JPAM
JOURNAL OF PURE AND APPLIED MICROBIOLOGY

RESEARCH ARTICLE   **OPEN ACCESS**

Plasmodium falciparum Drug Resistance Single Nucleotide Polymorphism Analysis Pipeline Based on Southeast Asian Genetic Samples

Daniel R. Fugaha[1] iD **, Rio Hermantara[2]** iD **, Arli A. Parikesit[1,2]*** iD
and Elizabeth Sidhartha[2] iD

[1]Department of Bioinformatics, School of Health and Life Sciences, i3L University, Jakarta 13210, Indonesia.
[2]Department of Biomedicine, School of Health and Life Sciences, i3L University, Jakarta 13210, Indonesia.

Abstract

Drug resistance in pathogens represents an escalating global health challenge, particularly in *Plasmodium falciparum*, the causative agent of malaria. The emergence of drug-resistant strains complicates treatment and highlights the need for rapid and accurate detection methods. In *P. falciparum*, resistance has been largely attributed to single nucleotide polymorphisms (SNPs) in key genes such as *PfMDR1*, *PfKelch13*, and *PfCRT*, which are associated with resistance to artemisinin-based combination therapies and chloroquine—two mainstays of antimalarial treatment. This study aimed to develop a bioinformatics pipeline capable of analyzing *P. falciparum* genomic sequences to detect and annotate SNPs that may confer drug resistance. The pipeline was implemented on the Galaxy online analysis platform using its workflow function. It processes both reference and sample sequences through alignment, mutation detection, SNP selection, and annotation based on a reference general feature format. Although no SNPs were identified directly within known drug resistance genes in the analyzed samples, the developed pipeline successfully detected and annotated SNPs across the chromosomes containing these genes. This approach provides a practical framework for future applications in point-of-care detection and surveillance of drug-resistant *P. falciparum* strains.

Keywords: Drug Resistance, *Plasmodium falciparum*, Single Nucleotide Polymorphisms (SNPs), Bioinformatics Pipeline, Artemisinin-based Combination Therapies

*Correspondence: arli.parikesit@i3l.ac.id

INTRODUCTION

Drug resistance is rapidly emerging in many pathogens, including multidrug-resistant *Mycobacterium tuberculosis*, methicillin-resistant *Staphylococcus aureus*, and *Neisseria gonorrhoeae*.[1,2] This rise in resistance has been linked to the extensive use of antimicrobials in medicine and agriculture, which enables surviving microbes to transmit their resistance traits to subsequent generations.[3,4] Drug resistance is typically assessed in laboratories by exposing pathogens to various antimicrobial agents and monitoring their proliferation rates.[5] These tests, however, are time-consuming, and early administration of effective treatment is crucial for improving patient outcomes.[6] Rapid, on-site detection of antimicrobial resistance would therefore help reduce the severity and lethality associated with drug-resistant infections.[7,8]

Drug resistance arises in microbes when a subset of the population exposed to antimicrobials survives, eventually leading to the evolution of resistance-associated genes.[9,10] Because these are genetic changes, alterations in the microbial genome can be detected and used to develop point-of-care diagnostic tests.[11] Drug resistance genes acquire various types of mutations, including insertions, deletions, and single nucleotide polymorphisms (SNPs).[12] Accordingly, several studies have sought to identify these genetic mutations to determine the drug resistance potential of pathogens.[13] For example, computational simulations using clustered regularly interspaced short palindromic repeats (CRISPR) and loop-mediated isothermal amplification protocols have been employed to detect treatment-resistant *Plasmodium falciparum* by targeting SNP loci through *in silico* methods.[14,15] Targeting suspected drug resistance genes with nucleases can further confirm their role in conferring resistance. CRISPR, in particular, is widely used for genetic modification because a single guide RNA directs the nuclease to create precise double-stranded breaks at the desired genomic site.[16]

P. falciparum is the parasite responsible for malaria in humans, with an estimated 263 million cases and 597,000 malaria-related deaths reported in 2023.[17,18] The disease is commonly transmitted through mosquito vectors.[19] However, *P. falciparum* has shown a steady rise in various forms of drug resistance, leading to increased treatment complexity and poorer patient outcomes.[20] Previous studies have demonstrated that drug resistance genes can be rapidly isolated and detected to enable point-of-care testing and guide more effective therapy.[14,15] To contribute to these efforts, the present study aims to develop a detection pipeline capable of identifying SNPs within chromosomes 5, 7, and 13 of *P. falciparum*.

MATERIALS AND METHODS

Study design

This study was designed in three main stages to develop a SNP detection pipeline: (1) data acquisition of both the reference genome and sample sequences, (2) SNP identification through sequence alignment, and (3) annotation of the detected SNPs, if present. The analysis and pipeline development were performed using existing tools available on Galaxy, a web-based data analysis platform from which the workflow was extracted.[21] The Galaxy genome annotation pipeline provides an accessible, automated, and scalable framework for genome analysis. It includes all the necessary tools to help with both functional and structural gene annotation, enabling researchers with minimal bioinformatics experience to perform accurate functional and structural gene annotation. An illustration of the workflow is presented in Figure, and the complete pipeline can be accessed through the Galaxy Server.[22] Because real laboratory samples were not available, dummy samples were generated by combining known mutation sites of *P. falciparum* with the full genomic sequences of chromosomes 5, 7, and 13.

Genomic sequence sample acquisition

Genomic sample data were obtained from previously published research with publicly available datasets accessible at http://www.malariagen.net/resource/29. The samples originated from regions in Africa, South Asia, and Southeast Asia.[23] Data from outside these regions were excluded through filtering. The dataset was provided in spreadsheet format and processed using the Python package *pandas*. Given the

large dataset size, a subset of 100 samples was randomly selected while ensuring representation from different geographic locations.

Reference sequence sample acquisition

The complete reference genome sequence of *P. falciparum* was obtained from the National Center for Biotechnology Information under accession number GCF_000002765.3.[24] The reference genome was provided in FASTA format. Genomic information from chromosomes 5, 7, and 13 was specifically used, as these regions contain the loci of the *PfMDR1*, *PfCRT*, and *PfKelch13* genes, respectively.

Sequence alignment

The acquired sample sequences were aligned against the reference genome to identify variations indicative of genomic mutations. The reference genome sizes for chromosomes 5, 7, and 13 were 1.3 Mbp, 1.4 Mbp, and 2.9 Mbp, respectively. Sequence alignment was performed using the *nucmer* tool from the MUMmer alignment suite, which is specifically designed to align nucleotide sequences between two genomes.[25] The resulting alignments were then sorted and indexed using SAMTools sort, which organizes the detected differences according to their genomic coordinates. The final output of this step was stored in Binary Alignment Map (BAM) format to enhance storage efficiency and facilitate subsequent analyses.

BAM and Variant Call Format (VCF) processing

The sorted BAM output from SAMTools sort was processed using FreeBayes to identify mutations and generate a VCF file for downstream analysis. The detected variants included insertions and deletions (indels), SNPs, multiple nucleotide polymorphisms, and other complex combinations of mutations. Because not all detected variants are of high quality, additional filtering criteria were applied based on Phred quality scores and read depth to ensure accuracy. Variant quality control was performed using SnpSift, which filters variants according to these parameters.[26] To isolate SNPs specifically, bcftools view was employed. The final output of this step was a curated VCF file containing only high-confidence SNPs.

SNP annotation

SnpEff was used to annotate the identified SNPs. SnpEff is a prediction tool that annotates detected mutations based on the general feature list of a reference genome.[26] To perform the annotation, a local database was first created by acquiring complete annotation files for *P. falciparum* chromosomes 5, 7, and 13 from PlasmoDB (plasmodb.org). These files, based on the latest Pf3D7 genome release, were compiled using the SnpEff database build utility. Annotation was then performed separately for each of the three chromosomes.

RESULTS AND DISCUSSION

SNP annotation

A random sample of 100 genomic sequences was acquired and verified to ensure representation from diverse geographic regions and countries. These sequences were aligned with the reference sequences for chromosomes 5, 7, and 13, which correspond to *P. falciparum* drug resistance genes. A total of 313, 3,321, and 4,188 variants were initially identified on chromosomes 5, 7, and 13, respectively. Because these variants included low-quality mutations and non-SNP variations, filtering was performed. After applying the quality threshold of a Phred score greater than 30, only 12, 103, and 79 high-confidence SNPs were retained for chromosomes 5, 7, and 13, respectively. These SNPs were subsequently annotated against the *P. falciparum* Pf3D7 version 3 reference genome to determine their genomic locations and potential functional impacts. The *SnpEff* tool generated annotation results that included information such as the number of effects categorized by impact, functional class, mutation type, and genomic region, as well as details on specific base substitutions. The complete annotation dataset, including positional and functional details for each mutation, is available in the appendix and has been archived as a data repository on Zenodo (https://zenodo.org).[27]

Chromosome 5 SNP annotations

The annotation results for chromosome 5 are presented in Tables 1-3. Table 1 summarizes the functional impact of the identified mutations

Table 1. Number of variants by impact of the annotated mutations in chromosome 5

Type	Count
Moderate	1
Modifier	11

Table 2. Number of variants by type of variance and where they occur in chromosome 5

Type		Region	
Type	Count	Type	Count
5' UTR variant	1	Exon	1
Intergenic Region	7	Intergenic	7
Intron Variant	3	Intron	3
Missense Variant	1	5' UTR	1

Table 3. Base changes count for each SNPs with the row being the original reference bases, meanwhile the column representing what base they changed into for each SNP in chromosome 5

	A	C	G	T
A	0	0	2	3
C	0	0	0	1
G	1	0	0	0
T	3	2	0	0

Table 4. Number of variants by impact of the annotated mutations in chromosome 7

Type	Count
Low	12
Moderate	19
Modifier	74

and their frequency of occurrence. Table 2 details the genomic regions where the SNPs were detected and their variant types, whereas Table 3 lists the observed base substitutions and their frequencies.

According to Table 1, a total of 12 SNPs were identified and annotated on chromosome 5, the majority of which were classified as modifiers. Only one SNP exhibited a moderate impact, whereas the remaining 11 were predicted to have modifier effects on gene translation within this chromosome.

As shown in Table 2, the single SNP with a moderate effect was identified as a missense mutation. Most of the remaining SNPs were located in intergenic regions, with only one occurring within an exon.

The base substitutions for these SNPs are summarized in Table 3. The most frequent nucleotide changes were adenine to thymine (A→T) and thymine to adenine (T→A), together accounting for six of the 12 SNPs. Other substitutions included adenine to guanine (A→G), thymine to cytosine (T→C), guanine to adenine (G→A), and cytosine to thymine (C→T). A single missense mutation occurred, resulting in a codon change from GAA to AAA. This substitution led to an amino acid change from glutamic acid to lysine.
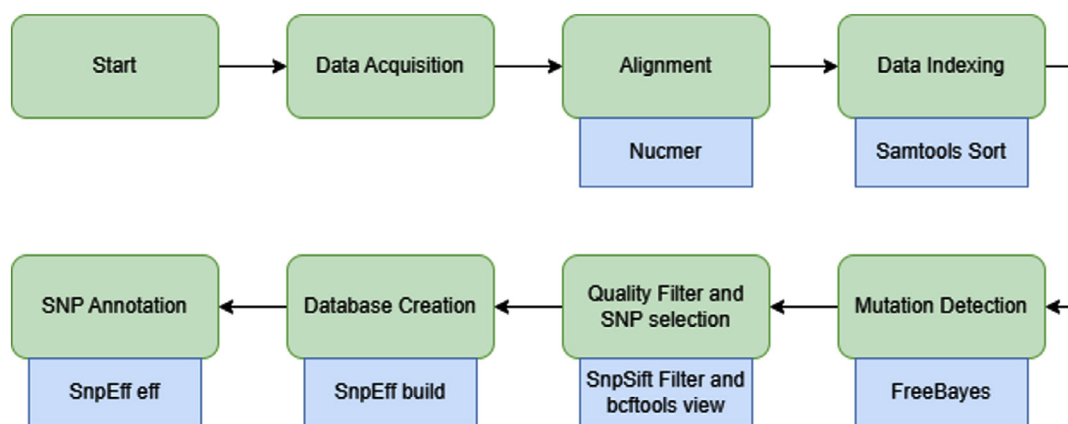

Figure. Schematic representation of the analysis pipeline. Green shapes indicate the main workflow steps, whereas blue shapes denote the specific tools used within the Galaxy platform

Table 5. Number of variants by type of variants and where they occur in chromosome 7

Type		Region	
Type	Count	Type	Count
Intergenic Region	59	Exon	33
Intron Variant	11	Intergenic	59
Missense Variant	19	Intron	11
Non-coding transcript exon variant	2	Transcript	2
Non-coding transcript variant	2		
Synonymous Variant	12		

Table 6. Base changes count for each SNPs with the row being the original reference bases, meanwhile the column representing what base they changed into for each SNP in chromosome 7

	A	C	G	T
A	0	10	15	9
C	2	0	2	10
G	15	2	0	6
T	13	17	2	0

Table 7. Number of variants by impact of the annotated mutations in chromosome 13

Type	Count
Low	4
Moderate	10
Modifier	65

Table 8. Number of variants by type of variance and where they occur in chromosome 13

Type		Region	
Type	Count	Type	Count
Intergenic Region	61	Exon	14
Intron Variant	4	Intergenic	61
Missense Variant	10	Intron	4
Splice region variant	1		
Synonymous Variant	4		

Table 9. Base changes count for each SNPs with the row being the original reference bases, meanwhile the column representing what base they changed into for each SNP in chromosome 13

	A	C	G	T
A	0	4	19	11
C	2	0	2	13
G	7	1	0	6
T	8	12	0	0

Chromosome 7 SNP annotations

The annotation results for chromosome 7 are presented in Tables 4-6. Table 4 summarizes the functional impact of the identified mutations and their frequency of occurrence. Table 5 details the genomic regions where the SNPs were detected and their variant types, whereas Table 6 lists the observed base substitutions and their frequencies.

According to Table 4, a total of 105 SNPs were identified and annotated on chromosome 7, the majority of which were classified as modifiers. Nineteen SNPs exhibited a moderate impact, whereas 74 were predicted to have modifier effects on gene translation within this chromosome. Moreover, 12 SNPs were categorized as having a low impact.

As show in Table 5, most SNPs were located in intergenic regions. Nineteen of the identified SNPs resulted in missense mutations, whereas 12 were synonymous. The remaining variants were distributed across intronic and transcript regions of the chromosome.

The base substitutions for these SNPs are summarized in Table 6. The most frequent nucleotide change was thymine to cytosine (T→C), observed 17 times. Both transition and transversion mutations were detected across all nucleotide bases. Codon changes (Appendix 1) associated with these SNPs resulted in multiple amino acid substitutions, most commonly asparagine to aspartic acid (three occurrences), and lysine to asparagine and glutamine substitutions, each occurring twice (Appendix 2). In total, 19 amino acid residues were altered owing to missense SNPs, while 12 synonymous mutations were also identified (Appendix 3).

Chromosome 13 SNP annotations

The annotation results for chromosome 13 are presented in Tables 7-9. Table 7 summarizes the functional impact of the identified mutations and their frequency of occurrence. Table 8

details the genomic regions where the SNPs were detected and their variant types, whereas Table 9 lists the observed base substitutions and their frequencies.

According to Table 7, a total of 79 SNPs were identified and annotated on chromosome 13, the majority of which were classified as modifiers. Ten SNPs exhibited a moderate impact, whereas 65 were predicted to have modifier effects on gene translation within this chromosome. Moreover, four SNPs were categorized as having a low impact.

As shown in Table 8, most SNPs were located in intergenic regions. Ten of the identified SNPs resulted in missense mutations, whereas four were synonymous. Additional variants were observed within intronic regions of the chromosome.

The base substitutions for these SNPs are summarized in Table 9. The most frequent nucleotide change was adenine to guanine (A→G), followed by cytosine to thymine (C→T) (Appendix 2). In terms of amino acid substitutions, nine residues were altered owing to missense SNPs, while four synonymous mutations were also identified (Appendix 4).

In total, 100 sequence samples were analyzed for existing SNPs, with variant quality determined using the Phred quality score (Q value). A Q value of 10 corresponds to a 1-in-10 probability of an incorrect base call, whereas a Q value of 20 indicates a 1-in-100 probability.[28] Under the parameters used in this study, the expected error rate was approximately 1 in 1,000. SNPs were annotated to assess their potential impact on amino acid translation, transcription, and gene expression.[29] Nucleotide substitutions can be categorized as either transversions or transitions.[30] Transversions involve the exchange of a purine (adenine or guanine) for a pyrimidine (cytosine or thymine) or vice versa, whereas transitions occur where a purine replaces another purine, or a pyrimidine replaces another pyrimidine.[31] Depending on their genomic context, these mutations may alter gene expression when occurring in regulatory regions or modify protein structure and function by changing the encoded amino acids.[32]

SNPs in chromosome 5 and their implications for *P. falciparum*

Analysis of the sample revealed 12 SNPs on chromosome 5, with only one located within an exon and affecting translation and transcription. The remaining 11 SNPs showed modifier effects, as expected since they occur in intergenic regions.[33] Modifier mutations can influence transcriptional regulation, potentially leading to overexpression or underexpression of specific genes.[34] Known mutations within the *PfMDR1* gene include N86Y, Y184F, and D1246Y;[20] these changes are known to alter the absorption and activity of artemisinin-based combination therapies in *P. falciparum*.[35] The *PfMDR1* gene is located between positions 957,890 and 962,149 on chromosome 5.[36] The single annotated exon SNP, found at position 172,801 (Appendix 5), lies outside this region, within an ATP-dependent helicase gene (*PF3D7_1202000*), suggesting a potential regulatory effect on transcription.[37] The absence of SNPs directly within the *PfMDR1* locus indicates that, within this dataset of 100 samples, no high-confidence mutations met the stringent filtering criteria. However, several modifier SNPs were detected adjacent to the *PfMDR1* gene (Appendix 5), demonstrating the capacity of the pipeline to accurately detect SNPs from genomic sequence input.[38]

SNPs in chromosome 7 and their Implications for *P. falciparum*

A total of 103 SNPs were identified and annotated on chromosome 7. Most were classified as modifiers, although several low- and moderate-impact SNPs were also detected. Low-impact SNPs are typically synonymous mutations, where base substitutions do not alter the amino acid sequence,[39] yet they may still affect gene regulation or downstream expression.[40] In contrast, moderate-impact SNPs cause amino acid substitutions that can alter protein structure or function.[41] Most SNPs were located in intergenic regions, suggesting possible regulatory effects on transcription and translation.[33] In total, 33 SNPs were detected within exons, 19 of which led to amino acid substitutions. The drug resistance gene located on chromosome 7 is *PfCRT*, which

encodes the protein responsible for exporting chloroquine from the digestive vacuole of the parasite.[42] Canonical *PfCRT* mutations—K76T, C101F, and L272F—occur within positions 403,222 to 406,317.[20] The annotated SNPs in this study did not correspond to those specific sites (Appendix 6), suggesting that no known chloroquine resistance mutations were present. Nevertheless, other detected SNPs may still influence parasite survival. For instance, several were located in genes involved in erythrocyte adhesion and immune evasion, which can reduce parasite clearance rates.[43] Furthermore, the abundance of intergenic SNPs could potentially increase *PfCRT* expression, enhancing chloroquine efflux and promoting resistance under drug pressure.[44]

SNPs in chromosome 13 and their implications for *P. falciparum*

Seventy-nine SNPs were detected and annotated on chromosome 13, the majority of which were located in intergenic regions. These variants are likely to exert modifier effects on gene translation and expression.[33] Synonymous mutations were also identified, alongside 10 SNPs that resulted in amino acid substitutions. Drug resistance associated with chromosome 13 is primarily linked to mutations in the *PfKelch13* gene, with common variants including N458Y, Y493H, and R539T.[20] The *PfKelch13* gene is located between positions 172,4817 and 172,6997, whereas the annotated SNPs in this study were found near—but not directly within—this region (Appendix 7).[45] Although none of the detected SNPs corresponded precisely to known resistance sites, their proximity suggests potential roles in regulatory modification or emerging variants that warrant further investigation.

CONCLUSION

P. falciparum remains a persistent challenge in malaria treatment owing to the emergence and spread of drug resistance. These resistances arise from genetic mutations driven by multiple factors, highlighting the importance of detecting and understanding their effects. Most resistance in *P. falciparum* has been linked to SNPs, which formed the focus of this study.

The developed pipeline successfully detected and annotated SNPs within chromosomes 5, 7, and 13, which harbor the *PfMDR1*, *PfCRT*, and *PfKelch13* genes—key determinants of chloroquine and artemisinin-based combination therapy resistance. Although no SNPs directly within these genes were identified in the sample dataset, the results demonstrate that the pipeline can reliably identify and annotate SNPs when provided with genomic sequence data. Further refinement and validation of the pipeline could be achieved by incorporating a larger dataset and using targeted amplification of resistance-associated genes through polymerase chain reaction to enhance accuracy and detection sensitivity. Overall, this research establishes a practical framework for the detection and analysis of SNPs in *P. falciparum*, contributing to ongoing efforts to monitor and address antimalarial drug resistance.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORS' CONTRIBUTION

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

DATA AVAILABILITY

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

The pipeline can be accessed through the galaxy server: https://usegalaxy.org/published/workflow?id=8b4b62a55a93a258

ETHICS STATEMENT
Not applicable.

REFERENCES

1. Chaidir L, Sengstake S, de Beer J, et al. *Mycobacterium tuberculosis* genotypic drug resistance patterns and clustering in Jayapura, Papua, Indonesia. *Int J Tuberc Lung Dis.* 2015;19(4):478-484. doi: 10.5588/ijtld.14.0350

2. Vieira G, Leal N, Rodrigues A, Chaves C, Rodrigues F, Osorio N. MRSA MSSA causing infections: prevalence of mecA gene. *Eur J Public Health.* 2020;30(Suppl_2):ckaa040.052. doi: 10.1093/eurpub/ckaa040.052

3. Lai Y, Wei W, Du Y, Gao J, Li Z. Biomaterials for *Helicobacter pylori* therapy: therapeutic potential and future perspectives. *Gut Microbes.* 2022;14(1):2120747. doi: 10.1080/19490976.2022.2120747

4. Saha K, Sarkar D, Khan U, et al. Capsaicin inhibits inflammation and gastric damage during *H. pylori* infection by targeting NF-κB–miRNA axis. *Pathogens.* 2022;11(6):641. doi: 10.3390/pathogens11060641

5. Salam MA, Yusuf AAM, Pawar JS, Akhter N, Lucy IB. Conventional methods and future trends in antimicrobial susceptibility testing. *Saudi J Biol Sci.* 2023;30(3):103582. doi: 10.1016/j.sjbs.2023.103582

6. Chaisirin W, Wongkrajang P, Praphruetkit N, et al. Role of point of care testing in reducing time to treatment decision making in urgency patients: a randomized controlled trial. *West J Emerg Med.* 2020;21(2):386-393. doi: 10.5811/westjem.2019.10.43655

7. Harikumar G, Krishanan K. The growing menace of drug resistant pathogens and recent strategies to overcome drug resistance: a review. *J King Saud Univ Sci.* 2022;34(2):101979. doi: 10.1016/j.jksus.2022.101979

8. Timms VJ, Hassan KA, Mitchell HM, Neilan BA. Comparative genomics between human and animal associated subspecies of the *Mycobacterium avium* complex: a basis for pathogenicity. *BMC Genomics.* 2015;16(1):695. doi: 10.1186/s12864-015-1889-2

9. Hill AW, Guralnick RP, Wilson MJC, Habib F, Janies D. Evolution of drug resistance in multiple distinct lineages of H5N1 avian influenza. *Infect Genet Evol.* 2009;9(2):169-178. doi: 10.1016/j.meegid.2008.10.006

10. Tang L, Zhu SL, Fang X, et al. Distinct evolutionary origins of common multi drug resistance phenotypes in *Salmonella typhimurium* DT104: a convergent process for adaptation under stress. *Mol Genet Genomics.* 2019;294(3):785-796. doi: 10.1007/s00438-019-01531-5

11. Emerson JJ, Li WH. The genetic basis of evolutionary change in gene expression levels. *Philos Trans R Soc B Biol Sci.* 2010;365(1552):2581 2590. doi: 10.1098/rstb.2010.0005

12. Fojo T. Multiple paths to a drug resistance phenotype: mutations, translocations, deletions and amplification of coding genes or promoter regions, epigenetic changes and microRNAs. *Drug Resist Updat.* 2007;10(12):59 67. doi: 10.1016/j.drup.2007.02.002

13. Anderson MW, Schrijver I. Next generation DNA sequencing and the future of genomic medicine. *Genes.* 2010;1(1):38-69. doi: 10.3390/genes1010038

14. Parikesit AA, Hermantara R, Kevin G, Sidhartha E. Designing hybrid CRISPR Cas12 and LAMP detection systems for treatment resistant *Plasmodium falciparum* with in silico method. *Narra J.* 2023;3(3):e301. doi: 10.52225/narra.v3i3.301

15. Parikesit AA, Hermantara R, Sidhartha E, et al. Optimizing CRISPR Cas12 based detection methods for drug resistant *Plasmodium falciparum* with an advanced gene visualization pipeline. *J Res Pharm.* 2025;29(3). doi: 10.12991/jrespharm.1694230

16. Ansori ANM, Antonius Y, Susilo RJ, et al. Application of CRISPR Cas9 genome editing technology in various fields: a review. *Narra J.* 2023;3(2):e184. doi: 10.52225/narra.v3i2.184

17. World Health Organization. *Antimicrobial Resistance: Global Report on Surveillance.* Geneva, Switzerland: World Health Organization; 2014. https://www.who.int/publications/i/item/9789241564748

18. World Health Organization. World Malaria Report 2024. Geneva, Switzerland: World Health Organization; 2024. https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2024

19. Weiss DJ, Dzianach PA, Saddler A, et al. Mapping the global prevalence, incidence, and mortality of *Plasmodium falciparum* and *Plasmodium vivax* malaria, 2000–22: a spatial and temporal modelling study. *Lancet.* 2025;405(10483):1180-1196. doi: 10.1016/S0140-6736(25)00038-8

20. She D, Wang Z, Liang Q, et al. Polymorphisms of *pfcrt, pfmdr1*, and *K13* propeller genes in imported falciparum malaria isolates from Africa in Guizhou province, China. *BMC Infect Dis.* 2020;20(1):741. doi: 10.1186/s12879-020-05228-8

21. Abueg AL, Afgan E, Allart O, et al. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic Acids Res.* 2024;52(W1):w83-w94. doi: 10.1093/nar/gkae410

22. Fugaha DR. SNP Analysis. Galaxy; 2025. Accessed 2025. https: usegalaxy.org/ published workflow?id=8b4b62a55a93a258

23. Jacob CG, Thuy Nhien N, Mayxay M, et al. Genetic surveillance in the Greater Mekong subregion and South Asia to support malaria control and elimination. *eLife.* 2021;10:e62997. doi: 10.7554/eLife.62997

24. Hernandez SI, Berezin CT, Miller KM, Peccoud SJ, Peccoud J. Sequencing strategy to ensure accurate plasmid assembly. *ACS Synth Biol.* 2024;13(12):4099-4109. doi: 10.1021/acssynbio.4c00539

25. Marcais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. MUMmer4: A fast and versatile genome alignment system. *PLoS Comput Biol.* 2018;14(1):e1005944. doi: 10.1371/journal.pcbi.1005944

26. Cingolani P, Platts A, Wang LL, et al. A program

for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin).* 2012;6(2):80-92. doi: 10.4161/fly.19695

27. Fugaha DR. Appendix of *Plasmodium falciparum* drug resistance single nucleotide polymorphism analysis pipeline development based on Southeast Asian genetic barcode samples. 2025. doi: 10.5281/zenodo.15703757

28. Luo J, Meng Z, Xu X, et al. Systematic benchmarking of nanopore Q20+ kit in SARS CoV 2 whole genome sequencing. *Front Microbiol.* 2022;13:973367. doi: 10.3389/fmicb.2022.973367

29. Robert F, Pelletier J. Exploring the impact of single nucleotide polymorphisms on translation. *Front Genet.* 2018;9:507. doi: 10.3389/fgene.2018.00507

30. Stoltzfus A, Norris RW. On the causes of evolutionary transition: transversion bias. *Mol Biol Evol.* 2016;33(3):595-602. doi: 10.1093/molbev/msv274

31. Guo C, McDowell IC, Nodzenski M, et al. Transversions have larger regulatory effects than transitions. *BMC Genomics.* 2017;18(1):394. doi: 10.1186/s12864-017-3785-4

32. Rojano E, Seoane P, Juan AGR, Perkins JR. Regulatory variants: from detection to predicting impact. *Brief Bioinform.* 2019;20(5):1639-1654. doi: 10.1093/bib/bby039

33. Chen J, Tian W. Explaining the disease phenotype of intergenic SNP through predicted long range regulation. *Nucleic Acids Res.* 2016;44(18):8641-8654. doi: 10.1093/nar/gkw519

34. Campino S, Marin Menendez A, Kemp A, et al. A forward genetic screen reveals a primary role for *Plasmodium falciparum* reticulocyte binding protein homologue 2a and 2b in determining alternative erythrocyte invasion pathways. *PLoS Pathog.* 2018;14(11):e1007436. doi: 10.1371/journal.ppat.1007436

35. Okell LC, Reiter LM, Ebbe LS, et al. Emerging implications of policies on malaria treatment: genetic changes in the *Pfmdr-1* gene affecting susceptibility to artemether lumefantrine and artesunate amodiaquine in Africa. *BMJ Glob Health.* 2018;3(5):e000999. doi: 10.1136/bmjgh-2018-000999

36. Pulcini S, Staines HM, Lee AH, et al. Mutations in the *Plasmodium falciparum* chloroquine resistance transporter, PfCRT, enlarge the parasite's food vacuole and alter drug sensitivities. *Sci Rep.* 2015;5:14552. doi: 10.1038/srep14552

37. Chauhan M, Tuteja R. *Plasmodium falciparum* specific helicase 2 is a dual, bipolar helicase and is crucial for parasite growth. *Sci Rep.* 2019;9(1):1048. doi: 10.1038/s41598-018-38032-1

38. Pandit K, Surolia N, Bhattacharjee S, Karmodiya K. The many paths to artemisinin resistance in *Plasmodium falciparum*. *Trends Parasitol.* 2023;39(12):1117 1130. doi: 10.1016/j.pt.2023.09.011

39. Zhao Z, Fu YX, Hewett Emmett D, Boerwinkle E. Investigating single nucleotide polymorphism (SNP) density in the human genome and its implications for molecular evolution. *Gene.* 2003;312:207-213. doi: 10.1016/S0378-1119(03)00670-X

40. Crawford ED, Quan J, Horst JA, Ebert D, Wu W, DeRisi JL. Plasmid free CRISPR Cas9 genome editing in *Plasmodium falciparum* confirms mutations conferring resistance to the dihydroisoquinolone clinical candidate SJ733. *PLoS One.* 2017;12(5):e0178163. doi: 10.1371/journal.pone.0178163

41. Petrosino M, Novak L, Pasquo A, et al. Analysis and interpretation of the impact of missense variants in cancer. *Int J Mol Sci.* 2021;22(11):5416. doi: 10.3390/ijms22115416

42. Ocan M, Ashaba FK, Mwesigwa S, Edgar K, Kamya MR, Nsobya SL. Prevalence of *arps10, fd, pfmdr 2, pfcrt* and *pfkelch13* gene mutations in *Plasmodium falciparum* parasite population in Uganda. *PLoS One.* 2022;17(5):e0268095. doi: 10.1371/journal.pone.0268095

43. Goel S, Palmkvist M, Moll K, et al. RIFINs are adhesins implicated in severe *Plasmodium falciparum* malaria. *Nat Med.* 2015;21(4):314-317. doi: 10.1038/nm.3812

44. Njokah MJ, Kang'ethe JN, Kinyua J, Kariuki D, Kimani FT. In vitro selection of *Plasmodium falciparum Pfcrt* and *Pfmdr1* variants by artemisinin. *Malar J.* 2016;15(1):381. doi: 10.1186/s12936-016-1443-y

45. Matrevi SA, Tandoh KZ, Bruku S, et al. Novel *pfk13* polymorphisms in *Plasmodium falciparum* population in Ghana. *Sci Rep.* 2022;12(1):8312. doi: 10.1038/s41598-022-11790-9