Original Article

Population genetic analysis of *Plasmodium vivax vir* genes in Pakistan

Check for updates

Sylvatrie-Danne Dinzouna-Boutamba[1,†] (iD), Zin Moon[1,†], Sanghyun Lee[1,†], Sahib Gul Afridi[2] (iD), Hương Giang Lê[3], Yeonchul Hong[1] (iD), Byoung-Kuk Na[3,*] (iD), Youn-Kyoung Goo[1,4,*] (iD)

[1]*Department of Parasitology and Tropical Medicine, School of Medicine, Kyungpook National University, Daegu 41944, Korea;* [2]*Department of Biochemistry, Abdul Wali Khan University Mardan, Pakistan;* [3]*Department of Parasitology and Tropical Medicine, Department of Convergence Medical Science, and Institute of Health Science, Gyeongsang National University College of Medicine, Jinju 52727, Korea;* [4]*Institute for Veterinary Biomedical Science, College of Veterinary Medicine, Kyungpook National University, Daegu 41566, Korea*

Abstract

Plasmodium vivax variant interspersed repeats (*vir*) refer to the key protein used for escaping the host immune system. Knowledge in the genetic variation of *vir* genes can be used for the development of vaccines or diagnostic methods. Therefore, we evaluated the genetic diversity of the *vir* genes of *P. vivax* populations of several Asian countries, including Pakistan, which is a malaria-endemic country experiencing a significant rise in malaria cases in recent years. We analyzed the genetic diversity and population structure of 4 *vir* genes (*vir* 4, *vir* 12, *vir* 21, and *vir* 27) in the Pakistan *P. vivax* population and compared these features to those of the corresponding *vir* genes in other Asian countries. In Pakistan, *vir* 4 (S = 198, H = 9, Hd = 0.889, Tajima's D value = 1.12321) was the most genetically heterogenous, while the features of *vir* 21 (S = 8, H = 7, Hd = 0.664, Tajima's D value = -0.63763) and *vir* 27 (S = 25, H = 11, Hd = 0.682, Tajima's D value = -2.10836) were relatively conserved. Additionally, *vir* 4 was the most genetically diverse among Asian *P. vivax* populations, although within population diversity was low. Meanwhile, *vir* 21 and *vir* 27 among all Asian populations were closely related genetically. Our findings on the genetic diversity of *vir* genes and its relationships between populations in diverse geographical locations contribute toward a better understanding of the genetic characteristics of *vir*. The high level of genetic diversity of *vir* 4 suggests that this gene can be a useful genetic marker for understanding the *P. vivax* population structure. Longitudinal genetic diversity studies of *vir* genes in *P. vivax* isolates obtained from more diverse geographical areas are needed to better understand the function of *vir* genes and their use for the development of malaria control measures, such as vaccines.

Keywords: *Plasmodium vivax*, Pakistan, *vir*, population structure, genetic diversity

Introduction

Pakistan is a malaria-endemic country that has experienced, due to frequent floods and weak disaster preparedness, a significant rise in malaria cases in recent years. In 2022, the number of suspected malaria cases increased to > 3.4 million, compared to 2.6 million estimated cases in 2021. Two *Plasmodium* species, *P. falciparum* and *P. vivax*, have been reported in Pakistan, with the latter predominating (> 80%) [1].

The variant surface antigens (VSAs) of many *Plasmodium* spp. are considered the key proteins used by the parasite to escape the host immune system [2]. These proteins are encoded by multigene families located on telomeric and subtelomeric regions of the parasite's

ORCID
Sylvatrie-Danne Dinzouna-Boutamba
(https://orcid.org/0000-0001-6127-5356)
Sahib Gul Afridi
(https://orcid.org/0000-0001-9411-4823)
Yeonchul Hong
(https://orcid.org/0000-0002-2806-188X)
Byoung-Kuk Na
(https://orcid.org/0000-0002-6734-1673)
Youn-Kyoung Goo
(https://orcid.org/0000-0002-8784-2387)

chromosomes. Similar to the *var* gene superfamily in *P. falciparum*, the multigene super-family of variant interspersed repeats (*vir*) in *P. vivax* encodes proteins used for avoiding the host immune system [3,4]. The *vir* gene superfamily consists of 346 genes divided into 12 subfamilies A–L [5]. VSAs are potential targets for vaccine development; however, the lack of an in vitro culture system for *P. vivax* and the variant nature of VIRs have presented challenges to their application as a vaccine or diagnostic tool [6].

Despite the variant nature of VAR proteins, vaccine trials have been partially established [7]. VIR proteins can protect against sequestration through the use of disrupting rosettes [8]. Additionally, these proteins induce moderate natural acquisition of antibody and cellular responses in *P. vivax*-endemic countries [9]. Like *var* genes, *vir* genes are potent candidates for vaccination against vivax malaria [5]. Since researching the genetic variation of VSAs is important, the genetic diversity of *vir* genes in the Republic of Korea (Korea), India, and Myanmar has been previously evaluated [10-13]. Here, we analyzed the genetic characteristics of *vir* genes in the *P. vivax* population of Pakistan and evaluated the population genetics of *vir* genes among different Asian populations. Our aim was to obtain in-depth information on the population structure of *P. vivax*, which can then be used for the development of vaccines or diagnostic tools.

Materials and Methods

Ethics statement
The Ethics Committee of Abdul Wali Khan University Mardan approved this study protocol (AWKUM/Biochem/Dept/Commit/ECR/808).

Sample collection and genomic DNA extraction
We collected a total of 89 blood samples from Khyber Pakhtunkhwa province of Pakistan between 2020 and 2021. *P. vivax* infection was diagnosed by rapid diagnostic test kits and microscopic examination. The patients' blood samples were spotted on filter paper discs (Whatman 3 mm, GE Healthcare, Chicago, IL, USA) prior to drug treatment, and then they were air-dried and stored in sealed plastic bags at ambient temperature until use. Genomic DNA of *P. vivax* was extracted from each blood spot using a QIAamp DNA Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. The extracted DNA was preserved at -70℃ until subsequent use. Whether infections were due to only *P. vivax* or mixed with *P. falciparum* was determined by performing nested polymerase chain reaction (PCR) targeting the 18S ribosomal RNA (rRNA) gene as previously described [14].

PCR amplification and sequencing of *vir* genes
Four *vir* genes (*vir* 4, *vir* 12, *vir* 21, and *vir* 27) were selected based on subfamilies, sizes, and potent functions. These genes were amplified as previously described with minor modifications (Supplementary Table S1) [11]. Each amplicon was purified from the gel using an AccuPrep Gel Purification Kit (Bioneer, Daejeon, Korea), cloned into a pGEM-T TA cloning vector (Promega, Madison, WI, USA), and sequenced via the pUC/M13 forward and reverse primers using the Big-Dye Terminator ver. 3.1 sequencing kit (Applied Biosystems, Waltham, MA, USA) and an automatic ABI3730 DNA Analyzer sequencer

(Applied Biosystems).

Diversity of *vir* genes and population genetic analysis

All of the *vir* gene sequences were aligned by the Clustal Omega multiple alignment function (https://www.ebi.ac.uk/jdispatcher/msa/clustalo). The segregating sites (S), number of haplotypes (H), haplotype diversity (Hd), average number of pairwise nucleotide differences (π), total number of segregating sites (Θ_w), and Tajima's D value were evaluated using DnaSP ver. 6 [15]. When comparing between means, *P*-values < 0.05 were considered statistically significant. Population genetic analyses, including pairwise fixation index (F_{ST}), analysis of molecular variance (AMOVA), haplotype frequencies, and nucleotide diversity based on Bei's net distance (DA), were performed for *vir* genes from Pakistan, Myanmar, India, and Korea using Arlequin v.3.5.2.2. Global *vir* gene sequences for population genetic analyses were obtained from GenBank under the following accession numbers: Sal-1: AAKM01000104; Indian isolates: JQ733948-JQ33952; Korea isolates: KY608341-KY608363; Myanmar isolates: MN436008-436011 for *vir* 4; Sal-1: AAKM01000016; Indian isolates: JQ733953-JQ733971; Korea isolates: KY608364-KY08419; Myanmar isolates: MN436012-MN436028 for *vir* 12; Sal-1: AAKM01000003; Indian isolates: JQ733972-JQ733988; Korea isolates: KY608420-KY608471; MN436029-MN436044 for *vir* 21; Sal-1: AAKM01000041; Indian isolates: JQ733915-JQ733947; Korea isolates: KY608472-KY608510; Myanmar isolates: MN436045-MN436067 for *vir* 27 [10-13]. We analyzed the haplotype networks and generated network plots using PopArt version 1.7 [16].

Results

Genetic polymorphism of Pakistan *vir* genes

Among 4 *vir* genes amplified from *P. vivax* samples collected in Pakistan, *vir* 27 had the highest PCR-positivity rate (19.1%, 17/89), followed by *vir* 4 (18.0%, 16/89), *vir* 21 (9.0%, 8/89), and *vir* 12 (6.7%, 6/89) (Table 1).

Compared with the Sal-1 reference sequence, the 16 Pakistan *vir* 4 sequences featured 198 single nucleotide polymorphisms (SNPs), 34 singleton variable sites, and 164 parsimony-informative sites. This number of SNPs detected in the Pakistan collection of *vir* 4 is remarkably higher than that in Myanmar (9 SNPs), India (3 SNPs), and Korea (6 SNPs). The

Table 1. Summary of genetic diversity of the 4 *vir* genes in Pakistan *P. vivax*

VIR genes (No. of analyzed genes)	Sequence length (bp)	Segregating sites (S)	Segregating sites (S) in Myanmar[a] (No. of analyzed genes)	Segregating sites (S) in India[b] (No. of analyzed genes)	Segregating sites (S) in Korea[c] (No. of analyzed genes)	No. of haplotypes (H)	Haplotype diversity (Hd)	Nucleotide diversity π	Nucleotide diversity Θ_w	Tajima's D test
vir 4 (16)	1,314	198	9 (15)	3 (6)	6 (26)	9	0.889	0.02905	0.04211	1.12321
vir 12 (6)	1,323	34	168 (29)	85 (19)	74 (59)	3	0.733	0.01816	0.01712	0.38820
vir 21 (8)	998	8	150 (23)	155 (22)	143 (43)	7	0.664	0.00244	0.00282	-0.63763
vir 27 (17)	1,268	25	32 (32)	22 (45)	10 (62)	11	0.682	0.00278	0.00584	-2.10836*

[a]Ref. [10].
[b]Ref. [13].
[c]Ref. [11].
*$P < 0.05$.

198 *vir* 4 SNPs results from 129 non-synonymous substitutions in this gene. The *vir* 4 sequences from Pakistan have been deposited in GenBank under accession numbers PP5066 44–PP506659. Analysis of Pakistan *vir* 4 sequences showed 9 haplotypes, with an Hd value of 0.889; high values of π and Θ_w of 0.02905 and 0.04211, respectively; and a positive Tajima's D value of 1.12321 (Table 1).

Analysis of 6 Pakistan *vir* 12 sequences identified a total of 34 SNPs, which is a number lower than those previously reported in Myanmar (168 SNPs), India (85 SNPs), and Korea (74 SNPs) (Table 1). We also observed 17 each of singleton sites with 2 variants and parsimony-informative sites with 2 variants in Pakistan *vir* 12. Compared to *vir* 12 of Sal-1, the Pakistan gene contains 2 nucleotide deletions at s sites: 706–720 (GGATCTGAAGGGGCA encoding GSEGA), 762–852 (GTAAAACCTGCACCTGCAAAACCTGTAGCG encoding VKPAPAKPVA), and 762–837 (GTAAAACCTGCACCTGCAAAACCTGTAGCG-GCAAAACCTGTGGCGACAAGCCCAGCACCGTCAGAACGTGCACCT encoding VKPAPAKPVAAKPVATSPAPSERAP). The sequences of Pakistan *vir* 12 have been deposited in GenBank under accession numbers PP505464–PP505469. The H, Hd, π, Θ_w, and Tajima's D values for Pakistan *vir* 12 are 3, 0.733, 0.01816, 0.01712, and 0.38820, respectively (Table 1).

Analysis of *vir* 21 sequences found 8 SNPs, with 5 singleton and 3 parsimony-informative sites, which are much lower than those of other regions (Table 1). Four SNPs were non-synonymous mutations. The *vir* 21 sequences have been deposited in GenBank under accession numbers PP505456–PP505463. The values of H, Hd, π, and Θ_w were 7, 0.664, 0.00244, and 0.00282, respectively, and Tajima's D value was -0.63763 (Table 1).

Compared to the Sal-1 reference sequence, analysis of Pakistan *vir* 27 sequences found 25 SNPs, with 22 singleton variable and 3 parsimony-informative sites. This number of segregating sites in Pakistan *vir* 27 is higher than those from India (22 SNPs) and Korea (10 SNPs) but lower than that from Myanmar (32 SNPs). Thirteen Pakistan *vir* 27 SNPs are non-synonymous mutations. The sequences of Pakistan *vir* 27 have been deposited in GenBank under accession numbers PP505470–PP505486. The H, Hd, π, Θ_w, and Tajima's D values for Pakistan *vir* 27 are 11, 0.682, 0.00278, 0.00584, and -2.10836 ($P < 0.05$), respectively (Table 1).

**Nucleotide variability of *vir* genes in Asian *P. vivax* isolates**

Analyses of 48, 85, 93, and 112 sequences of respective *vir* 4, *vir* 12, *vir* 21, and *vir* 27 from 4 Asian countries (Pakistan, India, Myanmar, and Korea) in comparison to the reference sequence of Sal-1 found 18, 35, 52, and 40 haplotypes, respectively.

Analysis of genetic variance of *vir* genes among and within the Asian populations found a higher level of genetic variance among different geographic populations than within each population (Table 2). The *vir* 4 gene showed the highest level of variation among geographic populations (96.1%) and the lowest level of variation within populations (3.9%). In contrast, *vir* 21 had the lowest level of variation among geographic populations (57.1%) and the highest level of variation within populations (42.9%). Fig. 1 shows the results of pairwise difference analysis among Asian populations, with the Pakistan population of *vir* 4 having the higher level of genetic distinction (πxx = 40.18, shown as red square in Fig. 1A) and *vir* 21 having the lower level of genetic distinction (πxx = 3.50, shown as ivory square

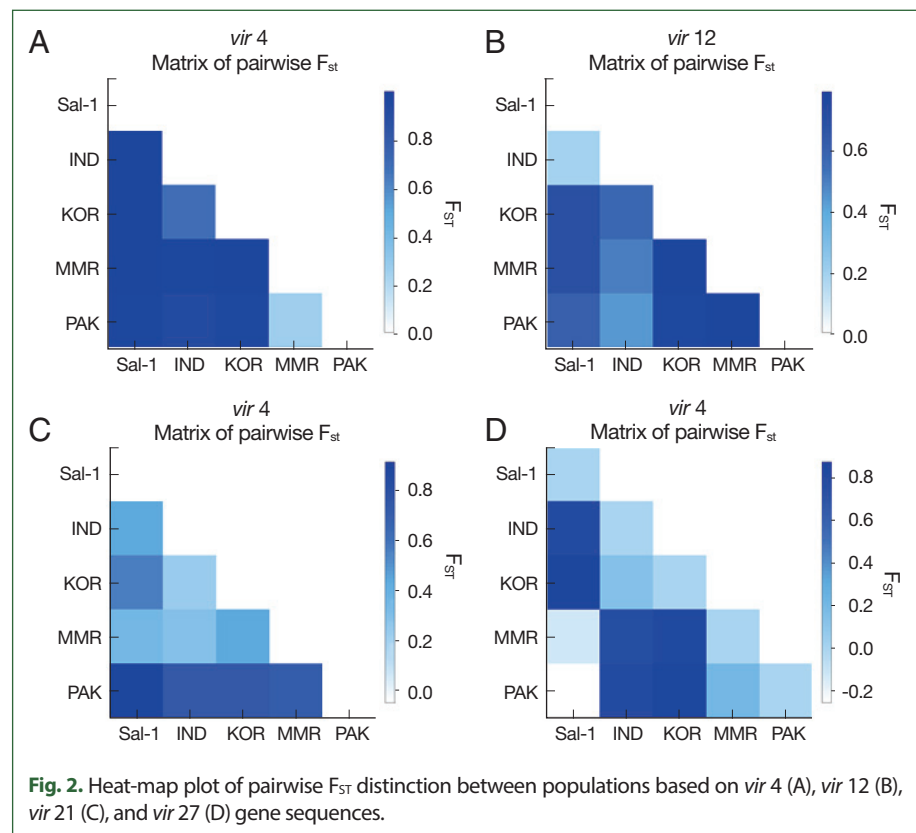**Table 2.** Analysis of molecular variance in populations of *vir* genes from Asian countries

vir genes (No. of analyzed genes)	Source of variation								F_{ST} value
	Among populations				Within populations				
	df	Sum of squares	Variance components	Percentage variation	df	Sum of squares	Variance components	Percentage variation	
vir 4 (48)	4	986.904	30.56922	96.1	46	54.402	1.23641	3.9	0.96113*
vir 12 (85)	4	3,899.622	62.20029	63.4	94	3,377.651	35.93245	36.6	0.63384*
vir 21 (93)	4	1,791.223	29.00720	57.1	89	1,940.288	21.80099	42.9	0.57092*
vir 27 (112)	4	370.749	4.40129	61.7	108	187.950	1.74028	28.3	0.61664*

*P < 0.05.
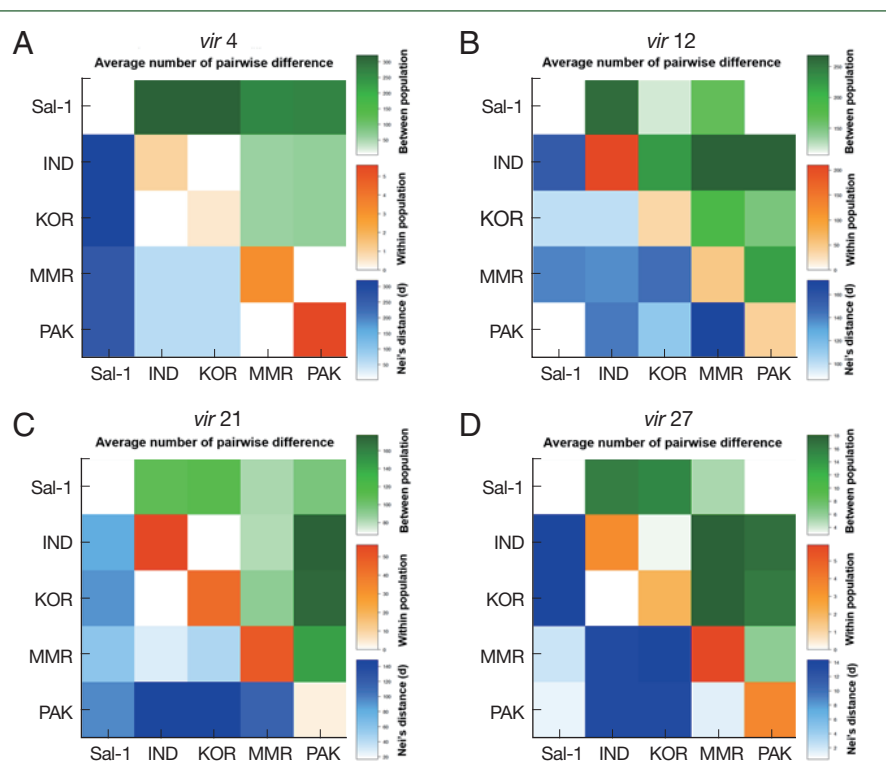
Fig. 1. Pairwise population genetic analyses. The average number of pairwise differences between populations (πxy, green above the diagonal), within populations (πxx, orange diagonal), and the net number of nucleotide differences between populations (Nei distance DA, blue below the diagonal) were analyzed based on variants of *vir* 4 (A), *vir* 12 (B), *vir* 21 (C), and *vir* 27 (D).

in Fig. 1C) compared to other populations (πxx of *vir* 4: 1.00, 0.51, and 3.17 for India, Korea, and Myanmar populations, respectively; and πxx of *vir* 21: 56.31, 43.29, and 90.28 for India, Korea, and Myanmar populations, respectively). The Indian population of *vir* 12 showed a greater within population genetic difference compared to those in Korea, Myanmar, and Pakistan (Fig. 1B).
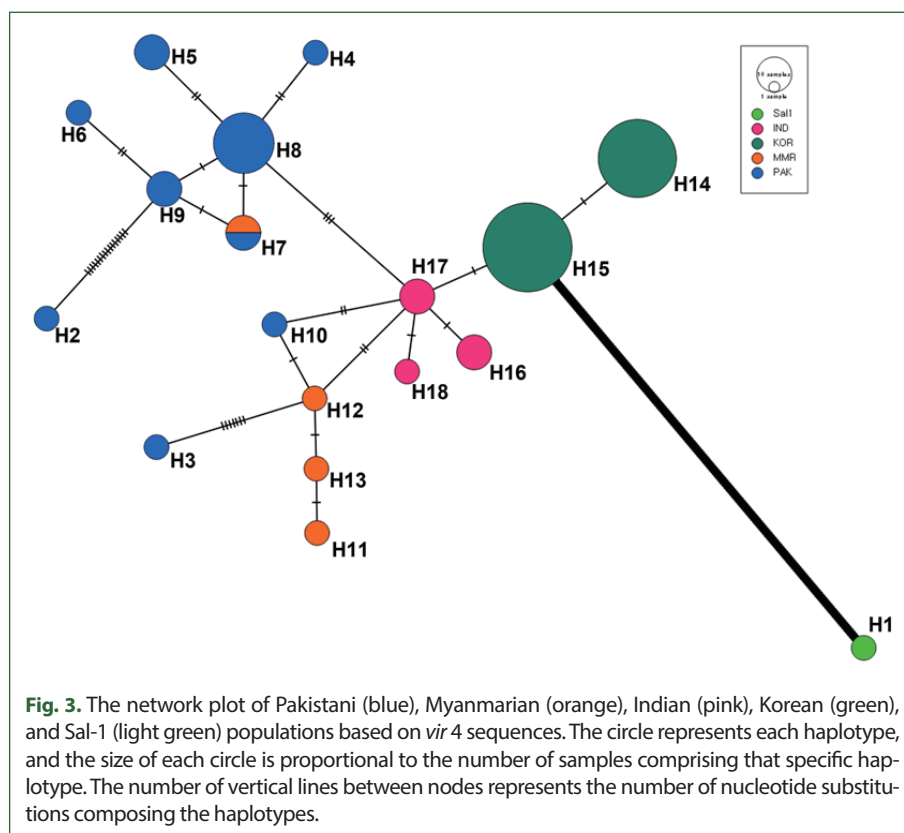
The fixation index F_{ST} was used to assess the genetic differentiation among the 4 *vir* genes collected from 4 different Asian populations. Fig. 2 shows different pairwise F_{ST} patterns in each *vir* gene. Compared to the Sal-1 reference sequence, the *vir* 4 gene demon-

Fig. 2. Heat-map plot of pairwise F_{ST} distinction between populations based on *vir* 4 (A), *vir* 12 (B), *vir* 21 (C), and *vir* 27 (D) gene sequences.

strated remarkable genetic differentiation among populations (Fig. 2A). The Pakistan population was moderately distinct from those of India and Korea; however, it was not distinct from that of Myanmar ($F_{ST} = -0.01812$, *P*-value < 0.05). Compared to the Sal-1 reference sequence, the F_{ST} values for *vir* 12 were comparable among all populations except for that of India ($F_{ST} = 0.19195$, *P* > 0.05, Fig. 2B). Comparisons for *vir* 21 showed the Pakistan population with a higher genetic difference to the other populations, especially with the Sal-1 reference sequence ($F_{ST} = 0.96392$, *P* > 0.05, Fig. 2C). Comparisons for *vir* 27 showed the Pakistan population with a noticeable genetic distinction from those of India and Korea, but not with the Sal-1 sequence (Fig. 2D).

The haplotype network of *vir* genes in Asian isolates

Based on the 18, 35, 52, and 40 haplotypes for *vir* 4, *vir* 12, *vir* 21, and *vir* 27, respectively, we constructed a haplotype network to characterize the frequencies and relationships of the different haplotypes. Haplotype networks, especially for *vir* 4, were geographically clustered (Fig. 3). All Pakistan *vir* 4 haplotypes were connected to the Indian and Myanmar populations, with a weaker connection to that of Korea, showing the geographical segregation of haplotypes. However, this pattern was not observed for the other *vir* genes. Haplotype networks of *vir* 12 and 21 showed that the Pakistan population was closely related to Korean and Indian populations. Meanwhile, haplotype network of *vir* 27 showed close relationships among all populations (Supplementary Fig. S1).

Fig. 3. The network plot of Pakistani (blue), Myanmarian (orange), Indian (pink), Korean (green), and Sal-1 (light green) populations based on *vir* 4 sequences. The circle represents each haplotype, and the size of each circle is proportional to the number of samples comprising that specific haplotype. The number of vertical lines between nodes represents the number of nucleotide substitutions composing the haplotypes.

Discussion

The *VIR* proteins are attractive vaccine candidates because they induce a long-lasting antibody response with similar antibody levels in immune sera between singly- and multiply-infected patients without clonal expression [2]. Moreover, their cell adherence abilities suggest that they play key roles in the pathophysiology of *P. vivax* infection [17,18]. The genetic diversity patterns in of *VIR* proteins in *P. vivax* isolates from several countries indicate that *vir* genes can be used as genetic markers for the evaluation of genetic polymorphism and transmission hotspots in malaria-endemic regions in the long term [19]. Therefore, assessing the genetic diversity and population structure in global isolates is important to provide new insights into vaccine development and genetic marker potential.

Consistent with the results of previous studies [10-13], we observed different amplification positivity rates among 4 *vir* genes, which may be due to their location on telomeres and subtelomeres of the *P. vivax* genome. Genes in telomeres are very difficult to amplify because of the repetitive nature of telomeric DNA and unusual terminal structure [20]. Previous analysis of the gene diversity of the 4 *vir* genes have found that *vir* 4 is the most conserved [11,14,15]. However, our analysis of Pakistan isolates found that *vir* 4 is the most polymorphic gene with many segregating sites compared to the Sal-1 reference sequence. Furthermore, *vir* 21, which is polymorphic in other populations, was conserved among the Pakistan isolates. We attribute our findings to the dynamic molecular evolution of *vir* genes due to the sharp increase in malaria cases in Pakistan. This explanation is supported by the

results of neutrality analyses of Pakistan *P. vivax* isolates, as demonstrated by the wide range of Tajima's D values in this population (from -2.10836 for *vir* 27 to 1.12321 for *vir* 4). A positive Tajima's D value indicates tendencies for an increase in polymorphism, a decrease in population size, and/or balancing selection. A negative Tajima's D value tendencies for low frequency of polymorphism, increased population size, and/or purifying selection. Additionally, a Tajima's D value below -2 or above 2 is generally a strong indication that a gene is not evolving neutrally. Genes with a Tajima's D value below -2 have an excess of rare alleles, indicating positive selection or a selective sweep. However, genes with a Tajima's D value above 2 have an excess of common alleles suggestive of balancing selection [21]. In contrast to *vir* 4, *vir* 27 has a Tajima's D test value below -2, which indicates unfavorable circumstances for survival and/or in the process of being purged from the gene pool via population expansion, an effect that increases with a lower mutation rate but decreases with a higher recombination rate [22]. Furthermore, the high diversity of *vir* genes, along with a dramatic population expansion, indicates that the *vir* genes may play a role in host immune evasion, although more research is needed to explore this possibility.

Analyses of pairwise differences and the haplotype network shows that a geographical relationship among global isolates based on the *vir* 4 gene, and the Sal-1 reference sequence had a remarkably high pairwise difference and low genetic relationship with other Asian isolates and even haplotypes from each of the Asian countries in the network map (Fig. 3). This result is supported by the results of AMOVA analysis, which demonstrates high (96.1%) and low (3.9%) percentage variation values among and within populations, respectively, based on *vir* 4 characteristics. These results suggest that *vir* 4 can be useful as a population marker to evaluate the geographical evolution and genetic structure of *P. vivax*. In contrast to *vir* 4, we observed the formation of one node consisting of *vir* 21 and *vir* 27 haplotypes among 5 Asian countries and even Brazil (Sal-1 strain). These close relations are consistent with the relatively low pairwise difference and percentage variation values in AMOVA analysis among populations.

There is currently no useful vaccine for *P. vivax* malaria; however, *vir* genes are vaccine candidates [23]. Despite this possibility, there are only limited studies on the genetic diversity and population relationships of *vir* genes at the global level. Longitudinal genetic diversity studies of *vir* genes in *P. vivax* populations from diverse countries are necessary to further understand *vir* gene function and its application to malaria control, e.g., as vaccines.

In the present study, we investigated the genetic diversity and relationships between populations of *P. vivax* from 5 countries (Pakistan, Myanmar, India, Korea, and Brazil) based on analysis of 4 *vir* genes. Among the 4 genes, *vir* 4 showed remarkable genetic differentiation among populations, although its characteristics were conserved within populations, suggesting that *vir* 4 can be a potent marker for understanding *P. vivax* population structure. In contrast, *vir* 21 and *vir* 27 are characterized by a low pairwise difference and percentage variation among populations; therefore, these can be considered vaccine candidates. A better understanding of the function of *vir* genes and their application to malaria control, e.g., as vaccines, requires longitudinal genetic diversity studies of *vir* genes in samples obtained from more diverse countries.

Acknowledgments

References

1. World Health Organization. World Malaria Report 2022. World Health Organization. Geneva, Switzerland. 2022.

2. Fernandez-Becerra C, Pein O, de Oliveira TR, Yamamoto MM, Cassola AC, et al. Variant proteins of Plasmodium vivax are not clonally expressed in natural infections. *Mol Microbiol* 2005;58(3):648-658. https://doi.org/10.1111/j.1365-2958.2005.04850.x

3. del Portillo HA, Lanzer M, Rodriguez-Malaga S, Zavala F, Fernandez-Becerra C. Variant genes and the spleen in *Plasmodium vivax* malaria. *Int J Parasitol* 2004;34(13-14):1547-1554. https://doi.org/10.1016/j.ijpara.2004.10.012

4. Pasternak ND, Dzikowski R. PfEMP1: an antigen that plays a key role in the pathogenicity and immune evasion of the malaria parasite *Plasmodium falciparum*. *Int J Biochem Cell Biol* 2009;41(7):1463-1466. https://doi.org/10.1016/j.biocel.2008.12.012

5. Fernandez-Becerra C, Yamamoto MM, Vêncio RZ, Lacerda M, Rosanas-Urgell A, et al. *Plasmodium vivax* and the importance of the subtelomeric multigene vir superfamily. *Trends Parasitol* 2009;25(1):44-51. https://doi.org/10.1016/j.pt.2008.09.012

6. Arévalo-Herrera M, Chitnis C, Herrera S. Current status of *Plasmodium vivax* vaccine. *Hum Vaccin* 2010;6(1):124-132. https://doi.org/10.4161/hv.6.1.9931

7. Bull PC, Abdi AI. The role of PfEMP1 as targets of naturally acquired immunity to childhood malaria: prospects for a vaccine. *Parasitology* 2016;143(2):171-186. https://doi.org/10.1017/S0031182015001274

8. Oliveira TR, Fernandez-Becerra C, Jimenez MC, Del Portillo HA, Soares IS. Evaluation of the acquired immune responses to *Plasmodium vivax* VIR variant antigens in individuals living in malaria-endemic areas of Brazil. *Malar J* 2006;5:83. https://doi.org/10.1186/1475-2875-5-83

9. Lee S, Choi YK, Goo YK. Humoral and cellular immune response to Plasmodium vivax VIR recombinant and synthetic antigens in individuals naturally exposed to *P. vivax* in the Republic of Korea. *Malar J* 2021;20(1):288. https://doi.org/10.1186/s12936-021-03810-2

10. Na BK, Kim TS, Lin K, Baek MC, Chung DI, et al. Genetic polymorphism of vir genes of *Plasmodium vivax* in Myanmar. *Parasitol Int* 2021;80:102233. https://doi.org/10.1016/j.parint.2020.102233

11. Son UH, Dinzouna-Boutamba SD, Lee S, Yun HS, Kim JY, et al. Diversity of vir genes in *Plasmodium vivax* from endemic regions in the Republic of Korea: an initial evaluation. *Korean J Parasitol* 2017;55(2):149-158. https://doi.org/10.3347/kjp.2017.55.2.149

12. Gupta P, Pande V, Das A, Singh V. Genetic polymorphisms in VIR genes among Indian *Plasmodium vivax* populations. *Korean J Parasitol* 2014;52(5):557-564. https://doi.org/10.3347/kjp.2014.52.5.557

13. Gupta P, Das A, Singh OP, Ghosh SK, Singh V. Assessing the genetic diversity of the vir genes in Indian *Plasmodium vivax* population. *Acta Trop* 2012;124(2):133-139.

14. Snounou G, Viriyakosol S, Zhu XP, Jarra W, Pinheiro L, et al. High sensitivity of detection of human malaria parasites by the use of nested polymerase chain reaction. *Mol Biochem Parasitol* 1993;61(2):315-320. https://doi.org/10.1016/0166-6851(93)90077-b

15. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 2009;25(11):1451-1452. https://doi.org/10.1093/bioinformatics/btp187

16. Leigh JW, Bryant D. PopART: full-feature software for haplotype network construction. *Methods Ecol Evol* 2015;6(9):1110-1116.

17. Goo YK. Vivax malaria and the potential role of the subtelomeric multigene vir superfamily. *Microorganisms* 2022;10(6):1083. https://doi.org/10.3390/microorganisms10061083

18. Fernandez-Becerra C, Bernabeu M, Castellanos A, Correa BR, Obadia T, et al. *Plasmodium vivax* spleen-dependent genes encode antigens associated with cytoadhesion and clinical protection. *Proc Natl Acad Sci U S A* 2020;117(23):13056-13065. https://doi.org/10.1073/pnas.1920596117

19. Wesolowski A, Taylor AR, Chang HH, Verity R, Tessema S, et al. Mapping malaria by combining parasite genomic and epidemiologic data. *BMC Med* 2018;16(1):190. https://doi.org/10.1186/s12916-018-1181-9

20. Bonnell E, Pasquier E, Wellinger RJ. Telomere replication: solving multiple end replication problems. *Front Cell Dev Biol* 2021;9:668171. https://doi.org/10.3389/fcell.2021.668171

21. Eckshtain-Levi N, Weisberg AJ, Vinatzer BA. The population genetic test Tajima's D identifies genes encoding pathogen-associated molecular patterns and other virulence-related genes in Ralstonia solanacearum. *Mol Plant Pathol* 2018;19(9):2187-2192. https://doi.org/10.1111/mpp.12688

22. Hudson RR, Kaplan NL. Deleterious background selection

with recombination. *Genetics* 1995;141(4):1605-1617.

23. Fantin RF, Coelho CH, Berhe AD, Magalhães LMD, Pereira DB, et al. Immunological characterization of a VIR protein family member (VIR-14) in *Plasmodium vivax*-infected sub-jects from different epidemiological regions in Africa and South America. *PLoS Negl Trop Dis* 2023;17(4):e0011229. https://doi.org/10.1371/journal.pntd.0011229