Phylogeny of *Plasmodium vivax* in Malaria-Endemic Regions of Khyber Pakhtunkhwa, Pakistan

Shehzad Zareen¹, Shahid Niaz Khan¹*, Muhammad Adnan², Noor Ul Akbar¹, Sadia Norin¹ and Rehman Ali³,4*

¹Molecular Parasitology and Virology Laboratory, Department of Zoology, Faculty of Biological Sciences, Kohat University of Science and Technology Kohat-26000, Khyber Pakhtunkhwa, Pakistan

²Department of Botany, Faculty of Biological Sciences, Kohat University of Science and Technology Kohat-26000, Khyber Pakhtunkhwa, Pakistan

³Aquatic Eco-Health Group, Fujian Key Laboratory of Watershed Ecology, Key Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China

⁴University of Chinese Academy of Sciences, Beijing 100049, China

ABSTRACT

Malaria is a big challenge in endemic areas, especially in developing countries throughout the world. Rapidly changing climatic conditions have brought about a major shift in the prevalence and spread of vector-borne diseases. Khyber Pakhtunkhwa province is known for seasonal outbreaks of malaria and shares borders with several malaria-endemic countries. The phylogeny of *Plasmodium vivax* is limited and poorly explored in this region. Therefore, this study was designed to address the limitation of the poorly explored phylogeny of *P. vivax* in three districts of Khyber Pakhtunkhwa. Blood samples were collected for seven months in 2018 from individuals who visited the district headquarters (DHQ) hospitals. The samples were successfully amplified with a band size of 945 bp for the merozoite surface protein-1 (pvmsp-1) gene. The phylogenetic analysis grouped *P. vivax* into two major clades with different subclades in each clade. The retrieved sequences from other countries were distributed in the subclades. The first clade was split up further into three subclades comprising isolates from China, India, Turkey, Azerbaijan, Iran, Myanmar, Afghanistan, South Korea, New Papua Guinea, and Pakistan. The second clade was comprised of isolates from D I Khan, Peshawar, and Kohat of Pakistan, Turkey, Bangladesh, Thailand, Korea, the UK, Australia, and Sal-I and Belem of Brazil. The phylogenetics of *P. vivax* represented a close relationship with the *P. vivax* population of Turkey, Bangladesh, Thailand, and Korea among others, suggesting a possible introduction of distinct variants in Khyber Pakhtunkhwa. Further studies with larger datasets and broader geographical settings should be carried out to better understand the phylogeny of *P. vivax*.

INTRODUCTION

Malaria is one of the most severe health complications, especially in malaria-endemic areas throughout the world (Villena et al., 2022). *Plasmodium vivax* is the foremost reason of malaria illness (Kuesap et al., 2022) while *Plasmodium falciparum*, among other species, is the main cause of mortality worldwide (Carlton et al., 2008; Price et al., 2007). *P. vivax* shows the emergence of resistance to antimalarial drugs and is a major obstacle in malaria control strategies (Baird, 2004). *P. vivax* exhibits
extensive genetic diversity among all *Plasmodium* species (Prugnolle *et al.*, 2013) and is the most genetically diversified species in Asia (Hupalo *et al.*, 2016). This exceptional genetic diversity helped *Plasmodium* species adapt to human hosts (Liu *et al.*, 2014). The genetic diversity of malaria parasites traces the origin and spread of new variants and can be used to evaluate the effectiveness of malaria control measures (Li *et al.*, 2022).

Pakistan is a malaria-endemic country with extreme weather conditions and four seasons a year, hence providing a suitable environment for vectors and vector-borne diseases. Malaria cases have been reported throughout the country in all four provinces (Hassan *et al.*, 2019). Khyber Pakhtunkhwa province is highly endemic to malaria due to the low socioeconomic status of the people in this province (Khan *et al.*, 2014, 2021). Moreover, military operations against terrorism by the Pakistan Army in FATA, North and South Waziristan Agencies led to a huge influx of internally displaced persons (IDPs) into various districts of Khyber Pakhtunkhwa (Khan, 2012). The IDP influx resulted in the mass spread of various infectious diseases including malaria in the province (Ahmad *et al.*, 2022; Khan *et al.*, 2021). Khyber Pakhtunkhwa also shares a border with the malaria-endemic country Afghanistan and cross-border movement of people greatly facilitates the transmission of the disease (Khan *et al.*, 2021). A massive migration of Afghan refugees took place to Pakistan in 1978, 1979, and 1992 after the communist coup, soviet invasion, and Najibullah government’s collapse, respectively. Consequently, Afghan refugees (approx. 3.2 million) came to Pakistan and 75%–80% settled down in Khyber Pakhtunkhwa and continuous cross-border shifts in population have been occurring (Kazmi and Pandit, 2001; Rowland, 1999, 2001). Rowland *et al.*, 1996). Therefore, it is one of the key malaria-endemic areas of Pakistan and several studies have reported the existence of malaria in the province (Karim *et al.*, 2021; Khan *et al.*, 2021; Qureshi *et al.*, 2020).

Several antimalarial drugs are ineffective against mutant malaria species in Pakistan and an estimated 50000 deaths occur due to malaria annually. Poor diagnosis by untrained laboratory personnel and improper medication, among others, are some of the common reasons for malaria persistence in Pakistan (Khattak *et al.*, 2021). Furthermore, synthetic drugs are generally perceived to have side effects and therefore, people in remote areas also rely on traditional therapies to treat diseases (Musarrat *et al.*, 2021). Some of these medicinal plants can be abundantly found in the studied province and employed as traditional recipes for the treatment of malaria (Tariq *et al.*, 2016).

Phylogenies are usually used to represent the relationships among species on the tree of life. However, in this era of advanced DNA sequencing technologies, phylogenies are simultaneously used to describe the relationships among different genes, the demographic changes and migration patterns of species, and the evolutionary and epidemiological patterns of pathogens (Yang and Rannala, 2012). Phylogenetic analysis helps in monitoring the genetic variations occurring in pathogens and this information can be used to investigate various outbreaks and epidemics. Monitoring of genetic variations can also be helpful to determine appropriate public health interventions for the control and prevention of infectious diseases (Villabona-Arenas *et al.*, 2020; Wang *et al.*, 2015). Therefore, understanding the prevailing *Plasmodium* species in endemic areas is always important and helpful in implementing effective control strategies. The prevalence and distribution of *P. vivax* and *P. falciparum* genotypes have been previously reported in different districts of the province (Karim *et al.*, 2021; Khan *et al.*, 2021; Qureshi *et al.*, 2020). This study explores the phylogeny of *P. vivax* in three endemic districts of the province which will deepen our understanding about the phylogenetic relationships of *P. vivax* present in the province.

**Fig. 1.** Map of Pakistan and geographic locations (sampling sites) of Peshawar, Kohat, and D.I. Khan (ArcGIS Desktop 10.8 Version: 10.7.0.10450).

**MATERIALS AND METHODS**

**Study area**

Khyber Pakhtunkhwa districts, namely, Peshawar, Kohat, and D.I. Khan (D.I. Khan) were chosen for sampling (Fig. 1). These are the areas, where mostly Afghan refugees and internally displaced persons (IDPs) are located and disease outbreaks of several infectious diseases have been observed over time (Ahmad *et al.*, 2022; Khan *et al.*, 2021). Most IDPs migrated to these regions after the Pakistan military operation (Khan, 2012). The province’s geographical area is 74, 521 km² and comprised of 35 million people as per the latest census (https://www.
Sample collection

Blood samples were collected to isolate *P. vivax* for culturing and in vitro evaluation of selected medicinal plants (Zareen et al., 2021). The study population comprised individuals who were classified under the following categories: Gender (male and female), different age groups, and areas (rural/urban), and who visited the district headquarters (DHQ) hospitals for medical examination. For data collection, hospitals were visited twice a week for a period of seven months (April to October) in 2018. Patients with clinical signs and symptoms of malaria (fatigue, chills, sweats, headaches, and fever) were included. The exclusion criteria of Raza et al. (2013) were followed where pregnant women, children of age <3 years and those who did not provide consent to participate were excluded.

Intravenous blood (3 ml) in a sterile syringe was randomly collected in ethylene diamine tetra-acetic acid (EDTA) (BD, USA) vacutainers from patients who fulfilled the inclusion criteria. All blood samples were collected in the presence of medical practitioners. The samples were transported to the Molecular Parasitology and Virology Laboratory, Kohat University of Science and Technology Kohat (KUST), Khyber Pakhtunkhwa, Pakistan.

Consent and ethical approval

Patients/guardians were orally convinced to obtain written informed consent before taking the blood sample. The Research and Ethical Committee of KUST guaranteed approval for the study vide Ref. No. KUST/Ethical Committee/17-06.

Sample processing

Leishman’s staining was used to initially detect malaria infection as previously described (Warhurst and Williams, 1996). Blood smears were examined for further *P. vivax* confirmation and exclusion of mixed infections. The smears (thick and thin) were prepared using Giemsa stain for 2 min, air dried, and examined at an oil immersion magnification of 100X under a microscope. The slides were gently rinsed with distilled water by using a dropper, air dried, and examined at an oil immersion magnification of 100X under a microscope (Olympus Binocular microscope) (Ndao et al., 2004). Samples after initial analysis were stored for future study at a temperature of –20 °C (Raza et al., 2013).

Whole blood 250 µl was processed for DNA isolation using a Thermo Scientific DNA extraction kit. The *pvmsp*-1 gene sequence was amplified using the forward and reverse primers *pvmsp*-1-F 5′GCCAAGACGTTGACTCTGACCTG3′ and *pvmsp*-1-R 5′CTTGTCAATTTCCCTTTTGAGGAC3′, respectively. The PCR amplification proceeded with an initial denaturation of 5 min at 95 °C, denaturation for 35 sec at 95 °C, annealing for 34 sec at 63 °C, and extension for 1 min at 72 °C, followed by a final extension for 5 min at 72 °C. The PCR reaction was comprised of a total volume of 25 µl including the master mix (12 µl), PCR water (10 µl), primers (0.5 µl each), and template DNA (2 µl). Blood samples from healthy individuals were used as a negative control. A UV transilluminator (Extra Gene®, USA) was used for PCR amplicon confirmation by using 2% agarose gel with ethidium bromide.

Sequenceing and phylogenetic analysis

The target *pvmsp*-1 gene yielded a 945 base pairs (bp) band size after amplification. A gene-GET PCR purification kit (Thermo Fisher Scientific, USA Lithuania) was used for DNA purification and the amplicons were sent to Advance Bioscience International (Lahore, Pakistan) for Sanger sequencing and were sequenced in both forward and reverse directions. The obtained sequences were visually scanned, analyzed to remove ambiguous base pairs, and finally trimmed by using the software BioEdit (version 7.2) (https://bioedit.software.informer.com/7.2/). Consensus sequences were obtained for BLAST analysis and Nucleotide BLAST (BLASTn) (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to analyze the sequenced samples. BLAST results of the highest similarity and query coverage were retrieved for further downstream process. Multiple sequence alignment was performed by using the MUSCLE algorithm in MEGA11 (version 11.0.11) software (Tamura et al., 2021). Phylogenetic tree analysis was performed using the Neighbor-joining statistical method and bootstrap replications of 1000 (Felsenstein, 1985; Saitou and Nei, 1987). The substitution model was Maximum Composite Likelihood, gaps were treated as partial deletion and all positions with less than 75% site coverage were eliminated (Tamura et al., 2004). The sequence data was submitted to the gene bank under the following accession numbers (OR452727, OR452728, OR452729).

RESULTS

All the samples were successfully amplified and produced a band size of 945 bp for the *pvmsp*-1 gene. The top hits and highly similar sequences (97%-99%) were retrieved for downstream phylogenetic tree construction. A dendrogram was constructed for *pvmsp*-1 and comprised sequences from 3 districts of Khyber Pakhtunkhwa and 69...
Fig. 2. Phylogenetic tree of *Plasmodium vivax* based on *pvmsp-1* gene. Evolutionary history was inferred using the Neighbor-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. Evolutionary analyses were conducted in MEGA11. The red-colored text represents sequences of the present study.

published sequences of different geographical localities including China, India, Turkey, Azerbaijan, Iran, New Papua Guinea, Myanmar, Afghanistan, Bangladesh, Pakistan, Thailand, Korea, Brazil, UK, and Australia. The phylogenetic analysis grouped *P. vivax* into two major clades with different subclades in each (Fig. 2). The retrieved sequences from other countries were distributed in the subclades. The first clade is split up further into three subclades comprising isolates from Hainan, Liaoning, Fujian, Anhui, Yunnan of China, India, Turkey, Azerbaijan, Iran, Myanmar, Afghanistan, South Korea, New Papua Guinea, and Pakistan. The second clade has isolates from D.I. Khan, Peshawar, and Kohat of Pakistan, Turkey, Bangladesh, Thailand, Korea, UK, Australia, and Sal-I and Belem of Brazil.

**DISCUSSION**

Malarial infection is a big challenge in endemic areas throughout the world, especially in developing countries. Rapidly changing climatic conditions have brought about a major shift in the prevalence and spread of vector-borne diseases. Khyber Pakhtunkhwa province shares borders with several malaria-endemic countries and seasonal outbreaks of *P. vivax* malaria always remained a major health problem. The increased movement of people across borders dramatically changes the *P. vivax* population and transmission pattern in the area. However, the phylogeny of *P. vivax* is limited and poorly explored in Khyber Pakhtunkhwa. Therefore, this study was designed to address the limitation of the poorly explored phylogeny of *P. vivax* in three malaria-endemic districts of Khyber Pakhtunkhwa. Therefore, this study was designed to address the limitation of the poorly explored phylogeny of *P. vivax* in three malaria-endemic districts of Khyber Pakhtunkhwa. In Pakistan, the overall estimated cases of malaria were reduced between 2019-2020, however, 18% of cases were due to *P. vivax* in 2020 (WHO, 2021). *P. vivax* and *P. falciparum* populations previously reported, from Khyber Pakhtunkhwa, are highly polymorphic and diverse allelic variants circulating in this region (Khan et al., 2014, 2021). Provincial estimates indicated that the highest percentage (30%) of malaria was reported from Khyber Pakhtunkhwa in comparison to other provinces of Pakistan. *P. vivax* is regarded as the leading cause of malarial infection in the Khyber Pakhtunkhwa and Punjab provinces (Karim et al., 2016; Qureshi et al., 2019).

The *P. vivax* genetic variations could appropriately be measured by using a promising vaccine candidate *pvmsp-1* gene (Cui et al., 2003a). This gene can differentiate the geographic distribution of the parasite by grouping it in different clades and further clusters according to its geographic origin (Cui et al., 2003b). *P. vivax* genetic diversity has also been explored in the province of Punjab, Pakistan using distinct genetic markers (Bibi et al., 2021;
Phylogeny of Plasmodium vivax in Malaria

Qureshi et al., 2019; Raza et al., 2013). Genetic studies of P. vivax are available in many countries such as Colombia, India, Myanmar, and Korea (Lim et al., 2000; Maestre et al., 2004; Moon et al., 2009; Zakeri et al., 2010a). P. vivax is assumed to possess a more diversified genetic makeup in comparison to other Plasmodium species (Qureshi et al., 2019).

The phylogenetic analysis grouped P. vivax into two major clades with different subclades in each. Phylogeny of P. vivax isolates from Punjab inferred two major clades and many subclades containing isolates from distinct countries of the world (Ali et al., 2016). Similarly, the P. vivax population from China was also separated into two major groups and many subdivisions in each group (Huang et al., 2014). While P. vivax of Afghanistan was separated into three major clusters with many sub-clusters (Zakeri et al., 2010b). The geographical differences in the distribution of mosquito vectors show susceptibility to different P. vivax variants which greatly influence the geographical distribution of the parasite. The genetic diversity of the parasite may be correlated to the people’s migration within endemic areas (Bibi et al., 2021), the influx of refugees from Afghanistan and the cross borders moments of people from other neighbouring countries as previously described by (Zakeri et al., 2010b). Individuals with different P. vivax variants can change the gene pool and thereby can lead to high genetic diversity in that area (Zakeri et al., 2010b).

Genetic diversity and gene mutations on different levels may probably lead to the emergence of antimalarial resistance in parasites. Among Plasmodium species, P. vivax exhibits high genetic diversity (Prugnolle et al., 2013) as well as the most diversified species in the Asian region (Hupalo et al., 2016). Plasmodium species have acquired the ability to show resistance to various trending drugs and multidrug-resistant strains have been reported (Price et al., 2007). However, scientists are actively searching for new drugs to overcome the problem of drug resistance and ethnomedicine is one of the best alternatives to cope with this problem. The people of Pakistan rely on the use of ethnomedicine to treat malaria and many plant species and compounds are being used in remote areas (Tariq et al., 2016). This mode of treatment suits the low socioeconomic conditions and could probably counter the problem of resistance to antimalarial drugs (Zareen et al., 2021). However, one of the major challenges is the lack of standardized protocols for the evaluation of the efficacy and safety of medicinal plants (Ali et al., 2020). More rigorous clinical trials are direly needed to evaluate the effectiveness of medicinal plants against P. vivax malaria. Furthermore, there is limited information on the mechanisms of action of medicinal plants against P. vivax. Previous studies have identified specific compounds in medicinal plants that have antimalarial activity (Tariq et al., 2016), but the underlying mechanisms of action are not well understood. Understanding these mechanisms is important for the development of new drugs and treatments (Ali et al., 2020, 2021). Finally, more comprehensive surveys are required of the traditional knowledge and use of medicinal plants for the treatment of P. vivax malaria. Many communities around the world use traditional medicinal plants to treat malaria, but there is limited information on the specific plants and their effectiveness (Mussarat et al., 2021). Collecting this information could help to identify new sources of antimalarial compounds and improve access to effective treatments for P. vivax malaria in developing and underdeveloped areas worldwide.

One of the major limitations of the study is the sample size; a small number of participants were selected, and the study duration was relatively shorter due to limited financial resources. Only representative samples from each district were processed for Sanger sequencing. Furthermore, the selection of only 3 districts out of 34 districts in the province is another limitation of the current study. We highly recommend the expansion of the study area by including other malaria-endemic districts and relatively large datasets to better understand the phylogeny of P. vivax.

CONCLUSION

In summary, the P. vivax population circulating in Khyber Pakhtunkhwa is highly heterogeneous. P. vivax was split into two major clades with many subclades comprising P. vivax variants from neighboring countries. The study demonstrates that cross-border movements and migration of internally displaced persons from war-affected areas could have resulted in the introduction of different variants of P. vivax. This may lead to a high emergence of antimalarial resistance in P. vivax.

ACKNOWLEDGEMENTS

The research work presented in this paper is part of the Doctor of Philosophy (PhD) thesis of Mr. Shehzad Zareen. We are indebted to all participants and lab technicians for their assistance during this study.

Funding

This study received no support from funding agencies in the public, commercial, or non-profit sectors.

IRB approval

The Research and Ethical Committee of KUST
guaranteed approval for the study vide Ref. No. KUST/ Ethical Committee/17-06.

Ethical statement
Patients/guardians were orally convinced to obtain written informed consent before taking blood samples. All participants were assured that the information would not be disclosed to any third party and would be used for research purposes only. All methods were performed following the relevant guidelines.

Statement of conflict of interest
The authors have declared no conflict of interest.

REFERENCES


Khan, S.N., Ali, R., Khan, S., Rooman, M., Norin,


Warhurst, D.C. and Williams, J.E., 1996. ACP broadsheet no 148. July 1996. Laboratory diagnosis of malaria. *J. clin. Pathol.*, **49**: 533-538. [https://doi.org/10.1136/jcp.49.7.533](https://doi.org/10.1136/jcp.49.7.533)


