ISSN: 2690- 912X

Research Article

Journal of Genetic Engineering and Biotechnology

Analysis of Antimalarial Drug Resistance Genes of *Plasmodium Falciparum* 3d7 to Understand Its Expression: A Bioinformatics Approach

Gemechis Waktole[1*], Donghee Cho[2]

[1]*Dambi Dollo University, Department of Biotechnology, College of Natural and Computational Sciences P.O.Box 260, Dambi Dollo, Ethiopia*

[2]*Adama Science and Technology University, Department of Applied Biology, School of Applied Natural Science, P.O.Box 1888, Adama, Ethiopia*

[*]Corresponding author
Gemechis Waktole, 1Dambi Dollo University, Department of Biotechnology, College of Natural and Computational Sciences P.O.Box 260, Dambi Dollo, Ethiopia

Submitted: 26 Jul 2022; **Accepted:** 01 Aug 2022; **Published:** 27 Sep 2022

Citation: *Waktole G, Cho D. (2022). Analysis of Antimalarial Drug Resistance Genes of Plasmodium Falciparum 3d7 to Understand Its Expression: A Bioinformatics Approach. J Gene Engg Bio Res, 4(3),322-327.*

Abstract

Background

Malaria, which poses a threat to half of the world's population, is one of the most serious infectious diseases. Ethiopia is a nation with a significant malaria burden. In Ethiopia, Plasmodium falciparum accounts for 64% of cases of malaria, with P. vivax causing the remaining instances (34 percent). The disease still claims the lives of countless children globally, mostly in sub-Saharan African nations, and malaria continues to be a significant public health issue in Ethiopia despite various advancements in malaria control measures. The potential for analyzing parasite genetics to support both national and international efforts to eradicate parasites is immense. To analyze regulatory components such as CpG islands, transcription factors (TFs), and their corresponding binding sites (TFBSs) involved in the control of gene expression of Plasmodium falciparum 3D7 isolate drug resistance genes.

Results

Nine drug resistance-related gene-coding sequences from the NCBI database were examined for this analysis. Only functional genes (protein-coding genes) were taken into consideration. Accordingly, genes affected by Plasmodium falciparum 3D7 drug resistance had 1-6 TSS, and five common candidate motifs (MPfI, MPfII, MPfIII, MPfIV, and MPfV) were found in the promoter prediction by neural network promoter prediction results. According to the study, CpG islands are poorly distributed in both the promoter and gene body regions, which may interfere with the accessibility of the promoter to transcription factors and, ultimately, the production of the genes.

Conclusion

This in silico analysis of genes encoding Plasmodium falciparum drug resistance-related genes may be useful for enhancing knowledge of the molecular data and supporting the identification of gene regulatory elements in the promoter regions.

Keywords: Drug resistance, CpG islands, Plasmodium falciparum 3D7, Promoter region

Background

Malaria is one of the most important infectious diseases that threatens half of the world's population. The World Malaria Report in 2019 estimated that there were 228 million cases of malaria in 2018 that caused deaths [18]. With 272,000 deaths, children less than 5 were the most vulnerable group worldwide. In Africa, 213 million people were affected by malaria, which made it the most vulnerable continent in 2018 [18]. Worldwide, there was a significant reduction in malaria cases during 2015–2017, but there is a major challenge to thoroughly eliminate malaria in many countries by 2030 [17].

Ethiopia is among countries with a high malaria burden. Plasmodium falciparum is the most common (64%) cause of malaria in Ethiopia, while *P. vivax* accounts for the remaining cases (34%) [19]. *Plasmodium falciparum* causes the most severe form of malaria; however, contrary to popular belief, *P. vivax* can also cause severe malaria and even death. Malaria morbidity and mortality have significantly decreased in Ethiopia and worldwide in the past decade [18]. Ethiopia's fight against malaria started many years ago, and transmission of this infectious disease has significantly decreased since 1959 [8].

Despite numerous advances in malaria control strategies, the disease still kills countless children worldwide, mainly in sub-Saharan African countries, and malaria remains a major public health problem in Ethiopia [18]. The World Health Organization (WHO) reported no significant progress in reducing the global

malaria burden during the period from 2015 to 2017[18].

The red efficacy of chloroquine (CQ) has forced a change in the selection of anti-malaria in the management of falciparum malaria. Since 2004, Ethiopia has adopted artemether-lumefantrine (AL) and CQ as first-line treatments for infection with P. falciparum and *P. vivax*, respectively. In cases of treatment failure of *P. falciparum*, quinine (QN) is the treatment of choice, and in cases of severe malaria, artemether, artesunate, or QN can be used [5]. Globally, artemisinin resistance in P. falciparum has emerged, especially in Southeast Asia, slowing therapeutic response and increasing rates of treatment failures [15]. Similarly, artemisinin resistance has been reported in Africa, although there is currently no evidence that it has taken hold [10].

Analyzing parasite genetics has enormous potential to aid both efforts at elimination and international efforts toward eradication [1]. For example, the use of molecular markers of antimalarial drug resistance has for some time been used to guide efficacy studies that define treatment policies; more recently, the detection of parasite mutants for the histidine-rich protein 2/3 (pfhrp2/pfhrp3) gene highlighted the need to develop rapid diagnostic tests (RDTs) with alternative targets of detection and to provide countries with guidance on the implications for case management [17]. Therefore, this *in silico* analysis aims to identify regulatory elements such as CpG islands, transcription factors (TFs), and their corresponding binding sites (TFBSs) involved in the regulation of gene expression to provide baseline information to revise existing eradication and elimination strategies and for designing target-specific drugs.

Methods

Determination of Transcription Start Sites and Promoter Regions

Plasmodium falciparum drug resistance affecting genes (*Pfcrt* (CQ monotherapy, AS-PY, and Amodiaquine), *Pfmdr1* (CQ *monotherapy, Melfoquine, Lumefantrine, Amodiaquine* and limited efficacy of ACT), *Pfdhfr* (ATQ-PGL, Sulfadoxine-Phyrimethamine), *Pfdhps* (Sulfadoxine-Phyrimethamine), *Pfarps10, Pferredoxin (Ferroquine), Pfexonuclease,* PfKpk-13 (AM-LF, AS monotherapy, AS-SP, AS-MQ, AS-PY) and Pfmdr2 (DHA-PPQ) genes.) were obtained from the NCBI genome data bank (https://www.ncbi.nlm.nih.gov). In this analysis, all drug resistance affecting gene-coding sequences, available in the NCBI database with the start codon at the beginning of the sequence, and only functional genes (protein coding) were deliberated. To determine their respective transcription, start sites (TSSs) and 1 kb sequences upstream of the start codon were excised from each gene [9]. All the TSSs of each of the selected Plasmodium falciparum 3D7 drug resistance-affecting genes were searched within this region by using Neural Network Promoter Prediction (NNPP) version 2.2 (https://www.fruitfly.org).

Tool set with the minimum standard predictive score (between 0 and 1) cutoff value of 0.8 [11]. This tool helps to locate the possible TSSs within the sequences upstream of the start codon where the RNA polymerases start their activity and transcription process. The NNPP tool can precisely recognize the position of a TSS for a given gene. For those regions containing more than one TSS, the one with the highest prediction score was considered to have a trustable and accurate prediction. Therefore, as previously done for SARS-CoV-2 gene promoter region determination, P. falciparum drug resistance affecting gene promoter sequences was defined as a 1 kb region upstream of each TSS [7].

Identification of Common Candidate Motifs and Transcription Factors

To find common candidate motifs that serve as binding sites for transcription factors that control the expression of Plasmodium falciparum drug resistance-affecting genes, promoter sequences from Plasmodium falciparum drug resistance-affecting genes that were identified based on the aforementioned criteria were analyzed using the MEME (Multiple Em for Motif Elicitation) version 5.3.3 searches via the web server hosted by the National Biomedical Computation Resource [2]. MEME searches for statistically significant candidate motifs in the input sequence set. The MEME output is in the form of XML, text, MAST HTML, MAST XML, MAST text, and HTML and shows the candidate motifs as local multiple alignments of the input promoter sequences. Briefly, MEME discovers novel, un-gapped motifs (recurring, fixed-length patterns) in sequences submitted in it. A motif is an approximate sequence pattern that occurs repeatedly in a group of related sequences. MEME represents motifs as position-dependent letter-probability matrices that describe the probability of each possible letter at each position in the pattern. MEME takes as input a group of sequences and outputs as many motifs as requested. MEME uses statistical modeling techniques to automatically choose the best width, number of occurrences, and description for each motif [2]. Buttons on the MEME HTML output allow one or all of the candidate motifs to be forwarded for further analysis to better characterize the identified candidate motifs by other web-based programs. In this case, the TOMTOM [6] web server was used to search for sequences matching the identified motif for its respective TF. The output of TOMTOM includes LOGOS representing the alignment of the candidate motif and TF with the p value and q-value (a measure of false discovery rate) of the match and links back to the parent transcription database for more detailed information about it [3, 4]

Search for CpG islands

Takai and Jones' rigorous search parameters, which include length 500 bp, ObsCpG/ExpCpG 0.65, and GC content 55%, will be applied [13]. The CpG island searcher program (CpGi130), which may be available at the website http://dbcat.cgm.ntu.edu.tw, was used to search for specific areas that were CpG-rich. Second, the MspI cutting sites of the restriction enzyme CLC bio's Aarhus, Denmark, were searched using the CLC Genomics Workbench (fragment sizes between 40 and 220 bps). Searching for MspI cutting sites is relevant for the identification of CGIs because CpG islands are isolated from short fragments after MspI digestion, which identifies CCGG sites. Studies using whole-genome CpG island libraries created for various species revealed that CpG islands are not randomly distributed but are clustered in particular regions.

Results

Identification of Transcription Start Sites (TSSs)

The nine (9) Plasmodium falciparum 3D7 drug resistance gene transcription start sites (TSSs) are listed below. The outcome of the promoter prediction by the neural network showed that genes affected by Plasmodium falciparum 3D7 drug resistance contain 1-6 TSS. As a result, 4/9 of the detected TSSs were less than -500 compared to the ATG start codon, or 44.44 percent of them. The sequences that were less than -500 distances from the start codon had prediction scores of 0.90, 0.93, 0.97, and 0.98. With projected score values of 0.84, 0.95, 0.96, 0.97, and 0.98, the remaining 55.55 percent of discovered TSSs were larger than -500. In Table 1, the relative positions of each TSS about the start codon are listed. The TSS for Exonuclease (-43) was the closest, followed by Pfdhfr (-186), while the outlying TSS was observed for PfArps-10 (-3381) upstream of the start codons of their respective genes.

Table 1: Number and Predictive Score Value For *Plasmodium Falciparum* Drug Resistance Affecting TSS Genes

Name/Gene ID	Corresponding promoter region name	Number of TSS identified	Predictive score at cutoff value of 0.8	Location of the best TSS from start codon
Pfcrt/2655199	*Pfcrt*	3	0.82,0.80,0.80,0.90	-203
Pfmdr1/813045	*Pfmdr1*	1	0.98	-760
Pfdhfr/9221804	*Pfdhfr*	2	0.84,0.93	-186
Pfdhps/2655294	*Pfdhps*	3	0.81,0.98,0.84	-331
Pfarps10/812163	*Pfarps10*	3	0.86,0.96,0.85	-3381
Pferredoxin/3885862	*Pferredoxin*	6	0.90, 0.95, 0.84, 0.97, 0.86, 0.90	-641
Pfexonuclease/811867	*Pfexonuclease*	5	0.97, 0.89, 0.92, 0.82, 0.92	-43
Pfmdr2/812037	*Pfmdr2*	4	0.92, 0.91, 0.83, 0.95	-517
PfKp-13/814205	*PfKp-13*	2	0.84,0.84	-994

Common Candidate Motifs and Associated Transcription Factors in the Promoter Regions of *Plasmodium Falciparum* Drug Resistance Affecting Genes

The goal of the in-silico research was to find the top candidate motifs for each of the nine genes regulating Plasmodium falciparum 3D7 drug resistance. As a result, the five frequently occurring candidate motifs MPfI, MPfII, MPfIII, MPfIV, and MPfV were found. Only 44.44 percent of the promoter sequences of the genes regulating Plasmodium falciparum 3D7 drug resistance were shared by the five potential motifs that were found. Every gene that is affected by Plasmodium falciparum 3D7 drug resistance, as shown in Table 2, however, shares the remaining four potential motifs. For Plasmodium falciparum 3D7 drug resistance-affected genes, the analysis was conducted using minimum and maximum motif widths of 6 and 50 residues, respectively. The maximum and minimum numbers of motifs were 45 and 11, respectively, which were utilized to identify likely promoter regulatory elements (motifs). Motifs that were prevalent in the majority of the promoter regions of the Plasmodium falciparum drug resistance-affecting genes were chosen to identify functionally significant motifs. Accordingly, MpfIV has been discovered as the common promoter motif for all (100%) genes that serve as binding sites for transcription factors involved in the expression regulation of these genes with the lowest E-value of 2.2e-002. The sequence logo for MpfIV generated by MEME is presented in Fig. 1 below. The relative positions of candidate motifs in the promoter region relative to TSSs are indicated in Fig. 2 below. The nucleotide positions are indicated at the bottom of the graph from+1 (beginning of TSSs) to the upstream 1000 bp in the promoter region for Plasmodium falciparum 3D7 drug resistance genes.

Table 2: Identified Common Candidate Motifs in *Plasmodium Falciparum* Drug Resistance Affecting Gene Promoter Regions

SN	Name	E-value	Sites	N (%) PCEOMs	Width
	MpfI	9.1e-008	9	9(100)	45
2	MpfII	3.2e-001	9	9(100)	11
3	MpfIII	4.0e-001	9	9(100)	29
4	MpfIV	2.2e-002	9	9(100)	26
5	MpfV	2.1e+001	4	4(44.4)	21

N (%) PCEOMs: Number (%) of promoters containing each one of the motifs

Figure 1: Sequence Logo for the Identified Common Motif Mpfiv Gene for *Plasmodium Falciparum* 3d7 Drug Resistance Affecting Genes

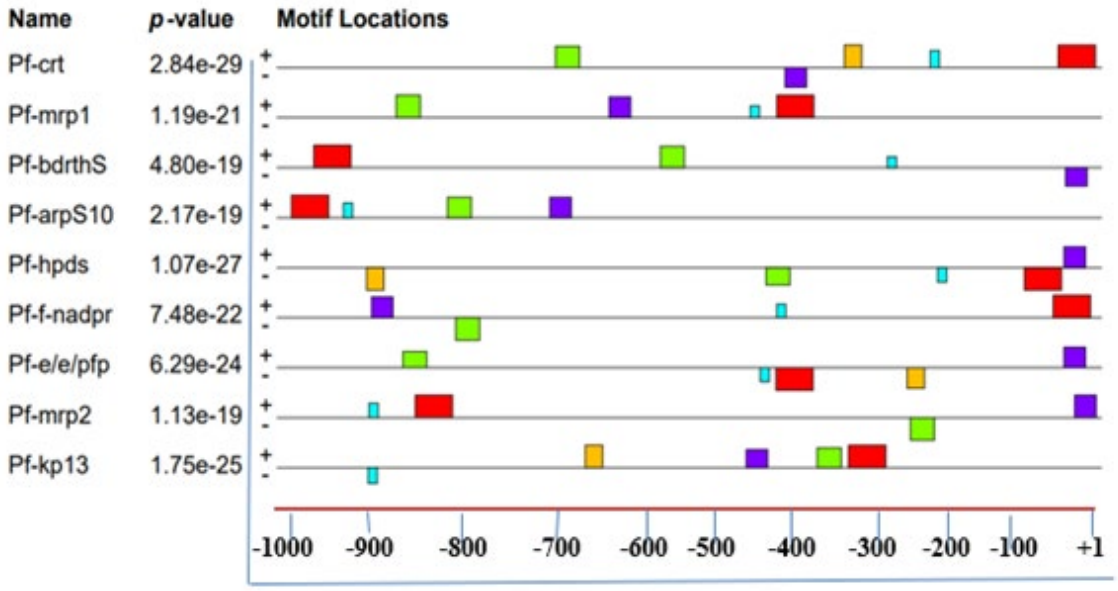

Figure 2: The Relative Positions of Motifs in Different *Plasmodium Falciparum* 3d7 Drug Resistance Affecting Gene Sequences Relative to TSSs

MEME represents motifs as position-dependent letter-probability matrices that describe the probability of each possible letter at each position in the pattern. MEME takes as input a group of sequences and outputs as many motifs as requested. MEME uses statistical modeling techniques to automatically choose the best width, number of occurrences, and description for each motif [2].

TOMTOM is a motif comparison algorithm that ranks the target motifs in a database based on the estimated statistical significance of the query-to-target match. TOMTOM, on the other hand, provides LOGOS, which represents the alignment of two motifs, a numeric score for the match between two motifs, and statistical significance [12]. Only 12 (30%) of the 40 motifs discovered in Plasmodium falciparum 3D7 promoters of drug resistance genes were found on the negative strands, whereas the other 70% were on the positive strand. Other motifs from publicly available databases, such as UniProt (https://www.UniProt.org) and uniPROBE (https://brain.bwh.harvard.edu/uniprobe), were compared to the lowest E-value motif (MpfIV). As a result, the data banks revealed that the pattern MpfIV corresponded to four (4) known transcription factors, as indicated in table 3 below.

Analysis of CpG islands (CGIs) in *Plasmodium falciparum* Drug Resistance Affecting Genes and Promoter Regions

In this study, two algorithms were used to analyze nine Plasmodium falciparum drug resistance genes influencing the promoter and body regions. First, Takai and Jones' method was used to search for CpG islands in both the promoter and gene body regions, but no such islands were located. Similarly, in silico digestion of *Plasmodium falciparum* 3D7 drug resistance influencing genes employing restriction enzyme MspI by CLC genomics workbench 3 software found no CpG islands in either promoter or gene body regions. This finding reveals a low frequency of CpG islands in both the gene body and promoter regions, which might influence gene promoter access to transcription factors and thus gene expression [13].

Table 3: The Lists of Candidate Transcription Factors That Could Bind To Motif *PfMIV*

Candidate transcription factors	Statistical significance			Source	Data Base	Overlap
	p value	*E-value*	*q-value*			
UP00029/TATA-box-binding protein	9.32e-05	1.68e-01	2.07e-01	*Homo sapiens*	UniPROBE	16
FOXB1/FOXB1_DBD_2	8.26e-04	1.49e+0	7.89e-01	*Homo sapiens*	UniProt	18
UP00150_1 (Irx6_2623.2)	2.92e-03	5.28e+0	1.00e+0	*Mus musculus*	UniPROBE	17
UP00094_2 (Zfp128_secondary)	5.40e-03	9.76e+0	1.00e+0	*Mus musculus*	UniPROBE	14

Discussions

Scientists have been attempting to comprehend the molecular underpinnings of drug action and resistance ever since chloroquine resistance first independently appeared in Southeast Asia and South America, later migrating from Asia to Africa. Precisely, it is yet unknown how processes such as the selection of the transcription start site (TSS) or the recruitment of RNA polymerase II (RNA PolII) regulate the commencement of transcription. We used the sequence data obtained from the NCBI database to identify and describe the promoter regions and regulatory elements of the Plasmodium falciparum 3D7 drug resistance affecting genes. There is a slight difference in the number of TSSs among the genes controlling Plasmodium falciparum 3D7 drug resistance, according to the analysis of their promoter regions (1-6 TSSs). The present investigation further demonstrated that TSSs of Plasmodium falciparum 3D7 drug resistance impacting genes were mainly found in the upstream region of -43 to -3381 bp relative to the ATG.

The current inquiry for the Plasmodium falciparum 3D7 drug resistance influencing genes has shown how crucial it is to identify all probable binding patterns for the same TF and co-factor binding motif. The study also discovered many binding sites in the candidate motifs' promoter regions, which may be used to improve binding interactions and have a variety of regulatory effects. By binding to cis-regulatory regions in DNA enhancers and promoters, these sites regulate the expression of genes. The promoter regions of the Plasmodium falciparum 3D7 drug resistance-related genes contain the majority of the recommended motifs, which are located less than 700 bp from the transcription start point. It is commonly accepted that the promoter regions of genes with comparable expression patterns share similarities [21]. In addition, the current study revealed that motif IV, the best common motif, shares characteristics with families of the fork-head box protein B1 transcription.

Although reports have indicated that CpG islands play a significant role in gene regulation mechanisms in this study, analysis of CpG revealed that the promoter and gene body regions of Plasmodium falciparum 3D7 drug resistance genes are devoid of CpG islands by using both algorithms. Contrary to the result, CpG islands are frequently found in the promoters of numerous tissue-specific genes and the majority of housekeeping genes, suggesting that they serve crucial regulatory roles in addition to their potential use as gene markers [22, 23].

The current in silico study used various techniques to examine the promoter and regulatory components of drug resistance genes in Plasmodium falciparum 3D7. We are unsure whether to fully endorse the direct participation of the examined Plasmodium falciparum 3D7 drug resistance genes and epigenetic regulators, although, because of diverse physiological and biological roles as well as the expression of drug resistance genes in tissues. Further in vitro or in vivo research is therefore required to confirm the results and determine the precise epigenetic controls on the drug-resistant Plasmodium falciparum 3D7 genes. Validation is crucial for any computationally based approach, including in silico study approaches. However, as wet molecular laboratories are very expensive and sophisticated, knowledge of bioinformatics approaches could also help to predict gene expression profiles in different infections. This is because the current study requires experimental confirmation for verification.

Conclusions

The result of this analysis could be critically important to understand the nature of promoter regions, the motif discovered in line with the transcription factor binding proteins of Plasmodium falciparum 3D7 drug resistance affecting genes. CpG islands are also regulatory elements in the promoter regions of the genome and are useful in the detection of promoters. However, in this analysis, either algorithm in Plasmodium falciparum 3D7 drug resistance-affecting genes identified no CpG islands. In general, this in silico analysis of genes encoding Plasmodium falciparum 3D7 drug resistance-affecting genes could be helpful to add knowledge about the molecular data and support the identification of gene regulatory elements in the promoter regions. It could also help to predict gene expression profiles in Plasmodium falciparum 3D7, which in turn could be helpful to improve present drug efficacy and to develop new drugs with high target specificity. Therefore, knowledge of bioinformatics methods is important to identify gene regulatory regions in the promoter regions, and gene body regions could help to predict gene expression profiles in various pathogens, since wet molecular laboratories are very expensive and sophisticated.

Funding

This research did not receive any grant from any funding agencies

Competing Interests

The authors have no conflict of interest to declare

Ethics Approval and Consent to Participate

Not applicable

Consent for Publication

Not applicable

Authors' Contributions

Gemechis Waktole designed the study, retrieved the data, analyzed the data and wrote the manuscript. Cho Donghee supervised and edited manuscript. Both authors reviewed and approved for publication

Availability of Data and Materials

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

Acknowledgements

Not applicable

References

1. Auburn, S., & Barry, A. E. (2017). Dissecting malaria biology and epidemiology using population genetics and genomics. International journal for parasitology, 47(2-3), 77-85.
2. Bailey, T. L., & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers.
3. Bailey, T. L., Williams, N., Misleh, C., & Li, W. W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic acids research, 34(suppl_2), W369-W373.
4. Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., ... & Noble, W. S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic acids research, 37(suppl_2), W202-W208.
5. World Health Organization. (2015). World health statistics 2015. World Health Organization.
6. Gupta, S., Stamatoyannopoulos, J. A., Bailey, T. L., & Noble, W. S. (2007). Quantifying similarity between motifs. Genome biology, 8(2), 1-9.
7. Dinka, H., & Milkesa, A. (2020). Unfolding SARS-CoV-2 viral genome to understand its gene expression regulation. Infection, Genetics and Evolution, 84, 104386.
8. McCann, J. C. (2015). The historical ecology of malaria in Ethiopia: Deposing the spirits. Ohio University Press.
9. Lenhard, B., Sandelin, A., & Carninci, P. (2012). Metazoan promoters: emerging characteristics and insights into transcriptional regulation. Nature Reviews Genetics, 13(4), 233-245.
10. Lu, F., Culleton, R., Zhang, M., Ramaprasad, A., von Seidlein, L., Zhou, H., & Cao, J. (2017). Emergence of indigenous artemisinin-resistant Plasmodium falciparum in Africa. New England Journal of Medicine, 376(10), 991-993.
11. Reese, M. G. (2001). Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome. Computers & chemistry, 26(1), 51-56.
12. Sandelin, A., Alkema, W., Engström, P., Wasserman, W. W., & Lenhard, B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic acids research, 32(suppl_1), D91-D94.
13. Takai, D., & Jones, P. A. (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proceedings of the national academy of sciences, 99(6), 3740-3745.
14. Takamiya, T., Hosobuchi, S., Noguchi, T., Paterson, A. H., Iijima, H., Murakami, Y., & Okuizumi, H. (2009). The Application of Restriction Landmark Genome Scanning Method for Surveillance of Non-Mendelian Inheritance in Hybrids. Comparative and functional Genomics, 2009.
15. Tun, K. M., Imwong, M., Lwin, K. M., Win, A. A., Hlaing, T. M., Hlaing, T., & Woodrow, C. J. (2015). Spread of artemisinin-resistant Plasmodium falciparum in Myanmar: a cross-sectional survey of the K13 molecular marker. The Lancet infectious diseases, 15(4), 415-421.
16. World Health Organization. (2019). Compendium of WHO malaria guidance: prevention, diagnosis, treatment, surveillance and elimination (No. WHO/CDS/GMP/2019.03). World Health Organization.
17. World Health Organization. (2016). World health statistics 2016: monitoring health for the SDGs sustainable development goals. World Health Organization.
18. World Health Organization. (2018). World Malaria Report Geneva WHO p 24.
19. World Health Organization. (2016). World malaria report 2015. World Health Organization.
20. World Health Organization. (2017). False-negative RDT results and implications of new reports of P. falciparum histidine-rich protein 2/3 gene deletions (No. WHO/HTM/GMP/2017.18). World Health Organization.
21. Wang Z, Dalkilic M, Kim S. Guiding motif discovery by iterative pattern refinement. In Proceedings of the 2004 ACM symposium on Applied computing 2004 Mar 14 (pp. 162-166).
22. Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. Genes & development, 25(10), 1010-1022.
23. Sujuan, Y., Asaithambi, A., & Liu, Y. (2008). CpGIF: an algorithm for the identification of CpG islands. Bio information, 2(8), 335.