	5485
3D7[a]	G	A	G	A	C	G	A	T	T	A	T	T	T	A	A	C	A	T	T	A	T	T	T
Blsp1[b]	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.
PfPH10[c]	A	T	A	G	T	C	G	C	A	G	A	C	C	G	T	G	T	A	A	G	C	C	A

Nt = nucleotide.
[a]Joy et al. (2003). [b]Present study. [c]Sharma et al. (2001).

the present observation on low mitochondrial genome diversity between a single *P. falciparum* isolate (Blsp1) and the reference 3D7 corroborates earlier opinion on the overall low variations among *P. falciparum mt* genomes [11]. Furthermore, the observation of 22 nucleotide changes between just 2 *P. falciparum mt* genomes from India (Table 2) reflects many-fold higher diversity in India, compared to only 30 SNPs found in 100 worldwide *P. falciparum* isolates of African, Asian, Papua New Guinean, and American origins [11]. Such observed high sequence variability between the Blsp1 and PfPH10 isolates might be the reason of failure for PCR amplification with primer information used to amplify and sequence the PfPH10 isolate [12]. This argument is further justified by the fact that, by using the 3D7 isolate (with only 1 nucleotide change), we could successfully amplify the Blsp1 isolate, justifying high sequence similarities. In order to nullify the role of PCR and sequencing errors, we have re-amplified and re-sequenced all the 19 DNA fragments of Blsp1 isolate using a different PCR thermal cycler (total 4x coverage), but the observed results were not different from the previously sequenced data. This observation essentially means that neither any inaccuracy in the PCR nor in sequencing techniques have contributed to our observed results.

With the whole *mt* genome sequence of a second *P. falciparum* isolate (Blsp1) in hand showing high sequence differences, we were interested to understand evolutionary interrelationships between the 2 Indian *mt* genomes with the reference 3D7 isolates by constructing NJ phylogenetic tree (Fig. 2). We have also included the published whole *mt* genome sequences of different *Plasmodium* species infecting an array of organisms. The tree topologies of the NJ phylogenetic tree justifies the evolutionary patterns of *Plasmodium* species according to their respective hosts [15,16]. For the *P. falciparum* isolates infecting humans, the Blsp1, 3D7, and C10 (GenBank accession no. M76611) form a single clade, whereas the PfPH10 isolate was placed away from this clade, justifying high genetic differentiation of this Indian isolates from the rest of *P. falciparum* isolates (Fig. 2). Whatever the case may be, the high sequence similarity between the whole *mt* genome sequences of the Blsp1 and 3D7 isolates justify the notion that intra-specific *mt* genome variations are in fact minimal [11] and therefore *mt* genomes remain conserved among the phylum Apicomplexa [8] as well as in *Plasmodium* species infecting different hosts (both humans and non-humans) [17], possibly due to very low recombination rate and uni-parental (maternal) inheritance [6].

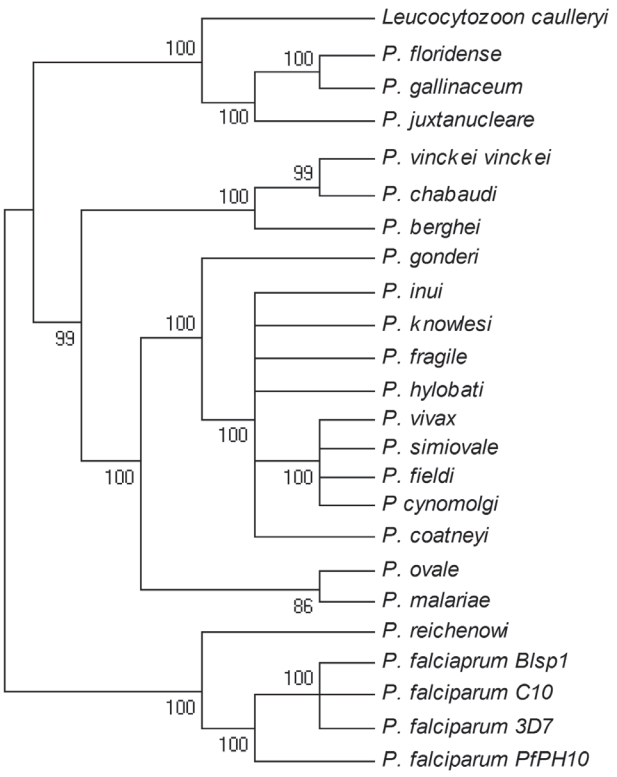

Fig. 2. Neighbour-joining (NJ) phylogenetic tree showing evolutionary inter-relationships among the Blsp1, 3D7, and PfPH10 isolates and also among different *Plasmodium* species infecting non-human hosts. The values in the internal nodes indicate bootstrapped values signifying the strength of each corresponding internal node of the NJ tree.

However, the observed high variability between the *mt* genomes of 2 Indian isolates suggests high genetic diversity in Indian *P. falciparum* [2] which can be further validated by sequencing isolates from more Indian populations. Such study would not only fill the gap of the existing knowledge about the worldwide *mt* genome diversity but also help to bring out important and so far not-fully-resolved evolutionary history of global *P. falciparum*. Moreover, whole *mt* genome sequence comparisons in multiple *P. falciparum* isolates from all over India would also inform the extent of genetic diversity of the *cytb* gene that is considered to be the target of an effective antimalarial, atovaquone [10]. As this antimalarial is currently not used in India, the knowledge of the extensive diversity through population genomic studies of *mt* genome of Indian *P. falciparum* would possibly help in deciding whether to incorporate atovaquone in malaria control programs in India.

ACKNOWLEDGMENTS

Financial assistance in the form of an extramural project by the Department of Biotechnology, Government of India to Aparup Das is thankfully acknowledged. Suchi Tyagi is a Senior Research Fellow of the Indian Council of Medical Research. We thank Prof. A.P. Dash, former Director of NIMR for encouragement and Dr. Neena Valecha, present Director of NIMR for facilities. The study has been approved by the ethical committee of NIMR. The paper bears NIMR publication committee clearance number 020/2013.

CONFLICT OF INTEREST

We have no conflict of interest related with this study.

REFERENCES

1. World Health Organization. World Malaria Report. Geneva, Switzerland. WHO. 2012; p 288.
2. Singh V, Mishra N, Awasthi G, Dash AP, Das A. Why is it important to study malaria epidemiology in India? Trends Parasitol 2009; 25: 452-457.
3. Das A, Bajaj R, Mohanty S, Swain V. Genetic diversity and evolutionary history of *Plasmodium falciparum* and *P. vivax*. Curr Sci 2007; 92: 1516-1524.
4. Avise JC, Walker D. Pleistocene phylogeographic effects on avian populations and the speciation process. Proc R Soc Lond B 1998; 265: 457-463.
5. Brown WM, George JRM, Wilson AC. Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA 1979; 76: 1967-1971.
6. Creasey A, Mendis K, Carlton J, Williamson D, Wilson I, Carter R. Maternal inheritance of extra-chromosomal DNA in malaria parasites. Mol Biochem Parasitol 1994; 65: 95-98.
7. Elson JR, Andrews RM, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Analysis of European *mt* DNAs for recombination. Am J Hum Genet 2001; 68: 145-153.
8. Feagin JE, Harrell MI, Lee JC, Coe KJ, Sands BH, Cannone JJ, Tami G, Schnare MN, Gutell RR. The Fragmented Mitochondrial Ribosomal RNAs of *Plasmodium falciparum*. PLoS ONE 2012; 7: e38320.
9. Krungkrai J. The multiple roles of the mitochondrion of the malarial parasite. Parasitology 2004; 129: 511-524.
10. Vaidya AB, Lashgari MS, Pologe LG, Morrisey J. Structural features of *Plasmodium* cytochrome b that may underlie susceptibility to 8-aminoquinolines and hydroxynaphthoquinones. Mol Biochem Parasitol 1993; 58: 33-42.
11. Joy DA, Feng X, Mu J, Furuya T, Chotivanich K, Krettli AU, Ho M, Wang A, White NJ, Suh E, Beerli P, Su X-z. Early origin and recent expansion of *Plasmodium falciparum*. Science 2003; 300: 318-321.
12. Sharma I, Rawat DS, Pasha ST, Biswas S, Sharma YD. Complete nucleotide sequence of the 6 kb element and conserved cytochrome b gene sequences among Indian isolates of *Plasmodium falciparum*. Int J Parasitol 2001; 31: 1107-1113.
13. Gupta B, Gupta P, Sharma A, Singh V, Dash AP, Das A. High proportion of mixed species *Plasmodium* infections in India revealed by PCR diagnostic assay. Trop Med Int Health 2010; 15: 819-824.
14. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011; 28: 2731-2739.
15. Prugnolle F, Ollomo B, Durand P, Yalcindag E, Arnathau C, Elguero E, Berry A, Pourrut X, Gonzalez J-P, Nkoghe D, Akiana J, Verrier D, Leroy E, Ayala FJ, Renaud F. African monkeys are infected by *Plasmodium falciparum* nonhuman primate-specific strains. Proc Natl Acad Sci USA 2011; 108: 11948-11953.
16. Duval L, Fourment M, Nerrienet E, Rousset D, Sadeuh SA, Goodman SM, Andriaholinirina NV, Randrianarivelojosia M, Paul RE, Robert V, Ayala FJ, Ariey F. African apes as reservoirs of *Plasmodium falciparum* and the origin and diversification of the Laverania subgenus. Proc Natl Acad Sci USA 2010; 107: 10561-10566.
17. Hikosaka K, Watanabe Y-I, Kobayashi F, Waki S, Kita K, Tanabe K. Highly conserved gene arrangement of the mitochondrial genomes of 23 *Plasmodium* species. Parasitol Int 2011; 60: 175-180.