

Genetic Characterization of Plasmodium vivax Isolates from Pakistan Using Circumsporozoite Protein (PvCSP) and Merozoite Surface Protein-1 (PvMSP-1) Genes as Genetic Markers

Zainab Bibi

Virtual University of Pakistan

Anam Fatima

Polyclinic Hospital Islamabad

Rehana Rani

Abasyn University

Ayesha Maqbool

Virtual University of Pakistan

Samea Khan

Virtual University of Pakistan

Shumaila Naz (✉ naz.pmasaaur@gmail.com)

National University of Medical Sciences <https://orcid.org/0000-0001-8088-0084>

Shahid Waseem

Alpha Genomics (Pvt) Ltd

Research

Keywords: Plasmodium vivax, Malaria, Genetic diversity, CSP, MSP-1

Posted Date: June 2nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-32064/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published on February 25th, 2021. See the published version at <https://doi.org/10.1186/s12936-021-03654-w>.

Abstract

Background *Plasmodium vivax* contribute over 70% malaria burden in Pakistan. Limited data exist on various aspects including genetic diversity of the parasite as compared to other parts of the world. The information about extent of genetic diversity assists to understand the transmission patterns of the parasite in human host. The current study was designed to understand population divergence of *Plasmodium vivax* in Pakistan using circumsporozoite protein and merozoite surface protein-I genes as molecular markers.

Methods *PvCSP* and *PvMSP-1* specific PCR and DNA sequencing were carried out for 150 blood samples collected from Islamabad and Rawalpindi, Pakistan. Genetic diversity was analysed using ChromasPro, ClustalW, MEGA7 and DnaSP v.5 programs.

Results The PCR for *PvCSP* and *PvMSP-1* genes was carried out for 150 *P. vivax* isolates resulting the PCR products ranging from 900 to 1100 bp for *PvCSP* gene and ~ 400 bp for *PvMSP-1* gene. Majority (93%; 121/150) of the *P. vivax* isolates were of VK210 variant type and only 9 isolates were found of VK247 variant type based on *PvCSP* gene. Out of the numerous peptide repeat motifs (PRMs) detected, GDRADGQPA (PRM1) and GDRAAGQPA (PRM2) were more extensively dispersed among the *P. vivax* isolates. Partial sequences (~ 400 bp) at the N-terminal of *PvMSP-1* gene depicted high level of diversity.

Conclusion High-level genetic diversity based on *PvCSP* and *PvMSP-1* genes was observed in clinical isolated in the study area. Parasite typing is essential in predicting pattern of antigenic variations, drug resistance and for effective drug and vaccine designing and development which can further evaluate for malaria control and eradication at individual and community level. The base-line data presented here warrant future studies to investigate more into the genetic diversity of *P. vivax* with large sample size from across the country for better understanding of the *vivax* malaria transmission patterns.

Background

Malaria remains a serious health concern worldwide. Globally, around 42% population is at risk [1] with 0.4 million deaths out of 228 million malaria cases [2]. Among five *Plasmodium* species, *Plasmodium vivax* (*P. vivax*) contributes around 70% malaria cases in Pakistan [3] with variable severity [4–6]. To circumvent the parasite load, there is a need to investigate the population structure, genetic diversity [7], parasite typing of the local *P. vivax* species, pattern of antigenic variations and drug resistance [8, 9]. It would lead to discontinue transmission cycle of the parasite in human host in endemic areas. Despite multiple academic research ventures, scanty data is available regarding the diverse genetic make-up of *P. vivax* in Pakistan [10]. Genetic sequences of *P. vivax* circum-sporozoite protein (*PvCSP*) and merozoite surface protein-I (*PvMSP-1*) are being used to understand the genetic diversity [11, 12]. *PvCSP*, a highly immunogenic sporozoite surface protein, hence a good vaccine candidate, is encoded by single copy gene [8]. *PvCSP* comprises of a central repeat domain that varies across *Plasmodium* species. It has two non-repetitive domains at N- and C-terminals [13, 14]. The varying degree of the amount of peptides in the central repeat region reveals three variants of *P. vivax* namely VK210, VK247, *P. vivax*-like [15, 16]. Across the globe, these variants exhibit certain spatial predispositions. With a GDRA(A/D)GQPA amino acid repetition, VK210 strain dominates in the endemic region [17, 18]. Originating in Thailand, the VK247 strain is mostly reported from the areas where mixed infections are prominent. VK247 depicts ANGAGNQPG amino acid repeat in the central region [19, 20]. In *PvCSP*, polymorphism is reported to be limited in central tandem repeat among the isolates from the regions of Sub-Saharan Africa [21]. *PvMSP-1* is expressed on the surface of the blood-stage parasite [22]. *PvMSP-1*, a large gene,

covers conserved and polymorphic regions [23] and has mosaic organization with 13 regions of variable blocks [24]. The three main regions of sequence divergence are block 2 (F1 region), 6–8 (F2 region) and 10 (F3 region) [25]. In the representative blocks, the genetically distinct *PvMSP-1* populations within the regions and polymorphism can be detected through PCR [4, 26]. Selective pressure of the host immune maintained the diversity of *PvMSP-1* gene. However, immunogenic properties can be affected by single-point mutation [27, 28]. The present study was designed to investigate the genetic diversity of *P. vivax* in Potohar region of Pakistan exploiting *PvCSP* and *PvMSP-1* genes as molecular markers. In addition, the study depicted base-line molecular epidemiology and phylogenetic characteristics of the local *P. vivax* isolates exploiting *PvCSP* and *PvMSP-1* gene-based diversity data. It would help to re-design malaria control or eradication measures in the area.

Methods

Study area

The study was conducted in Islamabad and Rawalpindi districts of the Punjab with longitudes 72°45' and 73°30' E and latitudes 33°30' and 33°50' N [29] (Fig. 1). The climate ranges from showery warm weather to chilly dry wintry weather with the attributes of the semi-arid region of Pakistan. The monsoon rains typically start in June, get peak in August, and finish by September. The rainfall is between 620 and 1,200 mm per year. The weather and geographical settings of this region is favorable for the mosquito breeding with highest frequency reported in Rawalpindi (25.5%) and lowest in Chakwal (15.9%) [30].

Sample Collection

Blood samples (n = 150) were collected during malaria transmission season (April and October, 2019) from Pakistan Institute of Medical Sciences (PIMS), Islamabad and Rawalpindi General Hospital, Rawalpindi, Pakistan. Venous blood (5 mL) was collected in ethylene diamine tetra-acetic acid (EDTA) (BD, USA) vacutainers from malaria patients for microscopy and molecular studies. Blood samples were stored at 4 °C until further analysis.

DNA extraction and PCR amplification of *PvCSP* and *PvMSP-1* genes

Genomic DNA from 150 malarial blood samples were extracted by using standard phenol-chloroform method as described elsewhere [31]. PCR primers for *PvCSP* and *PvMSP-1* genes were designed at Geneious software by using reference sequence of *P. vivax* (AB539044). The set of primers and cycling conditions used to amplify *PvCSP* gene is mentioned in Table 1. The PCR reactions were carried out in a 25 µL reaction mixture comprising of 2µL DNA template, 0.5 mM dNTPs, 1X PCR reaction buffer (SolarBio Life Sciences, China), 0.2 mM of each primer (BGI, China), 2.5 mM MgCl₂ and 1 unit of *Taq* DNA polymerase (BLIRT, Poland). The PCR products were visualized using 1% agarose gel (ThermoFisher, USA) stained with ethidium bromide (SolarBio Life Sciences, China) and visualized under UV-transilluminator (ThermoFisher, USA).

Table 1
Primers and PCR conditions

Gene	Primers	PCR Conditions
PvCSP	F: 5'-GGCCATAAATTTAAATGGAG-3' R: 5'-ATGCTAGGACTAACAATATG-3'	94 °C 10 min (94 °C 1 min, 52 °C 1 min, 72 °C 1 min), 35 cycles, 72 °C 10 min
PvMSP-1	F: 5ACATCATTAAGGACCCATACAAG-3' R: 5'-GCAATTTCTTTACAGTGATCTCG-3'	94 °C 10 min (94 °C 1 min, 56 °C 1 min, 72 °C 1 min), 35 cycles, 72 °C 10 min

Sanger Sequencing And Analysis

The amplified PCR products were purified and sent to BGI, China for sanger sequencing. DNA sequences of both *PvCSP* and *PvMSP-1* genes were read, assembled on both strands and analyzed by using ChromasPro (version 1.5) software (<http://technelysium.com.au/wp/chromaspro/>) and Bio Edit alignment editor (version 7.2) (<https://bioedit.software.informer.com/7.2/>). The sequenced samples were validated by BlastN (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) and alignment of top hit resulted sequences were done by ClustalW (<https://www.genome.jp/tools-bin/clustalw>). The number of haplotypes and nucleotide and haplotype diversity were calculated by DNAsp v5. Polymorphism of DNA sequence of *PvCSP* and *PvMSP-1* genes were also measured by Tajima's D, Fu and Li's D*, and Fu and Li's F* tests using DNAsp v5. The evolutionary relationships were established and phylogenetic trees of the both genes were constructed by the neighbor-joining method using Molecular Evolutionary Genetics Analysis (MEGA 7.0) software. The nucleotide sequences of *PvCSP* and *PvMSP-1* genes resulted from this study were submitted in NCBI database under accession number MT222296 to MT222330 and MT303819 to MT303848, respectively.

Results

A total of 150 blood samples were collected from *P. vivax* infected patients from twin cities of Islamabad and Rawalpindi. The blood samples were verified by microscopic examination and PCR and no mixed infections with other malaria parasites were found. The PCR for *PvCSP* and *PvMSP-1* genes were carried out in a total of 150 *P. vivax* isolates with PCR products ranging from 900 to 1100 bp for *PvCSP* gene and ~ 400 bp for *PvMSP-1* gene.

Sequence analysis of *PvCSP* gene

Multiple sequence alignment of the translated nucleotide sequences was carried out for the analysis of polymorphisms in the pre-, post- and central-repeats of the *PvCSP* gene. When compared with the reference sequence (KT588208.1), sequence analysis of *PvCSP* gene showed only the VK210 and VK247 variant types infection without the existence of the mixed infections. *PvCSP* gene sequence analysis revealed that majority (93%; 121/150) of the *P. vivax* isolates were of VK210 variant type and only 9 isolates were found to be VK247 variant type. All *PvCSP* gene-based *P. vivax* variants started with the same pre-repeat sequence (KLKQP region). In the central-repeat region (CRR), the VK210 sequences comprised of variable repeats of PRMs, GDRADGQPA (PRM1),

GDRAAGQPA (PRM2) which are abundant and found in all the isolates. It is followed through two conserved post-repeat sequence GNGAGGQAA (PRM3) and GGNAANK (PRM4) and one post-repeat insert i.e. KAEDA region. The one-time repeat region of GGNA was found after the CRR in all the analyzed sequences. The repeat units and differences in the order amino acid sequences in CRR have been listed and summarized in Table 2. The observed non-synonymous substitution based on diverse types of repetition in allotypes (RATs) which leads to different PRMs are mentioned in Table 3.

was analyzed and compared with *PvCSP* isolates of neighboring country Iran giving the values average dS and dN for each isolate. The haplotype diversity of *PvCSP* gene was similar between two countries ranging from 0.345 to 0.547. Adding to this, Tajima's D, Fu and Li's D* and F* tests also accepted occurrence of a neutral model of polymorphism with values for *PvCSP* (Pakistan) (Tajima's D = 0.54276, P < 0.10; Fu and Li's D* = 1.17870, P < 0.10; Fu and Li's F* = 1.12083, P < 0.10) and *PvCSP* (Iran) (Tajima's D = 0.43556, P < 0.10; Fu and Li's D* = 1.80767, P < 0.10; Fu and Li's F* = 0.43556, P < 0.10) respectively (Table 4).

Table 4
Estimates of nucleotide-, haplotype-diversity and DNA sequence polymorphisms of *P. vivax PvCSP* and *PvMSP-1* genes in Pakistan

Population	Type <i>P. vivax</i>	No. of samples	No. of Haplotypes	Haplotype diversity	Diversity ± SD	Fu & Li's D*	Fu & Li's F*	Tajima's D*
					Nucleotide Haplotype			
Pakistan	PvCSP	35	15	0.547	0.02371 ± 0.0005620	1.17870	1.12083	0.54276
Iran	PvCSP	28	10	0.345	0.084 ± 0.00701	1.01980	1.80767	0.43556
					0.02001 ± 0.0003146			
					0.057 ± 0.00478			
Pakistan	PvMSP-1	30	10	0.962	0.00162 ± 0.0000026	1.86276	2.13897	1.67790
Iran	PvMSP-1	32	10	0.954	0.012 ± 0.00014	1.78902	1.90878	1.56845
					0.00159 ± 0.0000023			
					0.012 ± 0.00014			

*P > 0.10

Phylogenetic Analysis

Based on sequence of *PvMSP-1* gene, a phylogenetic tree was constructed (Fig. 3). Two distinct clades can be inferred from the tree. The first clade is divided further into three sub-clades contained 11 Pakistani isolates and isolates belong to East Africa, Thailand and Mexico. Second clade is further divided into two sub-clades having Turkey, Iran and Southern Mexico isolates in addition of sequences of Pakistani isolates. The Neighbor-Joining method was used to deduce evolutionary history with the sum of branch length of tree = 0.50423062 was shown.

Discussion

Malaria, a protozoan parasitic disease, is one of the major public health concerns with limited data in Pakistan [32]. As Pakistan shares border with malaria endemic countries like Iran, India and Afghanistan so human migration across the border is inevitable which may facilitate substantial cross-border transmission of malaria. Human migration, immune responses exhibited by the human host, chemotherapy, genetic mutation and recombination

lead to genetic diversity and affects the frequency of new alleles in the parasite population [33, 34]. The study of genetic diversity can give significant indications about parasite response to different drugs and vaccines as positive selection will favor the fixation of important alleles in the population and can lead to reduction of genetic diversity [35]. Despite *P. vivax* contributing to 88% of malaria burden in Pakistan [4–6], data regarding genetic diversity of this key circulating species is lacking. There are limited studies from Pakistan which analyzed the diversity of local *P. vivax* in detail [4, 5, 23, 33]. So, the need for the detailed understanding of the extent and nature of *P. vivax* genetic diversity is obvious. *PvCSP* and *PvMSP-1* are among the other important genetic markers used by the researchers to understand population structure and evolutionary dynamics from different geographical regions [4, 25]. In the present study, genetic diversity of both *PvCSP* and *PvMSP-1* genes were estimated among clinical isolates collected from different hospitals of Federal and Punjab region of Pakistan.

The results of *PvCSP* uphold the findings of the previously published results showing that VK210 strain is predominant type with the prevalence rate ranging between 81–100% [9, 12, 14, 36]. There are few malaria endemic areas where VK247 isolates are commonly present [37, 38]. In the present study, the translated nucleotide sequences suggested that there were peptide repetitive motifs in which GDRADGQPA (PRM1) and GDRAAGQPA (PRM2) which found to be the two major PRMs. Earlier studies also pointed out the dominance of these two prime PRMs (GDRADGQPA, GDRAAGQPA) in the clinical isolates [7, 8, 12, 39]. All the isolates collected were composed of similar pre-repeat sequence (KLKQP) region and conserved post-repeat sequence GGNAANK (PRM4). At the end of the sequence the conserved post-repeat sequence is present as a last section in all of the VK210 isolates, consistent with the findings of aforementioned studies from India, Iran and Sri Lanka [8, 14, 20]. Another peptide repeat motif GNGAGGQAA (PRM3) was found at lower frequency (0.6%) in the isolates. The essential behind the development of genetic diversity found in the *PvCSP* gene, across various spatial locations, is the number of PRMs. Difference in the amino acid and nucleotide sequences of the *Plasmodium* antigens due to variations exist in the repeat unit numbers depicted the pressure of natural selection displayed by the host immune system [14, 40]. The arrangement of the main PRM1 and PRM2 factors lead to 15 different haplotypes in *PvCSP*. The pragmatic neutrality tests also accepted a neutral model of polymorphism which indicated that events of positive selection occurred in the complex group *P. vivax* isolate [14].

PvMSP-1 is one of the most promising vaccine candidates and is available for antigenic and genetic variation studies of *P. vivax* populations [41]. In this study, partial sequence (~ 400 bp) at N-terminal of *PvMSP-1* gene depicted a high-level of diversity. Such high level of genetic diversity is in concordance with what has already observed in neighboring country Iran [20] and in previous study from Pakistan [32]. In the northwestern region of Thailand, a high degree of mutational variety was observed in *PvMSP-1* genes of *P. vivax* isolates [28, 42]. The natural selection was determined by the neutrality tests on the *PvMSP-1* N-terminal fragment of *P. vivax*. Overflow of transitional frequency alleles were observed because of significantly positive values which may be the result of balancing the selection and population bottlenecks. The sequence diversity of the population is best studied by the intragenic recombination of *PvMSP-1* gene where the allelic recombination frequencies may aid as a character reference for understanding the parasitic population structure [27]. Kibria et al. [25] also indicated that the given variations suggest high genetic diversity in all areas under study, furthermore, that the *PvMSP-1* gene undergoes selective pressure for the existence and spread of the parasite. The *PvMSP-1* gene sequences were instructive in distinguishing the two central localities of sample origin in terms of geography as well as helping them to group in two different clades. These clades are further sub divided into different clusters and group these sequences on these sub clusters according to its geographic origin [43].

Other studies suggested that the mode of evolution in *PvCSP* gene can lead to cohort of variants that can elude the host immune response under the effect of both mitotic recombination and positive selection of *P. vivax* new variants [16]. It is safe to assume, therefore, that the wide variety of *P. vivax* is perhaps interrelated with multiple other variables including, but not limited to, genetic and biological characteristics, immunity of the host and the displacement of individuals within the boundaries of the endemic areas. Furthermore, the spread of *P. vivax* infections is also buttressed by relapse and early gametocytaemia, which in turn sustains local diversity, paving way for a more efficient transmission to the vector mosquitoes [12, 44].

In the research undertaken here, a broad range of genetic variety of *PvCSP* and *PvMSP-1* genes in *P. vivax* population was noticed. The localities of the study areas, inhabited by various ethnicities, suggest that the migration of people may carry diverse parasite entities that increase the variety of the gene-pool. Furthermore, individuals infected by the disease might carry various clones with varying degree of *PvCSP* derivatives, which may recombine during the sexual stage of the mosquito leading to the production of offspring with new *PvCSP* genotypes. This prevalent phenomenon could strengthen the introduction of new strains of *P. vivax* into the regions where conditions for malaria transmission are conducive [12, 25].

Conclusion

High-level genetic diversity based on *PvCSP* and *PvMSP-1* genes for *P. vivax* in clinical isolates was observed in Pakistan. The data will facilitate the documentation of areas targeted by the host

immune response and facilitate in better insight of the host-parasite relationship, will provide novel insights into the events/signaling mechanisms and recognition of novel vaccine and drug target sites. A better understanding of genetic variability will help to control malaria at individual and community level. In order to design and put new and more effective vaccinations into effect, these studies can provide substantial help. The base-line data presented here warrant future studies to investigate more into the genetic diversity of *P. vivax* with large sample size from across the country for better understanding of the vivax malaria transmission patterns.

Declarations

Ethics approval and consent to participate

The study was approved by institutional review board (IRB), Virtual University of Pakistan, (letter No. VU/ASRB/131-7). Blood samples were collected from diagnosed malaria patients after informed consent.

Consent for publication

Not applicable

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare no competing interests

Funding

There was no financial support for this research.

Authors' contributions

ZB conducted laboratory experiments and data analysis, AF helped in sample collection, clinical inputs and malaria identification, RR helped in primer designing and molecular data input, AM provided scientific inputs and critical review of manuscript, SK helped specifically in CSP gene analysis, SN supervised the wet labs, critically reviewed manuscript, helped in manuscript drafting and data analysis, SW conceived the idea and study design, provided all wet lab reagents/lab facilities and contributed to the critical revision of the manuscript, . All authors read and approved the final manuscript.

Acknowledgements

We would like to thank medical staff of hospital and all the study participants to participate in the study

Authors' information

¹Department of Molecular Biology, Virtual University of Pakistan, Lahore, Pakistan. ²Department of Medicine, Polyclinic Hospital, Islamabad, Pakistan. ³Department of Life Sciences, Abasyn University, Islamabad, Pakistan. ⁴Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan. ⁵Alpha Genomics (Pvt) Ltd., Islamabad, Pakistan.

References

1. Faburay B. The case for a 'one health' approach to combating vector-borne diseases. *Infection ecology epidemiology*. 2015;5(1):28132.
2. World Health Organization. World malaria report 2017. 2017. Reference Source. 2017.
3. Report PMalariaA. 2019. Malaria Control Program, Directorate of Malaria Control, Pakistan.
4. Raza A, Ghanchi NK, Thaver AM, Jafri S, Beg MA. Genetic diversity of *Plasmodium vivax* clinical isolates from southern Pakistan using *pvcsp* and *pvmSP1* genetic markers. *Malar J*. 2013;12(1):16.
5. Khan SN, Khan A, Khan S, Ayaz S, Attaullah S, Khan J, Khan MA, Ali I, Shah AH. PCR/RFLP-based analysis of genetically distinct *Plasmodium vivax* population of *PvmSP-3α* and *PvmSP-3β* genes in Pakistan. *Malar J*. 2014;13(1):355.
6. Ahmad T. *Plasmodium vivax* and *Plasmodium falciparum* are Common Malaria Species in Pakistan. *Biomedical Research Therapy*. 2016;3(6):1–7.
7. Zhang LL, Yao LN, Chen HL, Lu QY, Ruan W. Genetic diversity analysis of *PvCSP* and its application in tracking of *Plasmodium vivax*. *Experimental parasitology*. 2018;188:26–35.
8. Dias S, Wickramarachchi T, Sahabandu I, Escalante AA, Udagama PV. Population genetic structure of the *Plasmodium vivax* circumsporozoite protein (*Pvcsp*) in Sri Lanka. *Gene*. 2013;518(2):381–7.
9. Talha AA, Pirahmadi S, Mehrizi AA, Djadid ND, Nour BY, Zakeri S. Molecular genetic analysis of *Plasmodium vivax* isolates from Eastern and Central Sudan using *pvcsp* and *pvmSP-3α* genes as molecular markers. *Infection, Genetics and Evolution*. 2015;32:12–22.

10. Kaur H, Sehgal R, Goyal K, Makkar N, Yadav R, Bharti PK, Singh N, Sarmah NP, Mohapatra PK, Mahanta J, Bansal D. Genetic diversity of *Plasmodium falciparum* merozoite surface protein-1 (block 2), glutamate-rich protein and sexual stage antigen Pfs25 from Chandigarh, North India. *Tropical Med Int Health*. 2017;22(12):1590–8.
11. Simon B, Sow F, Al Mukhaini SK, Al-Abri S, Ali OA, Bonnot G, Bienvenu AL, Petersen E, Picot S. An outbreak of locally acquired *Plasmodium vivax* malaria among migrant workers in Oman. *Parasite*. 2017;24.
12. Liu Y, Zhou RM, Zhang YL, Wang DQ, Li SH, Yang CY, Qian D, Zhao YL, Zhang HW, Xu BL. Analysis of polymorphisms in the circumsporozoite protein gene of *Plasmodium vivax* isolates from Henan Province, China. *Malar J*. 2018;17(1):103.
13. Coppi A, Natarajan R, Pradel G, Bennett BL, James ER, Roggero MA, Corradin G, Persson C, Tewari R, Sinnis P. The malaria circumsporozoite protein has two functional domains, each with distinct roles as sporozoites journey from mosquito to mammalian host. *J Exp Med*. 2011;208(2):341–56.
14. Kaur H, Sehgal R, Kumar A, Sehgal A, Bharti PK, Bansal D, Mohapatra PK, Mahanta J, Sultan AA. Exploration of genetic diversity of *Plasmodium vivax* circumsporozoite protein (Pvcsp) and *Plasmodium vivax* sexual stage antigen (Pvs25) among North Indian isolates. *Malar J*. 2019;18(1):308.
15. Alves RT, Póvoa MM, Goldman IF, Cavasini CE, Rossit AR, Machado RL. A new polymerase chain reaction/restriction fragment length polymorphism protocol for *Plasmodium vivax* circumsporozoite protein genotype (VK210, VK247, and *P. vivax*-like) determination. *Diagn Microbiol Infect Dis*. 2007;59(4):415–9.
16. Patil A, Orjuela-Sánchez P, da Silva-Nunes M, Ferreira MU. Evolutionary dynamics of the immunodominant repeats of the *Plasmodium vivax* malaria-vaccine candidate circumsporozoite protein (CSP). *Infection Genetics Evolution*. 2010;10(2):298–303.
17. Rosenberg R, Wirtz RA, Lanar DE, Sattabongkot J, Hall T, Waters AP, Prasittisuk C. Circumsporozoite protein heterogeneity in the human malaria parasite *Plasmodium vivax*. *Science*. 1989;245(4921):973–6.
18. Leclerc MC, Durand P, Gauthier C, Patot S, Billotte N, Menegon M, Severini C, Ayala FJ, Renaud F. Meager genetic variability of the human malaria agent *Plasmodium vivax*. *Proceedings of the National Academy of Sciences*. 2004;101(40):14455-60.
19. Qari SH, Shi YP, Goldman IF, Udhaykumar V, Collins WE, Lal AA, Alpers MP. Identification of *Plasmodium vivax*-like human malaria parasite. *The Lancet*. 1993;341(8848):780–3.
20. Zakeri S, Abouie Mehrizi A, Djadid ND, Snounou G. Circumsporozoite protein gene diversity among temperate and tropical *Plasmodium vivax* isolates from Iran. *Tropical Med Int Health*. 2006;11(5):729–37.
21. Tahar R, Ringwald P, Basco LK. Heterogeneity in the circumsporozoite protein gene of *Plasmodium malariae* isolates from sub-Saharan Africa. *Mol Biochem Parasitol*. 1998;92(1):71–8.
22. Birkenmeyer L, Muerhoff AS, Dawson GJ, Desai SM. Isolation and characterization of the MSP1 genes from *Plasmodium malariae* and *Plasmodium ovale*. *Am J Trop Med Hyg*. 2010;82(6):996–1003.
23. Zakeri S, Raeisi A, Afsharpad M, Kakar Q, Ghasemi F, Atta H, Zamani G, Memon MS, Salehi M, Djadid ND. Molecular characterization of *Plasmodium vivax* clinical isolates in Pakistan and Iran using pvmsp-1, pvmsp-3a and pvcsp genes as molecular markers. *Parasitol Int*. 2010;59(1):15–21.
24. Bonilla JA, Validum L, Cummings R, Palmer CJ. Genetic diversity of *Plasmodium vivax* pvcsp and pvmsp1 in Guyana, South America. *The American Journal of Tropical Medicine Hygiene*. 2006;75(5):830–5.
25. Kibria MG, Elahi R, Mohon AN, Khan WA, Haque R, Alam MS. Genetic diversity of *Plasmodium vivax* in clinical isolates from Bangladesh. *Malar J*. 2015;14(1):267.

26. Imwong M, Pukrittayakamee S, Grüner AC, Rénia L, Letourneur F, Looareesuwan S, White NJ, Snounou G. Practical PCR genotyping protocols for *Plasmodium vivax* using Pvcs and PvmSP1. *Malar J.* 2005;4(1):20.
27. Putaporntip C, Jongwutiwes S, Sakihama N, Ferreira MU, Kho WG, Kaneko A, Kanbara H, Hattori T, Tanabe K. Mosaic organization and heterogeneity in frequency of allelic recombination of the *Plasmodium vivax* merozoite surface protein-1 locus. *Proceedings of the National Academy of Sciences.* 2002;99(25):16348-53.
28. Thanapongpichat S, Khammanee T, Sawangjaroen N, Buncherd H, Tun AW. Genetic Diversity of *Plasmodium vivax* in Clinical Isolates from Southern Thailand using PvMSP1, PvMSP3 (PvMSP3 α , PvMSP3 β) Genes and Eight Microsatellite Markers. *Korean J Parasitol.* 2019;57(5):469.
29. Sheikh IM, Pasha MK, Williams VS, Raza SQ, Khan KS. Environmental geology of the Islamabad-Rawalpindi area, northern Pakistan. *Regional Studies of the Potwar-Plateau Area, Northern Pakistan. Bull. G.* 2008;2078.
30. Qureshi NA, Fatima H, Afzal M, Khattak AA, Nawaz MA. Occurrence and seasonal variation of human *Plasmodium* infection in Punjab Province, Pakistan. *BMC Infect Dis.* 2019;19(1):935.
31. Sambrook J, Russell DW, Sambrook J. *The condensed protocols: from molecular cloning: a laboratory manual.* Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 2006.
32. Khatoon L, Baliraine FN, Bonizzoni M, Malik SA, Yan G. Prevalence of antimalarial drug resistance mutations in *Plasmodium vivax* and *P. falciparum* from a malaria-endemic area of Pakistan. *Am J Trop Med Hyg.* 2009;81(3):525–8.
33. Khatoon L, Baliraine FN, Bonizzoni M, Malik SA, Yan G. Genetic structure of *Plasmodium vivax* and *Plasmodium falciparum* in the Bannu district of Pakistan. *Malar J.* 2010;9(1):112.
34. Ord R, Polley S, Tami A, Sutherland CJ. High sequence diversity and evidence of balancing selection in the PvmSP3 α gene of *Plasmodium vivax* in the Venezuelan Amazon. *Mol Biochem Parasitol.* 2005;144(1):86–93.
35. Paul RE, Packer MJ, Walmsley M, Lagog M, Ranford-Cartwright LC, Paru R, Day KP. Mating patterns in malaria parasite populations of Papua New Guinea. *Science.* 1995;269(5231):1709–11.
36. Shabani SH, Zakeri S, Mehrizi AA, Mortazavi Y, Djadid ND. Population genetics structure of *Plasmodium vivax* circumsporozoite protein during the elimination process in low and unstable malaria transmission areas, southeast of Iran. *Acta Trop.* 2016;160:23–34.
37. González-Cerón L, Martínez-Barnetche J, Montero-Solís C, Santillán F, Soto AM, Rodríguez MH, Espinosa BJ, Chávez OA. Molecular epidemiology of *Plasmodium vivax* in Latin America: polymorphism and evolutionary relationships of the circumsporozoite gene. *Malar J.* 2013;12(1):243.
38. Chenet SM, Tapia LL, Escalante AA, Durand S, Lucas C, Bacon DJ. Genetic diversity and population structure of genes encoding vaccine candidate antigens of *Plasmodium vivax*. *Malaria journal.* 2012 Dec 1;11(1):68.
39. Lopez AC, Ortiz A, Coello J, Sosa-Ochoa W, Torres RE, Banegas EI, Jovel I, Fontecha GA. Genetic diversity of *Plasmodium vivax* and *Plasmodium falciparum* in Honduras. *Malar J.* 2012;11(1):391.
40. Hughes AL. The evolution of amino acid repeat arrays in *Plasmodium* and other organisms. *Journal of molecular evolution.* 2004;59(4):528–35.
41. Kim SH, Hwang SY, Shin JH, Moon CS, Kim DW, Kho WG. Molecular genetic characterization of the merozoite surface protein 1 Gene of *Plasmodium vivax* from reemerging Korean isolates. *Clin Vaccine Immunol.* 2009;16(5):733–8.
42. Putaporntip C, Hongsriruang T, Seethamchai S, Kobasa T, Limkittikul K, Cui L, Jongwutiwes S. Differential prevalence of *Plasmodium* infections and cryptic *Plasmodium knowlesi* malaria in humans in Thailand. *The Journal of infectious diseases.* 2009;199(8):1143–50.

43. Guimarães LO, Wunderlich G, Alves JM, Bueno MG, Röhe F, Catão-Dias JL, Neves A, Malafronte RS, Curado I, Domingues W, Kirchgatter K. Merozoite surface protein-1 genetic diversity in *Plasmodium malariae* and *Plasmodium brasilianum* from Brazil. *BMC Infect Dis.* 2015;15(1):529.
44. Barry AE, Waltmann A, Koepfli C, Barnadas C, Mueller I. Uncovering the transmission dynamics of *Plasmodium vivax* using population genetics. *Pathogens global health.* 2015;109(3):142–52.

Figures

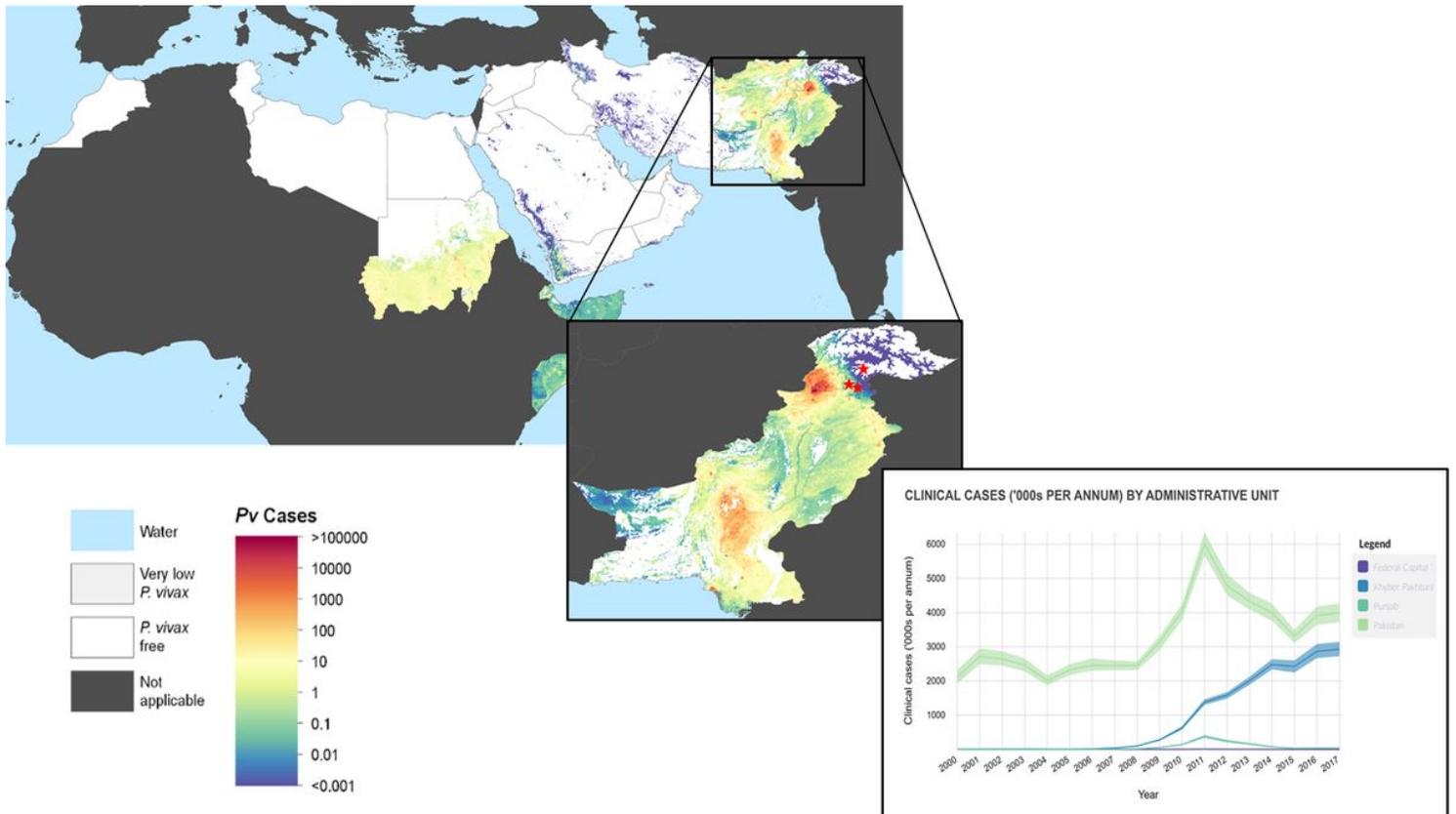


Figure 1

Map of *P. vivax* prevalence in EMRO region and Pakistan. The map showing all age clinical cases of *P. vivax* from year 2000-2017 of EMRO region and Pakistan. These maps were originally created by the Malaria Atlas Project, University of Oxford available at <https://malariaatlas.org/>. The estimated sites of the study sites described here are showed with red star: Rawalpindi and Islamabad.

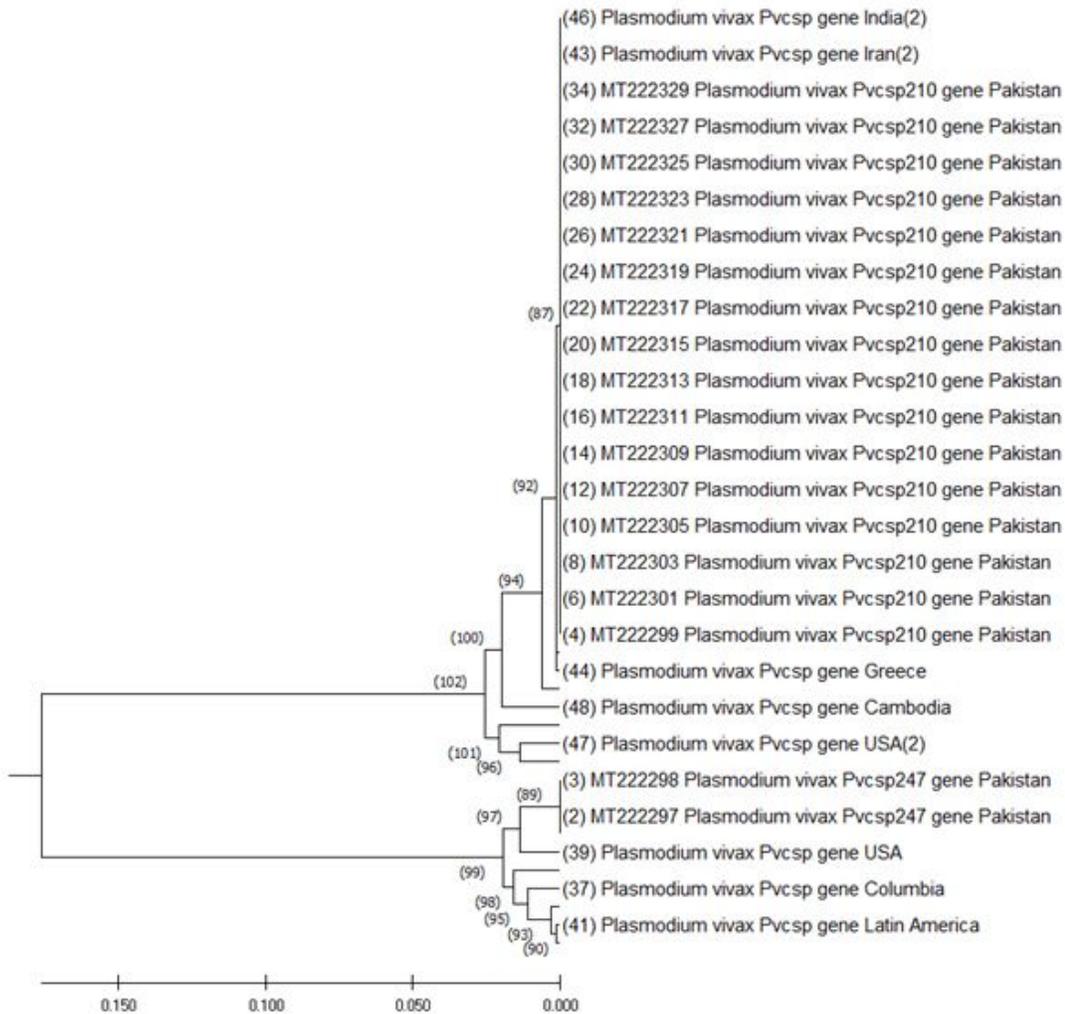


Figure 2

Phylogenetic relationships of VK210 and VK247 (based on PvCSP gene analysis) isolates from Pakistan. The Neighbor-Joining method was used to deduce evolutionary history with the branch length of =0.50423062. The p-distance method was used to compute the evolutionary distances and analysis involved 52 nucleotide sequences.

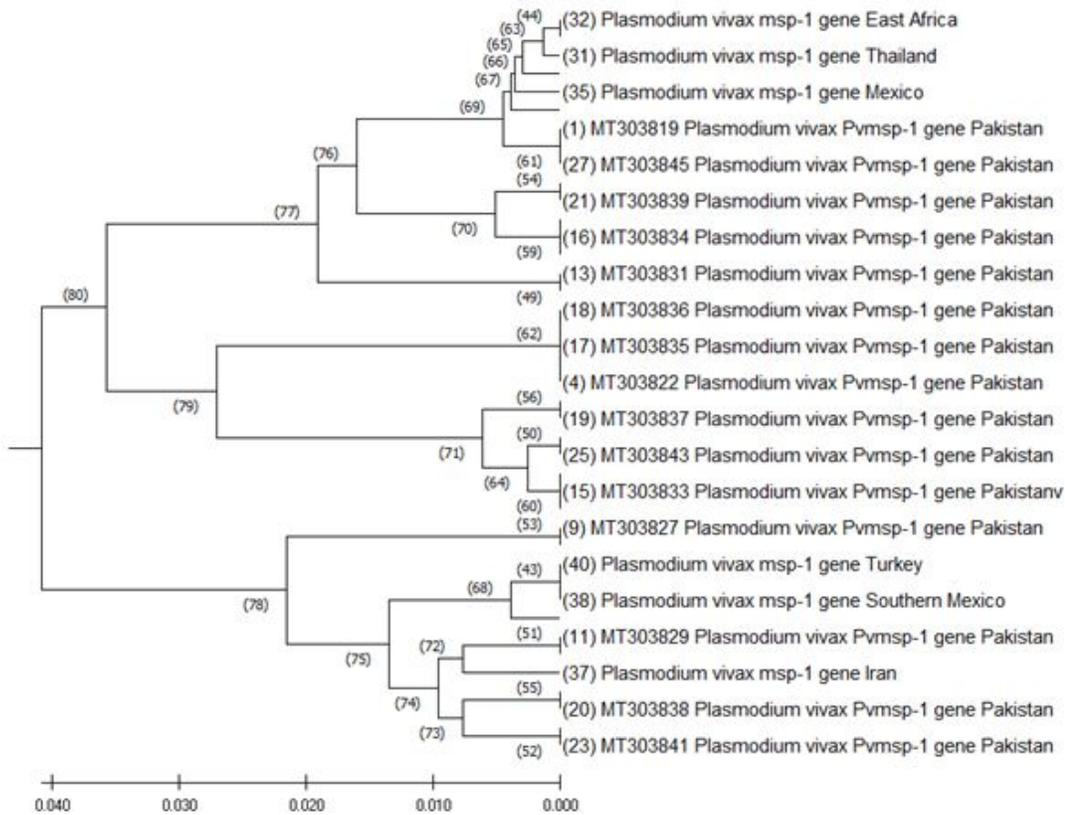


Figure 3

Phylogenetic relationships (based on PvMSP-1 gene) of *P. vivax* isolates from Pakistan. The Neighbor-Joining method was used to deduce evolutionary history with the branch length of =0.21993778. The p-distance method was used to compute the evolutionary distances and analysis involved 41 nucleotide sequences.